# CONSTRUCTION AND VALIDATION OF A RIGOROUS SURFACE HOPPING ALGORITHM FOR CONICAL CROSSINGS [*]

CAROLINE LASSER [†], TORBEN SWART [‡], AND STEFAN TEUFEL [§]

**Abstract.** This article presents and evaluates a surface hopping algorithm for time-dependent two-level Schrödinger systems with conically intersecting eigenvalues. The algorithm implements an asymptotic semigroup for approximating the solution's Wigner function, that was rigorously defined and derived from the Schrödinger equation by two of the authors in previous work. It is applied to two-dimensional isotropic systems, which include linear Jahn-Teller Hamiltonians, and Gaussian initial data. It reproduces energy level populations and expectation values with an accuracy of two to three percent.

**Key words.** Time-dependent Schrödinger system, eigenvalue crossing, trajectory surface hopping algorithm.

**1. Introduction.** We numerically validate a surface hopping algorithm, which stems from an asymptotic semigroup proposed in [13] for the time-evolution of two-level Schrödinger systems

$$
i\varepsilon\partial_t\psi(t) = \left(-\tfrac{\varepsilon^2}{2}\Delta_q + V(q)\right)\psi(t), \qquad \psi(0) = \psi_0 \in L^2(\mathbb{R}^d, \mathbb{C}^2), \qquad (1.1)
$$

whose real-symmetric potential matrix

$$
V(q) = \tfrac{1}{2}\operatorname{tr}V(q) + \begin{pmatrix} v_1(q) & v_2(q) \\ v_2(q) & -v_1(q) \end{pmatrix}
$$

has conically intersecting eigenvalues

$$
E^{\pm}(q) = \operatorname{tr}V(q) \pm \sqrt{v_1(q)^2 + v_2(q)^2}.
$$

We say that the two eigenvalues $E^{\pm}(q)$ intersect at a point $q_* \in \mathbb{R}^d$, if $E^+(q_*) = E^-(q_*)$. Such an intersection is called conical, if the vectors $\nabla_q v_1(q_*)$ and $\nabla_q v_2(q_*)$ are linearly independent. If all intersections are conical, the crossing set

$$
\Gamma = \left\{ q \in \mathbb{R}^d \mid E^+(q) = E^-(q) \right\}
$$

is a submanifold of codimension two in $\mathbb{R}^d$, and the matrix $V(q)$ is not smoothly diagonalizable even if

$$
V \in C^{\infty}(\mathbb{R}^d, \mathbb{R}^{2\times 2}_{\text{sym}}),
$$

see for example Chapter 2.5 in [11].

In [13] we proposed an asymptotic description for the Wigner function of the solution to the Cauchy problem (1.1) combining semiclassical propagation and surface

hopping. It has been shown rigorously in [13], that the description becomes asymptotically exact in the limit $\varepsilon \to 0$ for the special model system with linear isotropic potential

$$V_{\mathrm{lin}}(q) = \begin{pmatrix} q_1 & q_2 \\ q_2 & -q_1 \end{pmatrix}.$$

The goal of the present paper is to translate the asymptotic description to an algorithm and to understand the range of its applicability using a number of carefully controlled numerical experiments.

The analysis of time-dependent Schrödinger systems of the form (1.1) exploits the smallness of the semiclassical parameter

$$0 < \varepsilon \ll 1.$$

For molecular systems, it is comparable to the square root of the ratio of electronic and average nucleonic mass. This ratio is typically around one over two thousand, which means $\varepsilon \approx 10^{-2}$. Hence, it is appropriate to view the $\varepsilon$-scaled Laplacian as a singular perturbation of the potential matrix $V(q)$.

In absence of the Laplacian, the solution of the unperturbed system

$$\mathrm{i}\varepsilon \partial_t \widetilde{\psi} = V(q)\widetilde{\psi}, \qquad \widetilde{\psi}(0) = \widetilde{\psi}_0 \in L^2(\mathbb{R}^d, \mathbb{C}^2)$$

leaves the eigenspaces of the potential matrix $V(q)$ invariant. That is, if $\Pi^{\pm}(q)$ denote the orthogonal eigenprojectors of $V(q)$, then

$$\Pi^{\pm}\widetilde{\psi}_0 = 0 \qquad \Longrightarrow \qquad \forall t \in \mathbb{R} : \Pi^{\pm}\widetilde{\psi}(t) = 0.$$

While the $\varepsilon$-scaled Laplacian destroys the exact invariance of the eigenspaces, it is well known and at the basis of the time-dependent Born-Oppenheimer approximation, see for example [15], that the eigenspaces remain approximately invariant as long as the solution of (1.1) stays essentially supported away from the crossing set $\Gamma$. In short, away from the crossing manifold transitions between eigenspaces are of order $\varepsilon$, and this approximate decoupling is called adiabatic. However, once the solution reaches a certain $\varepsilon$-dependent neighborhood of the crossing manifold $\Gamma$, transitions between the eigenspaces are of order one. More precisely, for the Schrödinger system (1.1) there exist initial data $\psi_0^\varepsilon \in L^2(\mathbb{R}^d, \mathbb{C}^2)$ with

$$\Pi^{\pm}\psi_0^\varepsilon = 0 \qquad \& \qquad \exists t > 0 : \Pi^{\pm}\psi^\varepsilon(t) = \mathcal{O}(1), \quad \varepsilon \to 0.$$

Such transitions are called non-adiabatic. They are ubiquitous in the dynamics of polyatomic molecules involving excited electronic energy levels and explain ultrafast isomerization, radiationless decay as well as molecular collision processes, see [3] for an exposition of recent research results on conical crossings in the chemical physics' literature.

In the mentioned molecular application, the Schrödinger system (1.1) is a partial differential equation on a high-dimensional configuration space. $\mathbb{R}^d$ comprises the molecule's relevant nucleonic coordinates, and hence $d \gg 1$. Therefore, one aims at approximate descriptions of the dynamics, which allow an efficient algorithmical realization in high-dimensional situations. The approach followed in [13] avoids the direct approximation of the Schrödinger solution $\psi(t)$, but works with its Wigner function $W^\varepsilon(\psi(t))$. The Wigner function

$$W^\varepsilon(\psi(t))(q,p) = (2\pi)^{-d} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}x \cdot p} \psi(t, q - \tfrac{\varepsilon}{2}x) \otimes \overline{\psi}(t, q + \tfrac{\varepsilon}{2}x) \, \mathrm{d}x$$

is a square-integrable function on phase space $\mathbb{R}^{2d}$ with values in the space of Hermitian $2 \times 2$ matrices. At least heuristically, it is helpful to think of $\mathrm{tr}\,(W^\varepsilon(\psi(t))(q,p))$ as a density on phase space, although, in general, it attains negative values. All the physically relevant information contained in $\psi(t)$ is also contained in $W^\varepsilon(\psi(t))$, since, up to a global phase factor, the mapping $\psi \mapsto W^\varepsilon(\psi)$ is one-to-one. Also the energy level populations can be computed easily from the Wigner function through

$$\|\Pi^\pm \psi(t)\|_{L^2}^2 = \int_{\mathbb{R}^{2d}} \mathrm{tr}\,\left(W^\varepsilon(\psi(t))(q,p)\Pi^\pm(q)\right)\,\mathrm{d}q\,\mathrm{d}p. \tag{1.2}$$

We refer to appendix A for a brief summary of further properties.

