Applied Mathematics Research eXpress Advance Access published August 5, 2015

C. Bayer *et al.* (2015) "Computational Error Estimates for Molecular Dynamics,"
Applied Mathematics Research eXpress, abv007, 89 pages.
doi:10.1093/amrx/abv007

Computational Error Estimates for Born–Oppenheimer Molecular Dynamics with Nearly Crossing Potential Surfaces

Christian Bayer[1], Håkon Hoel[2], Ashraful Kadir[3], Petr Plecháč[4], Mattias Sandberg[3], and Anders Szepessy[3]

[1]Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstrasse 39, 10117 Berlin, Germany, [2]Department of Mathematics, University of Oslo, Postboks 1053 Blindern, 0316 Oslo, Norway, [3]Department of Mathematics, Kungl. Tekniska Högskolan, 100 44 Stockholm, Sweden, and [4]Department of Mathematical Sciences, University of Delaware, Newark, DE 19716, USA

Correspondence to be sent to: e-mail: szepessy@kth.se

The difference of the values of observables for the time-independent Schrödinger equation, with matrix-valued potentials, and the values of observables for ab initio Born–Oppenheimer molecular dynamics, of the ground state, depends on the probability to be in excited states, and the electron/nuclei mass ratio. The paper first proves an error estimate (depending on the electron/nuclei mass ratio and the probability to be in excited states) for this difference of microcanonical observables, assuming that molecular dynamics space-time averages converge, with a rate related to the maximal Lyapunov exponent. The error estimate is uniform in the number of particles and the analysis does not assume a uniform lower bound on the spectral gap of the electron operator and consequently the probability to be in excited states can be large. A numerical method to determine the probability to be in excited states is then presented, based on Ehrenfest molecular dynamics, and stability analysis of a perturbed eigenvalue problem.

1 Motivation for Error Estimates in ab initio Molecular Dynamics

Molecular dynamics is a computational method to study molecular systems in materials science, chemistry, and molecular biology. The simulations are used, for example, in

Received July 30, 2014; Revised April 28, 2015; Accepted July 7, 2015

designing and understanding new materials or for determining biochemical reactions in drug design [14].

The wide popularity of molecular dynamics simulations relies on the fact that in many cases it agrees very well with experiments. Indeed when we have experimental data it is often possible to verify the correctness of the method by comparing with experiments at certain parameter regimes. However, if we want the simulation to predict something that has no comparing experiment, we need a mathematical estimate of the accuracy of the computation. In the case of molecular systems with few particles, such studies are made by directly solving the Schrödinger equation. A fundamental and still open question in classical molecular dynamics simulations is how to verify the accuracy computationally, that is, when the solution of the Schrödinger equation is not a computational alternative.

The aim of this paper is to determine quantitative error estimates for molecular dynamics, including the case with nearly crossing electron potential surfaces that can yield large errors, without solving the Schrödinger equation but by combining mathematical stability analysis of eigenvalue problems with quantitative numerical Ehrenfest molecular dynamics computations of perturbations. Having molecular dynamics error estimates opens, for instance, the possibility of systematically evaluating which density functionals or empirical force fields are good approximations and under what conditions the approximation properties hold. Computations with such error estimates could also give improved understanding of when quantum effects are important and when they are not, in particular in cases when the Schrödinger equation is too computationally complex to solve.

The first step to check the accuracy of a molecular dynamics simulation is to know what to compare with. Here we compare the value of any *observable* $g : \mathbb{R}^{3N} \to \mathbb{R}$, of nuclei positions X, for the *time-independent Schrödinger* eigenvalue equation, $\hat{H}\Phi = E\Phi$, with the corresponding molecular dynamics observable, defined by a Hamiltonian $\hat{H}$, the wave function $\Phi : \mathbb{R}^{3N} \to \mathbb{C}^J$, and the eigenvalue $E \in \mathbb{R}$. The approximation error we study is therefore the microcanonical setting

$$\frac{\int_{\mathbb{R}^{3N}} \langle \Phi(X), g(X)\Phi(X) \rangle \, \mathrm{d}X}{\int_{\mathbb{R}^{3N}} \langle \Phi(X), \Phi(X) \rangle \, \mathrm{d}X} - \lim_{T\to\infty} \frac{1}{T} \int_0^T g(X_t) \, \mathrm{d}t, \tag{1}$$

based on a molecular dynamics path X_t, with total the energy equal to the Schrödinger eigenvalue E. The inner product $\langle \cdot, \cdot \rangle$ is defined in $\mathbb{C}^J$, where $\Phi \in \mathbb{C}^J$ corresponds to J discrete (electron) states. The observable can for instance be the local potential energy, used in [47] to determine phase field partial differential equations from molecular

dynamics simulations. For the sake of simplicity, we first explain the basic idea for observables depending only on position variables X. The general case of position and momentum dependent observables is treated in Section 2 using the Weyl quantization. The time-independent Schrödinger equation has a remarkable property of accurately predicting experiments in combination with no unknown data, thereby forming the foundation of computational chemistry. However, the drawback is the high-dimensional solution space for nuclei-electron systems with several particles, restricting numerical solution to small molecules. In this paper, we study the *time-independent* setting of the Schrödinger equation as the reference. The proposed approach has the advantage of not requiring any initial data as input, on the other hand, an assumption on convergence rates of time averages of molecular dynamics observables is needed.

The second step to check the accuracy is to derive error estimates. We have three types of error: time discretization error, sampling error, and modeling error. The time discretization error comes from approximating the differential equation for molecular dynamics with a numerical method, based on replacing time derivatives with difference quotients, and time steps Δt. The sampling error is due to truncating the infinite T and using a finite value of T in the integral in (1). The modeling error (also called coarse-graining error) originates from eliminating the electrons in the Schrödinger nuclei-electron system and replacing the nuclei dynamics with their classical paths; this approximation error was first analyzed by Born and Oppenheimer in their seminal paper [4].

The time discretization and truncation error components are in some sense simpler to handle by comparing simulations with different choices of Δt and T, although it can, of course, be difficult to know that the behavior does not change with even smaller Δt and larger T due to metastability, see [7, 31]. The modeling error is more difficult to check since a direct approach would require to solve the Schrödinger equation. Currently, the Schrödinger partial differential equation can only be solved with few particles, therefore it is not an option to solve the Schrödinger equation in general. The reason for using molecular dynamics is precisely to avoid solving the Schrödinger equation. Consequently, the modeling error requires mathematical error analysis. Egorov's theorem, cf. [5, 38], provides such error estimates and is used also here. However, in the literature there seems to be no error analysis that is precise, simple, and constructive enough so that a molecular dynamics simulation can use it in practice to assess the modeling error also in the case when the electron operator has eigenvalues (i.e., potential surfaces) which are not well separated and may cause large

modeling error. For instance, crossing eigenvalue surfaces can form so-called conical intersections, which provide the mechanism for many chemical reactions, see [44, 48]. If the excited electron energy levels are well separated from the ground state energy, molecular dynamics based on the ground state energy is a good approximation, as first analyzed by Born and Oppenheimer [4]. On the other hand, in a quantum system with two electron energy levels that are not well separated, the dynamic transition probability from one state to the other state for a moving particle can be substantial, as determined by Landau [28] and Zener [50], and generalized to a similar dynamic case in two space dimensions in [27]. We denote the transition probability to go from the ground state to excited states for a time-dependent problem "dynamic transition probability", as in the Landau–Zener model, to distinguish it from the "probability to be in excited states" for the time-independent Schrödinger eigenvalue problem.

Our alternative error analysis presented here relates computable dynamic transition probabilities to the probability to be in excited states for the time-independent Schrödinger equation and is developed with the aim to give a different point of view that could help to construct algorithms that estimate the modeling error in molecular dynamics computations. Our analysis differs from previous ones by combining analytical estimates with computations and it is based on three main steps:

- analyzing the time-independent Schrödinger equation as the reference model, including excited electron states with near crossing potential surfaces and the accuracy of observables as a function of the probability to be in excited states (in Section 2),
- studying stochastic molecular dynamics, constrained to the manifold of constant energy, based on small stochastic perturbations of the standard Born–Oppenheimer Hamiltonian dynamics,
- using stability analysis of a perturbed eigenvalue problem to estimate the probability to be in excited electron states, based on perturbations related to (Landau–Zener like) dynamic transition probabilities (in Section 3),
- applying Ehrenfest molecular dynamics to computationally estimate the dynamic transition probabilities (in Section 4).

The estimation method is tested on one- and two-dimensional problems but to conclude how useful the method is for realistic chemistry problems will require more work.

The next section introduces the Schrödinger and molecular dynamics models and ends with a more detailed outline.

1.1 The Schrödinger and molecular dynamics models with an outline of the main results

In deriving the approximation of the solutions to the full Schrödinger equation, the heavy particles are often treated within classical mechanics, that is, by defining the evolution of their positions and momenta by equations of motions of classical mechanics. Therefore, we denote by $X_t : [0,\infty) \to \mathbb{R}^{3N}$ the time-dependent function of the nuclei positions, with time derivatives denoted by

$$\dot{X}_t = \frac{\mathrm{d}X_t}{\mathrm{d}t}, \quad \ddot{X}_t = \frac{\mathrm{d}^2 X_t}{\mathrm{d}t^2}.$$

We denote the Euclidean scalar product on $\mathbb{R}^{3N}$ by

$$X \cdot Y = \sum_{n=1}^{N} \sum_{i=1}^{3} X_i^n Y_i^n.$$

Furthermore, we use the notation $\nabla\psi(X) = (\nabla_{X^1}\psi(X), \ldots, \nabla_{X^N}\psi(X))$, and as customary $\nabla_{X^n}\psi = (\partial_{X_1^n}\psi, \partial_{X_2^n}\psi, \partial_{X_3^n}\psi)$. The notation $\psi(X) = \mathcal{O}(M^{-\alpha})$ is also used for complex-valued functions, meaning that $|\psi(X)| = \mathcal{O}(M^{-\alpha})$ holds uniformly in X for $\psi(X) \in \mathbb{C}^J$.

The *time-independent Schrödinger* equation

$$\hat{H}(X)\Phi(X) = E\Phi(X) \tag{2}$$

models many-body (nuclei-electron) quantum systems and is obtained from minimization of the energy in the solution space of wave functions, see [2, 6, 30, 39, 40, 43]. It is an eigenvalue problem for the energy $E \in \mathbb{R}$ of the system, described by the Hamiltonian operator $\hat{H}(X)$

$$\hat{H}(X) = V(X) - \frac{1}{2}M^{-1}I\sum_{n=1}^{N}\Delta_{X^n}, \tag{3}$$

where I denotes the $J \times J$ identity matrix, and by the wave functions, $\Phi(X)$, belonging to a set of permissible electron states which we assume, for simplicity, to be finite, $\Phi(X) \in \mathbb{C}^J$. We use a normalized solution $\int_{\mathbb{R}^{3N}} |\Phi(X)|^2 \,\mathrm{d}X = 1$ and without loss of generality we assume that all nuclei have the same mass M. If this is not the case, we can introduce new coordinates $M_1^{1/2}\tilde{X}^k = M_k^{1/2}X^k$, which transform the Hamiltonian to the form we want $V(M_1^{1/2}M^{-1/2}\tilde{X}) - (2M_1)^{-1}I\sum_{k=1}^{N}\Delta_{\tilde{X}^k}$. In computational chemistry, the electron operator Hamiltonian $\bar{V}$, is independent of M and it is precisely determined by the sum of the kinetic energy of electrons and the Coulomb interaction between nuclei and electrons. The wave function is then in a subspace of $L^2(\mathrm{d}X\mathrm{d}x)$, with electron coordinates x, see [30]. By introducing a finite-dimensional basis $\{\phi_i\}_{i=1}^n$ of the electron solution space, minimization of the energy $\int_{\mathbb{R}^{3N}}\int_{\mathbb{R}^n}\Phi(X,x)^*(\bar{V}(X) - \frac{1}{2}M^{-1}I\sum_{n=1}^{N}\Delta_{X^n})\Phi(X,x)\,\mathrm{d}x\,\mathrm{d}X$ under

the constraint $\int_{\mathbb{R}^{3N}}\int_{\mathbb{R}^n}\Phi(X,x)\Phi(X,x)^*\,\mathrm{d}x\,\mathrm{d}X=1$ leads to the eigenvalue problem (2) for the solution $\Phi(X,x)=\sum_{j=1}^{J}\Phi_j(X)\phi_j(x)$, with the $J\times J$ matrix $V:=S^{-1}\bar{\bar{V}}$ defined by the matrix components $S_{ij}:=\int_{\mathbb{R}^n}\phi_i(x)\phi_j(x)^*\,\mathrm{d}x$ and $\bar{\bar{V}}_{ij}:=\int_{\mathbb{R}^n}\bar{V}(X,x)\phi_i(x)\phi_j(x)^*\,\mathrm{d}x$. To be able to compute with several electrons, that is, $n/3\gg 1$, the electron solution space is often simplified in the form of Hartree–Fock or density functional approximations, which also lead to eigenvalue problems which now are nonlinear, see [30]. We assume that the electron operator $V(X)$ is linear and self-adjoint, in a finite-dimensional complex-valued Euclidian space $\mathbb{C}^J$ (for simplicity), with the usual inner product $\langle\cdot,\cdot\rangle$, and acts as a matrix multiplication. An essential feature of the partial differential equation (2) is the high computational complexity of finding the solution in a subset of the Sobolev space $H^1(\mathbb{R}^{3N})^J$. The mass of the nuclei, M, are much greater than 1 (electron mass).

In contrast to the Schrödinger equation, a *molecular dynamics* model of N nuclei $\tilde{X}:[0,\infty)\to\mathbb{R}^{3N}$, with a given potential $V_p:\mathbb{R}^{3N}\to\mathbb{R}$, can be computationally studied for large N by solving the ordinary differential equations $M\mathrm{d}^2\tilde{X}_\tau/\mathrm{d}\tau^2=-\nabla V_p(\tilde{X}_\tau)$ in the fast time scale. We will use the slow time scale $t=M^{-1/2}\tau$ with positions $X_t=\tilde{X}_\tau$ and scaled momenta $P_t:=\dot{X}_t$ satisfying

$$\ddot{X}_t=-\nabla V_p(X_t). \tag{4}$$

In the slow time scale, the nuclei move $\mathcal{O}(1)$ in unit time, since $\dot{P}_t=-\nabla V_p(X_t)$. This computational and conceptual simplification motivates the study to determine the potential and its implied accuracy compared with the Schrödinger equation, as started already in the 1920s with the Born–Oppenheimer approximation [4]. The purpose of our work is to contribute to the current understanding of such derivations by showing convergence rates under new assumptions. The precise aim in this paper is to estimate the error

$$\frac{\int_{\mathbb{R}^{3N}}\langle\Phi(X),g(X)\Phi(X)\rangle\,\mathrm{d}X}{\int_{\mathbb{R}^{3N}}\langle\Phi(X),\Phi(X)\rangle\,\mathrm{d}X}-\lim_{T\to\infty}\frac{1}{T}\int_0^T g(X_t)\,\mathrm{d}t \tag{5}$$

for scalar smooth observables g of the time-indepedent Schrödinger equation (2) approximated by the corresponding molecular dynamics observable $\lim_{T\to\infty}T^{-1}\int_0^T g(X_t)\,\mathrm{d}t$, which is computationally cheaper to evaluate in the case with several nuclei; Section 2 includes also the general case with position and momentum dependent observables.

The main step to relate the Schrödinger wave function and the molecular dynamics solution is the so-called zero-order Born–Oppenheimer approximation, where X_t solves the classical ab initio molecular dynamics (4) with the potential $V_p:\mathbb{R}^{3N}\to\mathbb{R}$ determined as an eigenvalue of the electron Hamiltonian $V(X)$ for a given nuclei position

X. That is, $V_p(X) = \lambda_0(X)$ and

$$V(X)\Psi_0(X) = \lambda_0(X)\Psi_0(X), \tag{6}$$

for a normalized electron eigenvector $\Psi_0(X) \in \mathbb{C}^J$ (here the ground state). The initial data (X_0, P_0) is chosen to have its energy equal to the eigenvalue E, that is, $|P_0|^2/2 + \lambda_0(X_0)$. The Born–Oppenheimer expansion [4] is an approximation of the solution of the time-independent Schrödinger equation which is shown in [17, 25] to solve the time-independent Schrödinger equation approximately when the electron operator V have eigenvalues $\lambda_j(X)$ that are well separated, satisfying

$$\min_X(\lambda_1(X) - \lambda_0(X)) > c \text{ for a positive constant } c \text{ independent of } M. \tag{7}$$

This expansion, analyzed by the methods of multiple scales, pseudodifferential operators, and spectral analysis, for example, in [13, 16, 17, 25], can be used to study the approximation error (5). The work [16, 17, 25, 33] prove asymptotic expansions for the eigenfunction. In the literature, one can also find precise statements on the error for the setting of the time-dependent Schrödinger equation, for example, in [5, 26, 33, 35], which in some sense is easier since the stability issue is more subtle in the eigenvalue setting. The aim of our work is to present a method to analyze this stability of the eigenvalue problem, using the dynamic transition probability estimated from Ehrenfest dynamics simulations as quantitative input, without assuming an M-uniform lower bound c on the spectral gap.

We present error estimates for molecular dynamics approximation of time-independent observables in quantum mechanics, valid also when the eigenvalues λ_n of the electron operator V may nearly intersect and thereby increase the probability of being in excited states. For the Schrödinger solution Φ to (2), with the electron ground state Ψ_0 satisfying (6), the probability p_{ex} to be in excited states is

$$p_{\text{ex}} := \int_{\mathbb{R}^{3N}} \langle \Phi^\perp(X), \Phi^\perp(X) \rangle \, \mathrm{d}X, \tag{8}$$

where $\Phi^\perp(X) := \Phi(X) - \langle \Psi_0(X), \Phi(X) \rangle \Psi_0(X)$. The derivation of the approximation error is divided into three steps which defines the outline of the paper.

(I) The first step, in Section 2, proves that observables based on the time-independent Schödinger equation can be approximated by stochastic molecular dynamics on the constant energy manifold. The convergence rate depends on the nuclei/electron mass ratio M and the probability p_{ex} to

be in excited states. The proofs use Egorov's theorem and assume that the expected value of space-time averages of the molecular dynamics observable converge in distributional sense with a rate related to the maximal Lyapunov exponent.

(II) The second step, in Section 3, presents a numerical method to determine the probability to be in excited states from stability analysis of a perturbed eigenvalue problem. The perturbation is determined from a time-dependent excitation problem related to the Landau–Zener model (with p_d denoting the dynamic transition probability to go to the excited state). Resonances of the eigenvalue problem then determines the larger probability p_{ex} (to be in excited states for the Schrödinger eigenvalue problem) from the perturbations p_d in the dynamic problem.

(III) The final step, in Section 4, on numerical results, illustrate that the dynamic transition probability p_d can be determined from Ehrenfest molecular dynamics simulations applied to two simple model problems.

Section 5 proves a lemma on the regularity of the expected value of the molecular dynamics solution and establishes in three other lemmas estimates on remainder terms in Moyal expansions used in the theorems. Appendix 1 formulates WKB-solutions in the case with caustics.

The main ingredient in step I is Egorov's theorem which has been used extensively to study semiclassical limits both in the time-dependent and time-independent case, cf. [5, 38, 42]. We derive two estimates: Theorem 2.1 proves an estimate of a weighted difference of the Schrödinger and molecular dynamics observables and Theorem 2.2 derives an estimate of the difference of the Schrödinger and molecular dynamics observables, without the weight. The idea to assume ergodic classical dynamics to prove convergence of observables, in [29, 41], initiated the activity on quantum ergodicity [49] and related results on semiclassical limits for scalar potentials, [20], and matrix potentials with distinct eigenvalues [3]. The work [42] includes a recent review of semiclassical limit results in the case of well-separated electron eigenvalue surfaces. In this case with well-separated electron eigenvalues, satisfying (7), the probability to be in the excited state is proved to be small $p_{\text{ex}} = \mathcal{O}(M^{-1})$. We want to include the case when molecular dynamics simulations become inaccurate in practice. For an actual molecular dynamics simulation, the mass M is given and the electron potential V does not depend in M, so a positive spectral gap c does not depend on M. However, if the mass is not large enough the molecular dynamics approximation may be inaccurate. Our analysis is asymptotic for large M and to include the situation with a small gap for a certain mass

we therefore consider a model when the spectral gap

$$\min_X(\lambda_1(X) - \lambda_0(X)) =: \delta_M > 0, \tag{9}$$

can be close to zero, that is, when it is not uniformly bounded from below with respect to M, and then p_{ex} can be large as seen, for example, in the Landau–Zener probability (50) and Figure 4. There are four main new ingredients in our theorems compared with previous studies. It is well known that the standard proof of Egorov's theorem, in the time-dependent case, combined with an assumption on the molecular dynamics convergence rate toward its ergodic limit yields an error estimate for the molecular dynamics observables when compared with the Schrödinger observables. The precise form of the assumption on the convergence to the ergodic limit is important for the conclusion of the approximation error compared with Schrödinger observables: (1) for instance, assuming an algebraic convergence rate $\mathcal{O}(T^{-1/2})$, for molecular dynamics time averages of length T, would only give logarithmic error estimates $\mathcal{O}(1/\log^{1/2} M)$ (also in the case with a uniform spectral gap); (2) it is difficult to verify ergodicity for Hamiltonian dynamics, both theoretically and numerically. The *first new ingredient* here is to use stochastic dynamics constrained on the constant energy surface instead of Hamiltonian dynamics. Introducing small stochastic perturbations on the constant energy manifold have the advantage to guarantee ergodicity with respect to the microcanonical ensemble and to provide observables with higher regularity. Our *second ingredient* is to use a more precise convergence rate assumption based on space-time averages, which can be tested numerically (see Figure 10) and yields a better error estimate $\mathcal{O}(M^{-\bar{\gamma}})$, for a positive $\bar{\gamma}$, The *third ingredient* is to prove uniform convergence rates in the number of particles N by careful use of localized mollifiers. Previous convergence proofs require for instance, for a certain constant C, the quantity $\max_{z\in\mathbb{R}^{6N},|\alpha|\leq C\,N} |\partial_z^\alpha g(z)| M^{-1}$ to be small but typically this is a large number for systems with many particles unless the mass ratio M is very large and increasing with the number of particles N, see [20] and [51, Theorem 15.4]. The *fourth ingredient* is that we include nonuniformly spectral gap bounds (9): Theorems 2.1 and 2.2 prove the error estimate in step I in terms of the probability p_{ex} to be in excited states, which is not a typical a priori information. Consequently, the computational approximation of p_{ex} in steps II and III is an important complement. Although to evaluate the practical value of the suggested method to estimate the computational error in molecular dynamics these two steps need to be tested on more realistic chemistry problems.

The alternative to consider molecular dynamics at constant temperature using the canonical ensemble is important in practice and it also has the advantage of proved

ergodic limits. The canonical ensemble is in a sense the average of the microcanonical ensemble over all energies weighted with the Gibbs density. Therefore, the results here are useful for ongoing analysis in the constant temperature setting.

2 Molecular Dynamics Approximation of Schrödinger Observables for Matrix-valued Potentials

For every mass $M \gg 1$, we consider $L^2(\mathbb{R}^{3N})$ wave functions $\Phi : \mathbb{R}^{3N} \to \mathbb{R}^J$ and eigenvalues $E \in \mathbb{R}$, solving the Schrödinger eigenvalue equation

$$\hat{H}\Phi = E\Phi, \tag{10}$$

where

$$\hat{H} := -\frac{1}{2M}\Delta_X + V(X). \tag{11}$$

Here, the symmetric smooth matrix-valued potential $V : \mathbb{R}^{3N} \to \mathbb{R}^{J^2}$ has (electron) eigenvalues $\{\lambda_n\}_{n=0}^{J-1}$, ordered increasingly, with normalized eigenvectors $\{\Psi_n\}_{n=0}^{J-1}$ satisfying

$$V(X)\Psi_n(X) = \lambda_n(X)\Psi_n(X)$$

and the spectral gap condition (9). The eigenvalues $\{\lambda_n,\ n \geq 1\}$ may be degenerate but λ_0 is not. Both Φ and E depend on the mass M and we will only consider those eigenfunctions where the corresponding eigenvalue is bounded, that is, the Φ for which $E = \mathcal{O}(1)$ as $M \to \infty$. That is, we assume that for every $M \gg 1$ there are solutions $\Phi = \Phi_E \in L^2(\mathbb{R}^{3N})$ with eigenvalue $E = \mathcal{O}(1)$ to (10), as $M \to \infty$. On the other hand, the matrix V does not depend on M explicitly and consequently also the eigenvectors and eigenvalues Ψ_n and λ_n do not depend on M explicitly. However, to include the case of a small spectral gap, we assume that the positive spectral gap, δ, in (9) is not uniform in M and we study estimates depending on the two parameters M and δ. To handle ergodicity for the Born–Oppenheimer dynamics, we will below assume that the smallest eigenvalue $\lambda_0(X)$ is smooth and tends to infinity as $|X| \to \infty$. This assumption in fact also implies that the spectrum of $\hat{H}$ is discrete, see [9].

We consider a given smooth scalar observable $g : \mathbb{R}^{6N} \to \mathbb{R}$ in the Schwartz space constructed as

$$\begin{aligned} &g = \bar{g} * \phi_\eta, \\ &\sup_{z \in \mathbb{R}^{6N}} \sum_{|\beta| \leq 4} |\partial_z^\beta \bar{g}(z)|^2 \text{ is uniformly bounded with respect to } N, \end{aligned} \tag{12}$$

where for some $\eta > 4M^{-1/2}$ the mollifier $\phi_\eta : \mathbb{R}^{6N} \to \mathbb{R}$ is defined by $\phi_\eta(z) := (2\pi\eta)^{-3N} \mathrm{e}^{-|z|^2/(2\eta)}$ and $\bar{g} : \mathbb{R}^{6N} \to \mathbb{R}$ is smooth and compactly supported. We assume that each solution $\Phi = \Phi_E$ to (10) is normalized, that is, $\int_{\mathbb{R}^{3N}} \langle \Phi(X), \Phi(X) \rangle \, \mathrm{d}X = 1$, and define for each Φ the observable

$$g_S := \int_{\mathbb{R}^{3N}} \langle \Phi(X), \hat{g}\Phi(X) \rangle \, \mathrm{d}X, \tag{13}$$

based on the *Weyl quantization*, see [32],

$$\hat{g}\Phi(X) = \mathrm{Op}[g]\Phi(X) := (2\pi M^{-1/2})^{-3N} \int_{\mathbb{R}^{6N}} \mathrm{e}^{\mathrm{i}M^{1/2}(X-Y)\cdot P} g\left(\frac{X+Y}{2}, P\right) \Phi(Y) \, \mathrm{d}Y \, \mathrm{d}P,$$

using the inner products

$$\langle w, v \rangle := \sum_{j=1}^{J} w_j^* v_j \quad \text{for } w, v \in \mathbb{C}^J, \qquad X \cdot P := \sum_{n=1}^{3N} X^n P^n \quad \text{for } X, P \in \mathbb{R}^{3N}.$$

We note that in the case $g(X, P) = g(X)$ is a function of X only, then the Weyl quantization acts as a multiplication operator, that is, $\hat{g}\Phi(X) = g(X)\Phi(X)$. We also define the *Wigner transform*

$$W(X, P) := (2\pi M^{-1/2})^{-3N} \int_{\mathbb{R}^{3N}} \mathrm{e}^{\mathrm{i}M^{1/2}Y\cdot P} \left\langle \Phi\left(X + \frac{Y}{2}\right), \Phi\left(X - \frac{Y}{2}\right) \right\rangle \mathrm{d}Y \tag{14}$$

and

$$W_{jk}(X, P) := (2\pi M^{-1/2})^{-3N} \int_{\mathbb{R}^{3N}} \mathrm{e}^{\mathrm{i}M^{1/2}Y\cdot P} \Phi_j^*\left(X + \frac{Y}{2}\right) \Phi_k\left(X - \frac{Y}{2}\right) \mathrm{d}Y,$$

and note that $\Phi = \Phi_E$ is different for different eigenvalues E. The change of variables

$$X' = X + \frac{Y}{2}, \quad Y' = X - \frac{Y}{2}$$

shows that for any $g : \mathbb{R}^{6N} \to \mathbb{R}$ and $A : \mathbb{R}^{6N} \to \mathbb{R}^{J^2}$

$$\int_{\mathbb{R}^{3N}} \langle \Phi(X), \hat{g}\Phi(X) \rangle \, \mathrm{d}X = \int_{\mathbb{R}^{6N}} g(X, P) W(X, P) \, \mathrm{d}X \, \mathrm{d}P,$$
$$\int_{\mathbb{R}^{3N}} \langle \Phi(X), \hat{A}\Phi(X) \rangle \, \mathrm{d}X = \int_{\mathbb{R}^{6N}} \sum_{j=1}^{J} \sum_{k=1}^{J} A_{jk}(X, P) W_{jk}(X, P) \, \mathrm{d}X \, \mathrm{d}P.$$

Our estimates of remainder terms use the corresponding function $W^{(s)}$ defined for $s \in [0, 1]$ by

$$W^{(s)}(X, P) := (2\pi M^{-1/2})^{-3N} \int_{\mathbb{R}^{3N}} \mathrm{e}^{\mathrm{i}M^{1/2}Y\cdot P} \langle \Phi(X + sY), \Phi(X - (1-s)Y) \rangle \, \mathrm{d}Y, \tag{15}$$

which is the "Wigner"-function corresponding to the alternative quantization

$$(2\pi M^{-1/2})^{-3N} \int_{\mathbb{R}^{9N}} \langle \Phi(X), g(X + s(Y - X), P)\Phi(Y)\rangle \, \mathrm{e}^{\mathrm{i}M^{1/2}(X-Y)\cdot P} \, \mathrm{d}Y \, \mathrm{d}X \, \mathrm{d}P$$
$$= \int_{\mathbb{R}^{6N}} g(X, P) W^{(s)}(X, P) \, \mathrm{d}X \, \mathrm{d}P. \tag{16}$$

We will approximate the Schrödinger observable by dynamics related to ab initio molecular dynamics

$$\begin{aligned} \dot{X}_t &= P_t, \\ \dot{P}_t &= -\nabla_X \lambda_0(X_t), \end{aligned} \tag{17}$$

based on the Hamiltonian

$$H_0(X, P) := \frac{|P|^2}{2} + \lambda_0(X).$$

We assume that λ_0 is smooth and coercive in the sense that

$$\lim_{|X|\to+\infty} \lambda_0(X) = +\infty. \tag{18}$$

In order to obtain precise estimates of the approximation error which are uniform in N, we consider a mollified Hamiltonian

$$H_\eta(X, P) := |P|^2/2 + \lambda_\eta(X),$$
$$\lambda_\eta(X) := \lambda_0 * \phi_\eta(X),$$

where $\phi_\eta(X) = (2\pi\eta)^{-3N/2} \, \mathrm{e}^{-|X|^2/(2\eta)}$ is a standard mollifier on the small scale $\sqrt{\eta}$. To obtain good approximation of observables compared with the Schrödinger case, we use $\eta = 4M^{-1/2}$. To ensure ergodic molecular dynamics and still approximate the Born–Oppenheimer dynamics, we will compare the Schrödinger observable to the observable for stochastic dynamics restricted to the manifold

$$\Sigma_E := \{(X, P) \in \mathbb{R}^{6N} \mid H_\eta(X, P) = E\},$$

for the constant energy E equal to the Schrödinger eigenvalue, with a small stochastic perturbation of (17). Let $Z = (X, P) \in \mathbb{R}^{6N}$ denote the phase-space variable, then one example of such stochastic Stratonovich dynamics on Σ_E takes the form

$$\mathrm{d}Z_t = J\nabla H_\eta(Z_t) \, \mathrm{d}t + \sqrt{2\epsilon\tau}\mathbb{P}(Z_t) \circ \mathrm{d}\tilde{W}_t, \tag{19}$$

where $\epsilon < M^{-1/2}$ is a small positive parameter, $\tilde{W}$ is the standard Wiener process in $\mathbb{R}^{6N}$ with independent components, the skew symmetric matrix J is defined by $J\nabla_Z H_\eta(X,P) = (P, -\nabla_X \lambda_\eta(X))$, the parameter $\tau \sim 1$ is positive, and the projection $\mathbb{P}(Z)$ onto the tangent space of Σ_E at $Z \in \Sigma_E$ reads

$$\mathbb{P}(Z) := \mathrm{Id} - \hat{n}(Z) \otimes \hat{n}(Z)$$

with the normal $\hat{n}(Z) := \nabla H_\eta(Z)/|\nabla H_\eta(Z)|$ to Σ_E. We assume that $\nabla H_\eta|_{\Sigma_e} \neq 0$ for all $e \in \mathbb{R}$. Since also H_η is assumed to be smooth and coercive, the set Σ_e is a smooth compact codimension 1 manifold in $\mathbb{R}^{6N}$ for every $e \in \mathbb{R}$. The dynamics (19) is the projection of the Ito differential equation

$$\mathrm{d}z_t = (J\nabla H_\eta(z_t) - \epsilon \nabla H_\eta(z_t))\,\mathrm{d}t + \sqrt{2\epsilon\tau}\,\mathrm{d}\tilde{W}_t \tag{20}$$

to Σ_E and z_t has a unique equilibrium measure, which is the Gibbs measure

$$\mathrm{e}^{-H_\eta(z)/\tau}\,\mathrm{d}z \Big/ \int_{\mathbb{R}^{6N}} \mathrm{e}^{-H_\eta(z)/\tau}\,\mathrm{d}z.$$

The diffusion parameter ϵ can, for instance, also be a semipositive-definite constant diagonal matrix, so that (20) includes the Langevin equation. The work [12] proves that the projected dynamics (19) also is ergodic and the equilibrium measure is the micro-canonical measure

$$\mathrm{d}\nu(Z) := \left(\int_{\Sigma_E} \frac{\mathrm{d}\Sigma}{|\nabla H_\eta|}\right)^{-1} \frac{\mathrm{d}\Sigma(Z)}{|\nabla H_\eta(Z)|},$$

that is, $\lim_{T\to\infty} T^{-1}\int_0^T g(Z_t)\,\mathrm{d}t = \int_{\Sigma_E} g(z)\,\mathrm{d}\nu(z)$ where $\mathrm{d}\Sigma$ is the surface measure on Σ_E induced by the Lebesgue measure in $\mathbb{R}^{6N}$. Alternative ergodic dynamics on Σ_E sampling the microcanonical measure are presented in [12].

The Ito dynamics corresponding to (19) reads, see [8, 12],

$$\mathrm{d}Z_t = \underbrace{\mathbb{P}(Z_t)(J\nabla H_\eta(Z_t) - \epsilon\nabla H_\eta(Z_t))}_{=J\nabla H_\eta(Z_t)}\,\mathrm{d}t - \epsilon\tau\kappa(Z_t)\hat{n}(Z_t)\,\mathrm{d}t + \sqrt{2\epsilon\tau}\mathbb{P}(Z_t)\,\mathrm{d}\tilde{W}_t, \tag{21}$$

where $\kappa(z) := \mathrm{div}\,\hat{n}(z)$ is the mean curvature at z on Σ_E. The Kolmogorov equation for the expected value $u(t,z) := \mathbb{E}[g(\hat{Z}_T) \mid Z_t = z]$, for $t \le T$, of paths Z satisfying (19) becomes

$$\begin{aligned} \partial_t u(t,z) + \underbrace{(J\nabla H_\eta(z) - \epsilon\tau\kappa(z)\hat{n}(z))\cdot\nabla_z u(t,z) + \epsilon\tau\,\mathrm{trace}(\mathbb{P}(z)\nabla^2 u(t,z))}_{=Lu(t,z)} &= 0 \quad t < T, \\ u(T,\cdot) &= g, \end{aligned} \tag{22}$$

where $\nabla^2 u(t,Z)$ denotes the Hessian with respect to z, so that

$$\operatorname{trace}(\mathbb{P}(z)\nabla^2 u(t,z)) = \sum_{i,j} \mathbb{P}_{ij}(z)\partial_{z_i}\partial_{z_j} u(t,z) \quad \text{and} \quad \mathbb{P}_{ij} = \delta_{ij} - \hat{n}_i\hat{n}_j.$$

This Kolmogorov equation has an intrinsic definition based only on the coordinates on Σ_E, see [12]. The expected value u is smooth and the compact support of $\bar{g}$ implies that u decays rapidly as $|Z| \to \infty$, so that u is in the Schwartz class. The work [12] also proves an exponential convergence rate to equilibrium: let the initial data g be smooth in a neighborhood of Σ_E then there is a positive constant $\gamma_0 \sim \epsilon$ such that

$$\left\| u(t,\cdot) - \int_{\Sigma_E} u(T,Z)\,\mathrm{d}\nu(Z) \right\|_{L^2(\mathrm{d}\nu(\Sigma_E))} = \mathcal{O}(\mathrm{e}^{-\gamma_0(T-t)}), \quad t < T. \tag{23}$$

The upper and lower bounds in [21] on the convergence rate exponent γ_0 for Langevin dynamics in the full $\mathbb{R}^{6N}$ phase-space prove that the rate is bounded from below by a constant times ϵ and from above by a constant times $\log(1/\epsilon)$ where $\epsilon \ll 1$ is the damping factor, as in (20), for fixed positive temperature τ.

Let $S_{ts}(Z) := (Z_t \mid Z_s = Z)$, for $t > s$ be the stochastic flow of the dynamics (19), that is, the solution Z_t that starts in Z at time s, for a realization $\tilde{W}(\cdot,\omega)$ of the Wiener process. Then we have $u(t,Z) = \mathbb{E}[g \circ S_{Tt}(Z)]$. We write $S_t := S_{t0}$ so that $u(0,Z) = \mathbb{E}[g \circ S_T(Z)]$. This function satisfies the exponential growth

$$\sup_{Z\in\mathbb{R}^{6N}} \left(\sum_{|\beta|\le n} |\partial_{Z_0}^{\beta} \mathbb{E}[g \circ S_t(Z_0)]|^2 \right)^{1/2} \le \mathrm{e}^{\hat{C}(t+1)} \delta^{\min(0,-n+1)}, \quad n \le 4, \tag{24}$$

where $\hat{C}$ is independent of M, δ, and ϵ, as derived in Lemma 5.1 under assumption (97) including weak near crossing of eigenvalues.

Define the molecular dynamics microcanonical observable

$$g_{\mathrm{MD}}(Z_0) := \lim_{T\to\infty} \frac{1}{T} \int_0^T g(S_t(Z_0))\,\mathrm{d}t = \int_{\Sigma_E} g(z)\,\mathrm{d}\nu(z), \tag{25}$$

depending on the initial energy $E = H_\eta(Z_0)$. Our main assumption on the dynamics is that we assume that for some $\epsilon < M^{-1/2}$ there is a positive constants γ, independent of ϵ and M, such that

$$\begin{aligned} &\int_{\mathbb{R}^{6N}} \frac{2}{T} \int_{T/2}^{T} (\mathbb{E}[g \circ S_s](Z_0) - g_{\mathrm{MD}}(Z_0))\,\mathrm{d}s W(Z_0)\,\mathrm{d}Z_0 \\ &\quad \le \mathrm{e}^{-\gamma T} + \mathcal{O}(M^{-1}) \quad \text{for } T \le \log M/(\hat{C} + \gamma). \end{aligned} \tag{26}$$

The constant $\hat{C}$ depends on the observable g so that an observable related to a macroscopic quantity may have better convergence rate in the case of a large system.