In [13], an asymptotic semigroup $(\mathcal{L}_\varepsilon^t)_{t\geq 0}$ is constructed, whose action on the initial Wigner function $W^\varepsilon(\psi_0)$ approximates the Wigner function $W^\varepsilon(\psi(t))$ of the solution of the Schrödinger system (1.1) with linear isotropic potential $V_{\mathrm{lin}}$. Let us briefly sketch the analogous construction for general conical intersections. To obtain a good approximation to the $V$-diagonal components of the Wigner function

$$w_t^\pm(q,p) := \mathrm{tr}\,\left(W^\varepsilon(\psi(t))(q,p)\,\Pi^\pm(q)\right),$$

one transports the initial densities $w_0^+(q,p)$ and $w_0^-(q,p)$ independently along the flows $\Phi_+^t$ resp. $\Phi_-^t$ of the two Hamiltonian systems

$$\dot{q} = p, \qquad \dot{p} = -\nabla_q E^\pm(q), \tag{1.3}$$

as long as the trajectories stay sufficiently far away from the crossing manifold $\Gamma \times \mathbb{R}^d \subset \mathbb{R}^{2d}$. Near $\Gamma \times \mathbb{R}^d$ an effective non-adiabatic transfer of weight between $w_t^+$ and $w_t^-$ occurs. More precisely, whenever along a trajectory $(q(t),p(t))$ of the systems (1.3) the gap between the eigenvalues has a local minimum,

$$|E^+(q(t)) - E^-(q(t))| = \min!$$

a fraction $T_*$ of one density is transferred to the other level. If the point in phase space, where the minimum is attained, is denoted by $(q_*,p_*)$, the associated non-adiabatic transition rate is of the form

$$T_* = \exp\left(-\frac{\pi}{\varepsilon} \frac{(v(q_*) \wedge Dv(q_*)p_*)^2}{|Dv(q_*)p_*|^3}\right) \tag{1.4}$$

with

$$v(q) = (v_1(q),v_2(q))^T \in \mathbb{R}^2, \qquad Dv(q) = (\nabla_q v_1(q), \nabla_q v_2(q))^T \in \mathbb{R}^{2\times d}.$$

A heuristic derivation of the transition probabilities (1.4) is given in section 2, while a complete precise definition of $(\mathcal{L}_\varepsilon^t)_{t\geq 0}$ is provided in section 3.

For the case $V = V_{\mathrm{lin}}$, Theorem 3.2 of [13] proves the following leading order convergence result. If the initial data $(\psi_0^\varepsilon)_{\varepsilon>0}$ are a bounded sequence in $L^2(\mathbb{R}^d,\mathbb{C}^2)$ satisfying $\Pi^- \psi_0^\varepsilon = 0$, then

$$W^\varepsilon(\psi^\varepsilon(t)) - \mathcal{L}_\varepsilon^t W^\varepsilon(\psi_0^\varepsilon) \overset{*}{\rightharpoonup} 0, \qquad \varepsilon \to 0,$$

when evaluated against $V$-diagonal observables supported away from the crossing. $V$-diagonal observables are matrix-valued test functions $a \in C_0^\infty(\mathbb{R}^{2d},\mathbb{C}^{2\times 2})$, which

commute with the potential matrix: $[a(q,p),V(q)]=0$. A more precise formulation of the above approximation reads as

$$\sup_{t\in[0,T]}\int_{\mathbb{R}^{2d}}\operatorname{tr}\left(\left(W^{\varepsilon}(\psi^{\varepsilon}(t))(q,p)-\mathcal{L}_{\varepsilon}^{t}W^{\varepsilon}(\psi_{0}^{\varepsilon})(q,p)\right)a(q,p)\right)\mathrm{d}q\,\mathrm{d}p\to 0,\qquad\varepsilon\to 0,$$

for $T>0$ and $V$-diagonal $a\in C^{\infty}(\mathbb{R}^{2d},\mathbb{C}^{2\times 2})$ with $\operatorname{supp}(a)\cap(\Gamma\times\mathbb{R}^{d})=\emptyset$. In addition the sequence $(\psi_{0}^{\varepsilon})_{\varepsilon>0}$ must satisfy a certain tightness condition and must not concentrate in a $\sqrt{\varepsilon}$-neighborhood of the crossing manifold. For example, away from the crossing, the projections $\Pi^{\pm}(q)$ are $V$-diagonal observables and thus by (1.2) the level populations are well approximated by the action of the semigroup. For general conical intersections, the same type of convergence result can be proven [5].

The approximation by the semigroup $(\mathcal{L}_{\varepsilon}^{t})_{t\geq 0}$ has two inherent restrictions. First, only the dynamics of the $V$-diagonal components $w_{t}^{\pm}(q,p)$ of the Wigner function away from the crossing set are well approximated. The effective non-adiabatic transition rate (1.4) does not resolve the dynamics in a small $\sqrt{\varepsilon}$-vicinity of the crossing, but gives an asymptotically correct outcome when the solution has passed by. Second, neglecting the off-diagonal parts of the Wigner function, there are possible interferences of order one, which are not captured by the asymptotic semigroup description. If $w_{t}^{+}(q,p)$ and $w_{t}^{-}(q,p)$ arrive at the same time at the same phase space point $(q,p)$ close to the crossing, then a non-adiabatic transfer of weight without considering off-diagonal contributions might result in an incorrect approximation of the dynamics. For the system with linear isotropic potential $V=V_{\mathrm{lin}}$, the initial assumption $\Pi^{-}\psi_{0}=0$ guarantees that such phase space interlevel interferences do not occur: when $\Pi^{+}\psi(t)$ arrives near the crossing to perform the non-adiabatic transition, the part issued onto the lower level immediately moves away, since the lower eigenvalue $E^{-}(q)=-|q|$ is repulsive. For general conical crossings, conditions on the initial data ruling out phase space interferences are more complicated.

The semigroup $(\mathcal{L}_{\varepsilon}^{t})_{t\geq 0}$ combines classical transport and non-adiabatic transitions. The algorithmical realization, which will be proposed in section 4, is a surface hopping algorithm. The first algorithm of this type has been suggested by Tully and Preston in 1971 for the reaction of a hydrogen ion with a deuterium molecule [19]. The number of chemical publications on these algorithms is by now unmanageable, and we refer to part IV of the review [16] as a pointer to the chemical literature.

The numerical experiments to be presented in section 5 explore the performance of our algorithm in the isotropic setting, where the convergence result of [13] holds. To this end we compare energy level populations and expectation values obtained by numerically converged solutions of the original equation (1.1) to those derived from numerical expressions for $\mathcal{L}_{\varepsilon}^{t}W^{\varepsilon}$, while varying a number of parameters. More explicitly, the Schrödinger system has an isotropic potential with linear trace-free part

$$V(q)=\beta|q|^{2}+\gamma V_{\mathrm{lin}}(q),\qquad q\in\mathbb{R}^{2}.$$

Linear $E\otimes e$ Jahn-Teller Hamiltonians, which model the displacement of triatomic molecules from the equilateral triangle configuration, have potentials of exactly this form, see section 5.7. The initial data are associated with the plus eigenspace. They are products of a scalar Gaussian function,

$$\psi_{0}^{+}(q)=(\pi\varepsilon)^{-1/2}\exp\left(-\tfrac{1}{2\varepsilon}|q-q_{0}^{\varepsilon}|^{2}+\tfrac{\mathrm{i}}{\varepsilon}p_{0}\cdot(q-q_{0}^{\varepsilon})\right),$$

whose support has negligible overlap with the crossing region, and a real-valued eigenfunction of $V(q)$ associated with the eigenvalue $E^{+}(q)$. Such initial conditions are

typical for molecular systems, when the initial data result from the vertical laser excitation of an electronic Gaussian ground state function up to an excited electronic level, see for example [14]. The time interval under consideration is chosen small enough, such that the solution passes the crossing only once. The core parameter set of the experiments fixes the semiclassical parameter $\varepsilon$, the position and momentum center $(q_0^\varepsilon, p_0) \in \mathbb{R}^4$ of the initial Gaussian, the quadratic confinement $\beta$ and the width of the eigenvalue cones $\gamma$ as

$$\varepsilon = 10^{-2}, \qquad q_0^\varepsilon = (5\sqrt{\varepsilon}, \tfrac{1}{2}\sqrt{\varepsilon}), \qquad p_0 = (-1,0), \qquad \beta = 0, \qquad \gamma = 1.$$

The five sets of experiments vary these five parameters separately, while leaving the other ones fixed. Note that the $\varepsilon$-dependence of the average position $q_0^\varepsilon$ of the initial Gaussian $\psi_0^+$ is complemented by its width being $\varepsilon$-dependent as well. The $\varepsilon$-dependent width appears naturally for an laser-excited molecular ground state. Moreover, this combination of $\varepsilon$-dependent position and width brings forward non-adiabatic transitions of leading order. For all parameter variations, the solution keeps an average momentum of order one. Hence, the adiabatic decoupling is only violated by the conical crossing and not by the occurence of large momenta.

The solutions for the Schrödinger equation (1.1) are computed by a Strang splitting scheme with Fourier differencing, which is a standard grid based discretization scheme. Then, we compare expectation values and populations obtained from $\mathcal{L}_\varepsilon^t W^\varepsilon$ and the direct solution. Since both solvers are numerically converged, we only observe the methodical error of the surface hopping algorithm.

In a nutshell, the outcomes of our experiments are the following: For the computations with $\varepsilon = 10^{-2}$, the absolute error for the level populations is around 0.02, and the relative error for the expectation values is around one to three percent. When varying the semiclassical parameter $\varepsilon$, the error behaves like $\sqrt{\varepsilon}$ for all quantities. A decrease of the initial momentum leads to bigger errors, but even with zero initial momentum and $\varepsilon = 10^{-2}$ the error in the level-population is only 0.025. Other trends are the increase in error when starting closer to the crossing manifold and the decrease in error when increasing the quadratic confinement $\beta$ or decreasing the cone width $\gamma$. In all experiments the surface hopping algorithm underestimates the amount of non-adiabatic transfer of weight. The detailed outcomes are discussed in section 5.

In view of real-life molecular simulations, we add two remarks: First, the model parameters for the potential matrix usually stem from difficult electronic structure calculations and typically cannot be determined with high accuracy. Therefore, a methodical resolution of two to three percent seems appropriate in this context. Second, when aiming at an accuracy of two percent for our two-dimensional experiments, the computational effort of both algorithms is comparable. (On a 3 GHz Pentium 4 computer, our *Matlab 7.0* implementations take roughly one minute computing time.) However, conventional grid based discretizations cannot be set up in high-dimensional situations, while surface hopping algorithms with their combination of classical transport and non-adiabatic transitions are practicable also for systems with many degrees of freedom. Moreover, the numerical computation of uncoupled classical trajectories can easily be done in parallel. Aiming at the validation against a practicable consistent discretization scheme, however, this article can only present two-dimensional numerical experiments.

The article is organized as follows. In the next sections, we give a heuristic argument for asymptotically correct non-adiabatic transition rates (§2), define the semigroup for general conical intersections (§3), and formulate a corresponding surface

hopping algorithm (§4). Then, we present and discuss the numerical experiments (§5). The appendix summarizes basic properties of Wigner functions, Weyl quantization, and of the discretizations used.

**2. Non-adiabatic transitions: a heuristic picture.** The following heuristic derivation of the correct non-adiabatic transition rate $T_*(q_*, p_*)$ slightly generalizes an similar argument given in [13]. The basic idea is to insert the trajectories $(q(t), p(t))$ of the Hamiltonian systems (1.3) into the trace-free part of the potential matrix and obtain an ordinary differential system

$$\mathrm{i}\varepsilon \tfrac{\mathrm{d}}{\mathrm{d}t}\varphi(t) = \begin{pmatrix} v_1(q(t)) & v_2(q(t)) \\ v_2(q(t)) & -v_1(q(t)) \end{pmatrix} \varphi(t). \tag{2.1}$$

The crucial points $(q_*, p_*)$ in phase space are those, at which the gap between the eigenvalues

$$g(q(t)) = |E^+(q(t)) - E^-(q(t))| = 2|v(q(t))|$$

attains a local minimum along the classical trajectory $(q(t), p(t))$. A necessary condition for a trajectory attaining its minimal gap is thus given by

$$\tfrac{\mathrm{d}}{\mathrm{d}t}|v(q(t))|^2 = 0,$$

or equivalently by $v(q(t)) \cdot Dv(q(t))p(t) = 0$. Hence, a submanifold of phase space containing these points of minimal gap is

$$S = \big\{ (q,p) \in \mathbb{R}^{2d} \,|\, v(q) \cdot Dv(q)p = 0 \big\}.$$

Though we will refer to $S$ as the submanifold of minimal gap, it consists of those phase space points, along which trajectories attain their extremal gap. Whether some point $(q,p) \in S$ corresponds to a minimal gap, may be checked by differentiating the function $t \mapsto |v(q(t))|^2$ twice. This gives an additional explicit phase space condition, depending on the underlying flow $\Phi_\pm^t$. For the trace-free linear isotropic problem in two dimensions, that is for $V = V_{\mathrm{lin}}$, this condition reads as $|p|^2 \mp |q| > 0$.

Let a trajectory hit the submanifold $S$ at a point $(q_*, p_*) \in S$ at some time $t_* \geq 0$. For notational simplicity, we set $t_* = 0$. Taylor expanding,

$$v_j(q(t)) = v_j(q_*) + t\nabla_q v_j(q_*) \cdot p_* + \mathcal{O}(t^2), \qquad j = 1, 2,$$

and plugging the linear part into (2.1), one obtains

$$\mathrm{i}\varepsilon \tfrac{\mathrm{d}}{\mathrm{d}t}\varphi(t) = \begin{pmatrix} v_1(q_*) + t\nabla_q v_1(q_*)p_* & v_2(q_*) + t\nabla_q v_2(q_*)p_* \\ v_2(q_*) + t\nabla_q v_2(q_*)p_* & -v_1(q_*) - t\nabla_q v_1(q_*)p_* \end{pmatrix} \varphi(t). \tag{2.2}$$

After conjugation with the orthogonal matrix

$$\begin{pmatrix} \cos\vartheta_* & -\sin\vartheta_* \\ \sin\vartheta_* & \cos\vartheta_* \end{pmatrix}, \qquad \vartheta_* = \tfrac{1}{2}\arccos\left( \frac{\nabla_q v_1(q_*) \cdot p_*}{|Dv(q_*)p_*|} \right),$$

the system (2.2) turns into

$$\mathrm{i}\frac{\varepsilon}{|Dv(q_*)p_*|} \tfrac{\mathrm{d}}{\mathrm{d}t}\varphi(t) = \begin{pmatrix} t & \frac{v(q_*) \wedge Dv(q_*)p_*}{|Dv(q_*)p_*|^2} \\ \frac{v(q_*) \wedge Dv(q_*)p_*}{|Dv(q_*)p_*|^2} & -t \end{pmatrix} \varphi(t) \tag{2.3}$$

with $v \wedge w = v^{\perp} \cdot w = v_2 w_1 - v_1 w_2$ for vectors $v, w \in \mathbb{R}^2$. We note, that the calculations transforming (2.2) to (2.3) rely on the identities

$$\cos^2 \vartheta_* - \sin^2 \vartheta_* = \nabla_q v_1(q_*) \cdot p_* / |Dv(q_*)p_*|,$$
$$2 \cos \vartheta_* \sin \vartheta_* = \nabla_q v_2(q_*) \cdot p_* / |Dv(q_*)p_*|.$$

Setting

$$\widetilde{\varepsilon} = \varepsilon / |Dv(q_*)p_*|, \qquad \delta = v(q_*) \wedge Dv(q_*)p_* / |Dv(q_*)p_*|^2,$$

system (2.3) is clearly seen as a well known linear Landau-Zener problem,

$$i\widetilde{\varepsilon} \tfrac{\mathrm{d}}{\mathrm{d}t} \varphi(t) = \begin{pmatrix} t & \delta \\ \delta & -t \end{pmatrix} \varphi(t).$$

For such a Landau-Zener problem, the non-adiabatic transitions between the eigenspaces of the eigenvalues $\pm\sqrt{t^2 + \delta^2}$ occur in a neighbourhood of the minimal gap, that is at $t = 0$. Let $\chi^{\pm}(t)$ denote normalized eigenvectors of the Landau-Zener matrix and decompose the solution $\varphi(t)$ as $\varphi(t) = \varphi^+(t)\chi^+(t) + \varphi^-(t)\chi^-(t)$. If for large negative times $|\varphi^{\pm}(-\infty)| = 1$ and $|\varphi^{\mp}(-\infty)| = 0$ hold, then

$$\begin{pmatrix} |\varphi^+(+\infty)|^2 \\ |\varphi^-(+\infty)|^2 \end{pmatrix} = \begin{pmatrix} 1 - T_{\mathrm{LZ}} & T_{\mathrm{LZ}} \\ T_{\mathrm{LZ}} & 1 - T_{\mathrm{LZ}} \end{pmatrix} \begin{pmatrix} |\varphi^+(-\infty)|^2 \\ |\varphi^-(-\infty)|^2 \end{pmatrix}$$

with

$$T_{\mathrm{LZ}} \sim \exp\left(-\frac{\pi}{\widetilde{\varepsilon}} \delta^2\right) \qquad (\widetilde{\varepsilon} \to 0).$$

Retranslating this as a non-adiabatic transition rate for conical intersections, one obtains

$$T_* = \exp\left(-\frac{\pi}{\varepsilon} \frac{(v(q_*) \wedge Dv(q_*)p_*)^2}{|Dv(q_*)p_*|^3}\right).$$

We note, that the heuristically derived transition rates $T_*$ coincide with those resulting from a rigorous microlocal normal form construction, see [5].