From (23), we know that there is a constant $\gamma \sim \epsilon$ satisfying (26). Assumption (26) says that the decay with respect to time T of

$$\frac{2}{T}\int_0^{T/2}\left(u(t,Z_0)-\int_{\mathbb{R}^{6N}}u(T,\cdot)\,\mathrm{d}\nu\right)\mathrm{d}t$$

integrated over the Wigner distribution is exponential with the decay rate $\gamma \sim 1$ (for some $\epsilon < M^{-1}$ and $T < \log M/(\hat{C}+\gamma)$) plus possibly an additional constant small term M^{-1} independent of t. The ergodic assumption

$$\lim_{T\to\infty}\frac{1}{T}\int_0^T g\circ S_{t0}(Z_0)\,\mathrm{d}t=\int_{\mathbb{R}^{6N}}g(z)\,\mathrm{d}\nu(z),$$

for the Hamiltonian dynamics with $\epsilon = 0$ is used in [3, 20, 29, 41] to prove convergence of Schrödinger observables. This ergodic assumption is difficult to verify numerically, since a numerical approximation of the dynamics will always perturb the dynamics and it is theoretically unclear how this perturbation effects the dynamics over infinite time, see [46]. To find a method to prove ergodicity for general Hamiltonian systems also remains a challenge, cf. [46]. An advantage with assumption (26) is that it can be tested numerically for some initial points Z_0, see Figure 10, since the time discretization error can be made small compared with the bound $\mathcal{O}(M^{-1})$, by taking a small time step $\Delta t < M^{-1}$, and the simulation time $T < \log M/\hat{C}$ is finite. Not all dynamics satisfy (26). For instance, there exists billiard dynamics in "stadium" domains in $\mathbb{R}^2$ that is proved to be ergodic for almost all initial data but is nonergodic for some data with corresponding concentrated Schrödinger eigenfunctions, see [19]. Small perturbations of these data with nonergodic dynamics will need very long time to reach asymptotically ergodic behavior and consequently also the stochastic regularization above is likely to have γ in (26) depending on ϵ so that our assumption would not hold.

Having introduced the necessary notation and terminology, we are now ready to present an estimate for the weighted difference between the Schrödinger and stochastic molecular dynamics observables that is expressed in terms of the spectral gap, the mass, and the excitation probability. Theorem 2.2, presented in the end of this section, then estimates the difference $g_S - g_{\mathrm{MD}}$ of the observables without the weight.

Theorem 2.1. For every $M \gg 1$, consider the Schrödinger eigenvalue problem (10) and (11) with a solution $\Phi \in L^2(\mathbb{R}^{3N})$, for a corresponding eigenvalue $E = \mathcal{O}(1)$, and assume:

(i) the observable g satisfies (12) and the potential V is smooth with $\sup_{X\in\mathbb{R}^{3N}} \sum_{1\le|\sigma|\le 3} |\partial_X^\sigma V(X)|^2$ bounded, uniformly in N, where $|\cdot|$ is the matrix 2-norm in $\mathbb{C}^J$,

(ii) the molecular dynamics limit g_{MD} in (25) satisfies the weak convergence rate (26),

(iii) the electron eigenvector Ψ_0 and eigenvalue λ_η are smooth and have bounded derivatives satisfying, for some $\delta = \delta_M > 0$ defined in (9) (which may depend on M) and any $|\beta| \le 2$, and for any $|\alpha| \le 4$,

$$|\partial_X^\beta \Psi_0(X)| = \mathcal{O}(\delta^{-|\beta|}),$$
$$|\partial_X^\alpha \lambda_\eta(X)| = \mathcal{O}(\delta^{-|\alpha|+1}),$$

(iv) the Hamiltonian H_η is smooth, satisfies the coercivity (18) and $\nabla H_\eta|_{\Sigma_e} \neq 0$ for all $e \in \mathbb{R}$,

(v) the exponential bound (24) holds with the constant $\hat{C}$ uniformly bounded in N.

Let the probability to be in excited states be denoted by $p_{\mathrm{ex}} = p_{\mathrm{ex}}(\Phi)$, as defined in (8), and define the Wigner transform $W = W(\Phi)$ by (14). Then, for $M^{-1/2}\delta^{-2} = \mathcal{O}(1)$, the molecular dynamics observable g_{MD} approximates the Schrödinger observable g_{S}, defined in (13), with the weighted error estimate

$$\left| \int_{\mathbb{R}^{6N}} (g_{\mathrm{S}} - g_{\mathrm{MD}}(Z_0)) W(Z_0)\, \mathrm{d}Z_0 \right| \le C\,(\mathrm{e}^{\hat{C}T}(M^{-1}\delta^{-4} + C_N\epsilon\delta^{-1} + \eta\delta^{-2} + p_{\mathrm{ex}}^{1/2}) + \mathrm{e}^{-\gamma T}), \tag{27}$$

for any $T > 0$, and by choosing T such that $\mathrm{e}^{-(\hat{C}+\gamma)T} = (M^{-1/2}\delta^{-2} + p_{\mathrm{ex}}^{1/2})$, $\eta = 4M^{-1/2}$, and ϵ sufficiently small

$$\left| \int_{\mathbb{R}^{6N}} (g_{\mathrm{S}} - g_{\mathrm{MD}}(Z_0)) W(Z_0)\, \mathrm{d}Z_0 \right| \le C\,(M^{-1/2}\delta^{-2} + p_{\mathrm{ex}}^{1/2})^{\frac{\gamma}{\hat{C}+\gamma}}, \tag{28}$$

for a constant C, independent of $M \gg 1$, $p_{\mathrm{ex}} \ll 1$, δ, γ, N, and $\hat{C}$. □

To prove this theorem, we modify the proof of Egorov's theorem [38] to include stochastic dynamics and the assumption (26) on the space-time convergence rate and to replace the uniform spectral gap bound for matrix-valued perturbation potentials $V(X)$, in [3, 42], with Assumption (iii). Assumption (iii) with $\delta > 0$ excludes a conical

intersection but avoided crossings are allowed with a spectral gap in (9) that is not uniformly bounded from below as a function of M. The Landau–Zener probability (50) in Section 3.2 (combined with (52) and (82)) illustrate for instance that $p_{\rm ex}$ can be close to one for an avoided crossing, with a small spectral gap $\min_X(\lambda_1(X)-\lambda_0(X)) = \delta = M^{-1/4}$. That $p_{\rm ex}^{1/2}$ can be of the same order as $M^{-1}\delta^{-4}$ in the estimate (27) motivates the study of the approximation error as a function of $p_{\rm ex}$ and M. To have $p_{\rm ex}$ large seems to require smaller gaps $\delta \lesssim M^{1/4}$, see Section 3. It would be desirable to have analytic estimates of $p_{\rm ex}$, since it is determined by M and V. Such estimates are known, $p_{\rm ex} = o(1)$ as $M\to\infty$, for the case when spectral gap $\delta > c > 0$ is uniformly bounded from below while in the case with no uniform spectral gap the required smoothness of projection to the ground state is assumed, see [45]. The lack of analytic estimates of $p_{\rm ex}$ is our motivation for using $p_{\rm ex}$ in addition to M as parameter in the theorems and presenting the computational method to determine $p_{\rm ex}$ in this work.

We note that stochastic dynamics has a regularizing effect since $\mathbb{E}[g\circ S_s]$, with $\epsilon > 0$, is in general more regular than $g\circ S_s$, with $\epsilon = 0$. In fact, we do not explicitly need the stochastic flow S: the proof only uses the deterministic value $u(t,\cdot) = \mathbb{E}[g\circ S_{Tt}(\cdot)]$ and its flow $u(t,\cdot) =: \bar{S}_{tT}u(T,\cdot)$, for $t<T$, with the generator $\partial_t u(t,\cdot) = -Lu(t,\cdot)$. The exponential bound in assumption (v) is proved in Lemma 5.1 under an assumption that allows weak near crossing of eigenvalues away from regions where P vanishes.

Proof. Consider the solution operator $\mathrm{e}^{\mathrm{i}M^{1/2}t\hat{H}}$ of the time-dependent Schrödinger equation

$$\mathrm{i}M^{-1/2}\partial_t\Psi(t,X) = \hat{H}\Psi(t,X)$$

defined by $\Psi(t,\cdot) = \mathrm{e}^{-\mathrm{i}M^{1/2}t\hat{H}}\Psi(0,\cdot)$. We will compare the observable

$$\begin{aligned} g_{\rm S} &= \int_{\mathbb{R}^{3N}} \langle \mathrm{e}^{-\mathrm{i}M^{1/2}t\hat{H}}\Phi(X), \hat{g}\,\mathrm{e}^{-\mathrm{i}M^{1/2}t\hat{H}}\Phi(X)\rangle\,\mathrm{d}X \\ &= \int_{\mathbb{R}^{3N}} \langle \Phi(X), \mathrm{e}^{\mathrm{i}M^{1/2}t\hat{H}}\hat{g}\,\mathrm{e}^{-\mathrm{i}M^{1/2}t\hat{H}}\Phi(X)\rangle\,\mathrm{d}X \end{aligned}$$

to the expected value

$$\frac{2}{T}\int_{T/2}^{T}\int_{\mathbb{R}^{6N}} \mathbb{E}[g\circ S_t(Z_0)]W(Z_0)\,\mathrm{d}Z_0\,\mathrm{d}t = \frac{2}{T}\int_{T/2}^{T}\int_{\mathbb{R}^{3N}} \langle \Phi(X_0), \widehat{\mathbb{E}[g\circ S_t]}\Phi(X_0)\rangle\,\mathrm{d}X_0\,\mathrm{d}t.$$

By construction, we have that $S_t(Z_0)$ is in the compact manifold $\Sigma_{H_\eta(Z_0)}$. The coercivity (18) of λ_0 and the decay of g in phase-space imply therefore that $g(S_t(Z_0))$ decays sufficiently fast for large $|Z_0|$. The integrals above are therefore well defined. Since V is smooth, elliptic regularity implies that the solution Φ to the Schrödinger eigenvalue

problem also is smooth. Let $\bar{u}$ be the solution to

$$\begin{aligned} \partial_t \bar{u}(t,z) - L\bar{u}(t,z) &= 0, \quad t > 0, \\ \bar{u}(0) &= \bar{g}, \end{aligned} \tag{29}$$

then $\mathbb{E}[\bar{g} \circ S_t(Z_0)] = \bar{u}(t, Z_0)$ and $g = \bar{g} * \phi_\eta$. We have, using the commutator $[\hat{A}, \hat{B}] := \hat{A}\hat{B} - \hat{B}\hat{A}$ and that $\hat{H}\Phi = E\Phi$,

$$\begin{aligned} &\int_{\mathbb{R}^{6N}} \bar{u}(t,\cdot) * \phi_\eta(Z) W(Z)\,\mathrm{d}Z - g_S \\ &= \int_{\mathbb{R}^{3N}} \left\langle \Phi(X), \widehat{\bar{u}(t,\cdot) * \phi_\eta} - \mathrm{e}^{\mathrm{i}M^{1/2}t\hat{H}} \widehat{\bar{u}(0,\cdot) * \phi_\eta}\, \mathrm{e}^{-\mathrm{i}M^{1/2}t\hat{H}} \Phi(X) \right\rangle \mathrm{d}X \\ &= \int_{\mathbb{R}^{3N}} \left\langle \Phi(X), \int_0^t \frac{\partial}{\partial s} (\mathrm{e}^{\mathrm{i}M^{1/2}(t-s)\hat{H}} \widehat{\bar{u}(s,\cdot) * \phi_\eta}\, \mathrm{e}^{-\mathrm{i}M^{1/2}(t-s)\hat{H}})\,\mathrm{d}s \Phi(X) \right\rangle \mathrm{d}X \\ &= \int_{\mathbb{R}^{3N}} \left\langle \Phi(X), \int_0^t \mathrm{e}^{\mathrm{i}M^{1/2}(t-s)\hat{H}} \left(\frac{\partial}{\partial s} \mathrm{Op}[\bar{u}(s,\cdot) * \phi_\eta] - \mathrm{i}M^{1/2} [\hat{H}, \widehat{\bar{u}(s,\cdot) * \phi_\eta}] \right) \right. \\ &\qquad \left. \times \mathrm{e}^{-\mathrm{i}M^{1/2}(t-s)\hat{H}} \Phi(X) \right\rangle \mathrm{d}s\,\mathrm{d}X \\ &= \int_0^t \int_{\mathbb{R}^{3N}} \left\langle \mathrm{e}^{\mathrm{i}M^{1/2}(t-s)\hat{H}} \Phi(X), \mathrm{Op}[((P \cdot \nabla_X - \nabla\lambda_\eta \cdot \nabla_P)\bar{u}(s,\cdot)) * \phi_\eta(X,P)] \right. \\ &\qquad - \mathrm{i}M^{1/2} \left[\hat{H}, \widehat{\bar{u}(s,\cdot) * \phi_\eta} \right] + \mathrm{Op}\left[\left(-\frac{\epsilon\tau\kappa}{|\nabla_Z H_\eta|} \nabla_Z H_\eta \cdot \nabla_Z \bar{u}(s,\cdot) * \phi_\eta \right. \right. \\ &\qquad \left. \left. + \epsilon\tau\, \mathrm{trace}(\mathbb{P}\nabla_Z^2)\bar{u}(s,\cdot) \right) * \phi_\eta(Z) \right] \left. \times \mathrm{e}^{-\mathrm{i}M^{1/2}(t-s)\hat{H}} \Phi(X) \right\rangle \mathrm{d}X\,\mathrm{d}s \\ &= \int_0^t \int_{\mathbb{R}^{3N}} \left\langle \mathrm{e}^{\mathrm{i}M^{1/2}(t-s)E} \Phi(X), \mathrm{Op}[((P \cdot \nabla_X - \nabla\lambda_\eta(X) \cdot \nabla_P)\bar{u}(s,\cdot)) * \phi_\eta(Z)] \right. \\ &\qquad + \mathcal{O}(\epsilon) - \mathrm{i}M^{1/2} \left[\hat{H}, \widehat{\bar{u}(s,\cdot) * \phi_\eta} \right] \left. \mathrm{e}^{-\mathrm{i}M^{1/2}(t-s)E} \Phi(X) \right\rangle \mathrm{d}X\,\mathrm{d}s \\ &= \int_0^t \int_{\mathbb{R}^{3N}} \left\langle \Phi(X), (\mathrm{Op}[((P \cdot \nabla_X - \nabla\lambda_\eta(X) \cdot \nabla_P)\bar{u}(s,\cdot)) * \phi_\eta(Z)] \right. \\ &\qquad \left. + \mathcal{O}(\epsilon) - \mathrm{i}M^{1/2}[\hat{H}, \mathrm{Op}[\bar{u}(s,\cdot) * \phi_\eta]]) \Phi(X) \right\rangle \mathrm{d}X\,\mathrm{d}s. \end{aligned} \tag{30}$$

In the fourth equality, we have used (29) to conclude that

$$\begin{aligned} \frac{\partial}{\partial s} \mathrm{Op}[\bar{u}(s,\cdot) * \phi_\eta(Z)] &= \mathrm{Op}\left[\frac{\partial}{\partial s} \bar{u}(s,\cdot) * \phi_\eta(Z) \right] = \mathrm{Op}[(L\bar{u}(s,\cdot)) * \phi_\eta(Z)] \\ &= \mathrm{Op}\left[\left(\left(J\nabla_Z H_\eta - \frac{\epsilon\tau\kappa}{|\nabla_Z H_\eta|} \nabla_Z H_\eta \right) \cdot \nabla_Z + \epsilon\tau\, \mathrm{trace}(\mathbb{P}\nabla_Z^2) \right) \bar{u}(s,\cdot) * \phi_\eta \right] \\ &= \mathrm{Op}[(J\nabla_Z H_\eta \cdot \nabla_Z \bar{u}(s,\cdot)) * \phi_\eta(Z)] + \mathcal{O}(C_N \epsilon \delta^{-1}\, \mathrm{e}^{\hat{C}t}). \end{aligned}$$

Here the estimate

$$\left|\int_{\mathbb{R}^{3N}}\left\langle\Phi,\mathrm{Op}\left[\left(-\frac{\epsilon\tau\kappa}{|\nabla_Z H_\eta|}\nabla_Z H_\eta\cdot\nabla_Z+\epsilon\tau\,\mathrm{trace}(P\nabla_Z^2)\bar{u}(s,\cdot)*\phi_\eta\right)\right]\Phi\right\rangle\mathrm{d}X\right|\leq C_N\epsilon\delta^{-1}\,\mathrm{e}^{\hat{C}t} \tag{31}$$

follows by Lemma 5.4 and we note that the term is negligible small by choosing ϵ sufficiently small, although the constant C_N typically is large as it depends on order N derivatives of the symbol.

To estimate terms in the error representation (30) is now the remaining four steps in the proof.

1. The Moyal expansion for the commutator of two Weyl operators in [5, 25] and Lemma 5.5 is stated for scalars and here we apply Lemma 5.5 to each matrix component jk of $\hat{H}$ to obtain

$$\begin{aligned}&\mathrm{i}M^{1/2}[\hat{H},\widehat{\bar{u}(s,\cdot)*\phi_\eta}]_{jk}\Phi_k=\mathrm{i}M^{1/2}[\hat{H}_{jk},\widehat{\bar{u}(s,\cdot)*\phi_\eta}]\Phi_k\\&=(\mathrm{Op}[(\nabla_P H_{jk}(X,P)\cdot\nabla_X-\nabla_X H_{jk}(X,P)\cdot\nabla_P)(\bar{u}(s,\cdot)*\phi_\eta)(X,P)]+\hat{R}_M)\Phi_k(X)\end{aligned} \tag{32}$$

for the matrix components

$$H_{jk}(X,P)=\frac{|P|^2}{2}\delta_{jk}+V_{jk}(X),$$

where the remainder takes the form, see [5] and Section 5,

$$R_M:=\mathrm{Op}\left[\sum_{n=1}^{m}2M^{-n}(2\mathrm{i})^{-n}\sum_{|\alpha|=2n+1}\frac{(-1)^{|\alpha|}}{\alpha!}\partial_X^\alpha V_{jk}(X)(\partial_P^\alpha\bar{u}(s,\cdot)*\phi_\eta)(X,P)+M^{-(m+1)}r_m\right] \tag{33}$$

and r_m is smooth. We will use $\hat{R}_M$ for $m=1$. Lemma 5.5 in Section 5 shows that

$$\int_{\mathbb{R}^{3N}}\langle\Phi,\hat{R}_M\Phi\rangle=\mathcal{O}(\mathrm{e}^{\hat{C}T}\delta^{-2}M^{-1}). \tag{34}$$

2. The normalization property $1=\int_{\mathbb{R}^{6N}}W(Z)\,\mathrm{d}Z$ implies that

$$g_S=\int_{\mathbb{R}^{6N}}g_S W(Z)\,\mathrm{d}Z$$

and we obtain by (30) and (32)

$$\begin{aligned}
&\int_{\mathbb{R}^{6N}} \left(g_S - \frac{2}{T} \int_{T/2}^{T} \bar{u}(t,\cdot) * \phi_\eta(Z)\, \mathrm{d}t \right) W(Z)\, \mathrm{d}Z \\
&\quad = \frac{2}{T} \int_{T/2}^{T} \left(g_S - \int_{\mathbb{R}^{6N}} \bar{u}(t,\cdot) * \phi_\eta(Z) \right) W(Z)\, \mathrm{d}Z\, \mathrm{d}t \\
&\quad = \frac{2}{T} \int_{T/2}^{T} \int_0^t \int_{\mathbb{R}^{3N}} \langle \Phi(X), \mathrm{Op}[((\nabla \lambda_\eta(\cdot) I - \nabla_X V(X)) \cdot \nabla_P \bar{u}(s,\cdot)) * \phi_\eta(X,P)] \Phi(X) \rangle\, \mathrm{d}X\, \mathrm{d}s\, \mathrm{d}t \\
&\qquad + \underbrace{\frac{2}{T} \int_{T/2}^{T} \int_0^t \int_{\mathbb{R}^{3N}} \langle \Phi(X), (\mathcal{O}(\epsilon) + \hat{R}_M) \Phi(X) \rangle\, \mathrm{d}X\, \mathrm{d}s\, \mathrm{d}t}_{= \mathcal{O}(\mathrm{e}^{\hat{C}T}(M^{-1}\delta^{-2} + \epsilon C_N \delta^{-1}))},
\end{aligned} \tag{35}$$

where we use the bounds (34) and (31) in last term.

3. We have the desired quantity $\int_{\mathbb{R}^{6N}} (g_S - g_{\mathrm{MD}}) W\, \mathrm{d}X\, \mathrm{d}P$ in the left-hand side above by adding and subtracting the molecular dynamics observable g_{MD}, to the second term in the left-hand side, and using its convergence rate $\mathrm{e}^{-\gamma T} + M^{-1}$ from (26) as follows:

$$\begin{aligned}
&\int_{\mathbb{R}^{6N}} \frac{2}{T} \int_{T/2}^{T} \bar{u}(t,\cdot) * \phi_\eta(Z)\, \mathrm{d}t W(Z)\, \mathrm{d}Z \\
&\quad = \int_{\mathbb{R}^{6N}} g_{\mathrm{MD}}(Z) W(Z)\, \mathrm{d}Z + \int_{\mathbb{R}^{6N}} \frac{2}{T} \int_{T/2}^{T} (\bar{u}(t,\cdot) * \phi_\eta(Z) - g_{\mathrm{MD}}(Z))\, \mathrm{d}t W(Z)\, \mathrm{d}Z \\
&\quad = \int_{\mathbb{R}^{6N}} g_{\mathrm{MD}}(Z) W(Z)\, \mathrm{d}Z + \mathcal{O}(\mathrm{e}^{-\gamma T} + M^{-1}),
\end{aligned}$$

which yields the last term in the error bound (27).

4. It remains to estimate the first term in the right-hand side of (35). The rule for the composition of Weyl quantizations, cf. [32] and Lemma 5.5 in Section 5,

$$\hat{A}\hat{B} = \widehat{AB} - \mathrm{i} M^{-1/2} \mathrm{Op}\left[\underbrace{\{A,B\}}_{:= \nabla_P A \cdot \nabla_X B - \nabla_X A \cdot \nabla_P B} \right] + \mathcal{O}(M^{-1}), \tag{36}$$

for smooth scalar functions $A, B : \mathbb{R}^{6N} \to \mathbb{R}$. The first term in the right-hand side of (35) can be written as follows:

$$\int_{\mathbb{R}^{3N}} \langle \Phi, \mathrm{Op}[((\nabla \lambda_\eta(\cdot) I - \nabla_X V(X)) \cdot \nabla_P \bar{u}(s,\cdot)) * \phi_\eta(X,P)] \Phi \rangle\, \mathrm{d}X$$

and to estimate this term we use (8) and write $\Phi = \Phi^{\perp} + \Pi_0\Phi$, where $\|\Phi^{\perp}\|_{L^2(\mathrm{d}X)} = p_{\mathrm{ex}}^{1/2}$ and $\Pi_0 w := \langle w, \Psi_0\rangle\Psi_0$, for $w \in \mathbb{R}^J$, defines the projection onto the ground state. The integral terms including the factor $\Phi^{\perp}$ are bounded using Lemma 5.5

$$\int_{\mathbb{R}^{3N}} \langle \Phi^{\perp}, \mathrm{Op}[-(\nabla\lambda_\eta(\cdot)I - \nabla_X V(X)) \cdot \nabla_P \bar{u}(s,\cdot) * \phi_\eta(X,P)]\Phi\rangle \,\mathrm{d}X$$
$$+ \int_{\mathbb{R}^{3N}} \langle \Phi, \mathrm{Op}[-(\nabla\lambda_\eta(\cdot)I - \nabla_X V(X)) \cdot \nabla_P \bar{u}(s,\cdot) * \phi_\eta(X,P)]\Phi^{\perp}\rangle \,\mathrm{d}X = \mathcal{O}(\mathrm{e}^{\hat{C}T} p_{\mathrm{ex}}^{1/2}).$$

The main term related to the consistency of the approximation can be written as follows:

$$\begin{aligned}
&\mathrm{Op}[-(\nabla\lambda_\eta(\cdot)I - \nabla_X V(X)) \cdot \nabla_P \bar{u}(s,\cdot) * \phi_\eta(X,P)] \\
&\quad = \mathrm{Op}[-(\nabla\lambda_\eta(X)I - \nabla_X V(X)) \cdot \nabla_P \bar{u}(s,\cdot) * \phi_\eta(X,P)] \\
&\qquad + \mathrm{Op}[-(\nabla\lambda_\eta(\cdot) - \nabla\lambda_\eta(X)) \cdot \nabla_P \bar{u}(s,\cdot) * \phi_\eta(X,P)],
\end{aligned}$$

where Lemma 5.3 shows that the last term is bounded by

$$\sup_{(X,P)\in\mathbb{R}^{6N}} \underbrace{|(\nabla\lambda_\eta(X) - \nabla\lambda_\eta(\cdot)) \cdot \nabla_P \bar{u}(s,\cdot) * \phi_\eta(X,P)|}_{=\mathcal{O}(\mathrm{e}^{\hat{C}s}\eta\delta^{-2})} + \mathcal{O}(\eta\,\mathrm{e}^{\hat{C}s}\delta^{-2}) = \mathcal{O}(\mathrm{e}^{\hat{C}s}\eta\delta^{-2})$$

since the symbol satisfies

$$\begin{aligned}
&\int_{\mathbb{R}^{6N}} (\nabla\lambda_\eta(X) - \nabla\lambda_\eta(X - X')) \cdot \nabla_P \bar{u}(s, X - X', P - P')\phi_\eta(X', P') \,\mathrm{d}X' \,\mathrm{d}P' \\
&= \int_{\mathbb{R}^{3N}} \sum_{j=1}^{3N} \partial_{X_j} \nabla\lambda_\eta(X) \cdot \nabla_P \bar{u}(s, X, P - P') \underbrace{\int_{\mathbb{R}^{3N}} X'_j \phi_\eta(X', P') \,\mathrm{d}X'}_{=0} \,\mathrm{d}P' \\
&\quad + \int_{\mathbb{R}^{6N}} \sum_{j,k=1}^{3N} \int_0^1 (1-t)(\partial_{X_j X_k} \nabla\lambda_\eta(X - tX') \cdot \nabla_P \bar{u}(s, X - tX', P - P') \\
&\quad + 2\partial_{X_j} \nabla\lambda_\eta(X - tX') \cdot \partial_{X_k} \nabla_P \bar{u}(s, X - tX', P - P') \\
&\quad + \partial_{X_j}(\nabla\lambda_\eta(X) - \nabla\lambda_\eta(X - tX')) \cdot \partial_{X_k} \nabla_P \bar{u}(s, X - tX', P - P'))X'_j X'_k \phi_\eta(X', P') \,\mathrm{d}X' \,\mathrm{d}P' \,\mathrm{d}t \\
&\leq \Bigg\| \sum_{j,k=1}^{3N} \int_0^1 (1-t)(|\partial_{X_j X_k} \nabla\lambda_\eta(X - tX') \cdot \nabla_P \bar{u}(s, X - tX', P - P')| \\
&\quad + 2|\partial_{X_j} \nabla\lambda_\eta(X - tX') \cdot \partial_{X_k} \nabla_P \bar{u}(s, X - tX', P - P')| \\
&\quad + |\partial_{X_j}(\nabla\lambda_\eta(X) - \nabla\lambda_\eta(X - tX')) \cdot \partial_{X_k} \nabla_P \bar{u}(s, X - tX', P - P')|) \,\mathrm{d}t \Bigg\|_{L^\infty} \\
&\quad \times \sup_{j,k} \|X'_j X'_k \phi_\eta(X', P')\|_{L^1} \,\mathrm{d}t.
\end{aligned}$$

The remaining term satisfies, using (36) and the notation $\check{V}(X) := V(X) - \lambda_\eta(X)$,

$$
\begin{aligned}
&\int_{\mathbb{R}^{3N}} \langle \Pi_0\Phi, \mathrm{Op}[\nabla_X \check{V}(X) \cdot \nabla_P \bar{u}(s,\cdot) * \phi_\eta(X,P)]\Pi_0\Phi\rangle \,\mathrm{d}X \\
&\quad= \int_{\mathbb{R}^{3N}} \langle \Phi, \Pi_0 \mathrm{Op}[\nabla_X \check{V}(X) \cdot \nabla_P \bar{u}(s,\cdot) * \phi_\eta(X,P)]\Pi_0\Phi\rangle \,\mathrm{d}X \\
&\quad= \int_{\mathbb{R}^{3N}} \Big\langle \Phi, \Pi_0 \frac{1}{2}(\nabla_X \check{V}(X) \cdot \mathrm{Op}[\nabla_P \bar{u}(s,\cdot) * \phi_\eta(X,P)] \\
&\qquad + \mathrm{Op}[\nabla_P \bar{u}(s,\cdot) * \phi_\eta(X,P)] \cdot \nabla_X \check{V}(X))\Pi_0\Phi \Big\rangle \,\mathrm{d}X + \mathcal{O}(\mathrm{e}^{\hat{C}T} M^{-1}\delta^{-4}) \\
&\quad= \int_{\mathbb{R}^{3N}} \langle \Phi, \underbrace{\Pi_0 \nabla_X \check{V}(X)\Pi_0}_{=\mathcal{O}(\eta)} \cdot \mathrm{Op}[\nabla_P \bar{u}(s,\cdot) * \phi_\eta]\Phi\rangle \mathrm{d}X \\
&\qquad + \int_{\mathbb{R}^{3N}} \langle \Phi, \Pi_0 \nabla_X \check{V}(X) \cdot (\mathrm{Op}[\nabla_P \bar{u}(s,\cdot) * \phi_\eta(X,P)] \underbrace{(\Pi_0 - \Pi_0 * \phi_\eta)}_{=\mathcal{O}(\eta\delta^{-2})} \Phi\rangle \,\mathrm{d}X \\
&\qquad - \int_{\mathbb{R}^{3N}} \langle \Phi, \Pi_0 \nabla_X \check{V}(X) \cdot \underbrace{(\Pi_0 - \Pi_0 * \phi_\eta)}_{=\mathcal{O}(\eta\delta^{-2})} \mathrm{Op}[\nabla_P \bar{u}(s,\cdot) * \phi_\eta(X,P)])\Phi\rangle \mathrm{d}X \\
&\qquad + \int_{\mathbb{R}^{3N}} \Big\langle \Phi, \Pi_0 \nabla_X \check{V}(X) \\
&\qquad \times \underbrace{(\mathrm{Op}[\nabla_P \bar{u}(s,\cdot) * \phi_\eta(X,P)]\Pi_0 * \phi_\eta - \Pi_0 * \phi_\eta \mathrm{Op}[\nabla_P \bar{u}(s,\cdot) * \phi_\eta(X,P)])}_{=\mathcal{O}(\mathrm{e}^{\hat{C}T} M^{-1}\delta^{-4})} \Phi \Big\rangle \,\mathrm{d}X, \qquad (37)
\end{aligned}
$$

where the first term is $\mathcal{O}(\eta)$ since $\check{V} = (V - \lambda_0) + \lambda_0 - \lambda_\eta$ and the gradient of the ground state condition $(V - \lambda_0)\Psi_0 = 0$ implies

$$
\langle \Psi_0, \nabla(V - \lambda_0)\Psi_0\rangle = -\langle \Psi_0, (V - \lambda_0)\nabla\Psi_0\rangle = -\langle (V - \lambda_0)\Psi_0, \nabla\Psi_0\rangle = 0,
$$

and the second term is estimated by Lemma 5.5. Therefore, the right-hand side in (35) yields the contribution $\mathcal{O}(\mathrm{e}^{\hat{C}T}(M^{-1}\delta^{-4} + p_{\mathrm{ex}}^{1/2}))$ to the error bound (27). ∎

The next result shows that the difference of the observables can be estimated without the weight W.

Theorem 2.2. Let $\bar{E} := \int_{\mathbb{R}^{6N}} H_\eta(Z) W(Z) \,\mathrm{d}Z$ and assume that

(i) the assumptions in Theorem 2.1 hold and

(ii) the function $\tilde{g}_{\mathrm{MD}} : \mathbb{R} \to \mathbb{R}$ is defined by

$$
g_{\mathrm{MD}}(Z) =: \tilde{g}_{\mathrm{MD}}(H_\eta(Z)) \quad \text{for } Z \in \mathbb{R}^{6N},
$$

then

$$\int_{\mathbb{R}^{6N}} g_{\mathrm{MD}}(Z)W(Z)\,\mathrm{d}Z = \tilde{g}_{\mathrm{MD}}(\bar{E}) + \mathcal{O}(M^{-1}\delta^{-4} + p_{\mathrm{ex}}^{1/2})$$

and

$$g_{\mathrm{S}} - \tilde{g}_{\mathrm{MD}}(\bar{E}) = \mathcal{O}((M^{-1/2}\delta^{-2} + p_{\mathrm{ex}}^{1/2})^{\frac{\gamma}{\bar{C}+\gamma}}),$$

where

$$\bar{E} = E + \mathcal{O}(p_{\mathrm{ex}} + \eta). \qquad \square$$

Proof. The function $\tilde{g}_{\mathrm{MD}}(E) = \int_{\Sigma_E} g(z)\,\mathrm{d}\nu(z) / \int_{\Sigma_E} \mathrm{d}\nu(z)$ can be written as the convolution

$$\tilde{g}_{\mathrm{MD}}(E) = \int_{\Sigma_E} g(z)\,\mathrm{d}\nu(z) = \int_{\mathbb{R}^{6N}} \int_{\Sigma_E} \bar{g}(z-z')\,\mathrm{d}\nu(z)\phi_\eta(z')\,\mathrm{d}z' \tag{38}$$

and as

$$\tilde{g}_{\mathrm{MD}}(E) = \frac{\int_{\mathbb{R}^{6N}} g(Z)\delta(H_\eta(Z)-E)\,\mathrm{d}Z}{\int_{\mathbb{R}^{6N}} \delta(H_\eta(Z)-E)\,\mathrm{d}Z}.$$

Using the assumption that $|\nabla_Z H_\eta(Z)||_{\Sigma_E}$ does not vanish, the function $\tilde{g}_{\mathrm{MD}}$ is smooth since we have for $Z = (X, P)$

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}E}\int_{\mathbb{R}^{6N}} g(Z)\delta(H_\eta(Z)-E)\,\mathrm{d}Z &= -\int_{\mathbb{R}^{6N}} g(Z)\delta'(H_\eta(Z)-E)\,\mathrm{d}Z \\
&= -\int_{\mathbb{R}^{6N}} \frac{g(Z)}{|\nabla H_\eta(Z)|}\frac{\nabla H_\eta(Z)}{|\nabla H_\eta(Z)|}\cdot\nabla H_\eta(Z)\delta'(H_\eta(Z)-E)\,\mathrm{d}Z \\
&= -\int_{\mathbb{R}^{6N}} \frac{g(Z)}{|\nabla H_\eta(Z)|}\frac{\nabla H_\eta(Z)}{|\nabla H_\eta(Z)|}\cdot\nabla_Z\delta(H_\eta(Z)-E)\,\mathrm{d}Z \\
&= \int_{\mathbb{R}^{6N}} \mathrm{div}\left(\frac{g(Z)\nabla H_\eta(Z)}{|\nabla H_\eta(Z)|^2}\right)\delta(H_\eta(Z)-E)\,\mathrm{d}Z
\end{aligned}$$

and similarly for higher order derivatives.

Taylor expansion implies

$$\tilde{g}_{\mathrm{MD}}(\tilde{E}) = \tilde{g}_{\mathrm{MD}}(\bar{E}) + \tilde{g}'_{\mathrm{MD}}(\bar{E})(\tilde{E}-\bar{E}) + \tfrac{1}{2}\tilde{g}''_{\mathrm{MD}}(\xi)(\tilde{E}-\bar{E})^2,$$

for some $\xi = \xi(\tilde{E})$ between $\tilde{E}$ and $\bar{E}$ satisfying $\tilde{g}''_{\mathrm{MD}}(\xi(\tilde{E})) := 2\int_0^1 \tilde{g}''_{\mathrm{MD}}(s\tilde{E} + (1-s)E)(1-s)\,\mathrm{d}s$, and integration with $\tilde{E} = H_\eta(Z)$ yields the moment relation

$$\begin{aligned}
\int_{\mathbb{R}^{6N}} g_{\mathrm{MD}}(Z)W(Z)\,\mathrm{d}Z &= \int_{\mathbb{R}^{6N}} \tilde{g}_{\mathrm{MD}}(H_\eta(Z))W(Z)\,\mathrm{d}Z \\
&= \tilde{g}_{\mathrm{MD}}(\bar{E})\underbrace{\int_{\mathbb{R}^{6N}} W(Z)\,\mathrm{d}Z}_{=1} + \tilde{g}'_{\mathrm{MD}}(\bar{E})\underbrace{\int_{\mathbb{R}^{6N}} (H_\eta(Z)-\bar{E})W(Z)\,\mathrm{d}Z}_{=0} \\
&\quad + \frac{1}{2}\int_{\mathbb{R}^{6N}} \tilde{g}''_{\mathrm{MD}}(\xi(H_\eta(Z)))(H_\eta(Z)-\bar{E})^2 W(Z)\,\mathrm{d}Z.
\end{aligned} \tag{39}$$

The definition of $\bar{E}$ shows the error estimate

$$\begin{aligned}\bar{E} &= \int_{\mathbb{R}^{6N}} H_\eta(Z) W(Z)\,\mathrm{d}Z = \int_{\mathbb{R}^{3N}} \langle \Phi(X), \hat{H}_\eta \Phi(X)\rangle\,\mathrm{d}X \\ &= \int_{\mathbb{R}^{3N}} \langle \Phi(X), \hat{H}\Phi(X)\rangle\,\mathrm{d}X - \int_{\mathbb{R}^{3N}} \langle \Phi(X), (V-\lambda_0)\Phi(X)\rangle\,\mathrm{d}X \\ &\quad - \int_{\mathbb{R}^{3N}} \langle \Phi(X), (\lambda_0-\lambda_\eta)\Phi(X)\rangle\,\mathrm{d}X \\ &= E + \mathcal{O}(p_{\mathrm{ex}}) + \mathcal{O}(\eta), \end{aligned} \tag{40}$$

since $\hat{H}_\eta = \hat{H} - \hat{\hat{V}} = \hat{H} - (V-\lambda_0) - (\lambda_0 - \lambda_\eta)$.

The rule for composition of Weyl quantizations, in (117) in Lemma 5.5, shows that

$$\hat{A}\hat{B} = \widehat{AB} - \mathrm{i}M^{-1/2}\mathrm{Op}\left[\underbrace{\{A,B\}}_{=:\nabla_P A\cdot\nabla_X B - \nabla_X A\cdot\nabla_P B}\right] + \mathcal{O}(M^{-1}\delta^{-4}),$$

for $A = H_\eta - \bar{E}$ and $B(Z) = \tilde{g}''_{\mathrm{MD}}(\xi(H_\eta(Z)))$, which implies

$$\mathrm{Op}[\tilde{g}''_{\mathrm{MD}}(\xi(H_\eta))(H_\eta-\bar{E})^2] = \mathrm{Op}[H_\eta-\bar{E}]\mathrm{Op}[\tilde{g}''_{\mathrm{MD}}(\xi(H_\eta))]\mathrm{Op}[H_\eta-\bar{E}] + \mathcal{O}(M^{-1}\delta^{-4}),$$

since we have

$$\begin{aligned}\{\tilde{g}''_{\mathrm{MD}}(\xi(H_\eta)), H_\eta - \bar{E}\} &= 0, \\ \{H_\eta - \bar{E}, H_\eta - \bar{E}\} &= 0.\end{aligned}$$

Therefore, the last term in (39) becomes

$$\begin{aligned}&\int_{\mathbb{R}^{6N}} \tilde{g}''_{\mathrm{MD}}(\xi(H_\eta(Z)))(H_\eta(Z)-\bar{E})^2 W(Z)\,\mathrm{d}Z + \mathcal{O}(M^{-1}\delta^{-4}) \\ &= \int_{\mathbb{R}^{3N}} \left\langle \widehat{(H_\eta-\bar{E})}\Phi(X), \mathrm{Op}[\tilde{g}''_{\mathrm{MD}}(\xi(H_\eta(Z)))]\widehat{(H_\eta-\bar{E})}\Phi(X)\right\rangle \mathrm{d}X \\ &= \int_{\mathbb{R}^{3N}} \left\langle (\hat{H}-\bar{E}-\check{V})\Phi(X), \mathrm{Op}[\tilde{g}''_{\mathrm{MD}}(\xi(H_\eta(Z)))](\hat{H}-\bar{E}-\check{V})\Phi(X)\right\rangle \mathrm{d}X \\ &= \int_{\mathbb{R}^{3N}} \left\langle (E-\bar{E}-\check{V})\Phi(X), \mathrm{Op}[\tilde{g}''_{\mathrm{MD}}(\xi(H_\eta(Z)))](E-\bar{E}-\check{V})\Phi(X)\right\rangle \mathrm{d}X \\ &= \int_{\mathbb{R}^{3N}} \left\langle \underbrace{(E-\bar{E})}_{=\mathcal{O}(p_{\mathrm{ex}}+\eta)}\Phi - \check{V}\Phi^{\perp}(X), \mathrm{Op}[\tilde{g}''_{\mathrm{MD}}(\xi(H_\eta(Z)))](E-\bar{E}-\check{V})\Phi(X)\right\rangle \mathrm{d}X \\ &= \mathcal{O}(p_{\mathrm{ex}}^{1/2} + \eta),\end{aligned}$$

which combined with (39) and Theorem 2.1 proves the theorem. ∎

3 Determining the Probability to be in Excited States

The purpose of this section is to describe a numerical method to approximate the probability p_{ex} to be in excited states for the Schrödinger eigenvalue problem, without solving a Schrödinger eigenvalue problem and instead analyze a certain eigenvalue problem with respect to perturbations related to the dynamic behavior for a time-dependent problem. This perturbation study is formal in the sense that precise conditions for its validity is not presented here. The perturbation model is first formulated in one spatial dimension and then in multiple dimensions in Section 3.3. The perturbation uses the WKB-method to formulate transitions and is presented in Section 3.1. The dynamic problem is related to the Landau–Zener model and Ehrenfest dynamics as described in Section 3.2.

3.1 Construction of WKB-solutions

To construct a numerical method for approximating the probability p_{ex} to be in excited states, without solving a Schrödinger equation, we will use a decomposition of the Schrödinger solution into WKB-functions (78). This section presents a construction of such WKB-functions.

The singular perturbation $-(2M)^{-1}\sum_k \Delta_{X^k}$ of the matrix-valued potential V in the Hamiltonian (3) introduces an additional small scale $M^{-1/2}$ of high frequency oscillations, as shown by a WKB-expansion, see [10, 23, 36]. We shall construct solutions to (2) in such a WKB-form

$$\Phi(X) = \phi(X)\, e^{iM^{1/2}\theta(X)}, \tag{41}$$

where the amplitude function $\phi : \mathbb{R}^{3N} \to \mathbb{C}^J$ is complex valued, the phase $\theta : \mathbb{R}^{3N} \to \mathbb{R}$ is real valued in the classically allowed region $\{X \in \mathbb{R}^{3N} \mid \lambda_0(X) \leq E\}$ (and purely imaginary elsewhere, see Remark 3.1), and the factor $M^{1/2}$ is introduced in order to have well-defined limits of ϕ and θ as $M \to \infty$. Note that it is trivially always possible to find functions ϕ and θ satisfying (41), even in the sense of a true equality. Of course, the ansatz only makes sense if ϕ and θ do not have strong oscillations for large M. The standard WKB-construction [10, 34] is based on a series expansion in powers of $M^{1/2}$, which solves the Schrödinger equation with arbitrarily high accuracy. Instead of an asymptotic solution, we introduce an actual local solution based on a time-dependent Schrödinger transport equation. This transport equation reduces to the formulation in [34].

3.1.1 *A WKB-solution*

The WKB-function (41) satisfies the Schrödinger equation (2) provided that

$$\begin{aligned} 0 &= (\hat{H} - E)\phi \; e^{iM^{1/2}\theta(X)} \\ &= \left(\left(\frac{1}{2}|\nabla\theta|^2 + V - E \right) \phi - \frac{1}{2M}\Delta\phi - \frac{i}{M^{1/2}} \left(\nabla\phi \cdot \nabla\theta + \frac{1}{2}\phi\Delta\theta \right) \right) e^{iM^{1/2}\theta(X)}. \end{aligned} \tag{42}$$

We shall see that only eigensolutions Φ that correspond to dynamics without caustics correspond to such a single WKB-mode, as for instance when the eigenvalue E is inside an electron eigenvalue gap. Solutions in the presence of caustics use a Fourier integral of such WKB-modes, see [34] and Appendix 1. The purpose of the phase function θ is to generate an accurate approximation in the limit as $M \to \infty$. A possible definition, see [34], is the *Hamilton–Jacobi equation*, also called the *eikonal equation*

$$\tfrac{1}{2}|\nabla\theta|^2 = E - \lambda_0. \tag{43}$$

The solution to the Hamilton–Jacobi eikonal equation can be constructed locally from the associated Hamiltonian system

$$\begin{aligned} \dot{X}_t &= P_t, \\ \dot{P}_t &= -\nabla\lambda_0(X_t), \end{aligned} \tag{44}$$

through the characteristics path (X_t, P_t) satisfying $\nabla\theta(X_t) =: P_t$. For the time being, we fix some $P_0 \in \mathbb{R}^{3N}$ and vary X_0 over a hyperplane (denoted by I) in $\mathbb{R}^{3N}$ *orthogonal to* P_0, see Figure 1. Hence, the trajectories X_t can (locally) cover the space. Later on, in Section 3.3 we shall argue that a *superposition* of such solutions actually allows us to reconstruct the true (local) solution of the Schrödinger equation. In particular, we will solve the system (44) not just once, but for all different X_0.

The amplitude function ϕ can be determined by requiring the ansatz (41) to be a solution to (42), which gives

$$\begin{aligned} 0 &= (\hat{H} - E)\phi \; e^{iM^{1/2}\theta(X)} \\ &= \Bigg(\underbrace{\left(\frac{1}{2}|\nabla\theta|^2 + \lambda_0 - E \right)}_{=0} \phi \\ &\quad - \frac{1}{2M}\Delta\phi + (V - \lambda_0)\phi - \frac{i}{M^{1/2}} \left(\nabla\phi \cdot \nabla\theta + \frac{1}{2}\phi\Delta\theta \right) \Bigg) e^{iM^{1/2}\theta(X)}. \end{aligned}$$

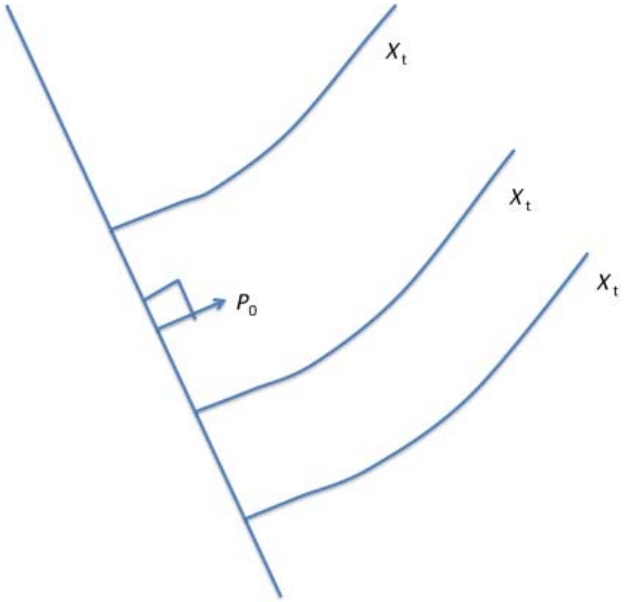

Fig. 1. Paths X_t with velocity $\dot{X}_t = P_t$ going out from the plane with fixed normal $P_0 = \text{constant}$.

Thus, by using (43), we have

$$-\frac{1}{2M}\Delta\phi + (V-\lambda_0)\phi - \frac{i}{M^{1/2}}\left(\nabla\phi\cdot\nabla\theta + \frac{1}{2}\phi\Delta\theta\right) = 0. \tag{45}$$

The usual method for determining ϕ from this so-called *transport equation* uses an asymptotic expansion $\phi \simeq \sum_{k=0}^{K} M^{-k/2}\phi_k$, see [17, 25]. An alternative is to write it as a Schrödinger equation, as in [34]: we apply the characteristics in (44) to write

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi(X_t) = \nabla\phi\cdot\dot{X}_t = \nabla\phi\cdot\nabla\theta,$$

and define the weight function G by

$$\frac{\mathrm{d}}{\mathrm{d}t}\log G_t = \frac{1}{2}\Delta\theta(X_t), \tag{46}$$

and the variable $\psi_t := \phi(X_t)G_t$. Then the transport equation (45) becomes a Schrödinger-type equation

$$iM^{-1/2}\dot{\psi}_t = (V-\lambda_0)\psi_t - \frac{G_t}{2M}\Delta_X\left(\frac{\psi_t}{G_t}\right) =: \tilde{V}\psi_t. \tag{47}$$

As (47) is (formally) a time-dependent Schrödinger equation, we need to impose initial conditions. Once more, we refer to Section 3.3 for a detailed discussion of this subtle issue.

The last step in the construction of the WKB system (θ, ϕ) is to patch together the different trajectories obtained in (44)–(47). A priori, X_t, P_t, G_t, and ψ_t are functions of (t, X_0), for $t \geq 0$ and $X_0 \in I$, where I was a hyperplane in $\mathbb{R}^{3N}$. By uniqueness of solutions to (44), the map $(t, X_0) \mapsto X_t$ is locally injective. Considering $-P_0$ in addition to P_0, we obtain that $(t, X_0) \mapsto X_t$ is locally invertible. Hence, we may also interpret the functions $P = P(t, X_0)$, $G = G(t, X_0)$, $\psi = \psi(t, X_0)$ as functions of the space variable alone. Abusing notation, we set $\nabla\theta(X) = P(t, X_0)$, $G(X) = G(t, X_0)$, $\psi(X) = \psi(t, X_0)$ for $X = X(t, X_0)$ in a neighborhood of I.

In conclusion, equations (43)–(47) determine the WKB-ansatz (41) to be a local solution to the Schrödinger equation (2) in the following sense. Assume that the Hamilton–Jacobi equation $\frac{|\nabla\theta(X)|^2}{2} + \lambda_0(X) = E$ has a C^2 solution $\theta : \mathcal{U} \to \mathbb{R}$ in a domain $\mathcal{U} \subseteq \mathbb{R}^{3N}$. Let $\dot{X}_t = \nabla\theta(X_t)$ and $P_t = \nabla\theta(X_t)$, then (X_t, P_t) solves the Hamiltonian system (44), for $t \in [0, t_*]$ such that $X_t \in \mathcal{U}$. Then

$$\Phi(X_t) = G^{-1}(X_t)\psi(X_t)\, \mathrm{e}^{iM^{1/2}\theta(X_t)} \tag{48}$$

solves (2) in $\mathcal{U}$, since both (43) and (45) are satisfied and consequently also (42). It is well known that Hamilton–Jacobi equations in general do not have global C^2 solutions, due to X-paths that collide and generate blow up in $\partial_{XX}\theta(X)$. However, if the domain is small enough and the data on the boundary is compatible (in the sense that $H_\eta(X, \nabla\theta(X)) = E$ on the boundary), noncharacteristic (in the sense that the normal derivative $\partial_n\theta(X) \neq 0$ on the boundary) and λ_0 is smooth, then the converse property holds, that is, the characteristics generate a local solution to the Hamilton–Jacobi equation, see [11]. Maslov's method to find a global asymptotic solution by patching together local solutions is described in [34].

Remark 3.1. In the classically forbidden region $\{X \in \mathbb{R}^{3N} \mid E < \lambda_0(X)\}$, we can apply the same WKB method by replacing the phase function θ by the imaginary phase $\mathrm{i}\theta$. The Eikonal equation becomes $-|P|^2/2 + \lambda_0(X) = E$, which then has a solution, and the transport equation becomes real valued

$$M^{-1/2}\dot{\psi}_t = -(V - \lambda_0)\psi_t + \frac{G_t}{2M}\Delta_X\left(\frac{\psi_t}{G_t}\right).$$

The matrix $V - \lambda_0$ is positive semidefinite, so that ψ remains bounded for bounded time and approaches the ground state Ψ_0. In the classical region, we instead use the oscillatory behavior to conclude that ψ tends to the ground state, see [1]. □

3.2 The Landau–Zener model, transition probabilities, and Ehrenfest dynamics

Note that the parameter t in (44)–(47) is just a numerical parameter, which, prima facie, has no connection with physical time. Indeed, we are working in a time-independent setting after all. On the other hand, if t *is* interpreted as time, (47) corresponds to a *time-dependent* Schrödinger equation, and we may ask for dynamic transition probabilities of the time-dependent model, denoted by $p_d(t) = p_d(X_t)$ and rigorously defined in (52). It turns out below that the time-dependent transition probabilities are easier to analyze than the time-independent ones, p_{ex}, at least under some simplifying assumptions. The link between the time-dependent transition probabilities p_d and the time-independent p_{ex} is not trivial and will be explored in detail in Section 3.3. To have some intuition, it might be helpful to think of $p_d(t)$ as a "local" excitation probability around X_t, with the idea that p_{ex}, in turn, is given as a time/space average of p_d.

We start our discussion by looking at a simple special case of (47), the Landau–Zener model, for which the first results on transition probabilities with crossing or nearly crossing electron potentials were obtained. It is given by

$$\mathrm{i}M^{-1/2}\dot{\phi}_t = \begin{bmatrix} P_0 t & \delta \\ \delta & -P_0 t \end{bmatrix} \phi_t \tag{49}$$

with a wave function $\phi : \mathbb{R} \to \mathbb{C}^2$, constant positive parameters (M, P_0, δ), and initial data $\lim_{t\to-\infty} \phi(t) = (1, 0)$. The transition probability

$$p_{\text{LZ}} := \mathrm{e}^{-\pi\delta^2 M^{1/2}/P_0} \tag{50}$$

is the so-called Landau–Zener probability, determined using Weber functions in [50], and illustrated in Figure 8. In this particular model, we note that $p_{\text{LZ}} = \lim_{t\to\infty} |\phi_2(t)|^2$.

In the context of (44)–(47), the Landau–Zener model can be seen as a special case of

$$\begin{aligned} \dot{X}_t &= P_t, \\ \dot{P}_t &= -\nabla\lambda_0(X_t), \\ \mathrm{i}M^{-1/2}\dot{\psi}_t &= \underbrace{(V(X_t) - \lambda_0(X_t))}_{=:\check{V}(X_t)}\psi_t - (2M)^{-1}G\Delta(\psi/G), \end{aligned} \tag{51}$$

which by the WKB-method (48) determines a Schrödinger WKB-solution Φ_Q locally and hence the transition probability

$$p_d(Q, X_t) := \langle\psi_t, \psi_t\rangle - |\langle\psi_t, \Psi_0(X_t)\rangle|^2, \tag{52}$$

using the initialization $\psi(X_t) = \Psi_0(X_t)$ for $X_t \in I$, where the inflow domain is given by

$$I = \{X : (X - Y) \cdot P = 0\} \tag{53}$$

for a given point $Q := (Y, P) \in \mathbb{R}^{6N}$ and the ground state Ψ_0 is defined in (6). By neglecting the small term $-G/(2M)\Delta(\psi/G)$ in the transport equation (51) and using the simplification with the potential $\lambda_0(X) = 0$ and

$$\check{V}(X) = \begin{bmatrix} X & \delta \\ \delta & -X \end{bmatrix}, \tag{54}$$

we obtain the Landau–Zener model. The Landau–Zener model was constructed to model and explain the dynamic transitions from the ground state to an excited state when the electron potential surfaces cross or nearly cross (with a minimal distance δ) and the eigenstates change rapidly (near $X = 0$ for small values of δ): the eigenvalues of $\check{V}$ in (54) are $\lambda_\pm(X) = \pm\sqrt{X^2 + \delta^2}$ and the eigenvectors

$$\Psi_\pm = \frac{1}{\sqrt{\delta^2 + (\lambda_\pm(X) - X)^2}} \begin{bmatrix} \delta \\ \lambda_\pm(X) - X \end{bmatrix}.$$

Remark 3.2. In the one-dimensional avoided crossing case (54) the electron eigenvectors satisfy $\Psi_+(X) = f(X/\delta)$ for a smooth function $f : \mathbb{R} \to \mathbb{R}^2$ with

$$\lim_{X\to\infty} \Psi_+(X) = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$
$$\lim_{X\to-\infty} \Psi_+(X) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

We see that

$$\|\partial_X \Psi_+\|_{L^\infty(\mathbb{R})} = \mathcal{O}(\delta^{-1}),$$

for $\delta \neq 0$. The other eigenvector Ψ_- satisfies the same bounds. Therefore, assumption (iii) in Theorem‘2.1 holds for this avoided crossing with any $\delta \neq 0$. □

3.2.1 *Ehrenfest dynamics*

As far as computations are concerned, the system (51) is still very demanding due to the Laplacian on the right-hand side. A further simplification mainly introduced for computational reasons (and used in the computational examples of this paper) is the

so-called Ehrenfest dynamics

$$\begin{aligned} \dot{X}_t &= P_t, \\ \dot{P}_t &= -\nabla\lambda_0(X_t) - \frac{\langle\psi_t, \nabla\check{V}(X_t)\psi_t\rangle}{\langle\psi,\psi\rangle}, \\ \mathrm{i}M^{-1/2}\dot{\psi}_t &= \check{V}(X_t)\psi_t, \end{aligned} \tag{55}$$

which is an approximate WKB solution, by neglecting the small term $-G/(2M)\Delta(\psi/G)$ in the transport equation and replacing the Eikonal equation with $|P|^2/2 + \lambda_0(X) + \langle\psi, \check{V}\psi\rangle/\langle\psi,\psi\rangle = E$, as in [1]. If we write the wave function in its real and imaginary parts $\psi = \psi^r + \mathrm{i}\psi^i$, and use initial data that satisfy $\langle\psi_0,\psi_0\rangle = 2M^{-1/2}$, then the Ehrenfest dynamics (55) is a Hamiltonian system with the Hamiltonian

$$H_E = |P|^2/2 + \lambda_0(X) + M^{1/2}\langle\psi, \check{V}(X)\psi\rangle/2, \tag{56}$$

using the primal variables X and ψ^r, the dual variables P and ψ^i. If the probability, $\langle\psi^\perp,\psi^\perp\rangle$, to be in the excited state is small (using the projection in (8)), the Ehrenfest Hamiltonian is $\mathcal{O}(\langle\psi^\perp,\psi^\perp\rangle)$ close to the Born–Oppenheimer Hamiltonian $|P|^2/2 + \lambda_0(X)$.

3.3 An estimate of the probability to be in excited states for matrix-valued potentials

This section presents a formal stability study of a perturbed eigenvalue problem that provides an approximation for the probability to be in excited states, p_{ex}. The two ingredients are first to determine the perturbation as a dynamic transition problem, related to the Landau–Zener model, and thereafter to use the stability analysis of a matrix eigenvalue problem to identify the probability to be in the excited state with the squared norm of the change of the eigenvector. The dynamic transition problem can be approximated numerically using Ehrenfest dynamics, as described in Sections 3.2 and 4. Our formulation of the perturbed eigenvalue problem is related to the transformation from local WKB solutions, which relate to the dynamic transition probability p_d, to a global solution of the Schrödinger equation. In dimension 1, we view a local WKB solution as a solution to the left of the hyperplane I (which is a point in dimension 1) and another WKB solution to the right of the hyperplane. The continuity condition to match these two WKB solutions to the right and left to one global differentiable solution forms the eigenvalue problem we will study. The eigenvalue problem is consequently not formulated directly for the unbounded Schrödinger operator. It is instead an eigenvalue problem with bounded solution operators obtained from WKB solutions. A reason we use this form of the perturbation analysis is that small dynamic transition probabilities p_d from

Landau–Zener like dynamics can be viewed as small regular perturbations while the corresponding perturbations of the potential in the Schrödinger equation is not small and not regular. The construction is extended to several dimensions in Section 3.3.4.

3.3.1 *The perturbed eigenvalue problem in one space dimension*

Consider first the scalar Schrödinger eigenvalue problem in one space dimension. Assume that we have solutions Φ in the domain $X>0$ and Φ in $X<0$, with right and left limits $\lim_{X\to 0+}\Phi(X)=:\Phi^r$, $\lim_{X\to 0-}\Phi(X)=:\Phi^\ell$, and $\lim_{X\to 0+}\Phi'(X)=:\Phi'^r$, $\lim_{X\to 0-}\Phi'(X)=:\Phi'^\ell$. In the one-dimensional case, the WKB method for an eigenvalue problem typically gives a caustic to the left, with phase θ_ℓ, and a caustic to the right, with phase θ_r (see Section 3.1, (A.5), and (A.19)) so that

$$\begin{aligned}\Phi(X)&\simeq\phi_r(X)\cos(M^{1/2}\theta_r(X)-\pi/4)\quad&&\text{for } X>0,\\ \Phi(X)&\simeq\phi_\ell(X)\cos(M^{1/2}\theta_\ell(X)-\pi/4)\quad&&\text{for } -X>0,\end{aligned}\tag{57}$$

for smooth functions $\theta_{r,\ell}$ and $\phi_{r,\ell}$. The continuity condition in order to have a global solution

$$\begin{aligned}\Phi^r&=\Phi^\ell,\\ \Phi'^r&=\Phi'^\ell,\end{aligned}$$

can with the notation $\Phi'^r=:R^r\Phi^r$ (and similarly for the limit to the left) be written

$$\begin{bmatrix}1&-1\\ R^r&-R^\ell\end{bmatrix}\begin{bmatrix}\Phi^r\\ \Phi^\ell\end{bmatrix}=0,\tag{58}$$

which implies that the 2×2 matrix

$$\begin{bmatrix}1&-1\\ R^r&-R^\ell\end{bmatrix}$$

must be singular, that is, $R^r=R^\ell$. For the WKB solutions, the derivative satisfies

$$\begin{aligned}R^r&=-M^{1/2}\theta_r'(0)\tan(M^{1/2}\theta_r(0)-\pi/4)+\mathcal{O}(M^0),\\ R^\ell&=-M^{1/2}\theta_\ell'(0)\tan(M^{1/2}\theta_\ell(0)-\pi/4)+\mathcal{O}(M^0).\end{aligned}$$

The two eigenvalues of the matrix

$$\begin{bmatrix}1&-1\\ R^r&-R^\ell\end{bmatrix}$$

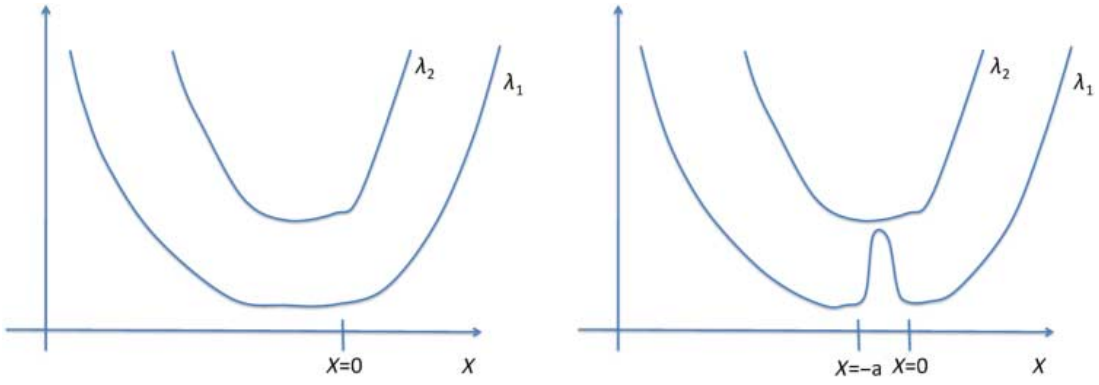

Fig. 2. The eigenvalues λ_1 and λ_2 of the potential V in the unperturbed (left) and perturbed (right) cases as a function of position X. For $X \notin [-a, 0]$, the potential $V(X)$ is diagonal in both the unperturbed and perturbed cases while for $X \in [-a, 0]$ the potential $V(X)$ is diagonal in the unperturbed case and nondiagonal in the perturbed case.

satisfy

$$\mu = \frac{1 - R^\ell}{2} \pm \sqrt{\left(\frac{1 - R^\ell}{2}\right)^2 + R^\ell - R^r}. \tag{59}$$

Consider the two decoupled scalar Schrödinger eigenvalue problems

$$\left(-\frac{1}{2M}\Delta + \lambda_1(X)\right)\Phi_1(X) = E\Phi_1(X) \quad \text{and} \quad \left(-\frac{1}{2M}\Delta + \lambda_2(X)\right)\Phi_2(X) = E\Phi_2(X). \tag{60}$$

The continuity condition corresponding to (58) for (60) will be used as our unperturbed eigenvalue problem, and in one space dimension this continuity condition becomes the unperturbed problem

$$\begin{bmatrix} \begin{bmatrix} 1 & -1 \\ R_1^r & -R_1^\ell \end{bmatrix} & 0 \\ 0 & \begin{bmatrix} 1 & -1 \\ R_2^r & -R_2^\ell \end{bmatrix} \end{bmatrix} \begin{bmatrix} \Phi_1^r \\ \Phi_1^\ell \\ \Phi_2^r \\ \Phi_2^\ell \end{bmatrix} = 0. \tag{61}$$

We consider a special case of the Schrödinger eigenvalue equation (2) in one space dimension where the potential is diagonal outside the interval $[-a, 0]$, that is, $V(X) = \begin{bmatrix} \lambda_1(X) & 0 \\ 0 & \lambda_2(X) \end{bmatrix}$, $X \notin [-a, 0]$, and view it as a perturbation of the diagonal Schrödinger equation (60), as illustrated in Figure 2. To have a diagonal potential in the unperturbed case is not necessary. What is important is that the electron eigenvalues are well separated so that the transition from ground state to excited states is negligible.

The WKB-method (48) shows that

$$\Phi^{\pm}(X) := \psi_{\pm}(X) G_{\pm}^{-1}(X)\, \mathrm{e}^{\pm \mathrm{i} M^{1/2}\theta(X)} \tag{62}$$

are local solutions to the Schrödinger equation (2), with positive and negative wave speeds $\pm\theta'(X)$. When the path X_t passes through a domain where the electron eigenvalues are close, the wave function ψ_+ (or ψ_-) determined by the transport equation (47) yields the probability p_d to transit from the ground state to an excited state, as in the Landau–Zener model (49). For a WKB-solution (62), with $\Phi = \Phi^+$ (or $\Phi = \Phi^-$), we can define the transition operator $\tilde{S}^+$ by

$$\Phi^+(0) = \tilde{S}^+ \Phi^+(-a) = \begin{bmatrix} \tilde{S}^+_{11} & \tilde{S}^+_{12} \\ \tilde{S}^+_{21} & \tilde{S}^+_{22} \end{bmatrix} \begin{bmatrix} \Phi^+_1(-a) \\ \Phi^+_2(-a) \end{bmatrix} \tag{63}$$

(and similarly for Φ^-) by solving the eikonal and transport equations (51) with data given on $X = -a$. The idea is that given the point $X = -a$ and the momentum P in this point the two first equations determine the path $(X(t), P(t))$. The third Schrödinger like transport equation in (51) then yields a linear solution operator along this path determining ψ at $X = 0$ from ψ at $X = -a$. Let us now determine $\tilde{S}$ more precisely. We want to determine the global effect of a small transit probability $|\tilde{S}_{21}|^2$ viewed as the perturbation to the eigenvalue problem (61) with $\tilde{S}^+\Phi^+(-a) + \tilde{S}^-\Phi^-(-a)$ replacing Φ^ℓ. To simplify the perturbation analysis, we factorize the transition

$$\begin{bmatrix} \tilde{S}_{11} & \tilde{S}_{12} \\ \tilde{S}_{21} & \tilde{S}_{22} \end{bmatrix} = \begin{bmatrix} 1 & S_{12} \\ S_{21} & 1 \end{bmatrix} \begin{bmatrix} \tilde{S}_{11} & 0 \\ 0 & \tilde{S}_{22} \end{bmatrix}$$

into a transition between the states $S := \left[\begin{smallmatrix} 1 & S_{12} \\ S_{21} & 1 \end{smallmatrix}\right]$ and a diagonal matrix $\left[\begin{smallmatrix} \tilde{S}_{11} & 0 \\ 0 & \tilde{S}_{22} \end{smallmatrix}\right]$, which does not contribute to the transition between states. Since we are only interested in estimates of the small components $\tilde{S}_{12} = S_{12}\tilde{S}_{22}$ and $\tilde{S}_{21} = \tilde{S}_{11}S_{21}$, we consider only perturbations generated by the transition matrix S. We note that the transition element $\tilde{S}_{21}$ by (63) can be written as the 2-component of $\Phi(0)$ if $\Phi(-a) = (1, 0)$. Therefore, (62) shows that the transition element

$$S^+_{21} = \frac{\tilde{S}^+_{21}}{\tilde{S}^+_{11}} = \frac{G_+(0)\, \mathrm{e}^{\mathrm{i}M^{1/2}\theta(0)}\psi_{2+}(0)}{G_+(0)\, \mathrm{e}^{\mathrm{i}M^{1/2}\theta(0)}\psi_{1+}(0)} = \frac{\psi_{2+}(0)}{\psi_{1+}(0)} \quad \text{where } |\psi_{1+}(-a)| = 1 \tag{64}$$

can be determined by the WKB-amplitude function ψ_+ and similarly we have the transition element $S^-_{21} = \psi_{2-}(0)/\psi_{1-}(0)$. The dynamic transition probability p_d, defined by Ehrenfest dynamics in (52), measures in this case the amplitude squared in the excited state at $X = 0$ for a wave starting in $X = -a$ in the ground state. Consequently, we have,

for $X_t = 0$ and Q corresponding to the + wave speed in (62), that

$$p_d(Q,0) = \langle \psi_+(0), \psi_+(0)\rangle - |\langle \psi_+(0), \Psi_0(0)\rangle|^2 = |\psi_{1+}(0)|^2 + |\psi_{2+}(0)|^2 - |\psi_{1+}(0)|^2 = |\psi_{2+}(0)|^2$$

so that by (64) $S_{21}^{\pm} = \mathcal{O}(p_d^{1/2})$, as $p_d \to 0$ (namely as the spectral gap becomes large). We assume for simplicity that the transition matrix $S = S^{\pm}$ is translation invariant, so that $\Phi'_{\pm}(0) = S^{\pm}\Phi'_{\pm}(-a)$. The total perturbation becomes the sum $S\Phi := S^+\Phi^+ + S^-\Phi^-$ of the perturbations of the two WKB-solutions $\Phi^{\pm}$, for the splitting $\Phi = \Phi^+ + \Phi^-$. Hence, we consider perturbations

$$\Phi^{\pm}(0) = S^{\pm}\Phi^{\pm}(-a) = \begin{bmatrix} 1-\mathcal{O}(p_d) & \mathcal{O}(p_d^{1/2}) \\ \mathcal{O}(p_d^{1/2}) & 1-\mathcal{O}(p_d) \end{bmatrix} \begin{bmatrix} \Phi_1^{\pm}(-a) \\ \Phi_2^{\pm}(-a) \end{bmatrix}$$

with the transition from the state 1 to state 2 determined by the matrix component $S_{21}^{\pm}$, which is of the (small) order $p_d^{1/2}$, where p_d, defined in (52), is related to the Landau–Zener probability. A reason to decompose the solution into WKB solutions is that their transition is determined by the amplitude functions $\psi_{\pm}$ with approximately conserved norm $|\psi_{\pm}(t)|$. The approximate conservation follows from the conservation $\frac{\mathrm{d}}{\mathrm{d}t}|\psi_{\pm}(t)|^2 = 0$ for the Ehrenfest dynamics (55) and the fact that the transport equation (47), determining $\psi_{\pm}$, becomes the Ehrenfest dynamics in the limit $M \to \infty$.

The perturbed eigenvalue condition becomes

$$(A+\beta)r_\beta = 0, \tag{65}$$

where

$$A := \begin{bmatrix} \begin{bmatrix} 1 & -1 \\ R_1^r & -R_1^\ell \end{bmatrix} & 0 \\ 0 & \begin{bmatrix} 1 & -1 \\ R_2^r & -R_2^\ell \end{bmatrix} \end{bmatrix} =: \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix},$$

$$r_\beta = \begin{bmatrix} \Phi_1(0+) \\ \Phi_1(-a) \\ \Phi_2(0+) \\ \Phi_2(-a) \end{bmatrix}, \tag{66}$$

$$A+\beta = \begin{bmatrix} 1 & -S_{11} & 0 & -S_{12} \\ R_1^r & -S_{11}R_1^\ell & 0 & -S_{12}R_1^\ell \\ 0 & -S_{21} & 1 & -S_{22} \\ 0 & -S_{21}R_2^\ell & R_2^r & -S_{22}R_2^\ell \end{bmatrix},$$

$$S\Phi(-a) = S^+\Phi^+(-a) + S^-\Phi^-(-a), \quad \beta \in \mathbb{C}^{4\times 4},$$

and we want to determine the change in the eigenvector $|r_\beta - r_0|^2$ (which measures the probability to be in the excited state) from the perturbed eigenvalue problem (65). For this, we use the result in the following lemma, whose derivation follows [15].

Lemma 3.3. Assume that the matrix $A \in \mathbb{C}^{n\times n}$ has n distinct eigenvalues, μ_j, $0 \le j \le n-1$. Let l_j and r_j be left and right eigenvectors of A, that is, $l_j A = \mu_j l_j$, $A r_j = \mu_j r_j$, satisfying

$$l_j \cdot r_k = \begin{cases} 1 & \text{if } j = k, \\ 0 & \text{if } j \neq k. \end{cases}$$

For sufficiently small matrices $\beta \in \mathbb{C}^{n\times n}$, the eigenvectors and eigenvalues of the perturbation matrix $A+\beta$ are differentiable functions of β. If r_β and μ_β denote the eigenvector and eigenvalue to $A+\beta$ that equal the eigenpair (r_0, μ_0) to the matrix A for $\beta = 0$, we have that

$$l_j \cdot (r_\beta - r_0) = -\frac{l_j \cdot \beta r_0}{\mu_j - \mu_0} + o(|\beta|) \tag{67}$$

and

$$\mu_\beta - \mu_0 = l_0 \cdot \beta r_0 + o(|\beta|). \tag{68}$$

□

Proof. That the eigenvectors and eigenvalues of $A+\beta$ are differentiable functions of β in a neighborhood of $\beta = 0$ follows directly by differentiation of the relations

$$(A+\beta)r = \mu r, \quad l(A+\beta) = \mu l.$$

Let $B := |\beta|^{-1}\beta$, where $|\beta|$ is a matrix norm, for example, the Euclidean operator norm. Let $r(\gamma)$ and $\mu(\gamma)$ be the perturbed normalized eigenvector and eigenvalue corresponding to r_0 and μ_0 given by $r(0) = r_0$ and $\mu(0) = \mu_0$, and

$$(A + \gamma B)r(\gamma) = \mu(\gamma) r(\gamma). \tag{69}$$

Differentiation of (69) at $\gamma = 0$ gives

$$Ar'(0) + Br(0) = \mu'(0)r(0) + \mu(0)r'(0). \tag{70}$$

Using that $r(0) = r_0$ and $\mu(0) = \mu_0$ and multiplying by l_j from the left, we get

$$(\mu_j - \mu_0) l_j \cdot r'(0) = l_j \cdot A r'(0) - \mu_0 l_j \cdot r'(0) = -l_j \cdot B r_0 + \mu'(0) l_j \cdot r_0 = -l_j \cdot B r_0 \quad \text{for } j \neq 0,$$

by the orthogonality between left and right eigenvectors. The Taylor expansion $r(|\beta|) = r(0) + r'(0)|\beta| + o(|\beta|)$ gives (67).