**3. An asymptotic semigroup.**    Given the transition rates $T_*$, one combines them with the classical flows $\Phi_{\pm}^t$ and the associated trajectories $(q^{\pm}(t), p^{\pm}(t))$ to an asymptotic description of propagation through conical crossings.

For this, one attaches to phase space points $(q, p) \in \mathbb{R}^{2d}$ a label $-1$ or $+1$, indicating reference to the eigenspace of the eigenvalue $E^-$ or $E^+$. We denote

$$\mathbb{R}_{\pm}^{2d} := \mathbb{R}^{2d} \times \{-1, +1\},$$

and define for labelled phase space points $(q, p, j) \in \mathbb{R}_{\pm}^{2d}$ random trajectories

$$\mathcal{J}_{\varepsilon}^{(q,p,j)} : [0, +\infty[ \to \mathbb{R}_{\pm}^{2d}, \qquad t \mapsto \mathcal{J}_{\varepsilon}^{(q,p,j)}(t),$$

such that $\mathcal{J}_{\varepsilon}^{(q,p,j)}(t) = (\Phi_j^t(q,p), j)$ as long as $v(q^j(t)) \cdot Dv(q^j(t))p^j(t) \neq 0$. A jump from $j$ to $-j$ occurs with probability $T_* = T_*(q_*, p_*)$, whenever $\Phi_j^t(q,p)$ hits the manifold of minimal gap $S$ at time $t = t_*$ in the point $(q_*, p_*) \in S$.

The random trajectories $\mathcal{J}_\varepsilon^{(q,p,j)}$ define a Markov process. The transition function

$$\mathbb{P}_\varepsilon^{(q,p,j)} : (t,\Omega) \mapsto \mathbb{P}_\varepsilon^{(q,p,j)}(t,\Omega)$$

of the Markov process gives the probability of being at time $t \in [0,+\infty[$ in the measurable set $\Omega \subset \mathbb{R}_\pm^{2d}$ having started in the state $(q,p,j)$. The associated backwards semigroup $(\mathcal{L}_\varepsilon^t)_{t \geq 0}$ is defined by its action on bounded measurable functions $a : \mathbb{R}_\pm^{2d} \to \mathbb{C}$,

$$(\mathcal{L}_\varepsilon^t a)(q,p,j) = \int_{\mathbb{R}_\pm^{2d}} a(x,\xi,k)\, \mathbb{P}_\varepsilon^{(q,p,j)}(t,\mathrm{d}(x,\xi,k)).$$

A key property of our semigroup is its independence on the choice of a diabatic basis. More precisely, the main constituents of the construction, namely the Hamiltonian systems (1.3), the manifold of minimal gap $S$, and the non-adiabatic transition rates $T_*$ remain unchanged, if the potential matrix $V(q)$ is replaced by

$$V(q) \longrightarrow U(q)V(q)U(q)^*$$

for any unitary matrix $U(q) \in \mathbb{C}^{2 \times 2}$ with smooth $q$-dependence. Hence, the semigroup $(\mathcal{L}_\varepsilon^t)_{t \geq 0}$ is invariant under unitary changes of the potential. Such an invariance property is important for molecular applications, when the potential is not uniquely determined by the electronic structure calculation but dependent on the choice of a diabatic basis, see [12].

**4. A surface hopping algorithm.** In this section we describe a surface hopping algorithm for numerically implementing the semigroup $(\mathcal{L}_\varepsilon^t)_{t>0}$. Recall that our goal is an accurate computation of the $V$-diagonal elements of $\mathcal{L}_\varepsilon^t W^\varepsilon$ to numerically quantify the systematic error of $(\mathcal{L}_\varepsilon^t)_{t>0}$, and not the design of an optimized algorithm. The principal algorithmic steps are the following, whereas the details of our implementation are given in section 5.

*1. Initial sampling.* Let

$$w_0^\pm(q,p) = \mathrm{tr}\left(W^\varepsilon(\psi_0)(q,p)\,\Pi^\pm(q)\right)$$

denote the $V$-diagonal components of the initial Wigner matrix. By an appropriate sampling of $w_0^\pm$ one chooses a set of sampling points

$$\left\{(q_k,p_k,j_k) \in \mathbb{R}^{2d} \times \{-,+\} \mid k=1,\ldots,N\right\}$$

with associated weights $w_k \in \mathbb{R}$ given by the values of the corresponding diagonal Wigner components

$$w_k = w_0^{j_k}(q_k,p_k).$$

*2. Hopping transport.* Away from the manifold $S$, each particle $(q_k,p_k,j_k)$ is transported by the associated classical flow $\Phi_{j_k}^t$ of the Hamiltonian system (1.3), i.e.

$$(q_k(t),p_k(t)) = \Phi_{j_k}^t(q_k,p_k) \quad \text{and} \quad j_k(t) = j_k, \quad w_k(t) = w_k.$$

However, when $(q_k(t),p_k(t))$ passes $S$, say at some time $t = t_* > 0$, a branching occurs, i.e. the weight $w_k$ is reduced and a new particle with index $l > N$ is created. More precisely, the hopping probability is

$$T_k(t_*) = T_*(q_k(t_*),p_k(t_*)).$$

Hence the weight $w_k$ changes as

$$w_k(t_*) \to (1 - T_k(t_*)) \, w_k(t_*)$$

and the new particle $l$ is created with

$$(q_l(t_*), p_l(t_*)) = (q_k(t_*), p_k(t_*)) \quad \text{and} \quad j_l(t_*) = -j_k(t_*), \quad w_l(t_*) = T_k(t_*) w_k(t_*).$$

Note that all particles are transported independently and therefore this part of the algorithm can easily be performed in parallel.

    *3. Final reconstruction.* At some final time $t_\mathrm{f} > 0$ one obtains $M \geq N$ points

$$\left\{ (q_k, p_k, j_k) \in \mathbb{R}^{2d} \times \{-, +\} \mid k = 1, \ldots, M \right\}$$

with associated weights $w_k$, which approximate the diagonal Wigner components

$$w_k \approx w_{t_\mathrm{f}}^{j_k}(q_k, p_k).$$

According to (1.2) these are for an approximation to the level populations,

$$P^\pm := \|\Pi^\pm \psi(t_\mathrm{f})\|_{L^2}^2 = \int_{\mathbb{R}^{2d}} w_{t_\mathrm{f}}^\pm(q, p) \, \mathrm{d}q \, \mathrm{d}p \approx \sum_{k=1}^M w_k \, \delta_{j_k, \pm} \omega_k =: P_\mathrm{hop}^\pm,$$

where $\delta_{j_k, \pm}$ is the Kronecker delta and $\omega_k$ an appropriate quadrature weight. Our numerical realization (see section 5) has a grid based initial sampling and uses the grid's volume element as quadrature weight for the final reconstruction. However, especially in high dimensional situations one might choose alternative quadrature approaches.