To prove (68), we take the scalar product with l_0 from the left in (70), which gives $\mu'(0) = l_0 \cdot Br_0$. The Taylor expansion $\mu(|\beta|) = \mu(0) + \mu'(0)|\beta| + o(|\beta|)$ gives (68). ■

The combination of (67) in Lemma 3.3 and the assumption that $r_\beta - r_0 = \mathcal{O}(1)$, as $|\beta| \to 0+$, also when $|\mu_j - \mu_0|$ is small, implies that

$$\begin{aligned}
\ell_j \cdot (r_\beta - r_0) &= \begin{cases} -\dfrac{\ell_j \cdot \beta r_0}{\mu_j - \mu_0} + o(|\beta|) & \text{if } |\mu_j - \mu_0|^{-1} = o(|\beta|^{-1}), \\ \mathcal{O}(1) & \text{otherwise}, \end{cases} \\
&= -\frac{\ell_j \cdot \beta r_0}{\mu_j - \mu_0 + c|\ell_j \cdot \beta r_0| \operatorname{sign}(\mu_j - \mu_0)} + o(|\beta|) \\
&\quad \text{[and by the definition } (\mu_j - \mu_0)^\sharp := \mu_j - \mu_0 + c|\ell_j \cdot \beta r_0| \operatorname{sign}(\mu_j - \mu_0)] \\
&=: -\frac{\ell_j \cdot \beta r_0}{(\mu_j - \mu_0)^\sharp} + o(|\beta|), \qquad (71)
\end{aligned}$$

for some positive constant c, which determines the perturbation orthogonal to r_0. We will also use that this representation holds separately for each perturbation S^+ and S^-. We do not determine the perturbation in the r_0 direction, which does not contribute to the transition between states. We denote by $r_\beta^\perp - r_0$ the projection of the perturbation $r_\beta - r_0$ on the hyperplane orthogonal to r_0. Since this plane is spanned by the left eigenvectors $l_1, \ldots, l_{n-1}$, we have by (67) in Lemma 3.3 that

$$r_\beta^\perp - r_0 = -\sum_{l_j \cdot r_0 \neq 0} \frac{l_j \cdot \beta r_0}{\mu_j - \mu_0} + o(|\beta|).$$

Returning to the perturbed problem (65), we recall from equation (59) that the unperturbed eigenvalues are given by

$$\begin{bmatrix} \mu_0 \\ \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} = \begin{bmatrix} \dfrac{1 - R_1^\ell}{2} - \sqrt{\left(\dfrac{1 - R_1^\ell}{2}\right)^2 + R_1^\ell - R_1^r} \\ \dfrac{1 - R_1^\ell}{2} + \sqrt{\left(\dfrac{1 - R_1^\ell}{2}\right)^2 + R_1^\ell - R_1^r} \\ \dfrac{1 - R_2^\ell}{2} - \sqrt{\left(\dfrac{1 - R_2^\ell}{2}\right)^2 + R_2^\ell - R_2^r} \\ \dfrac{1 - R_2^\ell}{2} + \sqrt{\left(\dfrac{1 - R_2^\ell}{2}\right)^2 + R_2^\ell - R_2^r} \end{bmatrix},$$

and the corresponding unperturbed eigenvectors are

$$\ell_j = \begin{cases} \dfrac{(1,-(1-\mu_j)/R_1^r,0,0)}{(1-(1-\mu_j)^2/R_1^r)} & \text{when } j=0,1, \\ \dfrac{(0,0,1,-(1-\mu_j)/R_2^r)}{(1-(1-\mu_j)^2/R_2^r)} & \text{when } j=2,3. \end{cases} \quad \text{and} \quad r_j = \begin{cases} (1,1-\mu_j,0,0) & \text{when } j=0,1, \\ (0,0,1,1-\mu_j) & \text{when } j=2,3. \end{cases}$$

Lemma 3.3 yields

$$\mu_\beta - \mu_0 = \ell_0 \cdot \beta r_0 + o(|\beta|) = (1-S_{11}) \frac{1-(1-\mu_0)^2 \frac{R_1^\ell}{R_1^r}}{1-(1-\mu_0)^2 \frac{1}{R_1^r}} + o(\sqrt{p_d}) = o(\sqrt{p_d}). \tag{72}$$

To have the perturbed eigenvalue equal to zero means that we start with an unperturbed eigenvalue $\mu_0 = o(\sqrt{p_d})$ in order to obtain the perturbed eigenvalue $\mu_\beta = 0$, satisfying $(A+\beta)r_\beta = 0$. Using the unperturbed eigenvector $r_0 = (1, 1+o(\sqrt{p_d}), 0, 0)$, with $\mu_0 = o(\sqrt{p_d})$, yields a perturbation in the Φ_2 component from zero (in the third component of r_0) to

$$S_{21} \sum_{j=2,3} \frac{1-(1-\mu_j)^2 \frac{R_2^\ell}{R_2^r}}{(\mu_j-\mu_0)^\sharp (1-(1-\mu_j)^2 \frac{1}{R_2^r})} + o(\sqrt{p_d}) = \mathcal{O}(\sqrt{p_d}) \tag{73}$$

and its squared absolute value measures the probability to be in the excited state, that is, the probability density $|\Phi_2|^2$, which integrated yields p_{ex} defined in (8).

Remark 3.4. Having a zero eigenvalue of A_1 in (66) means that the scalar Schrödinger eigenvalue problem

$$\left(-\frac{1}{2M}\Delta + \lambda_1(X)\right)\Phi_1(X) = E\Phi_1(X)$$

has a solution in the whole domain. We see from (73) that when also A_2 has a zero eigenvalue $\mu_2 = 0$ we obtain a resonance in the sense that a small perturbation β (where $|S_{21}| = \mathcal{O}(|\beta|) = \mathcal{O}(p_d^{1/2})$) yields a large change in the eigenvector $r_\beta = \Phi$, meaning that the probability to be in the excited state (which by (73) is of the order $|\beta|^2/|\mu_2|^2$) can be of order one even if the perturbation β is tiny. Resonance means that the two scalar Schrödinger eigenvalue problems have a coinciding eigenvalue E, see Figure 3. An approximation of this probability is

$$|\beta|^2/|\mu_2^\sharp|^2 \approx p_d/(|\mu_2-\mu_0|^2 + c^2 p_d). \tag{74}$$

□

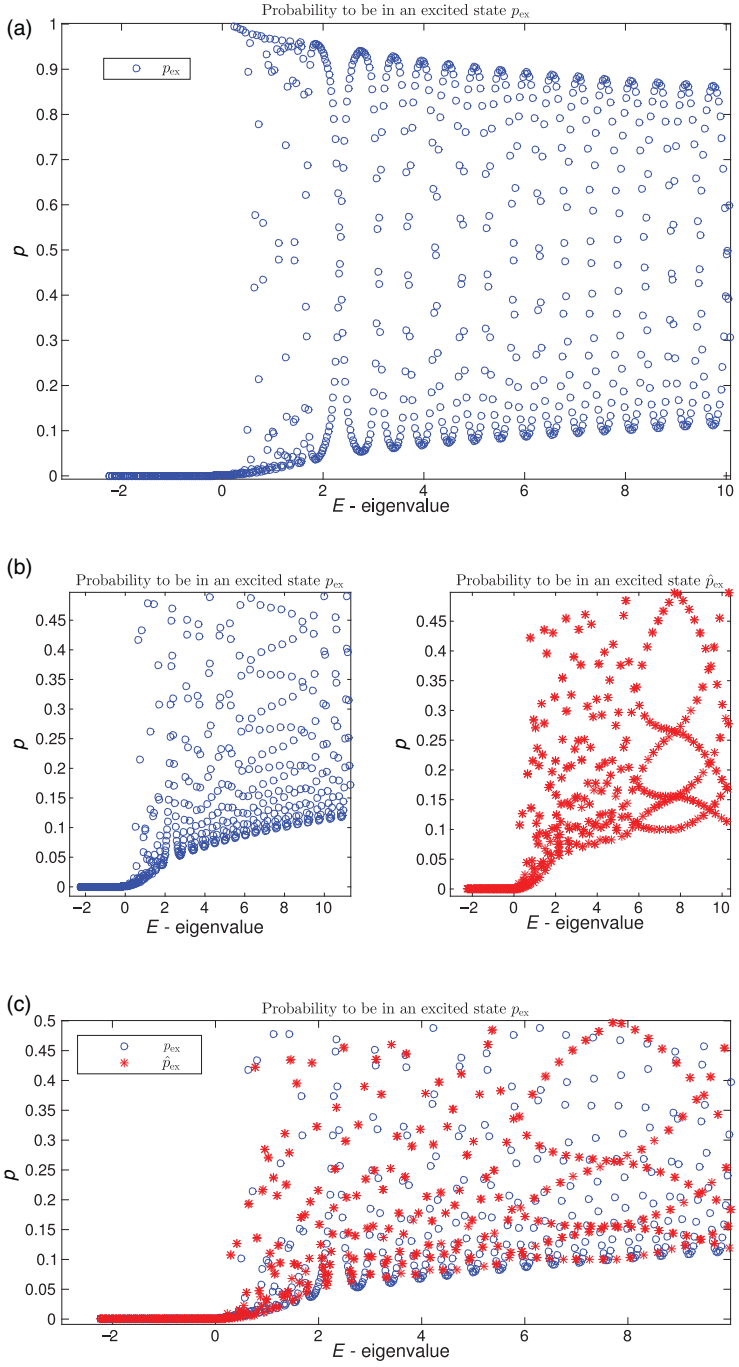

Fig. 3. Probabilities to be in the excited state, p_{ex}, using numerical approximation of (8), precisely defined in (87), and the estimations, $\hat{p}_{\text{ex}}$, using the formula (75), obtained by solving the discrete one-dimensional Schrödinger eigenvalue problem (86) with $M = 1,000$, $\delta = \sqrt{\frac{3}{5}} M^{-0.2} \approx 0.1946$, $C = 0.09$, and mesh size $h = 0.0001$. (a) The probability p_{ex}. (b) The probability p_{ex} (left panel) to be in an excited state compared with the estimate $\hat{p}_{\text{ex}}$ (right panel). (c) The probability p_{ex} ("o") to be in an excited state compared with the estimate $\hat{p}_{\text{ex}}$ ("$*$").

3.3.2 *Numerical test of the perturbation analysis*

A method to numerically test the validity of the perturbation analysis leading to the estimate (73), for the one-dimensional model, is to compare the probability to be in excited states p_{ex} in (8) (determined from numerical approximation of the Schrödinger equation (86)) with the expression (74) derived from (73) in the perturbation analysis. If p_{ex} shows some similarity with (74), the perturbation analysis is in some sense justified. The difference $\mu_2 - \mu_0$, of the eigenvalues in (74), can be obtained from numerical approximation of a difference $E_{2,k'} - E_{1,k}$, of the eigenvalues of the decoupled system (60), $(\lambda_i(X) - \frac{1}{2}M^{-1}\frac{\partial^2}{\partial X^2})\Phi_{i,k}(X) = E_{i,k}\Phi_{i,k}(X)$, $i = 1, 2$, using the same centered difference approximation as in (86). Let $\mathcal{E}_i = \{E_{i,k} \mid k = 1, 2, 3, \ldots\}$ be the set of all E_i eigenvalues, for $i = 1, 2$. By the approximation

$$|\mu_2 - \mu_0|^2 \approx C'|E_{1,k} - E_{2,k'}|^2 \approx C^{-1}|E_{1,k} - E_{2,k'}|^2/|E_{1,k} - E_{1,k+1}|^2,$$

where k and k' is chosen such that $E_{1,k}$ and $E_{2,k'}$ are the closest to E and using two constants C and $C' = C^{-1}/|E_{1,k} - E_{1,k+1}|^2$, we obtain the following estimate of the probability to be in the excited state:

$$\hat{p}_{\text{ex}} := \frac{p_d}{C^{-1}|E_1^l - E_2^l|^2/|E_1^l - E_1^m|^2 + p_d}, \tag{75}$$

with C being a constant, E being an eigenvalue of the discrete two-state Schrödinger equation corresponding to (75) and

$$E_1^l := \operatorname*{argmin}_{E_1 \in \mathcal{E}_1} |E_1 - E|, \quad E_1^m := \operatorname*{argmin}_{E_1 \in \mathcal{E}_1 \,\&\, E_1 \neq E_1^l} |E_1 - E_1^l|,$$
$$E_2^l := \operatorname*{argmin}_{E_2 \in \mathcal{E}_2} |E_2 - E_1^l|, \quad p_d := \exp(-\pi\delta^2\sqrt{M}/\sqrt{2(E - \lambda_-(0))}).$$

3.3.3 *Conclusions*

Figures 3 and 4 show that the estimation of the probabilities, $\hat{p}_{\text{ex}}$, to be in the excited state, obtained using the formula (75), and the numerically computed probabilities, p_{ex}, to be in the excited state, obtained using the formula (87) from the solution of the discrete one-dimensional Schrödinger eigenvalue problem (86), have similar qualitative behavior in the following two aspects: (1) the minimal value of p_{ex} and $\hat{p}_{\text{ex}}$ are similar, and (2) when E is close to a resonance, seen for $\hat{p}_{\text{ex}}$ in (75) as $|E_1^l - E_2^l| \ll |E_1^l - E_1^m|$, both p_{ex} and $\hat{p}_{\text{ex}}$ become larger. By also adjusting the constant C in (75), the behavior

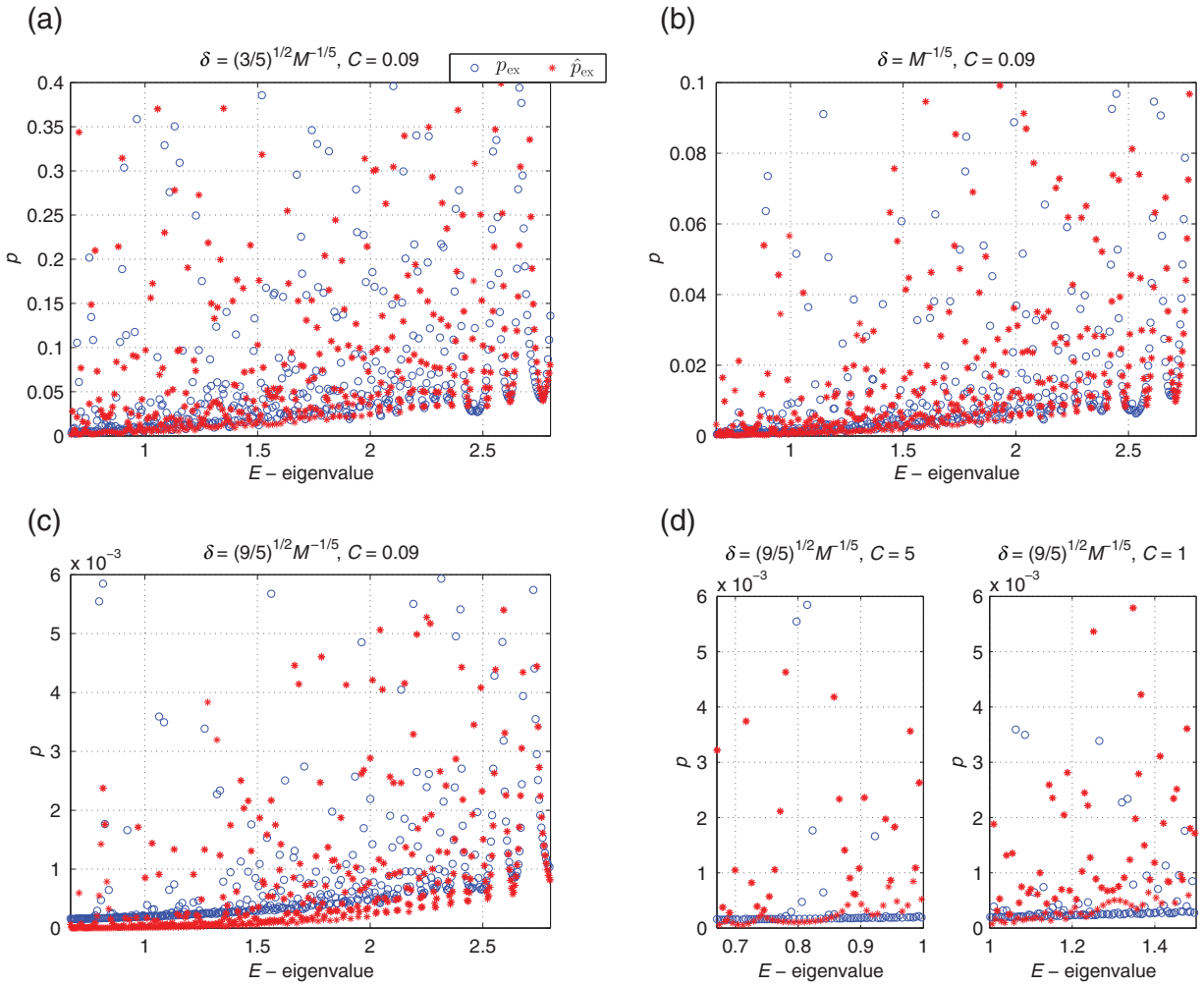

Fig. 4. Probabilities to be in the excited state, p_{ex}, using numerical approximation of (8), precisely defined in (87), and the estimations, $\hat{p}_{\text{ex}}$, using the formula (75), obtained by solving the discrete one-dimensional Schrödinger eigenvalue problem (86) with $M = 20,000$.

can also be quantitatively similar, see Figures 3 and 4. That is the perturbation analysis seems to capture this resonance phenomenon, at least qualitatively.

3.3.4 *The perturbed eigenvalue problem in the multi-dimensional case*

The main idea to extend the one-dimensional perturbed eigenvalue problem to multiple dimensions is to write the Schrödinger wave function as a phase-space integral of highly oscillatory functions using the Fourier–Bros–Iagolnitzer transform (FBI transform). This integral can be viewed as an integral over all hyperplanes I in $\mathbb{R}^{3N}$ defined in (53). Therefore, each FBI-mode can be initial data for a WKB solution (51) to the right

of a hyperplane I in $\mathbb{R}^{3N}$ and a WKB solution to the left of the hyperplane. The condition to have a differentiable WKB solution across the hyperplane then uses the continuity condition (65), as the one-dimensional setting, for each point in the hyperplane. The direction of the oscillations in the FBI-mode generates through the WKB-method a molecular dynamics path and an amplitude function. The amplitude function solves, along the path, a transport equation that is well approximated by Ehrenfest dynamics (55) and can consequently be computed similarly as the molecular dynamics paths but with smaller time steps, see Sections 3.2 and 4. The integral of WKB-solutions over all hyperplanes contribute through their amplitude functions to the transition, S. We will now explain this extension to multiple dimensions more precisely and then in the next section show how to computationally approximate the probability to be in excited states p_{ex}.

We will use the FBI transform. We could use the standard semiclassical Fourier transform as the initial data for the WKB-method but the FBI transform has the advantage of giving the high frequency content locally in X, that is, microlocally, which yields a more accurate WKB Ansatz. The important property of the FBI transform

$$T\varphi(Y,P) := \underbrace{(2^{1/3}\pi)^{-3N/4}M^{9N/8}}_{=:\alpha_M}\int_{\mathbb{R}^{3N}} \mathrm{e}^{\mathrm{i}M^{1/2}(Y-X)\cdot P - |Y-X|^2M^{1/2}/2}\varphi(X)\,\mathrm{d}X \tag{76}$$

is the identity

$$\varphi = T^*T\varphi \tag{77}$$

for $\varphi \in L^2(\mathbb{R}^{3N})$, where the adjoint operator T^* is defined by

$$T^*\phi(X) = \alpha_M \int_{\mathbb{R}^{6N}} \mathrm{e}^{-\mathrm{i}M^{1/2}(Y-X)\cdot P - |Y-X|^2M^{1/2}/2}\phi(Y,P)\,\mathrm{d}Y\,\mathrm{d}P$$

for example, for all Schwartz functions ϕ on $\mathbb{R}^{6N}$, see [25]. Therefore, the integral representation

$$\begin{aligned}\Phi(X) &= T^*T\Phi(X)\\ &= \alpha_M \int_{\mathbb{R}^{6N}} \mathrm{e}^{-\mathrm{i}M^{1/2}(Y-X)\cdot P - |Y-X|^2M^{1/2}/2}T\Phi(Y,P)\,\mathrm{d}Y\,\mathrm{d}P\end{aligned}$$

yields suitable boundary data on hyperplanes (e.g., $X_1 = 0$) for the WKB-method. We will, for each point $(Y,P) =: Q$ (in the classically allowed region), use a WKB-function (41). In the case of caustics, this single WKB-function is replaced by a finite

sum of WKB-functions based on the Legendre transform $\theta^*(P)$ of $\theta(X)$, as explained in Appendices 1 and 2,

$$\Phi_Q(X) \sim \sum_{\{P^*:\nabla_P\theta_Q^*(P^*)=X\}} \phi_Q(X;P^*)\,\mathrm{e}^{\mathrm{i}M^{1/2}\theta_Q(X;P^*)}. \tag{78}$$

This WKB solution Φ_Q solves the Schrödinger eigenvalue problem (2) in the two domains to the left and right of the hyperplane, with the boundary condition constructed so that the WKB-mode $\Phi_Q(X)$ is equal to the FBI-mode

$$\alpha_M\,\mathrm{e}^{-\mathrm{i}M^{1/2}(Y-X)\cdot P-|Y-X|^2M^{1/2}/2}T\Phi(Y,P)$$

for $X-Y$ in the hyperplane $L_P := \{q\in\mathbb{R}^{3N} \mid q\cdot P=0\}$ orthogonal to P:

$$\begin{aligned} &q\cdot\nabla\theta_Q(X) = q\cdot P = 0, \quad \text{for } q\in L_P \text{ and } X-Y\in L_P,\\ &\phi_Q(X;P^*) = \alpha_M T\Phi(Q)\,\mathrm{e}^{-|X-Y|^2M^{1/2}/2} \quad \text{for } X-Y\in L_P. \end{aligned} \tag{79}$$

The construction of the WKB-solutions in the case of caustics uses asymptotic convergence as $M\to\infty$, described by the asymptotically equal sign $\sim$, see [34, 51] and (A.19).

The obtained decomposition

$$\Phi(X) = \int_{\mathbb{R}^{6N}} \Phi_Q(X)\,\mathrm{d}Q \tag{80}$$

determines, for each Q, the corresponding operators $R^{r,\ell}$. The total perturbation comes from transitions from all ϕ_Q along the paths: for each $Q=(Y,P)$ we let the function $\phi_Q(X_t)$ solve the transport equation (47) along the WKB-path $\{X_t\}_{t=0}^{\infty}$ and initialize $\phi_Q(X_\tau)$ to the ground state $\Psi_0(X_\tau)$ each time X_τ is in the hyperplane $(X_\tau-Y)\cdot P=0$, generated by Q. The perturbed wave function is then as in (64) determined by

$$\langle S(Q,X),\Psi_i(X)\rangle = \frac{\langle\phi_Q(X),\Psi_i(X)\rangle}{\langle\phi_Q(X),\Psi_0(X)\rangle},$$

which depends on Q, by relating ψ_{1+} in (64) with $\langle\phi_Q,\Psi_0\rangle$ and ψ_{2+} with $\langle\phi_Q,\Psi_1\rangle$ in the case of two states. Integrating over all Q yields the total perturbation, as in (66),

$$\int_{\mathbb{R}^{6N}} S(Q,X)\Phi_Q(X)\,\mathrm{d}Q$$

at the point X. The corresponding perturbed eigenvalue problem (65) for the wave function Φ_Q implies, as in (73), the following larger change in the excited state:

$$\begin{aligned}\Phi^{\perp}(X) &= \Phi(X) - \langle \Psi_0(X), \Phi(X)\rangle \Psi_0(X)\\ &= \int_{\mathbb{R}^{6N}} \underbrace{S_{21}(Q,X)}_{=p_d^{1/2}(Q,X)} \Phi_Q(X) \sum_{j=2,3} \frac{1-(1-\mu_j)^2 \frac{R_2^{\ell}}{R_2^{r}}}{(\mu_j-\mu_0)^{\sharp}(1-(1-\mu_j)^2\frac{1}{R_2^{r}})}\,\mathrm{d}Q\\ &\quad + \mathcal{O}\left(\left|\int_{\mathbb{R}^{6N}} S_{21}(Q,X)\Phi_Q(X)\,\mathrm{d}Q\right|^2\right) \end{aligned} \tag{81}$$

at the point X, where μ_j and $R_2^{r,\ell}$ depend on Q. The index 1 in S now corresponds to the component in the ground state and the index 2 to be in an excited state, that is,

$$\Phi(X) = \gamma_1(X)\Psi_0(X) + \gamma_2(X)\Psi^{\perp}(X),$$

and $\gamma_j(X) \in \mathbb{C}$ with the orthogonality conditions

$$\langle \Psi_0, \Psi^{\perp}\rangle = 0, \quad \langle \Psi_0, \Psi_0\rangle = 1, \quad \langle \Psi^{\perp}, \Psi^{\perp}\rangle = 1.$$

The eigenvalues of the unperturbed problem (59) vary highly due the oscillatory functions $R^{r,\ell}$, with the phase $M^{1/2}\theta^{\pm}(X)$. Figure 4 illustrates a consequence of this variation and indicates that mean values of excitation probabilities corresponding to eigenvalues in a neighborhood are more stable than individual excitations probabilities. Let us therefore consider a local average of the probability to be in the excited state based on a simple model of the smallest eigenvalue $\mu_2 \sim \tan(M^{1/2}(\theta_2^{+} - \theta_2^{\ell}))$ (in the excited component) where we assume that $|\mu_2|$ has a bounded density, ρ_Q, defined on $\mathbb{R}$. In the case of one space dimension, we can fix a point $X \in \mathbb{R}$ (in the classically allowed region) where μ_2 is evaluated, that cuts the spatial domain into two. The average then corresponds to an ensemble of Schrödinger eigenvalues localized around E, related to a local mean value of excitation probabilities in Figure 4. We denote this local average by $\mathbb{A}$. In higher dimensions, we can similarly fix a codimension 1 cutting surface and evaluate μ_2 on the surface. Let us now study the average excitation probability in this model. Using that $(\mu_2 - \mu_0)^{\sharp}$ yields transition probabilities bounded by one for each Q as in (71) and that outside a compact set (corresponding to the classically allowed region) the factor $|S_{21}(Q,X)\Phi_Q(X)|$ is negligible small,

we have

$$\begin{aligned}
\mathbb{A}[\langle \Phi^{\perp}(X), \Phi^{\perp}(X)\rangle] &\le C\,\mathbb{A}\left[\left|\int_{\mathbb{R}^{6N}} \frac{|S_{21}(Q,X)\Phi_Q(X)|}{|(\mu_2-\mu_0)^{\sharp}|}\,\mathrm{d}Q\right|^2\right] \\
&\le C\,\mathbb{A}\left[\int_{\mathbb{R}^{6N}} \frac{|S_{21}(Q,X)\Phi_Q(X)|^2}{|\mu_2-\mu_0|^2+c^2|S_{21}(Q,X)|^2}dQ\right] \\
&= C\int_{\mathbb{R}^{6N}}\int_{\mathbb{R}} \frac{|S_{21}(Q,X)\Phi_Q(X)|^2}{|\mu_2-\mu_0|^2+c^2|S_{21}(Q,X)|^2}\rho_Q(\mu_2)\,\mathrm{d}\mu_j\,\mathrm{d}Q \\
&= C\int_{\mathbb{R}^{6N}} |S_{21}(Q,X)||\Phi_Q(X)|^2 \int_{\mathbb{R}} \frac{1}{\frac{|\mu_2-\mu_0|^2}{c^2|S_{21}(Q,X)|^2}+1}\rho_Q(\mu_2)\,\mathrm{d}\frac{\mu_2}{c^2|S_{21}(Q,X)|}\,\mathrm{d}Q \\
&\le C\int_{\mathbb{R}^{6N}} |S_{21}(Q,X)||\Phi(X)|^2\|\rho_Q\|_{L^\infty(\mathbb{R})}\,\mathrm{d}Q\int_{\mathbb{R}}\frac{\mathrm{d}x}{x^2+1} \\
&= \mathcal{O}\left(\int_{\mathbb{R}^{6N}} |S_{21}(Q,X)||\Phi(X)|^2\,\mathrm{d}Q\right).
\end{aligned}$$

We conclude that in a model of local ensemble averages of eigenvalues E, assuming that for each X the eigenvalue difference $|\mu_2-\mu_0|$ has a bounded density, then the average probability to be in the excited state has the bound

$$\mathbb{A}[p_{\mathrm{ex}}] = \mathcal{O}\left(\int_{\mathbb{R}^{3N}}\int_{\mathbb{R}^{6N}} |S_{21}(Q,X)||\Phi_Q(X)|^2\,\mathrm{d}Q\,\mathrm{d}X\right),$$

which, by (81) can be written as the average of the square root of the dynamic transition probability p_d

$$\mathbb{A}[p_{\mathrm{ex}}] = \mathcal{O}\left(\int_{\mathbb{R}^{3N}}\int_{\mathbb{R}^{6N}} p_d^{1/2}(Q,X)|\Phi_Q(X)|^2\,\mathrm{d}Q\,\mathrm{d}X\right). \tag{82}$$

3.3.5 *Ergodic computation of probabilities to be in excited states*

The FBI transform satisfies

$$\|T\Phi\|_{L^2(\mathbb{R}^{6N})} = \|\Phi\|_{L^2(\mathbb{R}^{3N})} = 1$$

and $T\Phi$ concentrates on the phase-space set $H_0(Y,P) = |P|^2/2 + \lambda_0(Y) = E$ in the limit as $M\to\infty$ and $p_{\mathrm{ex}}\to 0+$, see [25]. When the dynamics is ergodic, the phase-space measure is in addition uniform on the set $E < H_0(Q) < E+\gamma$ as $\gamma\to 0+$, cf. [37]. The WKB functions $\Phi_Q(X)$ behave similarly as $T\Phi$ locally, since their initial conditions are given by $T\Phi$ and they solve the Schrödinger equation. We may write $S_{21}(Q,X) =: S_{21}(Q;X,P)$ where P is the momentum for the path at the position X that started in the plane given by $Q = (Y',P')$, since this P is a function of Q and X. Therefore, we approximate p_{ex} by

the following time average of $p_d^{1/2}$:

$$p_{\mathrm{ex}} \approx \lim_{\gamma\to 0+} \int_{\mathbb{R}^{3N}} \frac{\int_{E<H_0(Q)<E+\gamma} |S_{21}(Q;X,P)|\,\mathrm{d}Q}{\int_{E<H_0(Q)<E+\gamma} \mathrm{d}Q}\,\mathrm{d}X = \lim_{\gamma\to 0+} \int_{\mathbb{R}^{3N}} \frac{\int_{E<H_0(Q)<E+\gamma} |S_{21}(X,P;Q)|\,\mathrm{d}Q}{\int_{E<H_0(Q)<E+\gamma} \mathrm{d}Q}\,\mathrm{d}X$$
$$= \lim_{T\to\infty} \int_{\mathbb{R}^{3N}} \int_0^T \underbrace{|S_{21}(X,P;X_t,P_t)|}_{=p_d^{1/2}(X;X_t)} \frac{\mathrm{d}t}{T}\,\mathrm{d}X, \tag{83}$$

Algorithm 1 Approximations of the probabilities, p_{ex}, to be in excited states using molecular dynamics in 2D based on (83)

Input: Energy E; potential functions V; mass M; time T; initial position X_0, initial momentum P_0.

Output: Approximated probability $\hat{\hat{p}}_{\mathrm{ex}}$ to be in excited states.

1. Sampling of initialization times $\{T_n\}_n$ and hyperplane coordinates $\{\overline{X}_{T_n}, \overline{P}_{T_n}\}_n$:
Set $t \leftarrow 0, n \leftarrow 0, T_n \leftarrow 0$ and $(\overline{X}_{T_n}, \overline{P}_{T_n}) \leftarrow (X_0, P_0)$ and define the hyperplane $\overline{P}_{T_n}^{\perp} := \{X \in \mathbb{R}^2 | X \cdot \overline{P}_{T_n} = 0\}$.
while $T_n < T$ **do**
Simulate the ground state Born-Oppenheimer molecular dynamics for (X_t, P_t) until X_t crosses the plane $\overline{P}_{T_n}^{\perp}$.
At the crossing time, set $n \leftarrow n+1, T_n \leftarrow t$ and $(\overline{X}_{T_n}, \overline{P}_{T_n}) \leftarrow (X_t, P_t)$ and define the hyperplane $\overline{P}_{T_n}^{\perp} := \{X \in \mathbb{R}^2 | X \cdot \overline{P}_{T_n} = 0\}$.
end while

2. Solve the Ehrenfest molecular dynamics using the Störmer-Verlet method with the entries in $\{T_k, \overline{X}_{T_k}, \overline{P}_{T_k}\}_{k=0}^n$ as input parameters and compute $\hat{\hat{p}}_{\mathrm{ex}}$:
$temp \leftarrow 0$
for $k = 0$ to $n-1$ **do**
$t_1 \leftarrow T_k$
$t_2 \leftarrow T_{k+1}$
$X_{t_1} \leftarrow \overline{X}_{T_k}$
$P_{t_1} \leftarrow \overline{P}_{T_k}$
$\psi_{t_1} = \Psi_-(X_{t_1})$
$\{X_t, P_t, \psi_t\} \leftarrow$ Ehrenfest dynamics path obtained from time t_1 to t_2 with initial data $X_{t_1}, P_{t_1}, \psi_{t_1}$
$temp \leftarrow temp + \int_{t_1}^{t_2} |\langle \psi_t, \Psi_+(X_t)\rangle|\,\mathrm{d}t$
end for
$\hat{\hat{p}}_{\mathrm{ex}} \leftarrow temp/T_n$

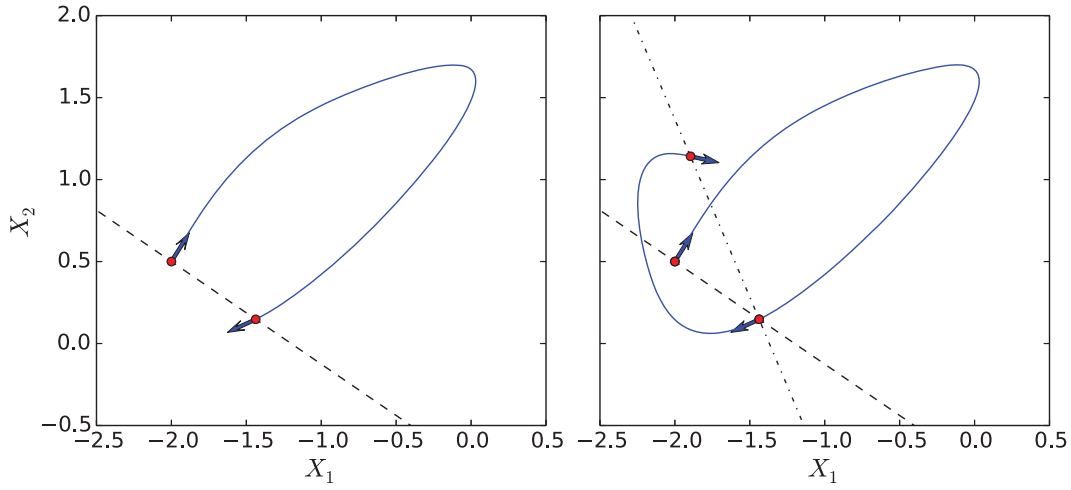

Fig. 5. Illustration of the sampling of hyperplanes and initialization times described in part 1 of Algorithm 1. Left figure: the two-dimensional Born–Oppenheimer molecular dynamics trajectory X_t (blue line) with initial conditions $X_{T_0} = (-2, 0.5)$ and $\dot{X}_{T_0} = P_{T_0}$ is simulated until the time $t = T_1$ when it crosses the hyperplane $P_{T_0}^{\perp}$ (dashed line). Right figure: a new hyperplane $P_{T_1}^{\perp}$ (dash-dotted line) is constructed and the trajectory of X_t is simulated further until the time $t = T_2$ when it crosses that hyperplane. The sampling procedure is iterated until the final time is reached.

where we used that the transition probability $S_{21}(Q; X, P) = S_{21}(X, P; Q)$ is symmetric by reversing time t to $-t$. In fact, along a given path $\{X_t\}_{t=0}^{T}$ also the momentum P', at position Y', is determined by the position Y' and X_t so that $p_d^{1/2}(Y'; X_t) = S_{21}(Y', P'; X_t, P_t)$ is well defined. The relation (83) means that we sample square roots, $p_d^{1/2}(X; X_t)$, of transition probabilities along the molecular dynamics path, normalized for each passage through a hyperplane, and then take the average of all hyperplanes. In our computations, we sample hyperplanes by means of the phase-space trajectory of ground state Born–Oppenheimer molecular dynamics, cf. Algorithm 1 and Figure 5.

4 Numerical Examples

The purpose of this section is to present two simple model problems, where the Schrödinger eigenvalue solution can be studied computationally and compared with the molecular dynamics approximation. In the following subsections, we

- show that the dynamic transition probability p_d in (52) can be determined by numerical solution of Ehrenfest dynamics,
- verify assumption (26) on exponential convergence rate in finite time,

- compare numerical approximations of observables from the Schrödinger equation and Born–Oppenheimer molecular dynamics,
- compare approximations of the probability to be in excited states p_{ex} from the Schrödinger equation with the molecular dynamics approximation $\hat{p}_{\text{ex}}$ from Algorithm 1.

4.1 **Model 1: a one-dimensional problem**

We consider the one-dimensional, time-independent Schrödinger equation (2) with the heavy-particle coordinate $X \in \mathbb{R}$, two electron states $J = 2$, and the Hamiltonian operator, $\hat{H}(X)$, defined by

$$\hat{H}(X) := V(X) - \frac{1}{2} M^{-1} \frac{\partial^2}{\partial X^2} \tag{84}$$

with the potential operator V defined by the matrix

$$V(X) := \begin{bmatrix} X + r(X) & \delta \\ \delta & -X + r(X) \end{bmatrix}, \tag{85}$$

where the parameter δ is a non-negative constant and the function $r : \mathbb{R} \to \mathbb{R}$ is given by

$$r(X) := \begin{cases} (a_l - X)^2 & \text{if } X < a_l, \\ (X - a_r)^2 & \text{if } X > a_r, \\ 0 & \text{otherwise.} \end{cases}$$

For each X, the potential V defines the eigenvalue problem $V(X)\Psi_\pm(X) = \lambda_\pm(X)\Psi_\pm(X)$, with the two eigenvalues $\lambda_\pm(X) \in \mathbb{R}$ and two eigenvectors $\Psi_\pm(X) \in \mathbb{R}^2$. The eigenvalues are given by $\lambda_\pm(X) = r(X) \pm \sqrt{X^2 + \delta^2}$, and $\Psi_-(X)$ and $\Psi_+(X)$, respectively, denote the ground and the excited state vector.

Choosing $\delta = 0$ gives a conical intersection at $X = 0$, and a positive value of δ gives a minimum gap 2δ between $\lambda_-(X)$ and $\lambda_+(X)$ at $X = 0$. A small value of δ corresponds to a large probability to be in excited states and a large value of δ corresponds to a small probability to be in excited states. Figure 6 illustrates examples of small and large spectral gaps between λ_+ and λ_- with two different values of δ, respectively.

4.1.1 *The one-dimensional discretized Schrödinger equation*

We use a central difference method to solve the one-dimensional two-state Schrödinger equation (2) in the domain $X \in (-2\pi, 2\pi)$ with $(a_l, a_r) = (-2, 3)$, mesh size $h = 4\pi/[10M^{3/4}]$, and partition $X^j = -2\pi + jh$, $j = 0, 1, 2, \ldots, [10M^{3/4}]$. The approximation of

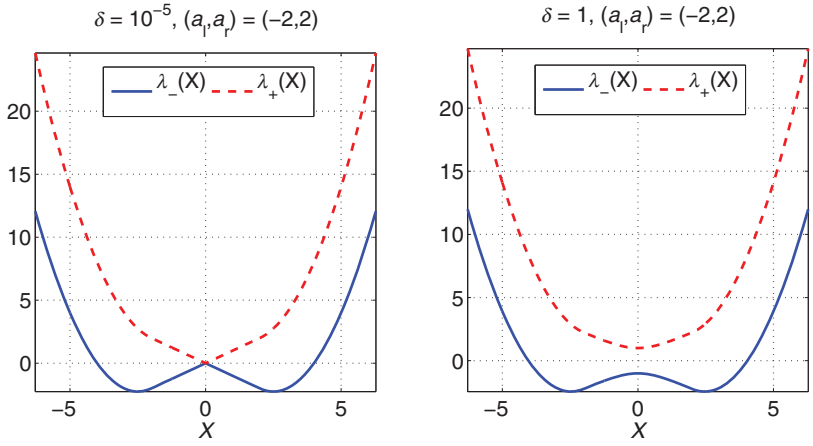

Fig. 6. Eigenvalues of the potential matrix (85) for small and large values of the parameter δ corresponding to the small and large minimum gaps between the eigenvalues, respectively.

the eigenvalue problem with the eigenvector components $\Phi_h^j \simeq \Phi(X^j)$ and corresponding eigenvalue $E_h \in \mathbb{R}$ becomes

$$-\frac{1}{2M}\frac{\Phi_h^{j-1}-2\Phi_h^j+\Phi_h^{j+1}}{h^2}+V(X^j)\Phi_h^j=E_h\Phi_h^j \quad \text{for } j=1,2,\ldots,[10M^{3/4}]-1, \tag{86}$$

using homogeneous Dirichlet boundary conditions $\Phi_h^0=\Phi_h^{10M^{3/4}}=0$.