    More generally than level populations, one can compute approximations to expectation values of semiclassical observables of the projected final wave functions

$$\psi^\pm(t_\mathrm{f}, q) = \Pi^\pm(q) \psi(t_\mathrm{f}, q).$$

That is, if $a(q, -i\varepsilon \nabla_q)$ denotes the Weyl quantized operator associated with the phase space function $a = a(q, p) \in C^\infty(\mathbb{R}^{2d}, \mathbb{C})$, see appendix A for details, then

$$(P^\pm)^{-1} \left\langle \psi^\pm(t_\mathrm{f}), a(q, -\mathrm{i}\varepsilon \nabla_q) \psi^\pm(t_\mathrm{f}) \right\rangle_{L^2(\mathbb{R}^d, \mathbb{C}^2)} \approx (P_\mathrm{hop}^\pm)^{-1} \sum_{k=1}^M a(q_k, p_k) \, w_k \, \delta_{j_k, \pm} \omega_k.$$

For example, the average position on the lower band at time $t_\mathrm{f}$ is

$$(P^-)^{-1} \left\langle \psi^-(t_\mathrm{f}), q \psi^-(t_\mathrm{f}) \right\rangle_{L^2(\mathbb{R}^d, \mathbb{C}^2)} \approx (P_\mathrm{hop}^-)^{-1} \sum_{k=1}^M q_k \, w_k \, \delta_{j_k, -} \omega_k,$$

or the average momentum on the upper band is

$$(P^+)^{-1} \left\langle \psi^+(t_\mathrm{f}), -\mathrm{i}\varepsilon \nabla_q \psi^+(t_\mathrm{f}) \right\rangle_{L^2(\mathbb{R}^d, \mathbb{C}^2)} \approx (P_\mathrm{hop}^+)^{-1} \sum_{k=1}^M p_k \, w_k \, \delta_{j_k, +} \omega_k.$$

    **5. Numerical experiments.** In this section, we first present the general setup of our numerical experiments and the precise implementation of the algorithm proposed in section 4. Then, we give the results of the experiments, which vary several problem parameters independently of each other. In particular, we discuss how the methodological error of the surface hopping algorithm depends on the choice of parameters and use the asymptotic semigroup for a qualitative explanation of the observed tendencies of non-adiabatic transfer.

**5.1. The setup.**    Our experiments cover a family of two-dimensional isotropic model problems,

$$i\varepsilon\partial_t\psi = \left(-\tfrac{\varepsilon^2}{2}\Delta_q + V(q)\right)\psi, \qquad \psi(0) = \psi_0 \in L^2(\mathbb{R}^2, \mathbb{C}^2), \qquad (5.1)$$

with quadratic confinement and linear trace-free part

$$V(q) = \beta|q|^2 + \gamma\begin{pmatrix} q_1 & q_2 \\ q_2 & -q_1 \end{pmatrix}, \qquad \beta, \gamma > 0,$$

and initial data associated with the upper level,

$$\psi_0(q) = \psi_0^+(q)\chi^+(q),$$

where

$$\chi^+(q) = (\cos(\vartheta(q)/2), \sin(\vartheta(q)/2))^T$$

denotes a normalized real-valued eigenvector of the potential matrix $V(q)$ associated with the eigenvalue $\beta|q|^2 + \gamma|q|$. Here, $\vartheta(q) \in (-\pi, \pi)$ is the polar angle of $q \in \mathbb{R}^2$. Therefore, $\chi^+(q)$ is discontinuous over the left half axis $\{q \in \mathbb{R}^2 \,|\, q_1 \leq 0, q_2 = 0\}$. The upper level function is an $\varepsilon$-scaled Gaussian

$$\psi_0^+(q) = (\pi\varepsilon)^{-1/2}\exp\left(-\tfrac{1}{2\varepsilon}|q - q_0^\varepsilon|^2 + \tfrac{i}{\varepsilon}p_0 \cdot (q - q_0^\varepsilon)\right)$$

centered in $(q_0^\varepsilon, p_0) \in \mathbb{R}^2 \times \mathbb{R}^2$ with $\|\psi_0^+\|_{L^2} = 1$. The center in position space $q_0^\varepsilon \in \mathbb{R}^2$ is chosen, such that the overlap of the Gaussian's support with the crossing at $q = 0$ and also with the discontinuities of $\chi^+(q)$ is negligible. Gaussian initial data have an explicitly computable Wigner transform,

$$W^\varepsilon(\psi_0^+)(q, p) = (\pi\varepsilon)^{-2}\exp\left(-\tfrac{1}{\varepsilon}|q - q_0^\varepsilon|^2 - \tfrac{1}{\varepsilon}|p - p_0|^2\right).$$

All our numerical experiments are variations of the basic parameter values

$$\beta = 0, \qquad \gamma = 1, \qquad \varepsilon = 10^{-2}, \qquad [t_i, t_f] = [0, 10\sqrt{\varepsilon}],$$
$$q_0^\varepsilon = (5\sqrt{\varepsilon}, \tfrac{1}{2}\sqrt{\varepsilon}), \qquad p_0 = (-1, 0). \qquad (5.2)$$

Plots of the dynamics of this prototypical system are given in figure 5.1, which shows the position density of the projected wave function $|\Pi^\pm(q)\psi(t, q)|^2$ at different points of time. Initially, the solution stays in the upper level. However, when it approaches the crossing point $q = 0$ around time $t = 0.4$, there are non-adiabatic transitions to the lower level, which are of leading order in the semiclassical parameter $\varepsilon$, see also table 5.1. Then, the remaining part of the wave function on the upper level approximately evolves with respect to the confining potential $E^+(q) = +|q|$, while on the lower level the repulsive potential $E^-(q) = -|q|$ quickly carries the wave function away from the crossing. In most of our numerical experiments, the semiclassical parameter is chosen as $\varepsilon = 10^{-2}$, which corresponds to the situation expected for a real life's molecular system. The time interval $[0, t_f]$ of our experiments has been chosen small enough, such that the solution stays localized in a fixed domain of position space and encounters the crossing set just once.
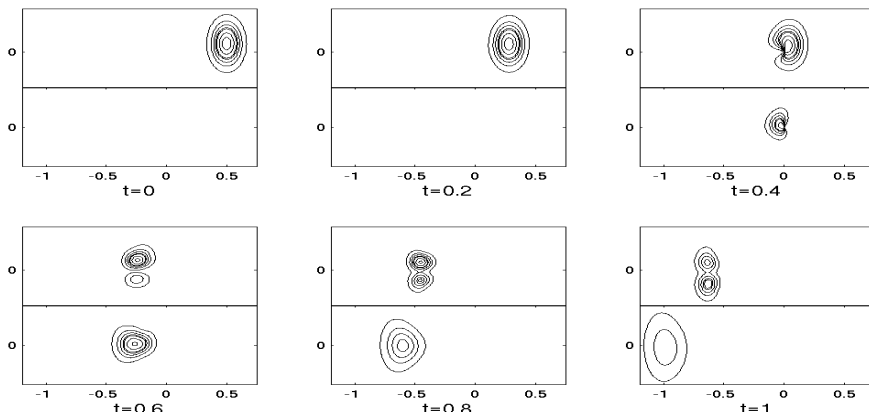
FIG. 5.1. *The propagation of the Schrödinger system* (5.1) *with parameters and initial values as given in* (5.2) *at different points of time* $t = 0, 0.2, 0.4, 0.6, 0.8, 1$. *Each of the six plots shows the position density of the projected wave function* $|\Pi^+(q)\psi(t,q)|^2$ *in the upper half and* $|\Pi^-(q)\psi(t,q)|^2$ *in the lower half. Around time* $t = 0.4$, *the wave function comes close enough to the crossing at* $(q_1, q_2) = (0,0)$ *to perform the non-adiabatic transition to the lower level.*

TABLE 5.1. *The energy level populations* $P_{\mathrm{Schr}}^-$ *and* $P_{\mathrm{hop}}^-$ *of the lower level at time* $t = t_{\mathrm{f}}$ *obtained by directly solving Schrödinger's equation* (5.1) *and by the surface hopping algorithm. The experiments vary the semiclassical parameter* $\varepsilon$, *the initial momentum* $p_0$ *and initial position* $q_0^\varepsilon$ *as well as the quadratic confinement* $\beta$ *and the cone width* $\gamma$. *The numbers illustrate non-adiabatic transitions to leading order in* $\varepsilon$. *The surface hopping algorithm underestimates the non-adiabatic transfer with an error ranging between* 0.001 *and* 0.034.

| $\varepsilon$ | $P_{\mathrm{Schr}}^-$ | $P_{\mathrm{hop}}^-$ | $-p_{0,1}$ | $P_{\mathrm{Schr}}^-$ | $P_{\mathrm{hop}}^-$ | $q_{0,1}^\varepsilon$ | $P_{\mathrm{Schr}}^-$ | $P_{\mathrm{hop}}^-$ |
|---|---|---|---|---|---|---|---|---|
| $10^{-4}$ | 0.432 | 0.431 | 0 | 0.752 | 0.729 | 0.4 | 0.557 | 0.537 |
| $10^{-3}$ | 0.479 | 0.471 | 0.5 | 0.645 | 0.621 | 0.5 | 0.578 | 0.559 |
| $10^{-2}$ | 0.578 | 0.559 | 1 | 0.578 | 0.559 | 0.6 | 0.595 | 0.578 |
| $10^{-1}$ | 0.640 | 0.606 | 2 | 0.595 | 0.586 | 0.7 | 0.607 | 0.592 |
|  |  |  | 3 | 0.643 | 0.638 | 0.8 | 0.617 | 0.602 |

| $\beta$ | $P_{\mathrm{Schr}}^-$ | $P_{\mathrm{hop}}^-$ | $\gamma$ | $P_{\mathrm{Schr}}^-$ | $P_{\mathrm{hop}}^-$ |
|---|---|---|---|---|---|
| 0 | 0.578 | 0.559 | $\frac{1}{3}$ | 0.673 | 0.659 |
| 1 | 0.644 | 0.629 | 0.5 | 0.629 | 0.613 |
| 2 | 0.695 | 0.684 | 1 | 0.578 | 0.559 |
| 3 | 0.736 | 0.727 | 2 | 0.561 | 0.542 |
| 4 | 0.770 | 0.762 | 3 | 0.566 | 0.547 |
| 5 | 0.796 | 0.791 |  |  |  |

**5.2. Implementation.**    The numerical experiments have been implemented along the following four steps.

*0. Schrödinger solution.* The initial value problem (5.1) is solved by a numerically converged Strang splitting scheme with Fourier differencing on a uniform $N_1 \times N_2$ grid with $N_1, N_2 \in \mathbb{N}$, see appendices B and C. For time $t = t_{\mathrm{f}}$, level populations or intralevel expectation values for position, momentum, and energy are computed. In the following, we denote them by

$$P_{\mathrm{Schr}}^\pm, \quad E_{\mathrm{p,Schr}}^\pm, \quad E_{\mathrm{m,Schr}}^\pm, \quad E_{\mathrm{E,Schr}}^\pm,$$

respectively. As in the splitting scheme, the action of the differential operators $-i\varepsilon\partial_j$ and $-\varepsilon^2\Delta_q$ is realized by Fourier multiplication.

*1. Initial sampling.* The Gaussian $\psi_0^+(q)$ and its $\varepsilon$-scaled Fourier transform

$$(\mathcal{F}^\varepsilon\psi_0^+)(p) = (\pi\epsilon)^{-1/2}\exp\left(-\tfrac{1}{2\varepsilon}|p-p_0|^2 - \tfrac{i}{\varepsilon}q_0^\varepsilon\cdot p\right)$$

are discretized on the domains

$$[q_{0,1}^\varepsilon - 5\sqrt{\varepsilon}, q_{0,1}^\varepsilon + 5\sqrt{\varepsilon}] \times [q_{0,2}^\varepsilon - 5\sqrt{\varepsilon}, q_{0,2}^\varepsilon + 5\sqrt{\varepsilon}],$$
$$[p_{0,1} - 5\sqrt{\varepsilon}, p_{0,1} + 5\sqrt{\varepsilon}] \times [p_{0,2} - 5\sqrt{\varepsilon}, p_{0,2} + 5\sqrt{\varepsilon}]$$

by uniform $m \times m$ grids. The number of grid points per direction of position and momentum space are called $m_\mathrm{p}$ and $m_\mathrm{m}$, respectively. The default choice is

$$m_\mathrm{p} = m_\mathrm{m} = 16.$$

The grid points $q_k$, $k=1,\ldots,m_\mathrm{p}^2$, and $p_k$, $k=1,\ldots,m_\mathrm{m}^2$ are ordered according to the size of the densities $|\psi_0^+(q)|^2$ and $|\mathcal{F}^\varepsilon\psi_0^+(p)|^2$, respectively. Let $\omega_\mathrm{p}$ and $\omega_\mathrm{m}$ denote the volume elements of the grids in position and momentum space. Let

$$0 < \mathrm{tol}_\mathrm{p}, \mathrm{tol}_\mathrm{m}, \mathrm{tol} \ll 1$$

be predescribed sampling tolerances. One determines the minimal number $N_\mathrm{p}, N_\mathrm{m} \in \mathbb{N}$ of points $q_1,\ldots,q_{N_\mathrm{p}}$ and $p_1,\ldots,p_{N_\mathrm{m}}$ such that

$$\sum_{k=1}^{N_\mathrm{p}}|\psi_0^+(q_k)|^2\omega_\mathrm{p} \geq 1-\mathrm{tol}_\mathrm{p}, \qquad \sum_{k=1}^{N_\mathrm{m}}|(\mathcal{F}^\varepsilon\psi_0^+)(p_k)|^2\omega_\mathrm{m} \geq 1-\mathrm{tol}_\mathrm{m}.$$

The resulting $N_\mathrm{p}\cdot N_\mathrm{m}$ phase space points $(q,p)$ are ordered according to the size of the Wigner function $W^\varepsilon(\psi_0^+)(q,p)$ as $(q_k,p_k)$ with $k=1,\ldots,N_\mathrm{p}\cdot N_\mathrm{m}$, and the minimal number $N \in \mathbb{N}$ is determined such that

$$\sum_{k=1}^{N}W^\varepsilon(\psi_0^+)(q_k,p_k)\omega_\mathrm{p}\omega_\mathrm{m} \geq 1-\mathrm{tol}$$

holds. Ensuring that the Wigner function has more than mass $(1-\mathrm{tol})$ on the points suggested by the presampling in position and momentum space, we always choose

$$\mathrm{tol}_\mathrm{p} = \mathrm{tol}_\mathrm{m} = \mathrm{tol}\cdot10^{-3}.$$

Finally, one attaches to each initial sampling point in

$$\left\{(q_k,p_k,+)\in\mathbb{R}^{2d}\times\{-,+\} \mid k=1,\ldots,N\right\}$$

the value $w_k = W^\varepsilon(\psi_0^+)(q_k,p_k)$. We note, that the Wigner function of the upper level function $\psi_0^+$ is an $\varepsilon$-approximation to the $V$-diagonal component $w_0^+$ of the Wigner matrix $W^\varepsilon(\psi_0)$, see appendix A. Since the methodical error of the hopping algorithm is larger than order $\varepsilon$, this approximation is negligible.

*2. Hopping transport.* The initial points $(q_k,p_k,+)$, $k=1,\ldots,N$, are transported by the Hamiltonian flow $\Phi_+^t$, which is discretized by the explicit Runge-Kutta `DOPRI45`

method of Dormand and Prince [2]. For checking, whether a trajectory has passed the manifold of minimal gap $S$, the event function

$$t \mapsto v(q(t)) \cdot Dv(q(t))p(t) = \gamma^2 (q(t) \cdot p(t))$$

is evaluated after each timestep, and possible zeros are detected by the Regula-Falsi method. This root search relies on the interpolation routine of `DOPRI45` and does not require additional evaluations of the Hamiltonian system's right hand side, see appendix B for details. At the points of extremal gap, the additional condition $|p|^2 - \gamma|q| \geq 0$ is checked for identifying the minima. Whenever a trajectory encounters a point $(q_*, p_*) \in S$ of minimal gap at time $t_*$, the associated weight changes from $w_k$ to $(1 - T_*)w_k$ with

$$T_* = \exp\left( -\frac{\pi}{\varepsilon} \frac{(v(q_*) \wedge Dv(q_*)p_*)^2}{|Dv(q_*)p_*|^3} \right) = \exp\left( -\frac{\pi}{\varepsilon} \frac{\gamma|q_*|^2}{|p_*|} \right). \tag{5.3}$$

Moreover, a new trajectory for the lower level flow $\Phi_-^t$ with $\Phi_-^{t_*}(q_*, p_*) = (q_*, p_*)$ and associated weight $T_* w_k$ is opened up. The time interval $[0, t_{\mathrm{f}}]$ under consideration is small enough, such that each upper level trajectory initiates such a non-adiabatic transfer of weight at most once.