The approximation of the probability p_{ex} to be in the excited state is given by

$$p_{\text{ex}}=\frac{\sum_{j=0}^{[10M^{3/4}]}\langle\Phi_h^j,\Psi_+(X^j)\rangle^2}{\sum_{j=0}^{[10M^{3/4}]}\langle\Phi_h^j,\Phi_h^j\rangle}. \tag{87}$$

4.1.2 *Approximation of the dynamic transition probability* p_d

We show in this computational example that the dynamic transition probability, p_d, defined in (52), can be obtained from Ehrenfest molecular dynamics simulations (55) using the following formula:

$$p_d(t)=\left|\frac{\langle\Psi_+(X_t),\psi_t\rangle}{\langle\Psi_+(X_t),\Psi_+(X_t)\rangle}\right|^2, \tag{88}$$

with $t\in\mathbb{R}$ denoting the time. We observe that numerical experiments illustrate that the transition probability $p_d(t)$ approximates, as time tends to infinity, the Landau–Zener probability p_{LZ} as defined in (50), given the Landau–Zener model (49), see Figure 8.

We approximate the solution of the transport equation (47) by the Ehrenfest molecular dynamics simulations (55) with $\lambda_0=\lambda_-$ (equal to the smallest eigenvalue

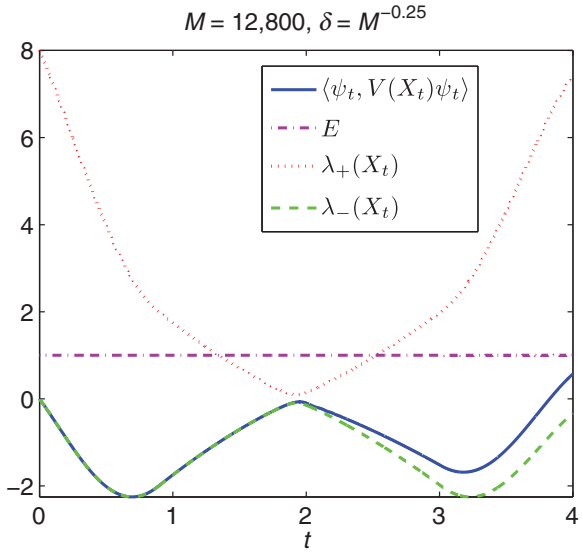

Fig. 7. The potential energy function, $t \mapsto \langle \psi_t, V(X_t)\psi_t \rangle$, approximated from the Ehrenfest molecular dynamics simulations (55), deviates away from the eigenvalues after the avoided crossing, if the gap is small and the mass is sufficiently small.

of V in(85)) in the Hamiltonian (56). For this computational example, we choose $E = 1$, $(a_l, a_r) = (-2, 2)$ and the initial conditions $t = 0$, $X_0 = -4.0$, $\psi_0 = \Psi_-(X_0)$, and $P_0 = \sqrt{2(E - \lambda_-(X_0))}$. We approximate the solution of the molecular dynamics using the Störmer–Verlet method, [18], based on the two symplectic Euler steps,

$$
\begin{aligned}
\psi^i_{k+\frac{1}{2}} &= \psi^i_k - \frac{\Delta t}{2}\sqrt{M}\check{V}(X_k)\psi^r_k,\\
P_{k+\frac{1}{2}} &= P_k - \frac{\Delta t}{2}\left(\lambda'_-(X_k) + \left\langle \psi^r_k, \check{V}'(X_k)\psi^r_k \right\rangle + \left\langle \psi^i_{k+\frac{1}{2}}, \check{V}'(X_k)\psi^i_{k+\frac{1}{2}} \right\rangle\right),\\
X_{k+1} &= X_k + \Delta t P_{k+\frac{1}{2}},\\
\psi^r_{k+1} &= \psi^r_k + \frac{\Delta t}{2}\sqrt{M}(\check{V}(X_k) + \check{V}(X_{k+1}))\psi^i_{k+\frac{1}{2}},\\
\psi^i_{k+1} &= \psi^i_{k+\frac{1}{2}} - \frac{\Delta t}{2}\sqrt{M}\check{V}(X_{k+1})\psi^r_{k+1},\\
P_{k+1} &= P_{k+\frac{1}{2}} - \frac{\Delta t}{2}\left(\lambda'_-(X_{k+1}) + \left\langle \psi^r_{k+1}, \check{V}'(X_{k+1})\psi^r_{k+1} \right\rangle + \left\langle \psi^i_{k+\frac{1}{2}}, \check{V}'(X_{k+1})\psi^i_{k+\frac{1}{2}} \right\rangle\right).
\end{aligned}
\tag{89}
$$

4.1.3 *Conclusions*

Figure 7 illustrates that the potential energy function $t \mapsto \langle \psi_t, V(X_t)\psi_t \rangle$, obtained from the Ehrenfest molecular dynamics simulations is equal to the ground state

energy, $\lambda_-(X_t)$, until the dynamics reaches the avoided conical intersection. Then it deviates from the ground state energy and continues in between the ground and excited state energies.

Figure 8 shows that the transition probability, $p_d(t)$, obtained using the formula (88) based on Ehrenfest molecular dynamics simulations, remains zero until it reaches the avoided conical intersection and then it oscillates near the avoided conical intersection region and approaches p_{LZ} asymptotically as $t \to \infty$.

4.2 Model 2: a two-dimensional problem

We consider the two-dimensional, time-independent Schrödinger equation (2) with the heavy-particle coordinate $X = (X_1, X_2) \in \mathbb{R}^2$, and two electron states $J = 2$. The Hamiltonian is

$$\hat{H}(X) = V(X) - \tfrac{1}{2} M^{-1} \Delta_X,$$

with

$$V(X) := \underbrace{(\tfrac{1}{2}(X_1^2 + \alpha X_2^2) + \beta \sin(X_1 X_2))}_{=:\lambda_s(X)} I + \eta \begin{bmatrix} v_1(X) & v_2(X) \\ v_2(X) & -v_1(X) \end{bmatrix}, \tag{90}$$

where I is the 2×2 identity matrix, the functions $v_1, v_2 : \mathbb{R}^2 \to \mathbb{R}$ are given by (A) or (B) below, and $\alpha = \sqrt{2}$, $\beta = 2$, and $\eta = 1/2$. As in Model 1, we have for each X the eigenvalue problem, $V(X)\Psi_\pm(X) = \lambda_\pm(X)\Psi_\pm(X)$, with the eigenvalues $\lambda_\pm(X) = \lambda_s(X) \pm \eta\sqrt{(v_1(X))^2 + (v_2(X))^2}$ and the ground state and excited state eigenvectors $\Psi_-(X)$ and $\Psi_+(X)$, respectively. We choose the energy, $E = 1.5$. In this example, we study the following two cases of interactions between the two potential surfaces λ_+ and λ_-:

(A) *A line intersection.* We choose $v_1(X) = \arctan(X_1/\eta)$ and $v_2(X) = \delta/\eta$, where δ is a non-negative constant that defines the minimum distance between the potential surfaces. In this example, the potential surfaces intersect each other (for $\delta = 0$) or have the minimum distance between each other at the line $X_1 = 0$. Here, choosing a small value for δ corresponds to a large probability to be in excited states and a large value for δ will give a small probability to be in excited states. The parameter δ in this two-dimensional example is analogous to the parameter δ in the one-dimensional example given in Section 4.1.

(B) *A conical intersection.* We choose $v_1(X) = \arctan((X_1 - a_1)/\eta)$ and $v_2(X) = \arctan((X_2 - a_2)/\eta)$ with $a = (a_1, a_2) \in \mathbb{R}^2$ being a chosen point in the two-dimensional space. If a is chosen such that $\lambda_\pm(a)$ are smaller than the

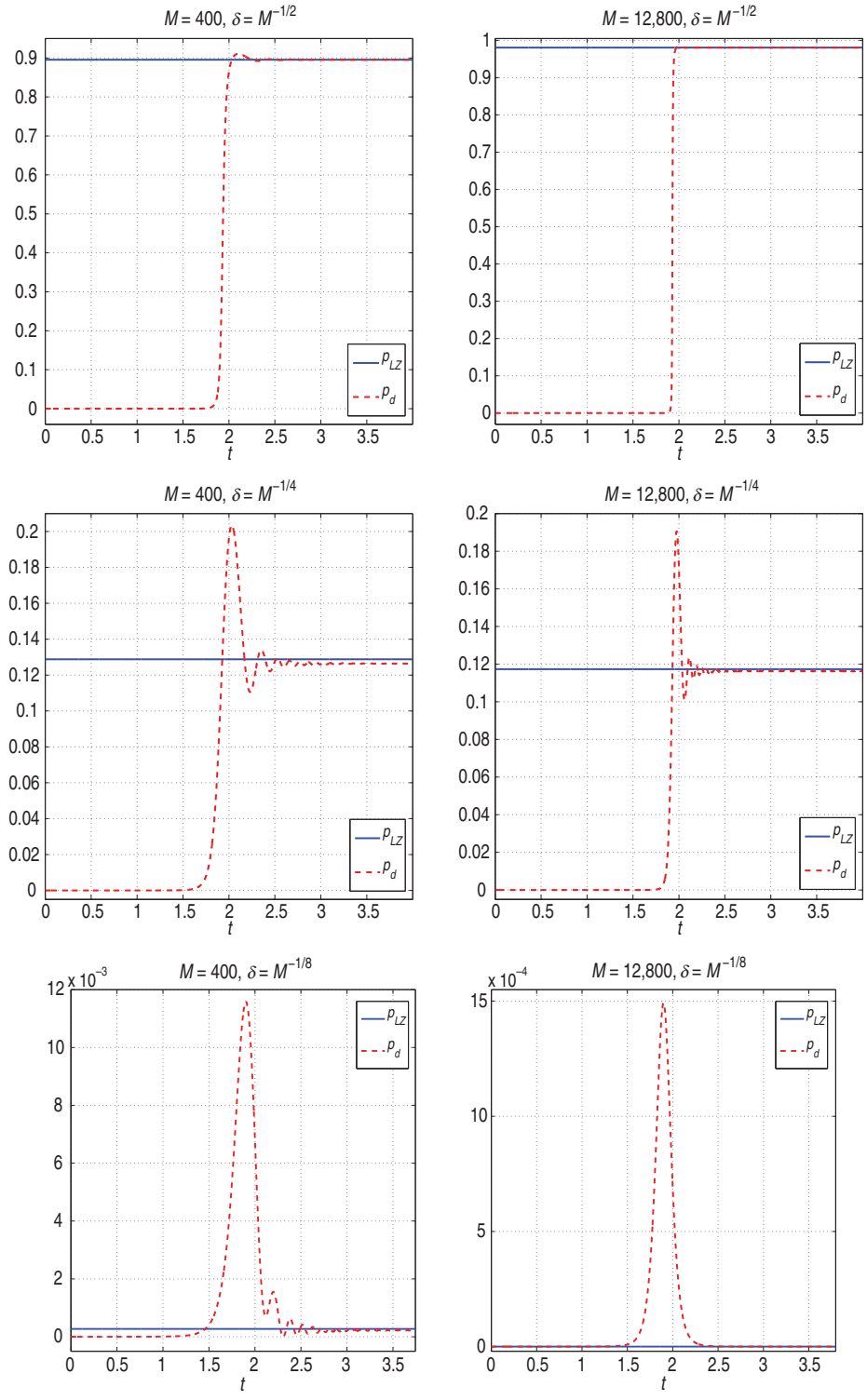

Fig. 8. Estimation of the Landau–Zener probability, p_{LZ}, from the Ehrenfest molecular dynamics simulations (55). The estimation, $p_d(t)$, overshoots near the avoided crossing and eventually stailizes around p_{LZ}. The overshoot period and the relative magnitude are more prominent for the cases of smaller Landau–Zener probabilities.

energy, E, we have a conical intersection between the potential surfaces at the point a in the classically allowed region $R := \{X : \lambda_-(X) \le E\}$, otherwise we have a positive gap between the potential surfaces in the domain R. Here, choosing a in the origin gives a larger absolute momentum, $|P|$, at the conical intersection whereas choosing a far from the origin will yield a smaller $|P|$, which yield larger and smaller probabilities to be in excited states, respectively. Figure 9 shows level curves of the eigenvalues, along

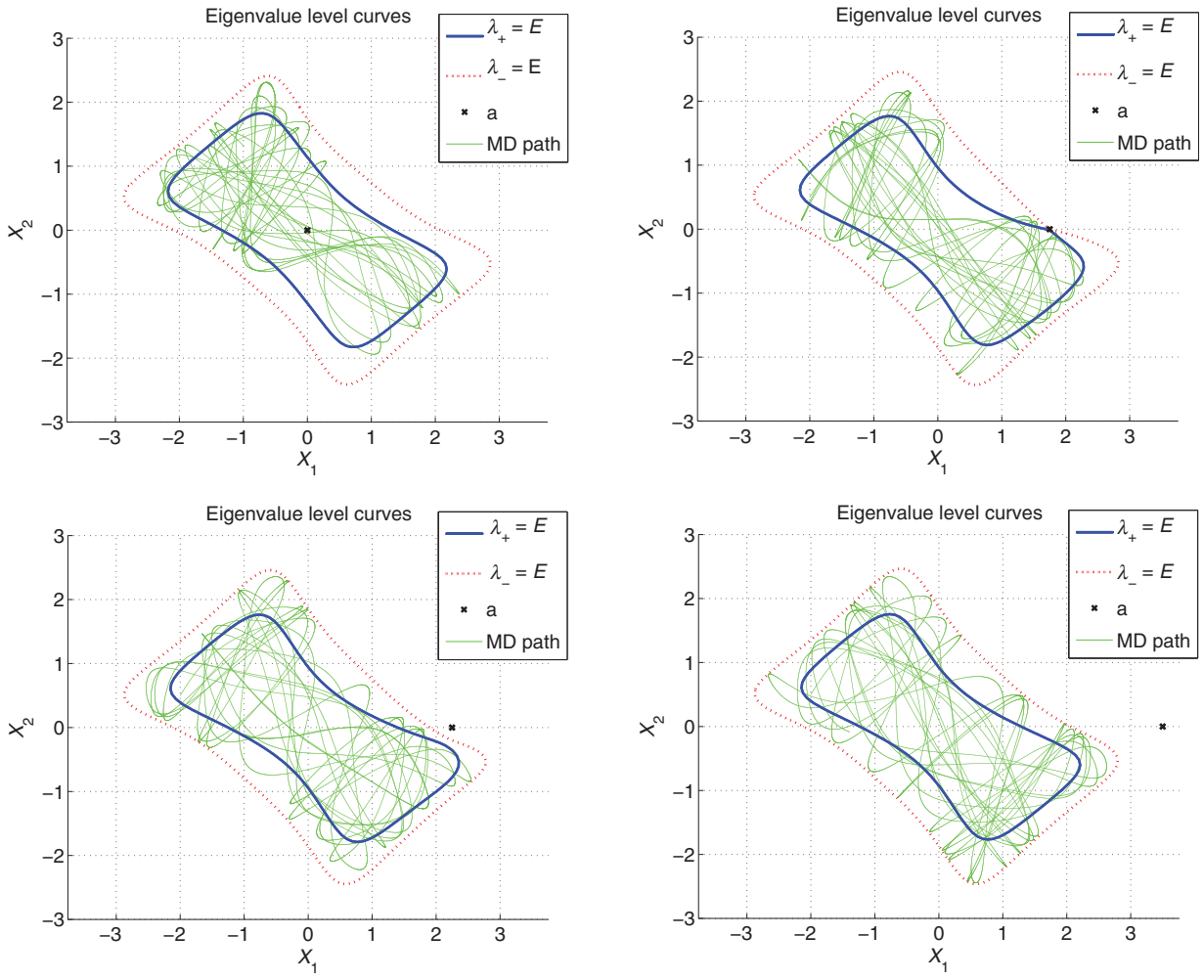

Fig. 9. Level curves of the eigenvalues $\lambda_\pm$ of the potential (90) at the energy $E = 1.5$ with conical intersection points $a = (0,0), (1.75,0), (2.25,0)$, and $(3.5,0)$, respectively. We illustrate examples of the Ehrenfest molecular dynamics (55) paths computed using Störmer–Verlet method (89) with mass $M = 100$, time $t \in [0,100]$, time steps $\Delta t = 1/(32\sqrt{M})$, and initial data $X_0 = [-2, 0.5]$, $P_0 = [1, \sqrt{2(E - \lambda_-(X_0)) - 1}]$, and $\psi_0 = \Psi_-(X_0)$.

with example molecular dynamics paths, for varying conical intersection points.

4.2.1 *The two-dimensional discretized Schrödinger equation*

We use a standard finite difference method to discretize the two-dimensional Schrödinger eigenvalue problem (2) with mesh size h in both the X_1 and X_2 directions in the computational domain $\Omega = [-4,4]^2$. The unknown eigenvalue components $\Phi_h^{j,k} \simeq \Phi(X^{j,k})$, with the nodal point $X^{j,k} = (-4+jh, -4+kh)$ for $j,k = 1,2,\ldots,8/h-1$, solve the discrete eigenvalue problem

$$-\frac{1}{2M}\left(\frac{\Phi_h^{j-1,k}-2\Phi_h^{j,k}+\Phi_h^{j+1,k}}{h^2}+\frac{\Phi_h^{j,k-1}-2\Phi_h^{j,k}+\Phi_h^{j,k+1}}{h^2}\right)+V(X^{j,k})\Phi_h^{j,k}=E_h\Phi_h^{j,k} \tag{91}$$

with homogeneous Dirichlet boundary conditions and $h = 1/(4\sqrt{M})$. We use the solution of the discrete Schrödinger eigenvalue problem to determine the approximate Schrödinger observables

$$g_S = \frac{\sum_{j,k} g(X^{j,k})\rho(X^{j,k})}{\sum_{j,k}\rho(X^{j,k})}, \tag{92}$$

with $g(X):\mathbb{R}^2\to\mathbb{R}$ and $\rho(X) = |\Phi_h(X)|^2$.

The value of the molecular dynamics observable, g_{MD}, is given by

$$g_{\mathrm{MD}} := \lim_{\delta\to 0+}\frac{\int_{E<H_0(X,P)<E+\delta} g(X,P)\,\mathrm{d}X\,\mathrm{d}P}{\int_{E<H_0(X,P)<E+\delta}\mathrm{d}X\,\mathrm{d}P},$$

which, for the two-dimensional case, can be written as (For a fixed $X\in\mathbb{R}^d$, P-integration using spherical coordinates in the shell $E<|P|^2/2+\lambda_0(X)<E+\delta$ yields the P-dependence $[|P|^d]_{\sqrt{E-\lambda_0(X)}}^{\sqrt{E-\lambda_0(X)+\delta}}$ which differentiated with respect to δ gives (93) for $d=2$ and the additional factor $(E-\lambda_0(X))^{d/2-1}$ for $d\neq 2$.)

$$g_{\mathrm{MD}} = \frac{\int_{H_0(X,0)\le E} g(X)\,\mathrm{d}X}{\int_{H_0(X,0)\le E}\mathrm{d}X}, \tag{93}$$

when g depends only on X.

4.2.2 *Verification of the ergodic rate condition* (26)

Figure 10 determines the convergence rate γ in assumption (26) based on numerical approximation of the ergodic projected stochastic dynamics (19). The numerical method

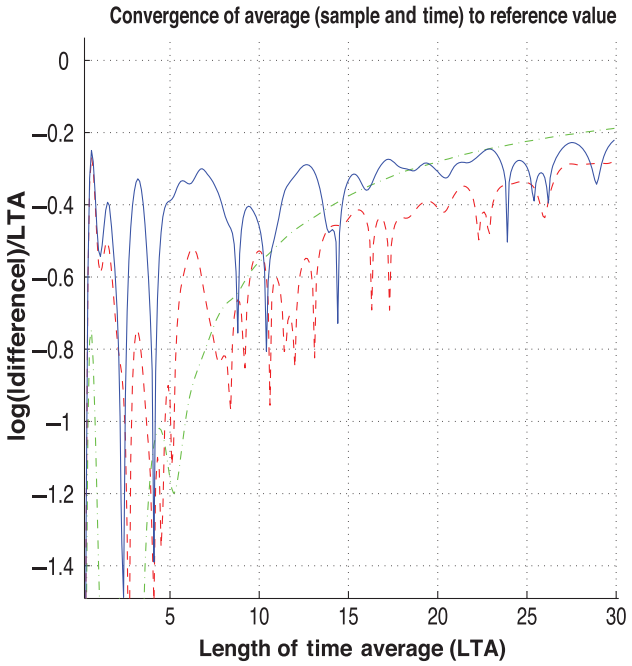

Fig. 10. The figure shows the function $T \mapsto T^{-1}\log(T^{-1}\int_T^{2T}(\bar{\mathbb{E}}[g\circ S_{t0}(Z_0)] - g_{\mathrm{MD}}(Z_0))\,\mathrm{d}t)$, which approximates the decay rate 2γ in assumption (26), for the projected dynamics (19) with the potential $\lambda(X) = X_1^2/2 + X_2^2/\sqrt{2} + 2\sin(X_1X_2)$ and energy $E = H(Z_0) = 1.5$. The empirical mean is $\bar{\mathbb{E}}[g\circ S_{t0}(Z_0)] := \sum_{n=1}^{N}\sin(\bar{X}_1(t;n)\bar{X}_2(t;n))/N$ and the number of stochastic paths are $N = 2\times 10^5$. The initial point is $Z_0 = (0,0,\sqrt{1.5},\sqrt{1.5})$ and $(\bar{X}(t;n),\bar{P}(t;n))$ is the phase-space point for sample path n at time t determined by the numerical method (94) for time steps $\Delta t = 0.01$. The solid curve is for diffusion coefficient $\epsilon = 10^{-4}$, the dashed for $\epsilon = 10^{-2}$, and the dashed dot for $\epsilon = 0.25$. Here the approximation $g_{\mathrm{MD}}(Z_0) = -0.4388$ is determined by an accurate Monte Carlo simulation in the microcanonical ensemble.

is given by the steps $\bar{Z}(n\Delta t)\mapsto \bar{Z}((n+1)\Delta t)$ satisfying

$$\begin{aligned}(X,P) &= \bar{Z}(n\Delta t),\\ X_* &= X + P\Delta t,\\ P_* &= P - \nabla\lambda(X_*)\Delta t,\\ Z^1 &= (X_*,P_*),\\ Z^2 &= Z^1 + \sqrt{\epsilon}\mathbb{P}(Z^1)\Delta W,\\ Z^3 &= Z^2 + \mu\nabla H(Z^2)/|\nabla H(Z^2)|,\\ \bar{Z}((n+1)\Delta t) &= Z^3,\end{aligned} \tag{94}$$

where $\mu\in\mathbb{R}$ is chosen so that $H(Z^2+\mu\nabla H(Z^2)/|\nabla H(Z^2)|) = E$. Here $\Delta W = \sqrt{\Delta t}(\xi_1,\xi_2,\xi_3,\xi_4)$ and ξ_i are all independent taking values ± 1 with probability

$1/2$. The temperature parameter is $\tau = 1/2$. This scheme with phase-space path $\{\bar{Z}(n\Delta t) = (\bar{X}(n\Delta t), \bar{P}(n\Delta t)) \mid n = 0, 1, 2, \ldots\}$ satisfies the convergence assumptions in [12] so that $\mathbb{E}[g(\bar{X}(T), \bar{P}(T)) - g(X(T), P(T))] = \mathcal{O}(\Delta t)$, uniformly in T. (Related to examples (4.15) and (4.7) in [12].) Figure 10 shows that in this case $2\gamma > 0.2$. The maximal Lyapunov constant is determined numerically to $\hat{C} = 0.35$ and the convergence rate exponent $\gamma/(\hat{C} + \gamma)$ in Theorem 2.1 becomes $0.1/0.45 \approx 0.2$.

4.2.3 *Convergence of observables in two dimensions*

Figures 11 and 12 show that the Schrödinger observables, g_{S}, are close to the ergodic molecular dynamics observables, g_{MD}. We can compare the numerical results in Figures 11 and 12 with Theorem 2.2 for the case when the probabilities to be in the excited state, p_{ex}, are small, as in Figure 13 for $a = (5, 0)$ in the conical intersection case. We see in Figures 11 and 12 that the standard deviation of $|g_{\mathrm{S}} - g_{\mathrm{MD}}|$ for $M = 3,200$ is roughly $1/2$ times of the standard deviation for $M = 200$, that is, the numerical results agree roughly with an $\mathcal{O}(M^{-1/4})$ estimate in the theorem, see Table 1.

4.2.4 *Approximation of the probability to be in excited states*

We approximate the probability to be in the excited state, p_{ex}, from the Schödinger equation by the approximation, $\hat{\bar{p}}_{\mathrm{ex}}$, from Algorithm 1, using Ehrenfest molecular dynamics simulations based on the Störmer–Verlet method as in (89). We choose $t_0 = 0$, $P_0 = [1, \sqrt{2(E - \lambda_-(X_0)) - 1}]$, $\psi_0 = \Psi_-(X_0)$, $X_0 = [2, -0.5]$ (line intersection) and $X_0 = [-2, 0.5]$ (conical intersection).

4.2.5 *Conclusions*

Figure 14 shows the probabilities, p_{ex}, to be in the excited state determined from the discrete Schrödinger eigenvalue problem (91) based on the formula (87). We note that for large gaps, in the right column, the probabilities to be in the excited state, p_{ex}, are either small or close to one, as expected, while for small gaps, in the left column, p_{ex} takes values in the whole interval $[0, 1]$. We also observe the lack of the structure present in the analogous Figure 4 for the one-dimensional case, due to not having the resonances ordered with the eigenvalues as in the one-dimensional case discussed in Remark 3.4.

Figure 13 shows p_{ex} determined from the discrete Schrödinger eigenvalue problem and from molecular dynamics simulations for both line and conical intersection cases computed by Algorithm 1. The shapes of the respective methods' solutions

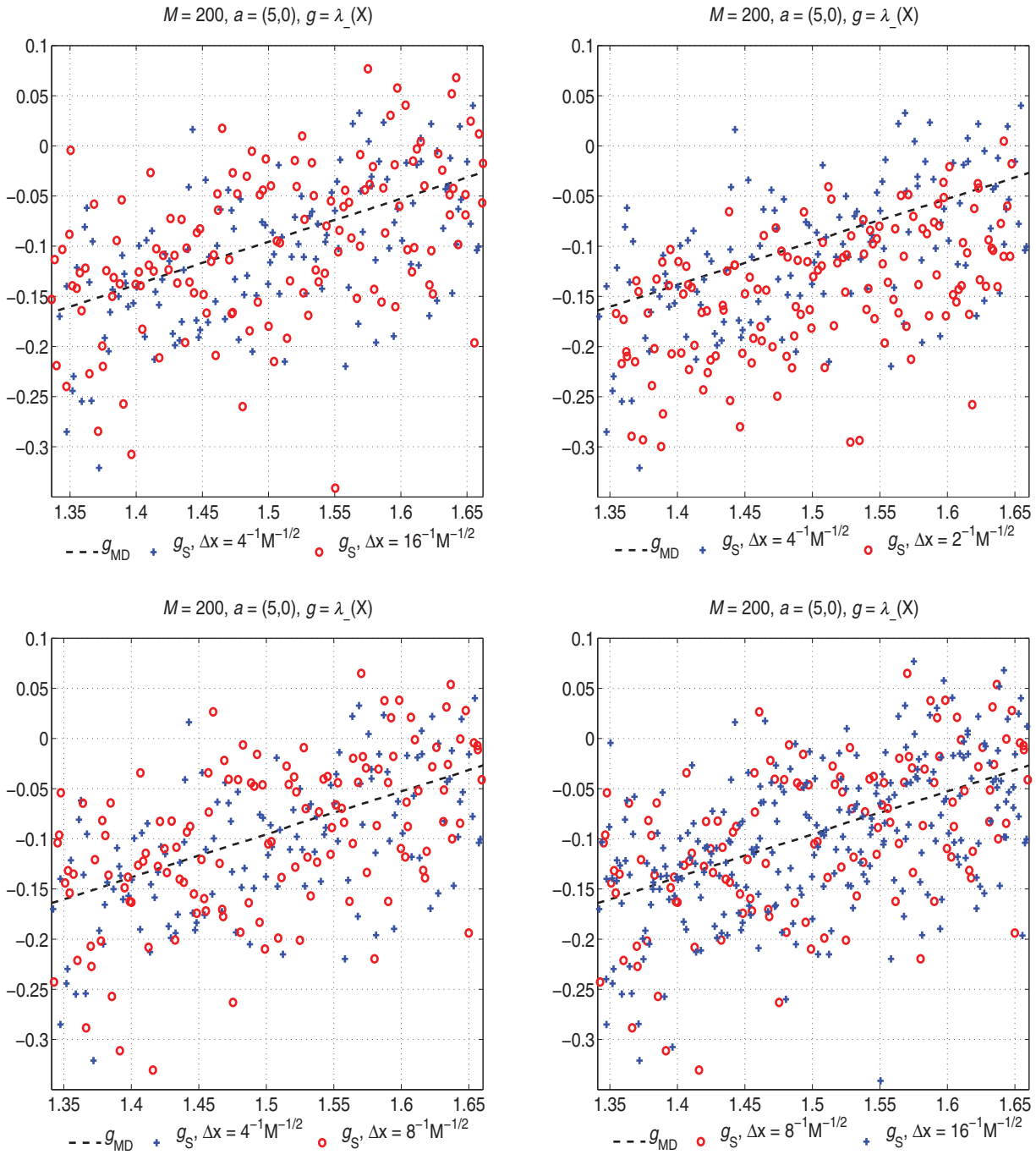

Fig. 11. Schrödinger observables, g_S, as a function of Schrödinger eigenvalues, E_h, for mass $M = 200$, and the conical intersection case with $a = (5, 0)$ outside the classically allowed region, compared with the corresponding molecular dynamics observables, g_{MD}, which we compute using Monte Carlo integrations based on the formula (93). The plots show that the solutions obtained with the mesh size $h = 1/(4\sqrt{M})$, $h = 1/(8\sqrt{M})$, and $h = 1/(16\sqrt{M})$ are comparable, whereas the solution obtained with the mesh size $h = 1/(2\sqrt{M})$ appears less accurate.

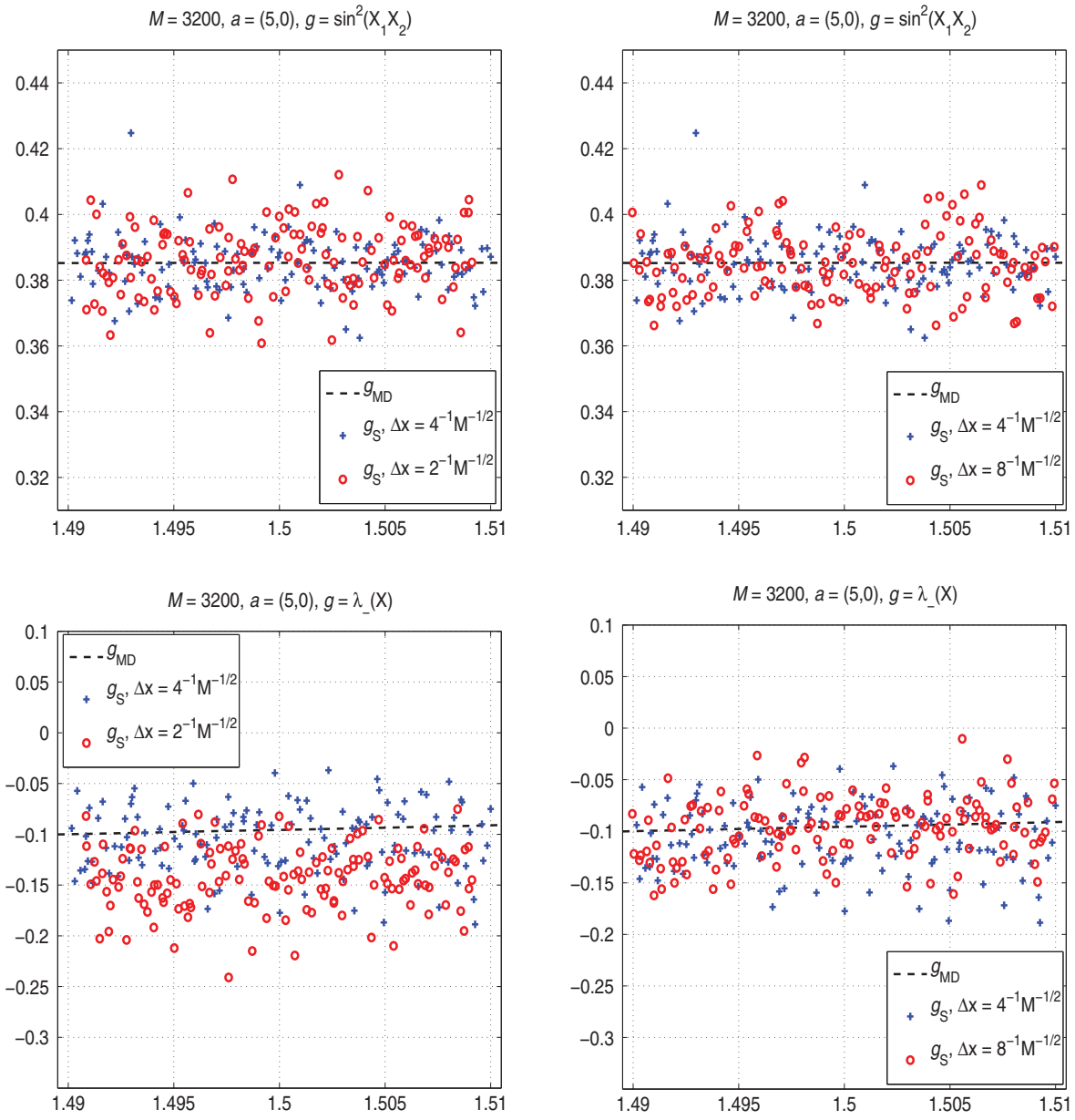

Fig. 12. Schrödinger observables, g_S, as a function of Schrödinger eigenvalues, E_h, for mass $M = 3,200$, and the conical intersection case with $a = (5, 0)$ outside the classically allowed region, compared with the corresponding molecular dynamics observables, g_{MD}, which we compute using Monte Carlo integrations based on the formula (93). The plots show that the solutions obtained with the mesh size $h = 1/(4\sqrt{M})$ and $h = 1/(8\sqrt{M})$ are comparable, whereas the solutions obtained with the mesh size $h = 1/(2\sqrt{M})$ appear less accurate.

are similar, especially when the probability to be in the exited state, p_{ex}, is small. Figure 15 presents an empirical probability density of inter-arrival hitting times of hyperplanes, cf. Algorithm 1. The empirical density is well approximated by a heavy

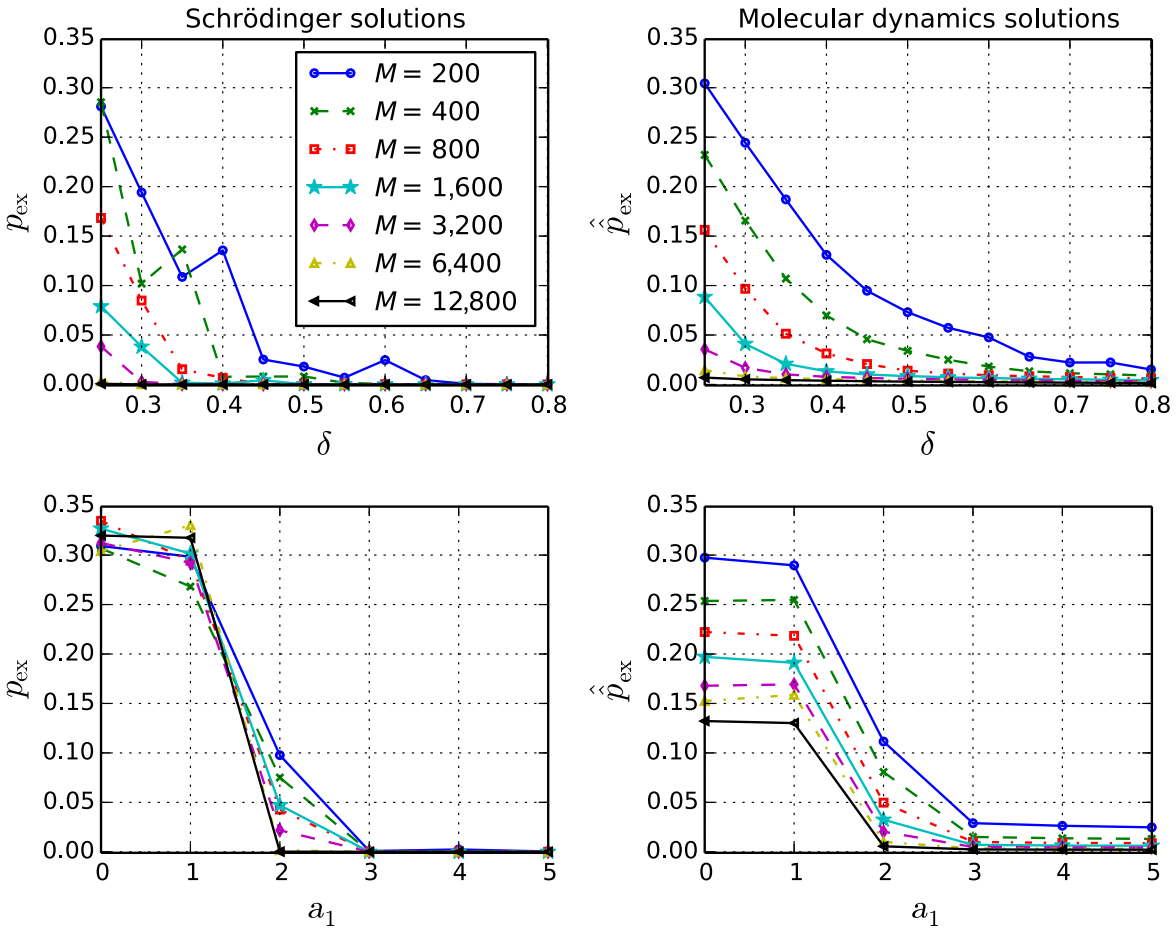

Fig. 13. Plots showing p_{ex} estimated from the solution of the discrete Schrödinger eigenvalue problems (left column) and the corresponding $\hat{p}_{\mathrm{ex}}$ estimated by the molecular dynamics Algorithm 1 (right column) for the line intersection cases (first row) and conical intersection cases (second row). The Schrödinger solution is computed by averaging the solutions for 20 eigenvalues around $E = 1.5$ for the line intersection and by averaging the solutions for 500 eigenvalues around $E = 1.5$ for the conical intersection.

tailed Burr XII probability density, which indicates that long inter-arrival hitting times are non-negligible. In Figure 16, we observe an expected decay of $\mathcal{O}(T^{-1/2})$ for the relative error $|\hat{\hat{p}}_{\mathrm{ex}}(T) - \hat{\hat{p}}_{\mathrm{ex}}|/\hat{\hat{p}}_{\mathrm{ex}}$ when studying the conical intersection problem with $a = (5, 0)$. In the more unstable parameter setting of $a = (0, 0)$, however, the error decay is slower with an observed rate of convergence slightly slower than the expected $-1/2$. This might be due to a longer correlation length in the molecular dynamics and hyperplane sampling of Algorithm 1 when the parameter values of a are close to $(0, 0)$.

Table 1 Numerical results for the order of convergence of the observables in two-dimensional as explained in Section 4.2.3

	$g = \sin^2(X_1 X_2)$	$g = \lambda_-(X)$
e_∞	−0.3079	−0.3159
e_W	−0.3052	−0.2506
e_1	−0.2744	−0.2950
e_2	−0.2671	−0.2890

The table shows ϱ where the order of convergence is $\mathcal{O}(M^\varrho)$. The convergence rate is computed as $\varrho = (\log e(M_1) - \log e(M_2))/(\log M_1 - \log M_2)$, with $e(M_i)$ being the norm of error corresponding to the mass M_i, where $M_1 = 200$, and $M_2 = 3200$. We use the norms $e_\infty = \max_i |(g_S - g_{MD})(E_i)|$, $e_W = |\sum_i (g_S - g_{MD})(E_i)|/N_E$, $e_1 = \sum_i |(g_S - g_{MD})(E_i)|/N_E$, and $e_2 = \sqrt{\sum_i ((g_S - g_{MD})(E_i))^2/N_E}$, with N_E being the number of eigenvalues E_i.

Remark 4.1 (Extension to higher dimension). The formation of a conical intersection for a 2×2 symmetric matrix V requires the eigenvalues to be equal, which means that the off diagonal element is zero and the diagonal elements are equal at an intersection point. To have a conical intersection point is therefore generic in dimension 2 and in higher dimensions the intersection is typically a codimension 2 set. An implementation of the algorithm for computing $\hat{\hat{p}}_{ex}$ in higher dimensions and for potentials used in chemistry is future work. The relevant geometry in the multi-dimensional case is the distance to this codimension 2 set, which makes the link to two-dimensional case studied here. □

5 Regularity and Weyl Calculus

This section establishes four lemmas where the last three are used in Theorems 2.1 and 2.2. The first lemma proves the bound (24) on derivatives of the solution $u(t, Z) = \mathbb{E}[g \circ S_{Tt}(Z)]$ to (22), which verifies that assumption (v) in Theorem 2.1 holds for a case with a near crossing of eigenvalues. The second lemma derives L^2-bounds on Weyl quantized operators obtained from mollified symbols, the third and fourth lemmas estimate remainder terms related to the Weyl quantization and establishes the Moyal expansion for the setting used in the proofs of Theorems 2.1 and 2.2.

We will use equation (21) written as the Ito equation

$$\mathrm{d}Z_t = a(Z_t)\,\mathrm{d}t + b(Z_t)\,\mathrm{d}\tilde{W}_t \tag{95}$$

for the molecular dynamics path $Z : [0, T] \times \Omega \to \mathbb{R}^{6N}$, where Ω is the sample space and assume that the Jacobians a' and b' and the higher order derivatives $\partial_Z^\alpha a(Z) \in \mathbb{R}^{(6N)^{(|\alpha|+1)}}$,

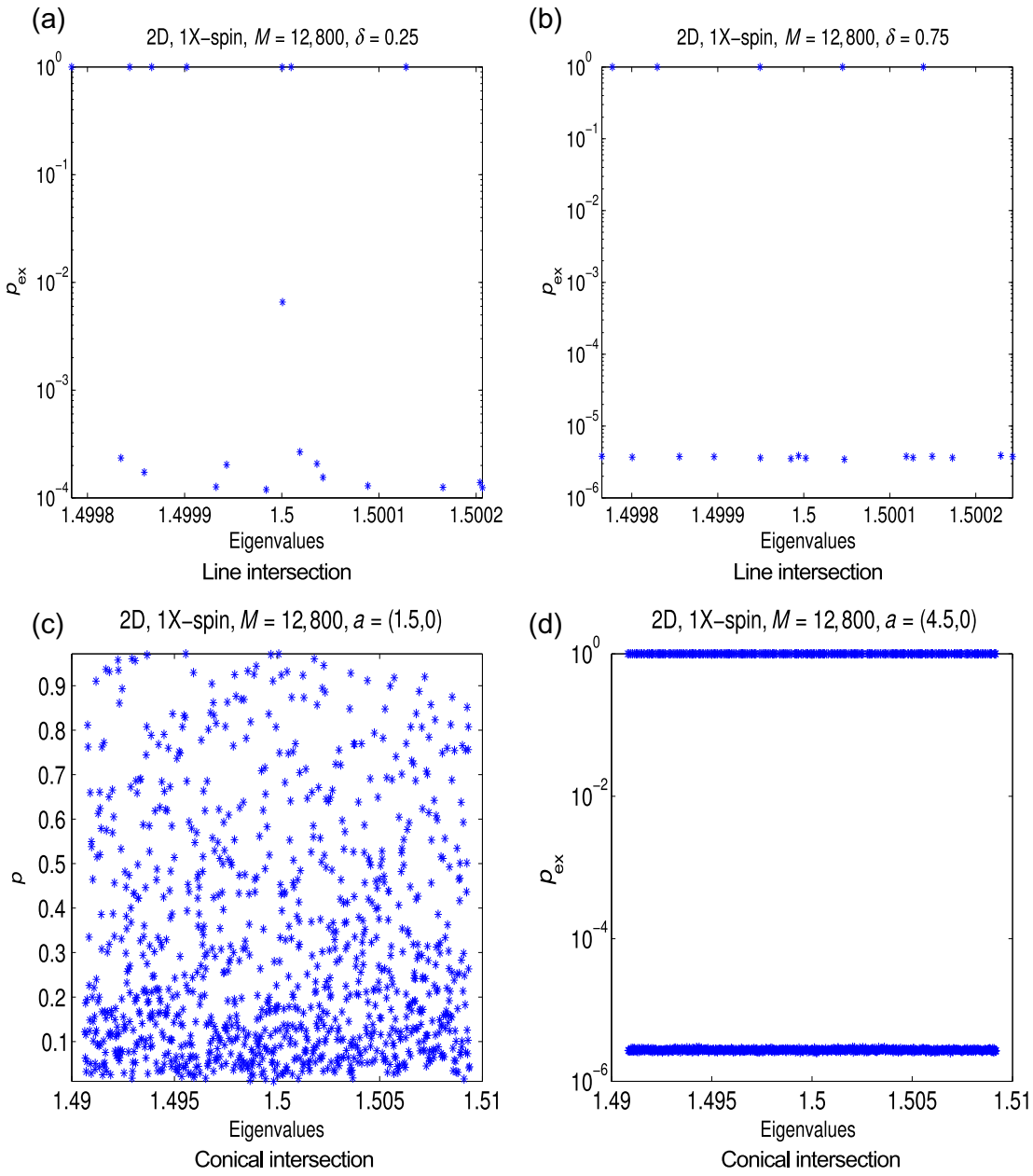

Fig. 14. Probabilities, p_{ex}, to be in the excited state as a function of the eigenvalues, E_h, computed from the solution of the two-dimensional Schrödinger equations and based on the formula (87). The first row corresponds to the line intersection for two different values of δ; the second row corresponds to the conical intersection case for the intersection point $a = (1.5, 0)$ inside and $a = (4.5, 0)$ outside the classically allowed region. We note that the plots contain the probabilities to be in the excited state, p_{ex}, for 20 and 1,000 eigenvalues for line and conical intersection cases, respectively. (a) Line intersection, (b) line intersection, (c) conical intersection, and (d) conical intersection.

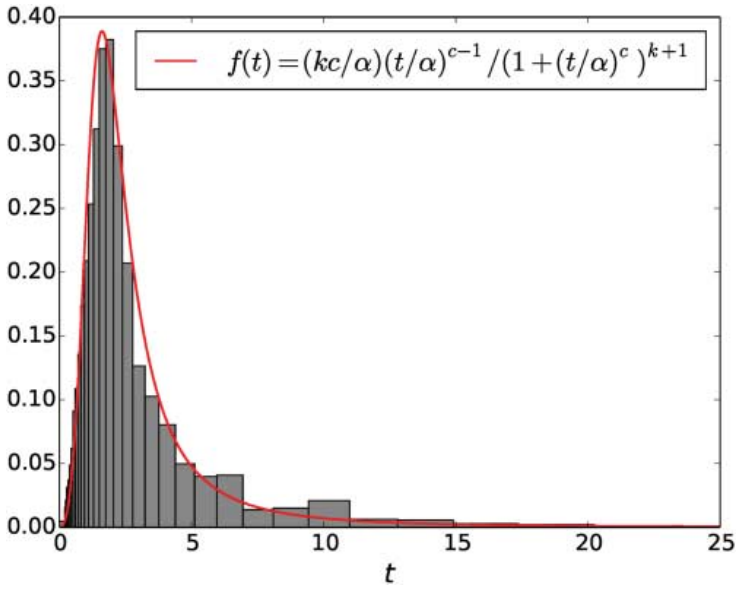

Fig. 15. Normalized histogram of the inter-arrival initialization times obtained by Algorithm 1 when applied to the conical intersection problem with $a = (5, 0)$ and $M = 12,800$, and, for comparison, the probability density function of the Burr XII distribution with parameter values $\alpha = 1.65$, $c = 3.6$, and $k = 0.5$. We do not know why this distribution fits reasonably well to the data.

with $|\alpha| = n$ for $n = 1, 2, 3, \ldots$, satisfy the ℓ^2 bounds

$$\begin{aligned} \|a'\|_{2,\infty} + \|b'\|_{2,\infty}^2 &= \mathcal{O}(\delta^{-1}), \\ \|\partial_Z^\alpha a\|_{2,\infty} + \|\partial_Z^\alpha b\|_{2,\infty}^2 &= \mathcal{O}(\delta^{-|\alpha|}), \end{aligned} \tag{96}$$

uniformly in N (but not in $|\alpha|$). Here we use the ℓ^2 (Frobenius) norm: assume that $|\alpha| = n$ then $\|\partial_Z^\alpha a\|_2 = \sqrt{\sum_{|\alpha|=n} \sum_{i=1}^{6N} |\partial_Z^\alpha a_i|^2}$ and the notation

$$\|a(Z_t)\|_{2,\infty} := \sup_{\omega \in \Omega} \|a(Z_t(\omega))\|_2.$$

Lemma 5.1. Assume (96) and that the near crossing is weak, with nonvanishing velocity P across the near crossing, namely that

$$\begin{aligned} \max_{X \in \mathbb{R}^{3N}} \|\lambda_\eta''(X)\|_2 &\le c_\delta \delta^{-1}, \\ \int_0^T \|\lambda_\eta''(X_t)\|_{2,\infty} \, \mathrm{d}t &\le C\,T, \\ \int_0^T (\mathbb{E}[\|\partial_X^\alpha \lambda_\eta''(X_t)\|_2^{2n}])^{1/(2n)} \, \mathrm{d}t &\le C_\alpha T \delta^{-|\alpha|}, \end{aligned} \tag{97}$$

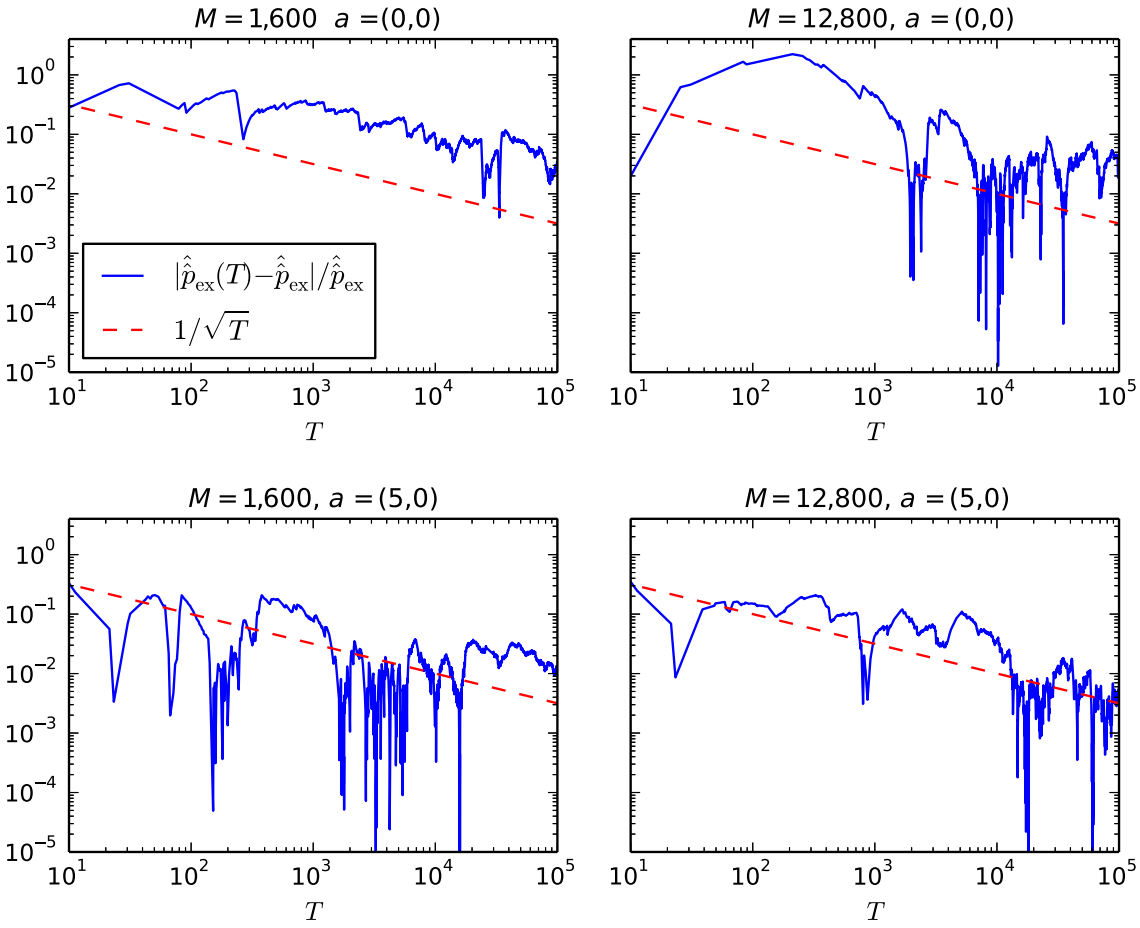

Fig. 16. Convergence plots for $\hat{\hat{p}}_{\mathrm{ex}}$ obtained for the hyperplane sampling and molecular dynamics simulation methods in Algorithm 1 applied to the conical intersection problem.

for a constant $c_\delta \leq 1/\log \delta^{-1}$ and constants C_α independent of δ, M, N, and ϵ. Then for each multi-index β there is a constant $\hat{C}$, independent of M, δ, N, and ϵ, such that

$$\sup_{Z_0 \in \mathbb{R}^{6N}} \|\partial_{Z_0}^\beta \mathbb{E}[g \circ S_T(Z_0) \mid Z_0]\|_2 \leq \mathrm{e}^{\hat{C}(T+1)} \delta^{\min(0,-|\beta|+1)}. \tag{98}$$

□

Proof. Remark 5.2 motives the assumptions. First we study solutions $w : [0,T] \times \Omega \to \mathbb{R}^{6N}$ to the linearized Ito equation

$$\mathrm{d}w_t = a'(Z_t) w_t \,\mathrm{d}t + b'(Z_t) w_t \,\mathrm{d}\tilde{W}_t + \alpha_t \,\mathrm{d}t + \beta_t \,\mathrm{d}\tilde{W}_t,$$

where Z_t solves (95)–(21) and the stochastic processes $\alpha : [0,T] \times \Omega \to \mathbb{R}^{6N}$ and $\beta : [0,T] \times \Omega \to \mathbb{R}^{6N \times 6N}$ satisfy for each $n \in \mathbb{N} \setminus 0$ and some positive number m the bound

$$\begin{aligned} \int_0^t \mathbb{E}[\|\beta_s\|_{2n}^{2n} \mid Z_0]\,\mathrm{d}s &= \mathcal{O}(\mathrm{e}^{Ct}), \\ \int_0^t (\mathbb{E}[\|\alpha_s\|_2^{2n} \mid Z_0])^{1/(2n)}\,\mathrm{d}s &= \mathcal{O}(\mathrm{e}^{Ct}\delta^{-m}). \end{aligned} \tag{99}$$

Ito's formula implies

$$\begin{aligned} \mathrm{d}\mathbb{E}[\|w_t\|_2^2 \mid Z_0] &= 2\mathbb{E}[w_t \cdot \mathrm{d}w_t \mid Z_0] + \mathbb{E}[\mathrm{d}w_t \cdot \mathrm{d}w_t \mid Z_0] \\ &\le 2\mathbb{E}[\|a'(Z_t)\|_2 \|w_t\|_2^2 \mid Z_0]\,\mathrm{d}t + \mathbb{E}[\|b'(Z_t)\|_2^2 \|w_t\|_2^2 \mid Z_0]\,\mathrm{d}t \\ &\quad + 2\mathbb{E}[\|w_t\|_2 \|\alpha_t\|_2 \mid Z_0]\,\mathrm{d}t + \mathbb{E}[\|\beta\|_2^2 + 2\|\beta_t\|_2 \|b_t'\|_2 \|w_t\|_2 \mid Z_0]\,\mathrm{d}t \\ &\le (2\|a'(Z_t)\|_{2,\infty} + \|b'(Z_t)\|_{2,\infty}^2)\mathbb{E}[\|w_t\|_2^2 \mid Z_0]\,\mathrm{d}t + \mathbb{E}[2\|w_t\|_2 \|\alpha_t\|_2 + \|\beta_t\|_2^2 \mid Z_0]\,\mathrm{d}t, \end{aligned}$$

where by assumption (96) we have $2\|a_s'\|_{2,\infty} + \|b_s'\|_{2,\infty}^2 = \mathcal{O}(\delta^{-1} + \epsilon\delta^{-2}) = \mathcal{O}(\delta^{-1})$ and by (97) there holds $\int_0^t 2\|a_s'\|_{2,\infty} + \|b_s'\|_{2,\infty}^2\,\mathrm{d}s = \mathcal{O}(t)$. Here $\|b'\|_{2,\infty}^2 = \sum_i \|b_{i\cdot\cdot}'\|_{2,\infty}^2$. We use $\epsilon \ll \delta$, with the arbitrarily small diffusion constant ϵ in (19), so that the contribution from b' and β are negligible compared with those from a' and α. Gronwall's inequality shows that

$$\begin{aligned} \mathbb{E}[\|w_t\|_2^2 \mid Z_0] &\le \left(\mathbb{E}[\|w_0\|_2^2 \mid Z_0] + \int_0^t \mathbb{E}[2\|w_s\|_2 \|\alpha_s\|_2 + \|\beta_s\|_2^2 \mid Z_0]\,\mathrm{d}s\right) \mathrm{e}^{\int_0^t 2\|a_s'\|_{2,\infty} + \|b_s'\|_{2,\infty}^2\,\mathrm{d}s} \\ &\le \left(\mathbb{E}[\|w_0\|_2^2 \mid Z_0] + \int_0^t 2\sqrt{\mathbb{E}[\|w_t\|_2^2 \mid Z_0]}\sqrt{\mathbb{E}[\|\alpha_s\|_2^2 \mid Z_0]} + \mathbb{E}[\|\beta_s\|_2^2 \mid Z_0]\,\mathrm{d}s\right) \mathrm{e}^{Ct}, \end{aligned} \tag{100}$$

which by (99) implies

$$\begin{aligned} \sup_{0<s<t} \sqrt{\mathbb{E}[\|w_s\|_2^2 \mid Z_0]} &\le \mathrm{e}^{Ct}\left(\sqrt{\mathbb{E}[\|w_0\|_2^2 \mid Z_0]} + \int_0^t 2\sqrt{\mathbb{E}[\|\alpha_s\|_2^2 \mid Z_0]}\,\mathrm{d}s + \mathcal{O}(\mathrm{e}^{Ct})\right) \\ &= \mathrm{e}^{Ct}\mathcal{O}(\delta^{-m}) \end{aligned} \tag{101}$$

provided this holds initially for $t = 0$. Analogously for $n > 1$, we obtain by Hölder's inequality

$$\begin{aligned} &\mathrm{d}\mathbb{E}[\|w_t\|_2^{2n} \mid Z_0] \\ &= n\mathbb{E}\left[\|w_t\|_2^{2(n-1)} w_t \cdot \mathrm{d}w_t + n(n-1)\|w_t\|_2^{2(n-2)} \frac{(w_t \cdot \mathrm{d}w_t)^2}{2} + n\|w_t\|_2^{2(n-1)} \frac{\mathrm{d}w_t \cdot \mathrm{d}w_t}{2} \mid Z_0\right] \\ &= n\mathbb{E}[\|w_t\|_2^{2(n-1)}(w_t \cdot a_t' w_t\,\mathrm{d}t + w_t \cdot \alpha_t\,\mathrm{d}t) + h.o.t. \mid Z_0] \\ &\le n\mathbb{E}[\|w_t\|_2^{2n}\|a_t'\|_2 \mid Z_0]\,\mathrm{d}t + n\mathbb{E}[\|w_t\|_2^{2n-1}\|\alpha_t\|_2 \mid Z_0]\,\mathrm{d}t + \mathrm{h.o.t.} \\ &\le n\mathbb{E}[\|w_t\|_2^{2n}\|a_t'\|_2 \mid Z_0]\,\mathrm{d}t + n(\mathbb{E}[\|w_t\|_2^{2n} \mid Z_0])^{1-1/(2n)}(\mathbb{E}[\|\alpha_t\|_2^{2n} \mid Z_0])^{1/(2n)}\,\mathrm{d}t + h.o.t. \end{aligned}$$

so that

$$\sup_{0<s<t}(\mathbb{E}[\|w_s\|_2^{2n}\,|Z_0])^{1/(2n)} \le \mathrm{e}^{Ct}\left((\mathbb{E}[\|w_0\|_2^{2n}\,|Z_0])^{1/(2n)} + \int_0^t 2(\mathbb{E}[\|\alpha_t\|_2^{2n}\,|\,Z_0])^{1/(2n)}\,\mathrm{d}s + \mathcal{O}(\mathrm{e}^{Ct})\right)$$
$$= \mathrm{e}^{Ct}\mathcal{O}(\delta^{-m}) \tag{102}$$

provided this holds initially for $t=0$.

The first variation $Z_t' = \partial Z_t/\partial Z_0$ satisfies

$$\mathrm{d}Z_t' = a'(Z_t)Z_t'\,\mathrm{d}t + b'(Z_t)Z_t'\,\mathrm{d}\tilde{W}_t,$$

so that (101), for $m=0$, implies

$$\mathbb{E}[\|Z_t'\|_2^{2n}\,|\,Z_0] = \mathcal{O}(\mathrm{e}^{Ct}).$$

The second variation satisfies

$$\mathrm{d}Z_t'' = a'(Z_t)Z_t''\,\mathrm{d}t + b'(Z_t)Z_t''\,\mathrm{d}\tilde{W}_t + a''(Z_t)Z_t'Z_t'\,\mathrm{d}t + b''(Z_t)Z_t'Z_t'\,\mathrm{d}\tilde{W}_t.$$

We have for $\alpha = a''Z'Z'$ that

$$\|a''Z'Z'\|_2^2 = \sum_{ijk}\left(\sum_{rq} a''_{irq}Z'_{rj}Z'_{qk}\right)^2$$
$$\le \sum_{ijk}\sum_{rq}(a''_{irq})^2\sum_{rq}(Z'_{rj}Z'_{qk})^2$$
$$= \sum_{irq}(a''_{irq})^2\sum_{rqjk}(Z'_{rj}Z'_{qk})^2$$
$$= \|a''\|_2^2\|Z'\|_2^4, \tag{103}$$

and similarly for β, which combined with Ito's formula and Gronwall's inequality as above imply

$$\sqrt{\mathbb{E}[\|Z_t''\|_2^2\,|Z_0]} = \mathcal{O}(\delta^{-1}\,\mathrm{e}^{Ct})$$

using (97), that (102) yields $(\mathbb{E}[\|Z_t'\|_2^8\,|\,Z_0])^{1/4} = \mathcal{O}(\mathrm{e}^{Ct})$ and

$$\int_0^t \sqrt{\mathbb{E}[\|a_s''Z_s'Z_s'\|_2^2\,|\,Z_0]}\,\mathrm{d}s \le \int_0^t (\mathbb{E}[\|a_s''\|_2^4\,|\,Z_0])^{1/4}(\mathbb{E}[\|Z_s'\|_2^8\,|\,Z_0])^{1/4}\,\mathrm{d}s \le \mathrm{e}^{Ct}\mathcal{O}(\delta^{-1}).$$

Also the higher variations satisfy

$$\mathrm{d}Z_t^{(m)} = a'(Z_t)Z_t^{(m)}\,\mathrm{d}t + b'(Z_t)Z_t^{(m)}\,\mathrm{d}\tilde{W}_t + \alpha_t\,\mathrm{d}t + \beta_t\,\mathrm{d}\tilde{W}_t,$$

where the stochastic processes α and β have the bound (99), since following (103) we have

$$\|a''Z''Z'\|_2^2 \le \|a''\|_2^2\|Z''\|_2^2\|Z'\|_2^2,$$
$$\|a'''Z'Z'Z'\|_2^2 \le \|a'''\|_2^2\|Z'\|_2^6,$$

and similarly for the higher order derivatives. Therefore, (101) and (102) yield

$$(\mathbb{E}[\|Z_t^{(m)}\|_2^{2n} \mid Z_0])^{1/(2n)} = \mathcal{O}(e^{Ct}\delta^{(1-m)}). \tag{104}$$

We have by Jensens inequality, for $|\beta| = n$,

$$\begin{aligned}\sup_{Z_0}\|\mathbb{E}[\partial_{Z_0}^\beta g(Z_t) \mid Z_0]\|_2 &\le \sup_{Z_0}\mathbb{E}[\|\partial_{Z_0}^\beta g(Z_t)\|_2 \mid Z_0]\\ &\le \sup_{Z_0}\sqrt{\mathbb{E}[\|\partial_{Z_0}^\beta g(Z_t)\|_2^2 \mid Z_0]}\\ &= \begin{cases}\sup_{Z_0}\sqrt{\mathbb{E}[\|g_t'Z_t'\|_2^2 \mid Z_0]}, & n=1,\\ \sup_{Z_0}\sqrt{\mathbb{E}[\|g_t'Z_t'' + g_t''Z_t'Z_t'\|_2^2 \mid Z_0]}, & n=2,\\ \sup_{Z_0}\sqrt{\mathbb{E}[\|g_t'Z_t''' + 2g_t''Z_t''Z_t' + g_t'''Z_t'Z_t'Z_t'\|_2^2 \mid Z_0]}, & n=3,\end{cases}\end{aligned}$$

which by (104) proves (98). ■

Remark 5.2. We motivate assumption (97) by the behavior for the example of the avoided conical intersection eigenvalues $\lambda_0(X) = \pm c_\delta\sqrt{|X|^2+\delta}$ for $X \in \mathbb{R}^2$ or $X \in \mathbb{R}^1$ discussed in Section 4.2 and (54). Although $\max_{X\in K}\|\lambda_0''(X_t)\|_2 = \mathcal{O}(\delta^{-1})$, we assume the near crossing is weak, that is, that $\|\lambda_0''(X)\|_2 \le c_\delta(\operatorname{dist}(X,a)^2+\delta^2|)^{-1/2}$ for some constant $c_\delta \le 1/\log\delta^{-1}$ and a codimension 2 set $a \subset \mathbb{R}^{3N}$. Then the exponent in Gronwall's inequality has the bound

$$\int_t^T \|\lambda_0''(X_s)\|_2\,\mathrm{d}s \le C(T-t) + c_\delta\int_0^1 \frac{\mathrm{d}s}{\sqrt{c^2s^2+\delta^2}} \le C(T-t) + c^{-1/2}c_\delta\log\delta^{-1}$$

provided the near crossing of potential surfaces is located away from the part of Σ_E where $|P| = 0$. Here we assume that the velocity through the avoided crossing domain is bounded from below by $c > 0$. Similarly, an assumption $\|\partial_X^\alpha\lambda_0''(X)\|_2 \le C(\operatorname{dist}(X,a)^2 + \delta^2|)^{-(|\alpha|+1)/2}$ satisfies the last estimate in (97). □

We will use the s-quantization and the notation

$$\int_{\mathbb{R}^{3N}} \langle \Phi(X), \mathrm{Op}^s[h]\Phi(X)\rangle \,\mathrm{d}X = \int_{\mathbb{R}^{6N}} h(Z)\cdot W^{(s)}(Z)\,\mathrm{d}Z$$

for any smooth symbol h and the Wigner functions $W = W^{(1/2)}$ and $W^{(s)}$ defined in (15).

Lemma 5.3. Assume that $h:\mathbb{R}^{6N}\to\mathbb{C}$ and $f:\mathbb{R}^{6N}\to\mathbb{C}$ can be written as $h=\bar{h}*\phi_\eta$ and $f=\bar{f}*\phi_\eta$, where $\phi_\eta(z) := (2\pi\eta)^{-3N}\,\mathrm{e}^{-|z|^2/(2\eta)}$, and that $\bar{h}$ and $\bar{f}$ have continuous second derivatives with polynomial growth, that is, for some $n\in\mathbb{Z}$ and $m\in\mathbb{Z}$ and $|\alpha|\le 2$ there is a constant C such that $\partial_z^\alpha \bar{h}(z) \le C(1+|z|^2)^m$ and $\partial_z^\alpha \bar{f}(z) \le C(1+|z|^2)^n$ uniformly in z, then for every $s\in[0,1]$

$$\left|\int_{\mathbb{R}^{3N}} \langle \Phi(X), \mathrm{Op}^s[h(z)f(z)]\Phi(X)\rangle \,\mathrm{d}X\right| = \sup_{z\in\mathbb{R}^{6N}} |h(z)f(z)| + \mathcal{O}(\eta),$$

uniformly in N. □

Proof. The proof to estimate the $L^2(\mathbb{R}^{3N})$ operator norm for a symbol r has three steps: to determine a representation of the solution of $\int_{\mathbb{R}^{3N}}\langle\Phi, \mathrm{Op}^s[r]\Phi\rangle\,\mathrm{d}X$ using the FBI transform $T\Phi$, to calculate a Fourier multiplicator applied to r using the representation and in the third step to verify that the product of two FBI transformed functions is the convolution of the Wigner function with a Gaussian.

1. We will use that the components of the Wigner function $W = W^{(1/2)}$ convolved with the Gaussian $\phi_{M^{-1/2}} = (2\pi M^{-1/2})^{-3N}\,\mathrm{e}^{-(|X|^2+|P|^2)M^{1/2}/2}$ is the product of the FBI transforms (76) of the components Φ_i and Φ_j

$$W_{ij} * \phi_{M^{-1/2}}(X,P) = T\Phi_i(X,P)\overline{T\Phi_j(X,P)} \tag{105}$$

as verified in Step 3. Consequently, the diagonal entries are non-negative

$$W_{ii} * \phi_{M^{-1/2}}(X,P) = |T\Phi_i^*(X,P)|^2 \ge 0.$$

An s_*-quantization remainder $\hat{r}^{s_*}$ can be related to a remainder $\hat{r} = \mathrm{Op}[h]$ in the Wigner quantization (with $s = 1/2$) by

$$\int_{\mathbb{R}^{6N}} r(X,P)\cdot W^{(s_*)}(X,P)\,\mathrm{d}X\,\mathrm{d}P = \int_{\mathbb{R}^{6N}} r^{s_*}(X,P)\cdot W^{(1/2)}(X,P)\,\mathrm{d}X\,\mathrm{d}P,$$

where for $s_*\in[0,1]$

$$r^{s_*}(X,P) = \mathrm{e}^{-\mathrm{i}M^{-1/2}(s_*-\frac{1}{2})\nabla_X\cdot\nabla_P} r(X,P),$$

which is proved in [32, Remark 2.7.3] and related to the expansion of the exponential in (112).

We will introduce the convolution with the Gaussian in the estimate of the remainder using the Taylor expansion of the exponential

$$\begin{aligned}\int_{\mathbb{R}^{3N}} \langle \Phi, \hat{r}^{s_*}\Phi\rangle \,\mathrm{d}X &= \int_{\mathbb{R}^{6N}} r^{s_*}(X,P)\cdot W^{(1/2)}(X,P)\,\mathrm{d}X\,\mathrm{d}P\\ &= \int_{\mathbb{R}^{6N}} \mathcal{F}r^{s_*}(\xi_X,\xi_P)\cdot \mathcal{F}W^{(1/2)}(\xi_X,\xi_P)\,\mathrm{d}\xi_X\,\mathrm{d}\xi_P\\ &= \int_{\mathbb{R}^{6N}} (\mathcal{F}\phi_{M^{-1/2}}(\xi_X,\xi_P))^{-1}\mathcal{F}r^{s_*}(\xi_X,\xi_P)\\ &\quad\cdot \mathcal{F}W^{(1/2)}(\xi_X,\xi_P)\mathcal{F}\phi_{M^{-1/2}}(\xi_X,\xi_P)\,\mathrm{d}\xi_X\,\mathrm{d}\xi_P\\ &= \int_{\mathbb{R}^{6N}} r^*(X,P)\cdot W^{(1/2)} * \phi_{M^{-1/2}}(X,P)\,\mathrm{d}X\,\mathrm{d}P,\end{aligned}$$

where the function r^* is defined by $r^* = \mathcal{F}^{-1}\{(\mathcal{F}\phi_{M^{-1/2}}(\xi_X,\xi_P))^{-1}\mathcal{F}r^{s_*}(\xi_X,\xi_P)\}$. Here we use the notation $r\cdot W := \sum_{ij} r_{ij}W_{ij}$. The tensor product property (105) of $W * \phi_{M^{-1/2}}$ implies the matrix norm estimate

$$\begin{aligned}\int_{\mathbb{R}^{3N}} \langle \Phi(X), \hat{r}^{s_*}\Phi(X)\rangle \,\mathrm{d}X &= \int_{\mathbb{R}^{6N}} r^*(X,P)\cdot W^{(1/2)} * \phi_{M^{-1/2}}(X,P)\,\mathrm{d}X\,\mathrm{d}P\\ &= \int_{\mathbb{R}^{6N}} \langle T\Phi(X,P), r^*(X,P)T\Phi(X,P)\rangle \,\mathrm{d}X\,\mathrm{d}P\\ &\le \int_{\mathbb{R}^{6N}} \|r^*(X,P)\|_2 |T\Phi(X,P)|^2\,\mathrm{d}X\,\mathrm{d}P\\ &\le \sup_{(X,P)\in\mathbb{R}^{6N}} \|r^*(X,P)\|_2 \int_{\mathbb{R}^{6N}} |T\Phi(X,P)|^2\,\mathrm{d}X\,\mathrm{d}P\\ &= \sup_{(X,P)\in\mathbb{R}^{6N}} \|r^*(X,P)\|_2 \int_{\mathbb{R}^{3N}} \langle T^*T\Phi,\Phi\rangle\,\mathrm{d}X\\ &= \sup_{(X,P)\in\mathbb{R}^{6N}} \|r^*(X,P)\|_2 \underbrace{\int_{\mathbb{R}^{3N}} \langle \Phi(X),\Phi(X)\rangle\,\mathrm{d}X}_{=1}\end{aligned} \tag{106}$$

using the L^2 identity $T^*T = I$ in (77) for the FBI transform and the ℓ^2 matrix norm $\|r^*(X,P)\|_2$ (or the less sharp Frobenius norm).

2. We have $\mathcal{F}r^* = \mathrm{e}^{|\omega|^2 M^{-1/2}/2}\,\mathrm{e}^{\mathrm{i}\omega_x\cdot\omega_p M^{-1/2}}\mathcal{F}r(\omega)$. The goal is to absorb this exponentially growing pre-factor $\mathrm{e}^{|\omega|^2 M^{-1/2}/2}\,\mathrm{e}^{\mathrm{i}\omega_x\cdot\omega_p M^{-1/2}}$ in $\mathcal{F}r(z)$.