  *3. Final reconstruction.* At time $t = t_{\mathrm{f}}$, one has a set of points

$$\left\{ (q_k, p_k, j_k) \in \mathbb{R}^{2d} \times \{-, +\} \mid k = 1, \ldots, M \right\},$$

with associated weights $w_k \in \mathbb{R}$, which are approximations to the $V$-diagonal components $w_{t_{\mathrm{f}}}^\pm(q, p)$ of the true final Wigner matrix evaluated in $(q, p) = (q_k, p_k)$. By Liouville's theorem, the Hamiltonian flows $\Phi_\pm^t$ preserve phase space volume. By construction, the non-adiabatic transfer of weight at the manifold of minimal gap $S$ is volume preserving, too. Hence, the approximate dynamics conserve the initial volume element of phase space $\omega_{\mathrm{p}} \omega_{\mathrm{m}}$. One computes approximations to the final energy level populations or the intralevel expectation values of position, momentum, and energy by the following Riemann sums:

$$P_{\mathrm{hop}}^\pm = \sum_{k=1}^M w_k \, \delta_{j_k, \pm} \, \omega_{\mathrm{p}} \omega_{\mathrm{m}}$$

or

$$E_{\mathrm{p,hop}}^\pm = (P_{\mathrm{hop}}^\pm)^{-1} \sum_{k=1}^M q_k \, w_k \, \delta_{j_k, \pm} \, \omega_{\mathrm{p}} \omega_{\mathrm{m}}, \quad E_{\mathrm{m,hop}}^\pm = (P_{\mathrm{hop}}^\pm)^{-1} \sum_{k=1}^M p_k \, w_k \, \delta_{j_k, \pm} \, \omega_{\mathrm{p}} \omega_{\mathrm{m}},$$

and

$$E_{\mathrm{E,hop}}^\pm = (P_{\mathrm{hop}}^\pm)^{-1} \sum_{k=1}^M \left( \tfrac{1}{2}|p_k|^2 \pm \gamma|q_k| \right) w_k \, \delta_{j_k, \pm} \, \omega_{\mathrm{p}} \omega_{\mathrm{m}}.$$

  **5.3. Initial sampling.**    Using the basic parameter set (5.2) as a starting point, we examine the dependence of the numerical error of the hopping algorithm on the sampling tolerance and the grid size. Let $P_{\mathrm{Schr}}^\pm$ and $P_{\mathrm{hop}}^\pm$ denote the level populations computed by the numerically converged Strang splitting scheme and by the surface hopping algorithm, respectively.
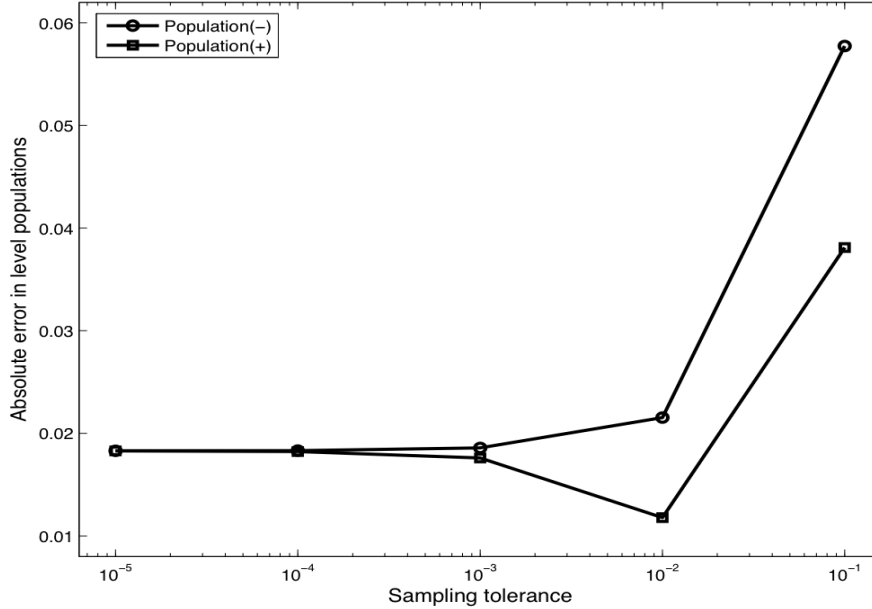
FIG. 5.2. *A semilogarithmic plot of the absolute errors of energy level populations with respect to the sampling tolerance* $\mathrm{tol} = 10^{-k}$, $k = 5, 4, 3, 2, 1$. *The line with circles corresponds to the lower level, the line with squares to the upper level. The sampling grids are of size* $(m_{\mathrm{p}}, m_{\mathrm{m}}) = (16, 16)$. *The plot shows, that the error is bounded by* $\frac{2}{3}\sqrt{\varepsilon} \approx 0.067$ *and stabilizes for tolerances smaller than* $\mathrm{tol} = 10^{-3} = \varepsilon/10$.

The semilogarithmic plot in figure 5.2 shows the absolute error of level populations $|P_{\mathrm{Schr}}^{\pm} - P_{\mathrm{hop}}^{\pm}|$ for the phase space sampling tolerance varying as

$$\mathrm{tol} = 10^{-k}, \qquad k = 5, 4, 3, 2, 1,$$

while the sampling grids are of size $(m_{\mathrm{p}}, m_{\mathrm{m}}) = (16, 16)$. One observes a stabilization of the error between $\mathrm{tol} = 10^{-3}$ and $\mathrm{tol} = 10^{-5}$ at a level of $0.02 = \sqrt{\varepsilon}/5$, which is the methodical error of the surface hopping algorithm. Table 5.2 gives the corresponding

TABLE 5.2. *The initial number of particles when sampling with grid size* $(m_{\mathrm{p}}, m_{\mathrm{m}}) = (16, 16)$ *for varying tolerance* $\mathrm{tol} > 0$. *The sampling with* $\mathrm{tol} = 10^{-3}$ *and* 2188 *particles results in an absolute error of level populations of size* 0.02.

| tol | $10^{-5}$ | $10^{-4}$ | $10^{-3}$ | $10^{-2}$ | $10^{-1}$ |
|---|---|---|---|---|---|
| ♯ particles | 5077 | 3531 | 2188 | 1159 | 410 |