Consider first the case $r = \bar{r} * \phi_\eta$, then $\mathcal{F}r^* = \mathrm{e}^{|\omega|^2 M^{-1/2}/2}\,\mathrm{e}^{\mathrm{i}\omega_x\cdot\omega_p M^{-1/2}}\mathcal{F}\phi_\eta(\omega)\mathcal{F}\bar{r}(\omega)$. We have $\mathcal{F}\phi_\eta(\omega) = \mathrm{e}^{-|\omega|^2\eta/2}$ so that

$$\mathrm{e}^{|\omega|^2 M^{-1/2}/2}\,\mathrm{e}^{-|\omega|^2\eta/2} = \mathrm{e}^{-|\omega|^2\nu/2} = \mathcal{F}\phi_\nu(\omega),$$

where $\nu = \eta - M^{-1/2} > 0$. The inverse Fourier transform of the pre-factor $\mathrm{e}^{\mathrm{i}s\omega_x\cdot\omega_p M^{-1/2}}\mathcal{F}\phi_\nu(\omega)$ is

$$\begin{aligned}&(2\pi)^{-6N}\int_{\mathbb{R}^{6N}} \mathrm{e}^{-\nu|\omega_x|^2/2-\nu|\omega_p|^2/2+\mathrm{i}\omega_x\cdot\omega_p sM^{-1/2}+\mathrm{i}\omega_x\cdot x+\mathrm{i}\omega_p\cdot p}\,\mathrm{d}\omega_x\,\mathrm{d}\omega_p\\ &\quad= \mathrm{e}^{-\frac{1}{2\nu(1+s^2M^{-1}\nu^{-2})}(|p|^2+|x|^2)}\mathrm{e}^{-\mathrm{i}p\cdot xsM^{-1/2}\frac{1}{s^2M^{-1}+\nu^2}}(2\pi\nu)^{-3N}\left(1+\frac{s^2}{\nu^2M}\right)^{-3N/2} =: \phi_{\nu,s},\end{aligned}$$

so that $r^* = \bar{r} * \phi_{\nu,s}$. The estimate (106) implies then in the case $r = \bar{r} * \phi_\eta$

$$\begin{aligned}\int_{\mathbb{R}^{3N}}\langle\Phi(X),\hat{r}^{s_*}\Phi(X)\rangle\,\mathrm{d}X &\le \sup_{z\in\mathbb{R}^{6N}}\|r^*(z)\|_2\\ &= \sup_{z\in\mathbb{R}^{6N}}\|\bar{r}*\phi_{\nu,s}(z)\|_2.\end{aligned}$$

If $\bar{r}$ is uniformly bounded in $\mathcal{C}^2(\mathbb{R}^{6N})$, we have by the two moments $\int_{\mathbb{R}^{6N}}\phi_\eta(z)\,\mathrm{d}z = 1$ and $\int_{\mathbb{R}^{6N}} z\phi_\eta(z)\,\mathrm{d}z = 0$ that

$$r - \bar{r} = \mathcal{O}(\eta)$$

and by the properties $\int_{\mathbb{R}^{6N}}\phi_{\nu,s}(z)\,\mathrm{d}z = 1$ and $\int_{\mathbb{R}^{6N}} z\phi_{\nu,s}(z)\,\mathrm{d}z = 0$ we obtain

$$\bar{r} * \phi_{\nu,s} - \bar{r} = \mathcal{O}(\eta)$$

so that

$$\int_{\mathbb{R}^{3N}}\langle\Phi(X),\hat{r}^{s_*}\Phi(X)\rangle\,\mathrm{d}X = \sup_{z\in\mathbb{R}^{6N}}|r(z)| + \mathcal{O}(\nu). \tag{107}$$

In the case $r = (\bar{h}*\phi_\eta)(\bar{f}*\phi_\eta)$, we have similarly

$$\begin{aligned}&(2\pi)^{-12N}\int_{\mathbb{R}^{12N}}\mathcal{F}\bar{h}(\omega')\mathcal{F}\bar{f}(\omega-\omega')\,\mathrm{e}^{-|\omega'|^2\eta/2-|\omega-\omega'|^2\eta/2}\,\mathrm{e}^{\mathrm{i}\omega_x\cdot\omega_p sM^{-1/2}+|\omega|^2/(2M^{-1/2})}\,\mathrm{e}^{\mathrm{i}\omega\cdot z}\,\mathrm{d}\omega' d\omega\\ &= (2\pi)^{-12N}\int_{\mathbb{R}^{12N}}\mathcal{F}\bar{h}(\omega')\mathcal{F}\bar{f}(\omega-\omega')\,\mathrm{e}^{-|\omega|^2\eta/4-|\omega'-\omega/2|^2\eta}\,\mathrm{e}^{\mathrm{i}\omega_x\cdot\omega_p sM^{-1/2}+|\omega|^2/(2M^{-1/2})}\,\mathrm{e}^{\mathrm{i}\omega\cdot z}\,\mathrm{d}\omega'\,\mathrm{d}\omega\\ &= (2\pi)^{-9N}\eta^{-3N}\int_{\mathbb{R}^{12N}}\mathcal{F}^{-1}\{\mathcal{F}\bar{h}\mathcal{F}\{\bar{f}(\omega-\cdot)\}\}(z')\,\mathrm{e}^{-|z'|^2/(4\eta)+\mathrm{i}\omega\cdot z'/2}\,\mathrm{e}^{\mathrm{i}\omega_x\cdot\omega_p sM^{-1/2}-\nu|\omega|^2/2}\,\mathrm{e}^{\mathrm{i}\omega\cdot z}\,\mathrm{d}\omega\,\mathrm{d}z'\\ &= \underbrace{(2\pi\eta)^{-3N}(2\pi\nu)^{-3N}\left(1+\frac{s^2}{\nu^2M}\right)^{-3N/2}}_{=:\sigma_{N,M}}\\ &\quad\times\int_{\mathbb{R}^{12N}}\bar{h}(z'-z'')\,\bar{f}(-z'')\,\mathrm{e}^{-|z'|^2/(4\eta)}\,\mathrm{e}^{-|z-z''+z'/2|^2/(4\nu(s))}\,\mathrm{e}^{\mathrm{i}\frac{(z_x-z''_x+z'_x/2)\cdot(z_p-z''_p+z'_p/2)sM^{-1/2}}{\nu^2+s^2M^{-1}}}\,\mathrm{d}z'\,\mathrm{d}z''\\ &= \sigma_{N,M}\int_{\mathbb{R}^{12N}}\bar{h}(v)\,\bar{f}(w)\,\mathrm{e}^{-|w-v|^2/(4\eta)}\,\mathrm{e}^{-|z+(w+v)/2|^2/(4\nu(s))}\,\mathrm{e}^{\frac{\mathrm{i}(z_x+(w_x+v_x)/2)\cdot(z_p+(w_p+v_p)/2)sM^{-1/2}}{\nu^2+s^2M^{-1}}}\,\mathrm{d}v\,\mathrm{d}w\\ &= \sigma_{N,M}\int_{\mathbb{R}^{12N}}\bar{h}(z-v)\,\bar{f}(z-w)\,\mathrm{e}^{-|w-v|^2/(4\eta)}\,\mathrm{e}^{-|(w+v)/2|^2/(4\nu(s))}\,\mathrm{e}^{\frac{\mathrm{i}(w_x+v_x)\cdot(w_p+v_p)sM^{-1/2}}{4(\nu^2+s^2M^{-1})}}\,\mathrm{d}v\,\mathrm{d}w\\ &=: Q(\bar{h}\bar{f})(z).\end{aligned}$$

We have by the Fourier transform

$$\sigma_{N,M}\int_{\mathbb{R}^{12N}} \mathrm{e}^{-|w-v|^2/(4\eta)}\,\mathrm{e}^{-|(w+v)/2|^2/(4\nu(s))}\,\mathrm{e}^{\frac{\mathrm{i}(w_X+v_X)\cdot(w_p+v_p)sM^{-1/2}}{4(\nu^2+s^2M^{-1})}}\,\mathrm{d}v\,\mathrm{d}w=1$$
$$\int_{\mathbb{R}^{12N}} v\,\mathrm{e}^{-|w-v|^2/(4\eta)}\,\mathrm{e}^{-|(w+v)/2|^2/(4\nu(s))}\,\mathrm{e}^{\frac{\mathrm{i}(w_X+v_X)\cdot(w_p+v_p)sM^{-1/2}}{4(\nu^2+s^2M^{-1})}}\,\mathrm{d}v\,\mathrm{d}w=0$$
$$\int_{\mathbb{R}^{12N}} w\,\mathrm{e}^{-|w-v|^2/(4\eta)}\,\mathrm{e}^{-|(w+v)/2|^2/(4\nu(s))}\,\mathrm{e}^{\frac{\mathrm{i}(w_X+v_X)\cdot(w_p+v_p)sM^{-1/2}}{4(\nu^2+s^2M^{-1})}}\,\mathrm{d}v\,\mathrm{d}w=0$$

so that

$$\begin{aligned}|Q(\bar h\bar f)(z)-\bar h(z)\,\bar f(z)| &= |Q(\bar h(\cdot)\,\bar f(\cdot)-\bar h(z)\,\bar f(z))|\\ &\le C\left(\left\|\frac{\bar h(z)}{1+|z|^m}\right\|^2_{\mathcal{C}^2(\mathbb{R}^{6N})}+\|\bar f(z)(1+|z|^m)\|^2_{\mathcal{C}^2(\mathbb{R}^{6N})}\right)\nu(s).\end{aligned}$$

Similarly using the moments $\int_{\mathbb{R}^{6N}}\phi_\eta(z)\,\mathrm{d}z=1$ and $\int_{\mathbb{R}^{6N}} z\phi_\eta(z)\,\mathrm{d}z=0$ and the second derivatives $\|\bar h\|^2_{\mathcal{C}^2(\mathbb{R}^{6N})}+\|\bar f\|^2_{\mathcal{C}^2(\mathbb{R}^{6N})}\le C$ we obtain

$$h(z)\,f(z)-\bar h(z)\,\bar f(z)=\mathcal{O}(\eta),$$

and we conclude that $|Q(\bar h\bar f)(z)-h(z)\,f(z)|=\mathcal{O}(\eta)$ which combined with (106), as in (107), proves the theorem.

3. This step verifies (105) following the proof of [24, Proposition 1]. We have

$$\begin{aligned}&W_{ij}*\phi_{M^{-1/2}}(X,P)\\ &\quad=(\pi M^{1/2}\times 2\pi M^{-1/2})^{-3N}\int_{\mathbb{R}^{9N}}\mathrm{e}^{\mathrm{i}M^{1/2}Y\cdot P'}\Phi_i\left(X'-\frac{Y}{2}\right)\Phi_j^*\left(X'+\frac{Y}{2}\right)\\ &\quad\times\mathrm{e}^{-(|X-X'|^2+|P-P'|^2)M^{1/2}}\,\mathrm{d}Y\,\mathrm{d}X'\,\mathrm{d}P'\\ &\quad=(\pi M^{1/2})^{-3N/2}(2\pi M^{-1/2})^{-3N}\int_{\mathbb{R}^{6N}}\mathrm{e}^{\mathrm{i}M^{1/2}Y\cdot P}\Phi_i\left(X'-\frac{Y}{2}\right)\Phi_j^*\left(X'+\frac{Y}{2}\right)\\ &\quad\times\mathrm{e}^{-(|X-X'|^2+\frac14|Y|^2)M^{1/2}}\,\mathrm{d}Y\,\mathrm{d}X'.\end{aligned}$$

The change of variables $X'-Y/2=v$ and $X'+Y/2=w$ implies

$$M^{1/2}\left(|X-X'|^2+\frac14|Y|^2\right)=\frac{M^{1/2}}{2}(|v-X|^2+|w-X|^2)$$

and we conclude that the Wigner function and the FBI transform have the relation

$$\begin{aligned}W_{ij}*\phi_{M^{-1/2}}(X,P) &= (2^{1/3}\pi)^{-3N/4}M^{9N/8}\int_{\mathbb{R}^{3N}}\mathrm{e}^{\mathrm{i}M^{1/2}(X-v)\cdot P}\,\mathrm{e}^{-|X-v|^2M^{1/2}/2}\Phi_i(v)\,\mathrm{d}v\\ &\quad\times(2^{1/3}\pi)^{-3N/4}M^{9N/8}\int_{\mathbb{R}^{3N}}\mathrm{e}^{-\mathrm{i}M^{1/2}(X-w)\cdot P}\,\mathrm{e}^{-|X-w|^2M^{1/2}/2}\Phi_j^*(w)\,\mathrm{d}w\\ &= T\Phi_i(X,P)(T\Phi_j(X,P))^*.\end{aligned}$$

■

Lemma 5.4. Assume that $r : \mathbb{R}^{6N} \to \mathbb{C}$ is smooth and for each $n \in \mathbb{N}$ and $|\alpha| \le n$ there is a constant C_n such that $\sup_{X\in\mathbb{R}^{3N}} \int_{\mathbb{R}^{3N}} |\partial_P^\alpha r(X,P)|\,\mathrm{d}P \le C_n$, then there is a constant C such that

$$\left| \int_{\mathbb{R}^{3N}} \langle \Phi(X), \mathrm{Op}^s[r]\Phi(X)\rangle \,\mathrm{d}X \right| \le C \sup_{n\le 3N+3} C_n. \tag{108}$$

If for $|\alpha| \le n$ there is a constant C_n such that $\sup_{z\in\mathbb{R}^{6N}} |\partial_z^\alpha r(z)| \le C_n \delta^{\min(0,-n+1)}$, then there is a constant C such that

$$\left| \int_{\mathbb{R}^{3N}} \langle \Phi(X), \mathrm{Op}^s[r]\Phi(X)\rangle \,\mathrm{d}X \right| \le C \left(\sup_{z\in\mathbb{R}^{6N}} |r(z)| + M^{-1/4}\delta^{-1}N^{1/2} \sup_{n\le CN} C_n \right). \tag{109}$$

□

Proof. A proof if this is given, for example, in [51, Theorem 4.21] and presented here for completeness.

The s-symbol r has the integral kernel

$$K(X,Y) = \left(\frac{1}{2\pi M^{-1/2}}\right)^{3N} \int_{\mathbb{R}^{3N}} \mathrm{e}^{\mathrm{i}M^{1/2}(X-Y)\cdot P} r((1-s)X+sY,P)\,\mathrm{d}P$$

so that

$$\mathrm{Op}^s[r]\Phi(X) = \int_{\mathbb{R}^{3N}} K(X,Y)\Phi(Y)\,\mathrm{d}Y$$

and

$$\begin{aligned} \|\mathrm{Op}^s[r]\Phi(X)\|^2_{L^2(\mathbb{R}^{3N})} &= \int_{\mathbb{R}^{9N}} \bar{K}(Z,Y)K(Z,X)\bar{\Phi}(Y)\Phi(X)\,\mathrm{d}X\,\mathrm{d}Y\,\mathrm{d}Z \\ &\le \frac{1}{2}\int_{\mathbb{R}^{9N}} |K(Z,Y)||K(Z,X)|(|\Phi(Y)|^2+|\Phi(X)|^2)\,\mathrm{d}X\,\mathrm{d}Y\,\mathrm{d}Z. \end{aligned} \tag{110}$$

Here

$$\begin{aligned} &\int_{\mathbb{R}^{9N}} |K(Z,Y)||K(Z,X)||\Phi(Y)|^2\,\mathrm{d}X\,\mathrm{d}Y\,\mathrm{d}Z \\ &\quad\le \sup_{Y\in\mathbb{R}^{3N}} \left(\int_{\mathbb{R}^{6N}} |K(Z,Y)||K(Z,X)|\,\mathrm{d}X\,\mathrm{d}Z \right) \underbrace{\int_{\mathbb{R}^{3N}} |\Phi(Y)|^2\,\mathrm{d}Y}_{=1} \\ &\quad\le \sup_{Y\in\mathbb{R}^{3N}} \int_{\mathbb{R}^{3N}} |K(Z,Y)|\,\mathrm{d}Z \sup_{Z\in\mathbb{R}^{3N}} \int_{\mathbb{R}^{3N}} |K(Z,X)|\,\mathrm{d}X \end{aligned} \tag{111}$$

and similarly for the term with Y and X replaced. We have

$$K(X,Y) = (M^{1/2})^{-3N}\mathcal{F}^{-1}\{r((1-s)X+sY,\cdot)\}((X-Y)M^{1/2})$$

and

$$\begin{aligned}
&\int_{\mathbb{R}^{3N}} |K(X,Y)|\,\mathrm{d}X \\
&\quad= \int_{\mathbb{R}^{3N}} |\mathcal{F}^{-1}\{r((1-s)X'M^{-1/2}+sY,\cdot)\}(X')|\,\mathrm{d}X' \\
&\quad= \int_{\mathbb{R}^{3N}} (1+|X'|^{3N+1})|\mathcal{F}^{-1}\{r((1-s)X'M^{-1/2}+sY,\cdot)\}(X')|\frac{1}{1+|X'|^{3N+1}}\,\mathrm{d}X' \\
&\quad= \int_{\mathbb{R}^{3N}} |\mathcal{F}^{-1}\{(1+\Delta_\cdot^{(3N+1)/2})r((1-s)X'M^{-1/2}+sY,\cdot)\}(X')|\frac{1}{1+|X'|^{3N+1}}\,\mathrm{d}X' \\
&\quad\le \|\mathcal{F}^{-1}\{(1+\Delta_\cdot^{(3N+1)/2})r((1-s)X'M^{-1/2}+sY,\cdot)\}(X')\|_{L^\infty}\int_{\mathbb{R}^{3N}}\frac{1}{1+|X'|^{3N+1}}\,\mathrm{d}X' \\
&\quad\le C\sup_{Y'}\int_{\mathbb{R}^{3N}} |(1+\Delta_P^{(3N+1)/2})r(Y',P)|\,\mathrm{d}P,
\end{aligned}$$

where $\Delta_\cdot$ means the Laplacian with respect to the second variable in r. The function r satisfies $\sup_{Y'\in\mathbb{R}^{3N}}\int_{\mathbb{R}^{3N}}|\partial_P^\alpha r(Y',P)|\,\mathrm{d}P\le C_n$ which together with (110) and (111) proves (108).

In the case $\sup_z|\partial_z^\alpha r(z)|\le C_n\delta^{-|\alpha|}$, [51, Theorem 5.1] applied to the symbol r proves (109). ■

Lemma 5.5. The composition $\hat{C}=\hat{A}\hat{B}$ of two Fourier integral operators, with smooth symbols $A(X)$ and $B(X,P)$ in the Schwartz space, has the Weyl symbol

$$\begin{aligned}
C(X,P) &= \mathrm{e}^{\frac{\mathrm{i}}{2}M^{-1/2}\sum_k(-\partial_{X_k}\partial_{P'_k})}A(X)B(X',P')|_{(X,P)=(X',P')} \\
&= \sum_{n=0}^{2}\frac{(\mathrm{i}M^{-1/2})^n}{2^n n!}(-\nabla_X\cdot\nabla_{P'})^n A(X)B(X',P')\Big|_{(X,P)=(X',P')} + M^{-3/2}r_2. \qquad (112)
\end{aligned}$$

The remainder r_2 is smooth and if $B(X_0,P_0)=\mathbb{E}[\bar{g}\circ S_t(\cdot)\mid X_0,P_0]*\phi_\eta(X_0,P_0)$, and $A(X_0)=\bar{A}*\phi_\eta(X_0)$ the remainder satisfies

$$\int_{\mathbb{R}^{3N}}\langle\Phi,\hat{r}_2\Phi\rangle\,\mathrm{d}X=\mathcal{O}(\mathrm{e}^{\hat{C}t}\delta^{-2}). \qquad (113)$$

If $A\in\mathcal{C}^\infty(\mathbb{R}^{6N})$ and $B\in\mathcal{C}^\infty(\mathbb{R}^{6N})$ are in the Schwartz Space, there holds

$$\begin{aligned}
C(X,P) &= \mathrm{e}^{\frac{\mathrm{i}}{2}M^{-1/2}\sum_k(\partial_{P_k}\partial_{X'_k}-\partial_{X_k}\partial_{P'_k})}A(X,P)B(X',P')|_{(X,P)=(X',P')} \\
&= \sum_{n=0}^{m}\frac{(\mathrm{i}M^{-1/2})^n}{2^n n!}(\nabla_P\cdot\nabla_{X'}-\nabla_X\cdot\nabla_{P'})^n A(X,P)B(X',P')\Big|_{(X,P)=(X',P')} \\
&\quad + M^{-(m+1)/2}r_m, \qquad (114)
\end{aligned}$$

where for $B(X_0, P_0) = \mathbb{E}[\bar{g} \circ S_t(\cdot) \mid X_0, P_0] * \phi_\eta(X_0, P_0)$ and $A(X_0, P_0) = \bar{A} * \phi_\eta(X_0, P_0)$

$$\int_{\mathbb{R}^{3N}} \langle \Phi, \hat{r}_m \Phi \rangle \, \mathrm{d}X = \mathcal{O}(\mathrm{e}^{\hat{C}t} \delta^{-m}), \tag{115}$$

for $A(X) = \nabla V(X) - \nabla \lambda_\eta * \phi_\nu(X)$ and $B(X_0, P_0) = \partial_P \mathbb{E}[\bar{g} \circ S_t(\cdot) \mid X_0, P_0] * \phi_\eta(X_0, P_0)$

$$\int_{\mathbb{R}^{3N}} \langle \Phi, \hat{r}_1 \Phi \rangle \, \mathrm{d}X = \mathcal{O}(\mathrm{e}^{\hat{C}t} \delta^{-4}), \tag{116}$$

and for $A(Z) = H_\eta(Z)$ and $B(Z) = 2 \int_0^1 \tilde{g}''_{\mathrm{MD}}(sH_\eta(Z) + (1-s)E)(1-s) \, \mathrm{d}s$

$$\int_{\mathbb{R}^{3N}} \langle \Phi, \hat{r}_1 \Phi \rangle \, \mathrm{d}X = \mathcal{O}(\mathrm{e}^{\hat{C}t} \delta^{-4}). \tag{117}$$

□

Proof. The proof has five steps. The first step uses the definition of the Weyl quantization to define an integral kernel for the product $\hat{C} = \hat{A}\hat{B}$ following Hörmander's work [22]. The next step formulates the Moyal expansion and identifies the remainder as an average of s-quantizations (16), which proves (112). The third step writes the s-quantized remainder as a Wigner quantization (with $s = 1/2$) and estimates the remainder, using that the product of two FBI transformed functions is the convolution of the Wigner function with a Gaussian. Step 4 proves (114) and (115). Step 5 proves (117).

1. To verify (112), we start with the definition of the Weyl operator

$$\hat{C}\Phi(X) = \underbrace{(2\pi M^{-1/2})^{-3N}}_{=:\gamma} \int_{\mathbb{R}^{6N}} C\left(\frac{X+Y}{2}, P\right) \mathrm{e}^{\mathrm{i}M^{1/2}(X-Y)\cdot P} \Phi(Y) \, \mathrm{d}Y \, \mathrm{d}P.$$

The L^2 inner product

$$\begin{aligned}\int_{\mathbb{R}^{3N}} \hat{C}\Phi(X)\Psi^*(X) \, \mathrm{d}X &= \gamma \int_{\mathbb{R}^{9N}} C\left(\frac{X+Y}{2}, P\right) \mathrm{e}^{\mathrm{i}M^{1/2}(X-Y)\cdot P} \Phi(Y)\Psi^*(X) \, \mathrm{d}P \, \mathrm{d}Y \, \mathrm{d}X \\ &= \int_{\mathbb{R}^{6N}} C_K(X, Y)\Phi(Y)\Psi^*(X) \, \mathrm{d}Y \, \mathrm{d}X\end{aligned}$$

defines the kernel

$$C_K(X, Y) := \gamma \int_{\mathbb{R}^{3N}} C\left(\frac{X+Y}{2}, P\right) \mathrm{e}^{\mathrm{i}M^{1/2}(X-Y)\cdot P} \, \mathrm{d}P, \tag{118}$$

and the inverse Fourier transform implies

$$C(U, P) = \int_{\mathbb{R}^{3N}} C_K\left(U + \frac{Z}{2}, U - \frac{Z}{2}\right) \mathrm{e}^{-\mathrm{i}M^{1/2}Z\cdot P} \, \mathrm{d}Z. \tag{119}$$

Our examples have A independent of P, that is, $A(X,P)=A(X)$, and the kernel of the composition $\hat{A}\hat{B}$ becomes a multiplication

$$C_K(X,Y)=\gamma\int_{\mathbb{R}^{3N}}A(X)B\left(\frac{X+Y}{2},P'\right)\mathrm{e}^{\mathrm{i}M^{1/2}(X-Y)\cdot P'}\,\mathrm{d}P'.$$

The definition $\hat{C}=\hat{A}\hat{B}$ and (119) yields

$$\begin{aligned}C(U,P)&=(2\pi M^{-1/2})^{-3N}\int_{\mathbb{R}^{6N}}A\left(U+\frac{Z}{2}\right)B(U,P+P')\,\mathrm{e}^{\mathrm{i}M^{1/2}P'\cdot Z}\,\mathrm{d}Z\,\mathrm{d}P'\\&=(\pi M^{-1/2})^{-3N}\int_{\mathbb{R}^{6N}}A(U+Z')B(U,P+P')\,\mathrm{e}^{2\mathrm{i}M^{1/2}P'\cdot Z'}\,\mathrm{d}Z'\,\mathrm{d}P'.\end{aligned}$$

The final step in Hörmander's derivation uses the standard Fourier transforms $\mathcal{F}(f)=\hat{f}$ of a Schwartz function $f(X,P)$ and of $\mathrm{e}^{\mathrm{i}M^{1/2}X\cdot P}$ combined with $L^2(\mathbb{R}^{6N})$ Fourier transform isometry and Taylor expansion function of the exponential. Here we modify this by identifying the remainder in the Moyal expansion for $f(X+X',P+P'):=A(X+X')B(X,P+P')$ as an s-quantization (16).

2. We use in this step the Fourier transform $\mathcal{F}_P$ in the P-direction defined for $f(X_0,P_0)\in\mathbb{R}$ by

$$\mathcal{F}_P\{f\}(X_0,Y):=\int_{\mathbb{R}^{3N}}f(X_0,\xi)\,\mathrm{e}^{-\mathrm{i}Y\cdot\xi}\,\mathrm{d}\xi.$$

The remainder is based on a Taylor expansion of the exponential function as follows:

$$\begin{aligned}&\int_{\mathbb{R}^{6N}}f(X+X',P+P')\,\mathrm{e}^{2\mathrm{i}M^{1/2}X'\cdot P'}\,\mathrm{d}X'\,\mathrm{d}P'\\&\quad=\left(\frac{1}{2\pi}\right)^{6N}\left(\frac{\pi}{\sqrt{M}}\right)^{3N}\int_{\mathbb{R}^{6N}}\mathcal{F}f(\xi_X,\xi_P)\,\mathrm{e}^{-\frac{\mathrm{i}}{2}M^{-1/2}\xi_X\cdot\xi_P}\,\mathrm{d}\xi_X\,\mathrm{d}\xi_P\\&\quad=\left(\frac{1}{2\pi}\right)^{6N}\left(\frac{\pi}{\sqrt{M}}\right)^{3N}\int_{\mathbb{R}^{6N}}\mathcal{F}f(\xi_X,\xi_P)\left(\sum_{n=0}^{m}\left(\frac{-\mathrm{i}\xi_X\cdot\xi_P}{2M^{1/2}}\right)^n\frac{1}{n!}\right.\\&\qquad\left.+\left(\frac{-\mathrm{i}\xi_X\cdot\xi_P}{2M^{1/2}}\right)^{m+1}\frac{1}{m!}\int_0^1(1-s)^m\,\mathrm{e}^{-\frac{\mathrm{i}s}{2}M^{-1/2}\xi_X\cdot\xi_P}\,\mathrm{d}s\right)\mathrm{d}\xi_X\,\mathrm{d}\xi_P\\&\quad=\left(\frac{\pi}{\sqrt{M}}\right)^{3N}\left(\sum_{n=0}^{m}\frac{1}{n!}\left(\frac{-\mathrm{i}\nabla_X\cdot\nabla_P}{2M^{1/2}}\right)^n f(X,P)\right.\\&\qquad\left.+\left(\frac{1}{2M^{1/2}}\right)^{m+1}\int_0^1\int_{\mathbb{R}^{6N}}\mathcal{F}((-\mathrm{i}\nabla_X\cdot\nabla_P)^{m+1}f)(\xi_X,\xi_P)\,\mathrm{e}^{-\frac{\mathrm{i}s}{2}M^{-1/2}\xi_X\cdot\xi_P}\frac{(1-s)^m}{(2\pi)^{6N}m!}\,\mathrm{d}\xi_X\,\mathrm{d}\xi_P\,\mathrm{d}s\right)\\&\quad=\left(\frac{\pi}{M^{1/2}}\right)^{3N}\left(\sum_{n=0}^{m}\frac{1}{n!}\left(\frac{-\mathrm{i}\nabla_X\cdot\nabla_P}{2M^{1/2}}\right)^n f(X,P)+\left(\frac{1}{2M^{1/2}}\right)^{m+1}\int_0^1\int_{\mathbb{R}^{3N}}\mathcal{F}_P((-i\nabla_X\cdot\nabla_P)^{m+1}\right.\\&\qquad\left.\times\ f(X+\cdot,P+\cdot))\left(\frac{s\xi_P}{2M^{1/2}},\xi_P\right)\frac{(1-s)^m}{(2\pi)^{3N}m!}\,\mathrm{d}\xi_P\,\mathrm{d}s\right).\end{aligned}\tag{120}$$

The remainder term can be written as follows:

$$\begin{aligned}
&\left(\frac{1}{2M^{1/2}}\right)^{m+1}\int_0^1\int_{\mathbb{R}^{3N}}\mathcal{F}_P((-\mathrm{i}\nabla_X\cdot\nabla_P)^{m+1}f(X+\cdot,P+\cdot))\left(\frac{s\xi_P}{2M^{1/2}},\xi_P\right)\frac{(1-s)^m}{(2\pi)^{3N}m!}\,\mathrm{d}\xi_P\,\mathrm{d}s\\
&\quad=\left(\frac{1}{2M^{1/2}}\right)^{m+1}\int_0^1\int_{\mathbb{R}^{6N}}(-\mathrm{i}\nabla_X\cdot\nabla_P)^{m+1}f\left(X+\frac{s\xi_P}{2M^{1/2}},P+P'\right)\\
&\qquad\times\mathrm{e}^{\mathrm{i}P'\cdot\xi_P}\frac{(1-s)^m}{(2\pi)^{3N}m!}\,\mathrm{d}P'\,\mathrm{d}\xi_P\,\mathrm{d}s\\
&\quad=:\left(\frac{1}{2M^{1/2}}\right)^{m+1}r_m,
\end{aligned}$$

where r_m is smooth since f is a Schwartz function. The change of variables $P+P'=P''$ and (118) shows that kernel of the remainder becomes

$$\begin{aligned}
r_{m,K}(X,Y)&=\gamma\int_0^1\int_{\mathbb{R}^{9N}}(-\mathrm{i}\nabla_X\cdot\nabla_P)^{m+1}f\left(\frac{X+Y}{2}+\frac{s\xi_P}{2M^{1/2}},P''\right)\\
&\qquad\times\mathrm{e}^{\mathrm{i}M^{1/2}(X-Y)\cdot P}\,\mathrm{e}^{\mathrm{i}(P''-P)\cdot\xi_P}\frac{(1-s)^m}{(2\pi)^{3N}m!}\,\mathrm{d}P''\,\mathrm{d}\xi_P\,\mathrm{d}P\,\mathrm{d}s\\
&=\gamma\int_0^1\int_{\mathbb{R}^{3N}}(-\mathrm{i}\nabla_X\cdot\nabla_P)^{m+1}f\left(\frac{X+Y}{2}+\frac{s(X-Y)}{2},P''\right)\mathrm{e}^{\mathrm{i}M^{1/2}(X-Y)\cdot P''}\frac{(1-s)^m}{m!}\,\mathrm{d}P''\,\mathrm{d}s\\
&=\gamma\int_0^1\mathcal{F}_P\{(-\mathrm{i}\nabla_X\cdot\nabla_P)^{m+1}f\}\left(\frac{X+Y}{2}+\frac{s(X-Y)}{2},M^{1/2}(X-Y)\right)\frac{(1-s)^m}{m!}\,\mathrm{d}s.
\end{aligned}$$

We note that the kernel $A_K^{s_*}(X,Y)$ for any symbols A in the s_*-quantization (16) is

$$\begin{aligned}
A_K^{s_*}(X,Y)&=\gamma\int_{\mathbb{R}^{3N}}A(X+s_*(Y-X),P)\,\mathrm{e}^{\mathrm{i}M^{1/2}(X-Y)\cdot P}\,\mathrm{d}P\\
&=\gamma\mathcal{F}_P\{A\}(X+s_*(Y-X),M^{1/2}(X-Y))\\
&=\gamma\mathcal{F}_P\{A\}\left(\frac{X+Y}{2}+\frac{s}{2}(Y-X),M^{1/2}(X-Y)\right)
\end{aligned}$$

for $s_*=(1-s)/2$. Consequently, we have

$$\int_{\mathbb{R}^{3N}}\langle\Phi,\hat{r}_m\Phi\rangle\,\mathrm{d}X=\int_0^1\int_{\mathbb{R}^{6N}}(\nabla_X\cdot\nabla_P)^{m+1}f(X,P)W^{(s_*)}(X,P)\,\mathrm{d}X\,\mathrm{d}P\,\frac{(1-s)^m}{m!}\,\mathrm{d}s \tag{121}$$

and the next step shows the final bound

$$\int_{\mathbb{R}^{3N}}\langle\Phi,\hat{r}_2\Phi\rangle\,\mathrm{d}X=\mathcal{O}(\mathrm{e}^{\hat{C}T}\delta^{-2}). \tag{122}$$

3. It remains to estimate the remainder for $r=r_2$ in (122). Lemmas 5.1 and 5.3 imply that we have

$$\sup_{(X,P)\in\mathbb{R}^{6N}}\|r^*(X,P)\|_2\le C\sup_{(X,P)\in\mathbb{R}^{6N}}\|r_2(X,P)\|_\infty+\mathcal{O}(\eta)=\mathcal{O}(\mathrm{e}^{\hat{C}T}\delta^{-2}),$$

uniformly in N, which proves (122).

4. The general case $\hat{C} = \hat{A}\hat{B}$ of two Fourier integral operators, with smooth symbols $A, B \in C^\infty(\mathbb{R}^{6N})$ in the Schwartz space yields

$$\begin{aligned} C(U,P) &= \gamma^2 \int_{\mathbb{R}^{12N}} A\left(\frac{U+Z}{2} + \frac{T}{4}, P''\right) B\left(\frac{U+Z}{2} - \frac{T}{4}, P'\right) \mathrm{e}^{\mathrm{i}M^{1/2}F}\, \mathrm{d}P''\, \mathrm{d}P'\, \mathrm{d}Z\, \mathrm{d}T \\ &= \gamma^2 \int_{\mathbb{R}^{12N}} A\left(\frac{U+Z}{2} + \frac{T}{4}, P''\right) B\left(\frac{U+Z}{2} - \frac{T}{4}, P'\right) \mathrm{e}^{\mathrm{i}M^{1/2}F}\, \mathrm{d}P''\, \mathrm{d}P'\, \mathrm{d}Z\, \mathrm{d}T, \end{aligned}$$

where

$$\begin{aligned} F &:= \left(U - Z + \frac{T}{2}\right) \cdot P'' + \left(Z - U + \frac{T}{2}\right) \cdot P' - T \cdot P \\ &= \left(U - Z + \frac{T}{2}\right) \cdot (P'' - P) + \left(Z - U + \frac{T}{2}\right) \cdot (P' - P). \end{aligned}$$

The change of variables $(P'' - P, P' - P, (Z - U + \frac{T}{2})/2, (Z - U - \frac{T}{2})/2)$ replacing (P'', P', Z, T) has the Jacobian 2^{6N} and implies

$$\begin{aligned} C(U,P) = (\pi M^{-1/2})^{-6N} \int_{\mathbb{R}^{12N}} & A(U+Z, P+P'') \\ & \times B(U+T, P+P')\, \mathrm{e}^{\mathrm{i}M^{1/2}(P'\cdot Z - T\cdot P'')}\, \mathrm{d}Z\, \mathrm{d}P''\, \mathrm{d}T\, \mathrm{d}P'. \end{aligned}$$

The same expansion as in (120) shows that

$$\begin{aligned} C(X,P) &= \mathrm{e}^{\frac{\mathrm{i}}{2}M^{-1/2}\sum_k (\partial_{P_k}\partial_{X'_k} - \partial_{X_k}\partial_{P'_k})} A(X,P)B(X',P')|_{(X,P)=(X',P')} \\ &= \sum_{n=0}^{m} \frac{(\mathrm{i}M^{-1/2})^n}{2^n n!} (\nabla_P \cdot \nabla_{X'} - \nabla_X \cdot \nabla_{P'})^n A(X,P)B(X',P') \Big|_{(X,P)=(X',P')} \\ &\quad + M^{-(m+1)/2} r_m \end{aligned}$$

based on the representation

$$\begin{aligned} &\int_{\mathbb{R}^{6N}} f(X+X', P+P', X+X'', P+P'')\, \mathrm{e}^{2\mathrm{i}M^{1/2}(X'\cdot P' - X''\cdot P'')}\, \mathrm{d}X'\, \mathrm{d}P'\, \mathrm{d}X''\, \mathrm{d}P'' \\ &\quad = \left(\frac{\pi}{M^{1/2}}\right)^{3N} \left(\sum_{n=0}^{m} \frac{1}{n!} (\mathrm{i}(\nabla_{X''} \cdot \nabla_{P'} - \nabla_{X'} \cdot \nabla_{P''}))^{m+1} f(X+X', P+P', X+X'', P+P'') \Big|_{\substack{X'=X''=0 \\ P'=P''=0}} \right. \\ &\qquad \left. + \left(\frac{1}{2M^{1/2}}\right)^{m+1} r_m \right), \\ &r_m := \int_0^1 \int_{\mathbb{R}^{12N}} \mathrm{i}(\nabla_{X''} \cdot \nabla_{P'} - \nabla_{X'} \cdot \nabla_{P''})^{m+1} f\left(X + \frac{s\xi_{P'}}{2M^{1/2}}, P+P', X - \frac{s\xi_{P''}}{2M^{1/2}}, P+P''\right) \\ &\qquad \times \mathrm{e}^{\mathrm{i}(P'\cdot\xi_{P'} - P''\cdot\xi_{P''})} \frac{(1-s)^m}{(2\pi)^{6N} m!}\, \mathrm{d}P'\, \mathrm{d}\xi_{P'}\, \mathrm{d}P''\, \mathrm{d}\xi_{P''}\, \mathrm{d}s. \end{aligned}$$

The kernel for $m=1$ becomes

$$r_{1,K}(X,Y)=\gamma\int_0^1 \mathcal{F}_{P'P''}\{(\mathrm{i}(\nabla_{X''}\cdot\nabla_{P'}-\nabla_{X'}\cdot\nabla_{P''}))^2 f\}\\ \times\left(\frac{X+Y}{2}-\frac{s(X-Y)}{2M^{1/2}},M^{1/2}(X-Y),\frac{X+Y}{2}-\frac{s(X-Y)}{2M^{1/2}},M^{1/2}(X-Y)\right)(1-s)\,\mathrm{d}s$$

and as in (121) we obtain

$$\int_{\mathbb{R}^{3N}}\langle\Phi,\hat{r}_1\Phi\rangle\,\mathrm{d}X\\ =\int_0^1\int_{\mathbb{R}^{6N}}(\nabla_{X''}\cdot\nabla_{P'}-\nabla_{X'}\cdot\nabla_{P''})^2 f(X+X',P+P',X+X'',P+P'')\Big|_{X'=X''=P'=P''=0}\\ \times W^{(s_*)}(X,P)\,\mathrm{d}X\,\mathrm{d}P(1-s)\,\mathrm{d}s.$$

We apply this to the special case $C(X,P)=H_\eta(X,P)B(X,P)$ where $B(Z)=\mathbb{E}[\bar{g}\circ S_t(\cdot)\,|\,X_0,P_0]*\phi_\eta(Z)$ is a Schwartz function. Then we have $f(X',P',X'',P'')=H_\eta(X',P')B(X'',P'')$ and for $B=\partial_P\mathbb{E}[\bar{g}\circ S_T(\cdot)\,|\,X_0,P_0]*\phi_\eta(X,P)$. Lemma 5.3 yields as in Step 3 the bound

$$\int_{\mathbb{R}^{3N}}\langle\Phi,\hat{r}_1\Phi\rangle\,\mathrm{d}X=\mathcal{O}(\mathrm{e}^{\hat{C}T}\delta^{-4}).$$

5. If $A=H_\eta$ and $B(Z)=2\int_0^1\tilde{g}''_{\mathrm{MD}}(sH_\eta(Z)+(1-s)E)(1-s)\,\mathrm{d}s$ the symbol $D:=AB$ is a smooth product of two convolutions, by (38), and $D(Z)$ tends to zero fast for large $|Z|$. We obtain as in Step 3

$$\int_{\mathbb{R}^{3N}}\langle\Phi,\hat{r}_1\Phi\rangle\,\mathrm{d}X=\mathcal{O}(\mathrm{e}^{\hat{C}T}\delta^{-4}).$$

■

Appendix 1. Fourier Integral WKB States Including Caustics

A.1 A preparatory example with the simplest caustic

As an example of a caustic, we study first the simplest example of a fold caustic based on the Airy function $\mathrm{A}:\mathbb{R}\to\mathbb{R}$ which solves

$$-\partial_{xx}\mathrm{A}(x)+x\mathrm{A}(x)=0. \tag{A.1}$$

The scaled Airy function

$$u(x)=C\;\mathrm{A}(M^{1/3}x)$$

solves the Schrödinger equation

$$-\frac{1}{M}\partial_{xx}u(x)+xu(x)=0, \tag{A.2}$$

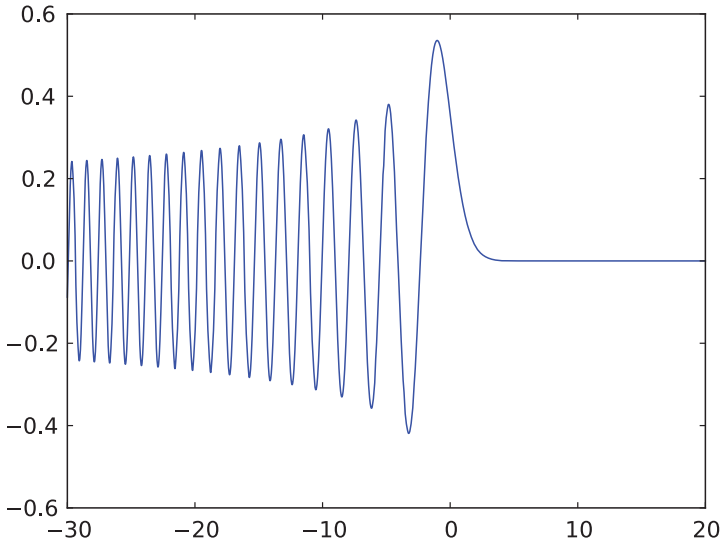

Fig. A1. The Airy function.

for any constant C. In our context, an important property of the Airy function is the fact that it is the inverse Fourier transform of the function

$$\hat{\mathrm{A}}(p) = \sqrt{\frac{2}{\pi}}\, e^{ip^3/3},$$

that is,

$$\mathrm{A}(x) = \frac{1}{\pi} \int_{\mathbb{R}} e^{i(xp+p^3/3)}\, \mathrm{d}p. \tag{A.3}$$

In the next section, we will consider a general Schrödinger equation and determine a WKB Fourier integral corresponding to (A.3) for the Airy function; as an introduction to the general case we show how the derive (A.3): by taking the Fourier transform of the ordinary differential equation (A.1)

$$0 = \int_{\mathbb{R}} (-\partial_{xx} + x)\mathrm{A}(x) e^{-ixp}\, \mathrm{d}x = (p^2 + i\partial_p)\hat{\mathrm{A}}(p), \tag{A.4}$$

we obtain an ordinary differential equation for the Fourier transform $\hat{\mathrm{A}}(p)$ with the solution $\hat{\mathrm{A}}(p) = C e^{ip^3}$, for any constant C. Then, by differentiation, it is clear that the scaled Airy function u solves (A.2). Furthermore, the stationary phase method, cf. Appendix 2, shows that to the leading order u is approximated by

$$u(x) \simeq C\,(-xM^{1/3})^{-1/4} \cos(M^{1/2}(-x)^{3/2} - \pi/4) \quad \text{for } x < 0, \tag{A.5}$$

and $u(x) \simeq 0$ to any order (i.e., $\mathcal{O}(M^{-K})$ for any positive K) when $x > 0$. The behavior of the Airy function is illustrated in Figure A1.

A.1.1 *Molecular dynamics for the Airy function*

The eikonal equation corresponding to (A.2) is

$$p^2 + x = 0$$

with solutions for $x \le 0$, which leads to the phase

$$p = \theta'(x) = \pm(-x)^{1/2} \quad \text{and} \quad \theta(x) = \mp\tfrac{2}{3}(-x)^{3/2}. \tag{A.6}$$

We have the Legendre transform

$$\theta^*(p) = \min_y (yp - \theta(y))$$

with the solution $p = \theta'(y)$. Consequently, we obtain by (A.6) that $x = y$ and

$$\theta^*(p) = xp - \theta(x) = -p^2 p + \frac{2}{3}p^3 = -\frac{p^3}{3}.$$

We note that this solution is also obtained from the eikonal equation

$$p^2 + \partial_p \theta^*(p) = 0,$$

which is solved by

$$\theta^*(p) = -p^3/3.$$

Thus, we recover the relation for the Legendre transform $-xp + \theta^*(p) = -\theta(x)$.

A.2 **A general Fourier integral ansatz**

In order to treat a more general case with a caustic of the dimension d, we use the Fourier integral ansatz

$$\Phi(X) = \int_{\mathbb{R}^d} \phi(X)\, e^{-iM^{1/2}\Theta(\check{X},\hat{X},\check{P})}\, \mathrm{d}\check{P}, \tag{A.7}$$

where

$$\begin{aligned}
&\phi(X) \in \mathbb{C}^J,\\
&\quad X = (\hat{X}, \check{X}) \in \mathbb{R}^{3N}, \quad P = (\hat{P}, \check{P}) \in \mathbb{R}^{3N},\\
&\check{X} \cdot \check{P} = \sum_{j=1}^{d} \check{X}^j \check{P}^j, \quad \hat{X} \cdot \hat{P} = \sum_{j=d+1}^{N} \hat{X}^j \hat{P}^j,\\
&\Theta(\check{X}, \hat{X}, \check{P}) = \check{X} \cdot \check{P} - \theta^*(\hat{X}, \check{P}),
\end{aligned}$$

based on the Legendre transform relationships (cf. [11])

$$\theta^*(\hat{X},\check{P}) = \min_{\check{X}}(\check{X}\cdot\check{P} - \theta(\hat{X},\check{X})),$$
$$\theta(\hat{X},\check{X}) = \min_{\check{P}}(\check{X}\cdot\check{P} - \theta^*(\hat{X},\check{P})).$$

If the function $\theta^*(\hat{X},\check{P})$ is not defined for all $\check{P}\in\mathbb{R}^d$, but only for $\check{P}\in\mathcal{U}\subset\mathbb{R}^d$ we replace the integral over $\mathbb{R}^d$ by integration over $\mathcal{U}$ using a smooth cut-off function $\chi(\check{P})$. The cut-off function is zero outside $\mathcal{U}$ and equal to one in a large part of the interior of $\mathcal{U}$. The ansatz (A.7) is inspired by Maslov's work [34], although it is not the same since our amplitude function ϕ depends on $(\hat{X},\check{X})$ but not on $\check{P}$.

A.2.1 *Making the ansatz for a Schrödinger solution*

In this section, we construct a solution to the Schrödinger equation from the ansatz (A.7). The constructed solution will be an *actual* solution and not only an asymptotic solution as in [34]. We consider first the case when the integration is over $\mathbb{R}^d$ and then conclude in Remark A.2 that the cut-off function $\chi(\check{P})$ can be included in all integrals without changing the property of the Fourier integral ansatz being a solution in the $\check{X}$-domain where $\check{X}=\nabla_{\check{P}}\theta^*(\hat{X},\check{P})$ for some $\check{P}$ satisfying $\chi(\check{P})=1$.

The requirement to be a solution means that there should hold

$$\begin{aligned}
0 &= (\hat{H}-E)\Phi\\
&= \int_{\mathbb{R}^d}\left(\frac{1}{2}|\nabla_{\hat{X}}\theta^*(\hat{X},\check{P})|^2 + \frac{1}{2}|\check{P}|^2 + V_0(X) - E\right)\phi(X)e^{-iM^{1/2}\Theta(\check{X},\hat{X},\check{P})}\,\mathrm{d}\check{P}\\
&\quad - \int_{\mathbb{R}^d}\left(iM^{-1/2}\left(\nabla_{\hat{X}}\phi\cdot\nabla_{\hat{X}}\theta^* - \nabla_{\check{X}}\phi\cdot\check{P} + \frac{1}{2}\phi\Delta_{\hat{X}}\theta^*\right) - (V-V_0)\phi + \frac{1}{2M}\Delta_X\phi\right)\\
&\quad \times e^{-iM^{1/2}\Theta(\check{X},\hat{X},\check{P})}\,\mathrm{d}\check{P}.
\end{aligned}\tag{A.8}$$

Comparing this expression to the previously discussed case of a single WKB-mode, we see that the zero-order term is now $\Delta_{\hat{X}}\theta^*$ instead of $\Delta_X\theta$ and that we have $-\nabla_{\check{X}}\phi\cdot\check{P}$ instead of $\nabla_{\check{X}}\phi\cdot\nabla_{\check{X}}\theta$. However, the main difference is that the first integral is not zero (only the leading order term of its stationary phase expansion is zero, cf. (A.19)). Therefore, the first integral contributes to the second integral. The goal is now to determine a function $F(\hat{X},\check{X},\check{P})$ satisfying

$$\begin{aligned}
&\int_{\mathbb{R}^d}\left(\frac{1}{2}|\nabla_{\hat{X}}\theta^*|^2 + \frac{1}{2}|\check{P}|^2 + V_0(X) - E\right)e^{-iM^{1/2}\Theta(\check{X},\hat{X},\check{P})}\,\mathrm{d}\check{P}\\
&\qquad = iM^{-1/2}\int_{\mathbb{R}^d}F(\hat{X},\check{X},\check{P})\,e^{-iM^{1/2}\Theta(\check{X},\hat{X},\check{P})}\,\mathrm{d}\check{P},
\end{aligned}\tag{A.9}$$

and verify that it is bounded.

Lemma A.1. There holds $F = F_0 + F_1$ where

$$F_0 = \frac{1}{2}\sum_{i,j} \partial_{\check{X}^i \check{X}^j} V_0(\nabla_{\check{P}}\theta^*(\check{P}))\partial_{\check{P}^j \check{P}^i}\theta^*(\check{P}),$$

$$F_1 = iM^{-1/2}\int_0^1\int_0^1\int_{\mathbb{R}^d}\sum_{i,j,k} t(1-t)\partial_{\check{P}^k}[\partial_{\check{X}^i \check{X}^j \check{X}^k} V_0(\nabla_{\check{P}}\theta^*(\check{P}) + s\,t\,\delta\theta^*(\check{P}))\partial_{\check{P}_j \check{P}_i}\nabla_{\check{P}}\theta^*(\check{P})]\,\mathrm{d}t\,\mathrm{d}s.$$

□

Proof. The function $\theta^*(\hat{X}, \check{P})$ is defined as a solution to the Hamilton–Jacobi (eikonal) equation

$$\tfrac{1}{2}|\nabla_{\hat{X}}\theta^*(\hat{X}, \check{P})|^2 + \tfrac{1}{2}|\check{P}|^2 + V_0(\hat{X}, \nabla_{\check{P}}\theta^*(\hat{X}, \check{P})) - E = 0 \tag{A.10}$$

for all $(\hat{X}, \check{P})$. Consequently, the integral on the left-hand side of (A.9) is

$$\int_{\mathbb{R}^d} (V_0(\hat{X}, \check{X})) - V_0(\hat{X}, \nabla_{\check{P}}\theta^*(\hat{X}, \check{P}))\, e^{-iM^{1/2}(\check{X}\cdot\check{P} - \theta^*(\hat{X}, \check{P}))}\,\mathrm{d}\check{P}.$$

Let $\check{P}_0(\check{X})$ be any solution to the stationary phase equation $\check{X} = \nabla_{\check{P}}\theta^*(\hat{X}, \check{P}_0)$ and introduce the notation

$$\Theta'(\check{X}, \hat{X}, \check{P}) := \nabla_{\check{P}}\theta^*(\hat{X}, \check{P}_0)\cdot\check{P} - \theta^*(\hat{X}, \check{P}).$$

Then by writing a difference as $V(y_1) - V(y_2) = \int_0^1 \partial_y V(y_2 + t(y_1 - y_2))\,\mathrm{d}t \cdot (y_1 - y_2)$, identifying a derivative $\partial_{\check{P}_i}$ and integrating by parts the integral can be written as follows:

$$\begin{aligned}
&\int_{\mathbb{R}^d} (V_0(\hat{X}, \nabla_{\check{P}}\theta^*(\hat{X}, \check{P}_0)) - V_0(\hat{X}, \nabla_{\check{P}}\theta^*(\hat{X}, \check{P})))\, e^{-iM^{1/2}\Theta'(\check{X}, \hat{X}, \check{P})}\,\mathrm{d}\check{P} \\
&= \int_0^1\int_{\mathbb{R}^d}\sum_i \partial_{\check{X}^i} V_0(\nabla_{\check{P}}\theta^*(\check{P}) + t[\nabla_{\check{P}}\theta^*(\check{P}_0) - \nabla_{\check{P}}\theta^*(\check{P})]) \\
&\quad\times \left(\partial_{\check{P}^i}\theta^*(\check{P}_0) - \partial_{\check{P}^i}\theta^*(\check{P})\right) e^{-iM^{1/2}\Theta'(\check{X}, \hat{X}, \check{P})}\,\mathrm{d}\check{P}\,\mathrm{d}t \\
&= -iM^{-1/2}\int_0^1\int_{\mathbb{R}^d}\sum_i \partial_{\check{X}^i} V_0(\nabla_{\check{P}}\theta^*(\check{P}) + t[\nabla_{\check{P}}\theta^*(\check{P}_0) - \nabla_{\check{P}}\theta^*(\check{P})])\partial_{\check{P}_i} e^{-iM^{1/2}\Theta'(\check{X}, \hat{X}, \check{P})}\,\mathrm{d}\check{P}\,\mathrm{d}t \\
&= iM^{-1/2}\int_0^1\int_{\mathbb{R}^d}\sum_i \partial_{\check{P}_i}\partial_{\check{X}^i} V_0(\nabla_{\check{P}}\theta^*(\check{P}) + t[\nabla_{\check{P}}\theta^*(\check{P}_0) - \nabla_{\check{P}}\theta^*(\check{P})])\, e^{-iM^{1/2}\Theta'(\check{X}, \hat{X}, \check{P})}\,\mathrm{d}\check{P}\,\mathrm{d}t.
\end{aligned}$$

Therefore, the leading order term in $F =: F_0 + F_1$ is

$$\begin{aligned}
F_0 &:= \int_0^1 \sum_{i,j}(1-t)\partial_{\check{X}^i \check{X}^j} V_0(\nabla_{\check{P}}\theta^*(\check{P}))\partial_{\check{P}^j \check{P}^i}\theta^*(\check{P})\,\mathrm{d}t \\
&= \frac{1}{2}\sum_{i,j}\partial_{\check{X}^i \check{X}^j} V_0(\nabla_{\check{P}}\theta^*(\check{P}))\partial_{\check{P}^j \check{P}^i}\theta^*(\check{P}).
\end{aligned}$$

Denoting $\delta\theta^*(\check{P}) := \nabla_{\check{P}}\theta^*(\check{P}_0) - \nabla_{\check{P}}\theta^*(\check{P})$, the remainder becomes

$$
\begin{aligned}
&- iM^{-1/2}\int_0^1\int_{\mathbb{R}^d}\sum_{i,j}[\partial_{\check{X}^i\check{X}^j}V_0(\nabla_{\check{P}}\theta^*(\check{P})) - \partial_{\check{X}^i\check{X}^j}V_0(\nabla_{\check{P}}\theta^*(\check{P}) + t\delta\theta^*(\check{P}))]\\
&\quad\times(1-t)\partial_{\check{P}^j\check{P}^i}\theta^*(\check{P})\,e^{-iM^{1/2}\Theta'(\check{X},\hat{X},\check{P})}\,\mathrm{d}\check{P}\,\mathrm{d}t\\
&\qquad= iM^{-1/2}\int_0^1\int_0^1\int_{\mathbb{R}^d}\sum_{i,j,k}t(1-t)\partial_{\check{X}^i\check{X}^j\check{X}^k}V_0(\nabla_{\check{P}}\theta^*(\check{P}) + st\delta\theta^*(\check{P}))\partial_{\check{P}^j\check{P}^i}\theta^*(\check{P})\\
&\qquad\quad\times(\partial_{\check{P}^k}\theta^*(\check{P}_0) - \partial_{\check{P}^k}\theta^*(\check{P}))\,e^{-iM^{1/2}\Theta'(\check{X},\hat{X},\check{P})}\,\mathrm{d}\check{P}\,\mathrm{d}t\,\mathrm{d}s\\
&\qquad= -\frac{1}{M}\int_0^1\int_0^1\int_{\mathbb{R}^d}\sum_{i,j,k}t(1-t)\partial_{\check{P}^k}[\partial_{\check{X}^i\check{X}^j\check{X}^k}V_0(\nabla_{\check{P}}\theta^*(\check{P}) + st\delta\theta^*(\check{P}))\partial_{\check{P}^j\check{P}^i}\theta^*(\check{P})]\\
&\qquad\quad\times e^{-iM^{1/2}\Theta'(\check{X},\hat{X},\check{P})}\,\mathrm{d}\check{P}\,\mathrm{d}t\,\mathrm{d}s,
\end{aligned}
$$

hence the function F_1 is purely imaginary and small

$$
F_1 = iM^{-1/2}\int_0^1\int_0^1\int_{\mathbb{R}^d}\sum_{i,j,k}t(1-t)\partial_{\check{P}^k}[\partial_{\check{X}^i\check{X}^j\check{X}^k}V_0(\nabla_{\check{P}}\theta^*(\check{P}) + st\delta\theta^*(\check{P}))\partial_{\check{P}_j\check{P}_i}\nabla_{\check{P}}\theta^*(\check{P})]\,\mathrm{d}t\,\mathrm{d}s
$$

and

$$
2\mathrm{Re}\,F = \sum_{i,j}\partial_{\check{X}^i\check{X}^j}V_0(\nabla_{\check{P}}\theta^*(\check{P}))\partial_{\check{P}^j\check{P}^i}\theta^*(\check{P}). \tag{A.11}
$$

■

The eikonal equation (A.10) and the requirement that $(\hat{H} - E)\Phi = 0$ in (A.8) then imply that

$$
\begin{aligned}
0 = \int_{\mathbb{R}^d}\Bigg[&iM^{-1/2}\left(\nabla_{\hat{X}}\phi\cdot\nabla_{\hat{X}}\theta^* - \nabla_{\check{X}}\phi\cdot\check{P} + \frac{1}{2}\phi(\Delta_{\hat{X}}\theta^* - 2F(X,\check{P}))\right)\\
&-(V - V_0)\phi + \frac{1}{2M}\Delta_X\phi\Bigg]e^{-iM^{1/2}\Theta(\check{X},\hat{X},\check{P})}\,\mathrm{d}\check{P}.
\end{aligned} \tag{A.12}
$$

The Hamilton–Jacobi eikonal equation (A.10), in the primal variable $(\hat{X},\check{P})$ with the corresponding dual variable $(\hat{P},\check{X})$, can be solved locally by the characteristics

$$
\begin{aligned}
\dot{\hat{X}} &= \hat{P},\\
\dot{\hat{P}} &= -\nabla_{\hat{X}}V_0(\hat{X},\check{X}),\\
\dot{\check{X}} &= -\check{P},\\
\dot{\check{P}} &= \nabla_{\check{X}}V_0(\hat{X},\check{X}),
\end{aligned} \tag{A.13}
$$

using the definition

$$\nabla_{\hat{X}}\theta^*(\hat{X},\check{P}) = \hat{P},$$
$$\nabla_{\check{P}}\theta^*(\hat{X},\check{P}) = \check{X}.$$

The characteristics give

$$\frac{\mathrm{d}}{\mathrm{d}t}\phi = \nabla_{\hat{X}}\phi\cdot\nabla_{\hat{X}}\theta^* - \nabla_{\check{X}}\phi\cdot\check{P},$$

so that the Schrödinger transport equation becomes, as in (47),

$$iM^{-1/2}\left(\dot{\phi} + \phi\frac{\dot{G}}{G}\right) = (V - V_0)\phi - \frac{1}{2M}\Delta_X\phi \tag{A.14}$$

and for $\psi = G\phi$

$$iM^{-1/2}\dot{\psi} = (V - V_0)\psi - \frac{G}{2M}\Delta_X\frac{\psi}{G} \tag{A.15}$$

with the complex-valued weight function G defined by

$$\frac{\mathrm{d}}{\mathrm{d}t}\log G_t = \frac{1}{2}\Delta_{\hat{X}}\theta^*(\hat{X}_t,\check{P}_t) - F(\hat{X}_t,\check{P}_t). \tag{A.16}$$

This transport equation is of the same form as the transport equation for a single WKB-mode, with a modification of the weight function G.

Differentiation of the second equation in the Hamiltonian system (A.13) implies that the first variation $\partial\check{P}_t/\partial\check{X}_0$ satisfies

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial\check{P}_t^i}{\partial\check{X}_0}\right) = \sum_{j,k}\partial_{\check{X}^i\check{X}^j}V_0(\hat{X},\check{X}_t)\partial_{\check{P}^j\check{P}^k}\theta^*(\check{P})\frac{\partial\check{P}_t^k}{\partial\check{X}_0},$$

which by the Liouville formula (A.18) and the equality

$$2\mathrm{Re}\,F = \sum_{i,j}\partial_{\check{X}^i\check{X}^j}V_0\partial_{\check{P}^j\check{P}^i}\theta^* = \mathrm{Tr}\left(\sum_j\partial_{\check{X}^i\check{X}^j}V_0\partial_{\check{P}^j\check{P}^k}\theta^*\right)$$

in (A.11) yields the relation,

$$e^{-2\int_0^t\mathrm{Re}\,F\,\mathrm{d}t'} = C\left|\det\frac{\partial\check{P}_t}{\partial\check{X}_0}\right|, \tag{A.17}$$

for the constant $C := |\det\frac{\partial\check{X}_0}{\partial\check{P}_0}|$.

Remark A.2. The conclusion in this section holds also when all integrals over $d\check{P}$ in $\mathbb{R}^d$ are replaced by integrals with the measure $\chi(\check{P})\,\mathrm{d}\check{P}$. Then there holds $2\mathrm{Re}\,F = \sum_{ij} \partial_{\check{X}^i\check{X}^j} V \partial_{\check{P}^i}(\chi\,\partial_{\check{P}^j}\theta^*)$. □

A.2.2 *Liouville's formula*

Here we state Liouville's formula

$$\frac{G_0^2}{G_t^2} = e^{-\int_0^t \mathrm{Tr}\,(\nabla P(X_t))\,\mathrm{d}t} = \left|\det \frac{\partial(X_0)}{\partial(X_t)}\right|, \tag{A.18}$$

given in [34].

Appendix 2. The Stationary Phase Expansion

Consider the phase function $\check{X}\cdot\check{P} - \theta^*(\hat{X},\check{P})$ and let $\check{P}_0(\hat{X})$ be any solution to the stationary phase equation $\check{X} = \nabla_{\check{P}}\theta^*(\hat{X},\check{P}_0)$. We rewrite the phase function

$$\check{X}\cdot\check{P} - \theta^*(\check{X},\check{P}) = \underbrace{\check{X}\cdot\check{P}_0 - \theta^*(\check{X},\check{P}_0)}_{=\theta(\hat{X},\check{X})} + (\check{P}-\check{P}_0)\cdot\int_0^1 (1-t)\partial_{PP}\theta^*(\check{P}_0 + t[\check{P}-\check{P}_0])\,\mathrm{d}t[\check{P}-\check{P}_0].$$

The relation

$$\frac{1}{2}Y\cdot\partial_{PP}\bar{\theta}(\check{P}_0)Y = (\check{P}-\check{P}_0)\cdot\int_0^1 (1-t)\partial_{PP}\bar{\theta}(\check{P}_0 + t[\check{P}-\check{P}_0])\,\mathrm{d}t[\check{P}-\check{P}_0]$$

defines the function $Y(\check{P})$, and its inverse $\check{P}(Y)$, so that the phase is a quadratic function in Y. The stationary phase expansion of an integral takes the form, see [51],

$$\begin{aligned}
&\int_{\mathbb{R}^d} w(\check{P})\,e^{-iM^{1/2}(\check{X}\cdot\check{P}-\theta^*(\hat{X},\check{P}))}\,\mathrm{d}\check{P} \\
&\quad\sim \sum_{\{\check{P}_0:\nabla_{\check{P}}\theta^*(\check{P}_0)=\check{X}\}} (2\pi M^{-1/2})^{d/2}\left|\det\frac{\partial(\check{P}_0)}{\partial(\check{X})}\right|^{1/2} e^{i\frac{\pi}{4}\mathrm{sgn}(\partial_{PP}\theta^*(\check{P}_0))}\,e^{-iM^{1/2}\theta(\hat{X},\check{X};\check{P}_0)} \\
&\qquad\times\sum_{k=0}^{\infty}\frac{M^{-k/2}}{k!}\left(\sum_{l,j} i(\partial_{P^lP^j}\theta^*)^{-1}(\check{P}_0)\partial_{Y^lY^j}\right)^k\left(w(\check{P}(Y))\left|\det\frac{\partial(\check{P})}{\partial(Y)}\right|\right)\Bigg|_{Y=0}.
\end{aligned} \tag{A.19}$$

Appendix 3. An Alternative Motivation for Assumption (26)

The double average (26) in time and phase-space is in Monte Carlo methods approximated by sampling several paths with random initial points (X_0, P_0). For a single path, one expects $\mathcal{O}(T^{-1/2})$ convergence rate, in the case of bounded correlation times.

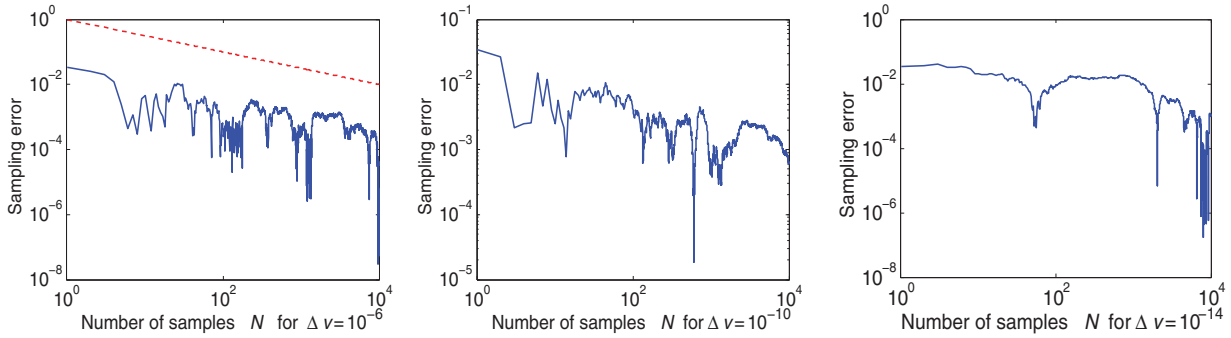

Fig. A2. Ergodicity test results for Born–Oppenheimer molecular dynamics simulations.

Many paths correspond in some sense to longer averaging time for a single path and one may ask how long. This answer is related to how the dynamics maps an initial phase-space neighborhood D to $S_{T/2}(D)$. The exponential $e^{\hat{C}T}$ growth of the first variation $|\partial^{\alpha}_{P_0} g \circ S_T(X_0, P_0)|$, for $|\alpha| = 1$, implies that an initial domain D, on the constant energy hyper-surface $F := \{(X, P) \mid H_0(X, P) = \text{constant}\}$, is stretched but mapped to the bounded energy surface F. If the dynamics is ergodic, the sampling density becomes uniform on this energy hyper-surface, asymptotically in time. Paths that are very close are not acting as independent samples for finite T. The assumption (26) means that the exponential growth of the first variation measures how small the initial distance, $e^{-\hat{C}T/2}$, of the paths can be in order to lead to approximately independent paths after time $T/2$. The paths that are initially $e^{-\hat{C}T/2}$ close will after time $T/2$ have a distance of order 1. The remaining $T/2$ time this $\mathcal{O}(e^{\hat{C}T/2})$ number of paths would act as approximately independent samples (over time $T/2$), as tested in Figure A2, and would give the expected decay $e^{-\hat{C}T/4}$, if the Lyapunov exponent is positive in only one direction, $d = 1$. If d Lyapunov exponents are similar to the maximal, we expect the decay $e^{-\hat{C}Td/4}$.

Figure A2 tests ergodicity for Born–Oppenheimer molecular dynamics simulations with the potential $\lambda_s(X) = X_1^2/2 + X_2^2/\sqrt{2} + 2\sin(X_1 X_2)$ and the double average over time and N initial points $\{(X_0[n], P_0[n]) = (0, 0, \sqrt{2(E - \lambda_s(0))}\cos(1.2 + n\Delta v), \sqrt{2(E - \lambda_s(0))}\sin(1.2 + n\Delta v)\}_{n=1}^{N}$ in (26). The vertical axis shows the sampling error $\sum_{n=1}^{N} \frac{2}{T}\int_{T/2}^{T}(g(X_t) - g_{\text{MD}})(X_0[n], P_0[n])\,\mathrm{d}t/N$ and the horizontal axis shows the number of samples N for the symplectic Euler method. The convergence rate $\mathcal{O}(N^{-1/2})$ for $\Delta v = 10^{-6}$ (left figure) and $\Delta v = 10^{-10}$ (middle figure) indicate independent sample paths while the closer paths with $\Delta v = 10^{-14}$ in (right figure) are not acting independently since the convergence is slower. The maximal Lyapunov exponent is in this case computed to be roughly 0.35 (the others are $(-0.35, 0, 0)$), so that the expansion of the path distance from

time zero to time 70 becomes $e^{0.35\times 70} \approx 10^{10}$ and the approximate $N^{-1/2}$ convergence rate for $\Delta v \geq 10^{-10}$ indicates that $\gamma \approx 1/4$ as suggested in the paragraph following (26). Here $T = 140$, $\Delta t = 0.01$, $E = 1.5$, $g(X) = \sin(X_1 X_2)$. The ergodic limit g_{MD} is approximated in two different ways: by a single symplectic Euler path to -0.4389 and by the Monte Carlo samples (93) to -0.4388. The dashed line shows $N^{-1/2}$.

Funding

The research was partially supported by the U.S. National Science Foundation under the grant NSF-DMS-0813893, U.S. DOD-ARO Grant Award W911NF-14-1-0247, Swedish Research Council grants 621-2010-5647 and 621-2014-4776, and the Swedish e-Science Research Center.

Acknowledgements

P.P. also thanks KTH and Nordita for their hospitality during his visit when the presented research was initiated.

References

[1] Bayer, C., H. Hoel, P. Plecháč, A. Szepessy, and R. Tempone. "How accurate is molecular dynamics?" (2011): preprint arXiv:11040953.

[2] Berezin, F. A. and M. A. Shubin. *The Schrödinger Equation*. Dordrecht: Kluwer Academic Publishers, 1991.

[3] Bolte, J. and R. Glaser. "A semiclassical Egorov theorem and quantum ergodicity for matrix valued operators." *Communications in Mathematical Physics* 247, no. 2 (2004): 391–419.

[4] Born, M. and R. Oppenheimer. "Zur Quantentheorie der Molekeln." *Annalen der Physik* 389, no. 20 (1927): 457–84.

[5] Bouzounia, A. and D. Robert. "Uniform semiclassical estimates for the propagation of quantum observables." *Duke Mathematical Journal* 111, no. 2 (2002): 223–52.

[6] Cances, E., M. Defranceschi, W. Kutzelnigg, C. LeBris, and Y. Maday. *Computational Chemistry: A Primer*. Handbook of Numerical Analysis X. Amsterdam: North-Holland, 2007.

[7] Cances, E., F. Legoll, and G. Stolz. "Theoretical and numerical comparison of some sampling methods for molecular dynamics." *Mathematical Modelling and Numerical Analysis* 41, no. 2 (2007): 351–89.

[8] Ciccotti, G., T. Lelièvre, and E. Vanden-Eijnden. "Projection of diffusions on submanifolds: application to mean force computation." *Communications on Pure and Applied Mathematics* 61, no. 3 (2008): 371–408.

[9] Dall'ara, G. M. "Discreteness of the spectrum of Schrödinger operators with non-negative matrix values potentials." *Journal of Functional Analysis* 268, no. 12 (2015): 3649–79.

[10] Dimassi, M. and J. Sjöstrand. *Spectral Asymptotics in the Semiclassical Limit*. LMS Lecture Note Series 268. Cambridge: Cambridge University Press, 1999.

[11] Evans, L. C. *Partial Differential Equation*. Providence, RI: American Mathematical Society, 1998.

[12] Faou, E. and T. Lelièvre. "Conservative stochastic differential equations: mathematical and numerical analysis." *Mathematics of Computation* 78, no. 268 (2009): 2047–74.

[13] Fefferman, C. and L. Seco. "Eigenvalues and eigenfunctions of ordinary differential operators." *Advances in Mathematics* 95, no. 2 (1992): 145–305.

[14] Frenkel, D. and B. Smith. *Understanding Molecular Simulation*. Orlando, FL: Academic Press, 2002.

[15] Golub, G. H. and C. F. Van Loan. *Matrix Computations*. Baltimore, MD: John Hopkins University Press, 1996.

[16] Hagedorn, G. A. "High order corrections to the time-independent Born–Oppenheimer approximation I: smooth potentials." *Annales de l'Institut Henri Poincaré Section A* 47, no. 1 (1987): 1–16.

[17] Hagedorn, G. A. "High order corrections to the time-independent Born–Oppenheimer approximation II: diatomic Coulomb systems." *Communications in Mathematical Physics* 116, no. 1 (1988): 23–44.

[18] Hairer, E., C. Lubich, and G. Wanner. "Geometric numerical integration illustrated by the Störmer–Verlet method." *Acta Numerica* 12 (2003): 399–450.

[19] Hassel, A. "Ergodic billiards that are not quantum unique ergodic (with an appendix by A. Hassell and L. Hillairet)." *Annals of Mathematics* 171 (2010): 605–18.

[20] Helffer, B., A. Martinez, and D. Robert. "Ergodicité et limite semi-classique." *Communications in Mathematical Physics* 109, no. 2 (1987): 313–26.

[21] Hérau, F. and F. Nier. "Isotropic hypellipticity and trend to the equilibrium for the Fokker–Planck equation with high degree potential." *Archive for Rational Mechanics and Analysis* 171, no. 2 (2004): 151–218.

[22] Hörmander, L. "Weyl calculus of pseudo-differential operators." *Communications on Pure and Applied Mathematics* 32, no. 2 (1979): 360–444.

[23] Jeffreys, H. "On certain approximate solutions of linear differential equations of the second order." *Proceedings of the London Mathematical Society* 23, no. 1 (1924): 428–36.

[24] Keller, J. and C. Lasser. "Propagation of quantum expectations with Husimi functions." *SIAM Journal on Applied Mathematics* 73, no. 4 (2013): 1557–81.

[25] Klein, M., A. Martinez, R. Seiler, and X. P. Wang. "On the Born–Oppenheimer expansion for polyatomic molecules." *Communications in Mathematical Physics* 143, no. 3 (1992): 607–39.

[26] Lasser, C. and S. Röblitz. "Computing expectations values for molecular quantum dynamics." *SIAM Journal on Scientific Computing* 32, no. 3 (2010): 1465–83.

[27] Lasser, C. and S. Teufel. "Propagation through conical crossings: an asymptotic semigroup." *Communications on Pure and Applied Mathematics* 58, no. 9 (2005): 1188–230.

[28] Landau, L. "Zur theorie der energieubertragung. II." *Physikalische Zeitschrift der Sowjetunion* 1 (1932): 46–51.

[29] Lazutkin, V. F. *KAM Theory and Semiclassical Approximations to Eigenfunctions*. Berlin: Springer, 1993. With an addendum by A. I. Shnirel'man.

[30] LeBris, C. *Computational Chemistry from the Perspective of Numerical Analysis*, 363–444. Acta Numerica 14. Cambridge: Cambridge University Press, 2005.

[31] Lelièvre, T., M. Rousset, and G. Stoltz. *Free Energy Computations: A Mathematical Perspective*. London: Imperial College Press, 2010.

[32] Martinez, A. *An Introduction to Semiclassical and Microlocal Analysis*. New York: Springer, 2002.

[33] Martinez, A. and V. Sordoni. "Twisted pseudodifferential calculus and application to the quantum evolution of molecules." *Memoirs of the American Mathematical Society* 200, no. 936 (2009).

[34] Maslov, V. P. and M. V. Fedoriuk. *Semi-classical Approximation in Quantum Mechanics*. D. Reidel Publishing Company, 1981; based on: V. P. Maslov, *Theory of perturbations and asymptotic methods*. Moscow: Moskov. Gos. Univ., 1965 (Russian).

[35] Panati, G., H. Spohn, and S. Teufel. "Space-adiabatic perturbation theory." *Advances in Theoretical and Mathematical Physics* 7, no. 1 (2003): 145–204.

[36] Rayleigh, L. "On the propagation of waves through a stratified medium, with special reference to the question of reflection." *Proceedings of the Royal Society (London) Series A* 86, no. 586 (1912): 207–26.

[37] Reed, M. and B. Simon. *Methods of Modern Mathematical Physics, I: Functional Analysis*. New York: Academic Press INC., 1980.

[38] Robert, D. *Autour de l'approximation semi-classique*. Boston: Birkhäuser, 1987.

[39] Schiff, L. *Quantum Mechanics*. New York: McGraw-Hill, 1968.

[40] Schrödinger, E. *Collected Papers on Wave Mechanics*. London: Blackie and Son, 1928.

[41] Schnirelman, A. I. "Ergodic properties of eigenfunctions (Russian)." *Uspekhi Matematicheskikh Nauk* 29, no. 6 (1974): 181–2.

[42] Stiepan, H.-M. and S. Teufel. "Semiclassical approximations for Hamiltonians with operator-valued symbols." *Communications in Mathematical Physics* 320, no. 3 (2013): 821–49.

[43] Tanner, D. J. *Introduction to Quantum Mechanics: A Time-dependent Perspective*. Sausalito, CA: University Science Books, 2006.

[44] Teller, E. "The crossing of potential surfaces." *Journal of Physics and Chemistry* 41, no. 1 (1937): 109–16.

[45] Teufel, S. *Adiabatic Perturbation Theory in Quantum Dynamics*. Lecture Notes in Mathematics 1821. Berlin: Springer, 2003.

[46] Tupper, P. F. "Ergodicity and the numerical simulation of Hamiltonian systems." *SIAM Journal on Applied Dynamical Systems* 4, no. 3 (2005): 563–87.

[47] von Schwerin, E. and A. Szepessy. "A stochastic phase-field model determined from molecular dynamics." *MathematicalModelling and Numerical Analysis* 44, no. 4 (2010): 627–46.

[48] Worth, G. A. and L. S. Cederbaum. "Beyond Born–Oppenheimer: molecular dynamics through a conical intersection." *Annual Review of Physical Chemistry* 55 (2004): 127–58.

[49] Zelditch, S. "Quantum ergodicity and mixing of eigenfunctions." In *Encyclopedia of Mathematical Physics*, edited by J.-P. Françoise, G. L. Naber, and T. S. Tsun, vols 1–5. Oxford: Academic Press/Elsevier Science, 2006.

[50] Zener, C. "Non-adiabatic crossing of energy levels." *Proceedings of The Royal Society of London. Series A* 137, no. 833 (1932): 696–702.

[51] Zworski, M. *Semiclassical Analysis*. Providence, RI: American Mathematical Society, 2012.