number of initial particles for the different sampling tolerances at fixed grid size $(m_{\mathrm{p}}, m_{\mathrm{m}}) = (16, 16)$. For reaching the methodical error, roughly two thousand initial particles are needed. We note, that within the time interval considered, every initial trajectory opens up at most one trajectory on the other level, such that the number of particles roughly doubles. On a 3 GHz Pentium 4 computer, our *Matlab 7.0* code requires approximately ten seconds for the propagation of thousand particles, which results in less than one minute computing time for the case $\mathrm{tol} = 10^{-3}$ and $(m_{\mathrm{p}}, m_{\mathrm{m}}) = (16, 16)$. In figure 5.3, the difference of the level populations computed with grid size
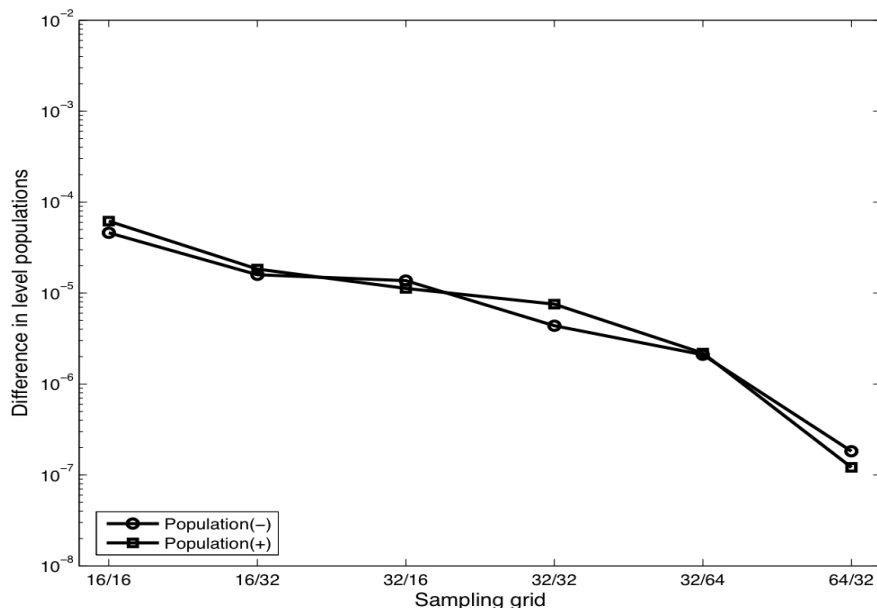
FIG. 5.3. *A semilogarithmic plot of the difference of energy level populations obtained by varying the initial grid size $(m_{\mathrm{p}}, m_{\mathrm{m}}) \in \{(16,16),(16,32),(32,16),(32,32),(32,64),(64,32)\}$ and subtracting from the result for $(m_{\mathrm{p}}, m_{\mathrm{m}}) = (64,64)$. The sampling tolerance is fixed at $\mathrm{tol} = 10^{-3}$. The plot shows differences smaller than $10^{-4}$, which are negligible compared to the methodical error of around 0.02.*

$(m_{\mathrm{p}}, m_{\mathrm{m}}) = (64,64)$ to those obtained for

$$(m_{\mathrm{p}}, m_{\mathrm{m}}) \in \{(16,16),(16,32),(32,16),(32,32),(32,64),(64,32)\}$$

is shown as a semilogarithmic plot. The sampling tolerance is fixed at $\mathrm{tol} = 10^{-3}$. The differences are bounded by $10^{-4}$, which is more than two orders of magnitude smaller than the methodical error of two percent. Table 5.3 presents the associated numbers

TABLE 5.3. *The initial number of particles when sampling with fixed tolerance $\mathrm{tol} = 10^{-3}$ for varying grid size $(m_{\mathrm{p}}, m_{\mathrm{m}})$. Finer sampling grids and hence more particles do not improve accuracy, if the methodical borderline is already reached, see figure 5.2.*

| $(m_{\mathrm{p}}, m_{\mathrm{m}})$ | $(16,16)$ | $(16,32)$ | $(32,16)$ | $(32,32)$ | $(32,64)$ | $(64,32)$ | $(64,64)$ |
|---|---|---|---|---|---|---|---|
| $\sharp$ particles | 2188 | 9080 | 9080 | 38960 | 160438 | 160439 | 662761 |

of initial particles, which reach from roughly two thousand for the $(m_{\mathrm{p}}, m_{\mathrm{m}}) = (16,16)$ grid to more than half a million for the case $(m_{\mathrm{p}}, m_{\mathrm{m}}) = (64,64)$. In summary, a sampling on finer grids and a corresponding increase in the particle number does not improve accuracy, if the methodical borderline of the algorithm is already reached. Hence, for the following experiments we choose initial grids of size $(m_{\mathrm{p}}, m_{\mathrm{m}}) = (16,16)$ and use $\mathrm{tol} = \varepsilon/10$ as sampling tolerance.

**5.4. Varying semiclassical parameter.**    The next series of experiments deviates from the basic parameter set by varying the semiclassical parameter as

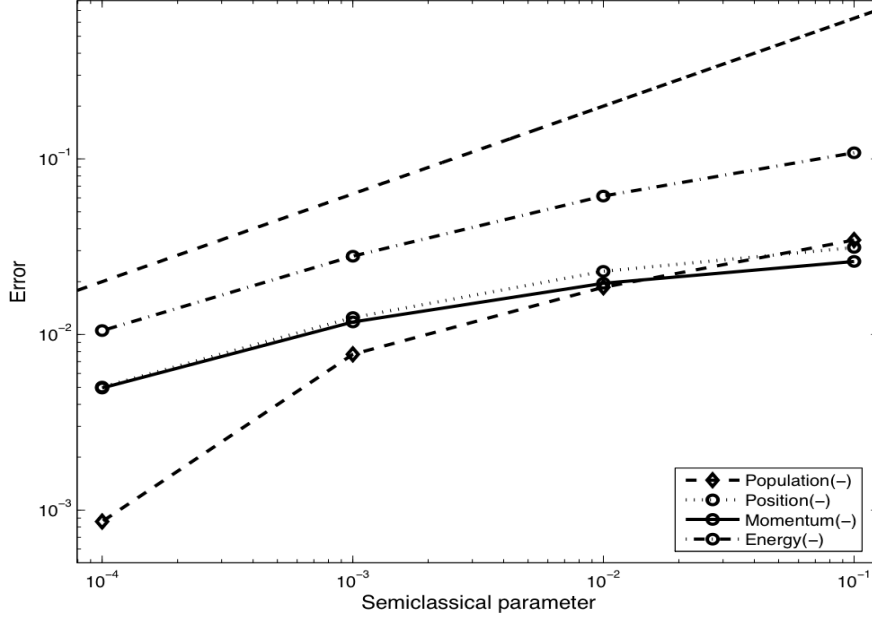$$\varepsilon = 10^{-k}, \qquad k = 4,3,2,1.$$

FIG. 5.4. *The error of level populations and expectation values for position, momentum, and energy, when varying the semiclassical parameter $\varepsilon = 10^{-k}$, $k = 4, 3, 2, 1$. There are two dashed lines: the lower one with the diamonds is the absolute error of level populations on the minus level, while the upper one without circles is the function $\varepsilon \mapsto 2\sqrt{\varepsilon}$. The three intermediate lines with circles correspond to relative errors of expectation values with respect to the final wave function projected on the minus-level. The dotted, solid, and dashdotted line refer to position, momentum, and energy expectation, respectively.*

The initial Gaussian $\psi_0^+$ is centered in $q_0^\varepsilon = (5\sqrt{\varepsilon}, \frac{1}{2}\sqrt{\varepsilon})$, $p_0 = (-1, 0)$, and the time interval considered is $[t_i, t_f] = [0, 10\sqrt{\varepsilon}]$. The sampling grid size and tolerance are chosen as $(m_p, m_m) = (16, 16)$ and $\text{tol} = \varepsilon/10$.

The data in table 5.1 show, that for increasing semiclassical parameter $\varepsilon$ the level population $P^-$ of the lower level increases, i. e. the non-adiabatic transfer grows. This tendency can be explained as follows: Writing the classical flow $\Phi_+^t(q, p) = (q(t), p(t))$ as

$$q(t) = q + pt - \int_0^t \int_0^s \frac{q(\tau)}{|q(\tau)|} \, d\tau \, ds, \qquad p(t) = p - \int_0^t \frac{q(\tau)}{|q(\tau)|} \, d\tau,$$

one observes, that on time-intervals of order $\sqrt{\varepsilon}$ the position of the flow $\Phi_+^t$ is $\varepsilon$-close to the position of the free flow $\Phi_0^t(q, p) = (q + pt, p)$. Hence, the larger $\varepsilon$, the more the $q_2$-component of the wave packet's center deviates from the straight motion $q_2(t) = \text{const}$ and moves towards the crossing at $(q_1, q_2) = (0, 0)$. As a consequence the non-adiabatic transfer of weight increases.

Here and in the following experiments, we consider the absolute error of level populations $|P_{\text{Schr}}^\pm - P_{\text{hop}}^\pm|$ and the relative errors of intralevel expectation values for position, momentum, and energy,

$$|E_{\text{p,Schr}}^\pm - E_{\text{p,hop}}^\pm|/|E_{\text{p,Schr}}^\pm|, \ |E_{\text{m,Schr}}^\pm - E_{\text{m,hop}}^\pm|/|E_{\text{m,Schr}}^\pm|, \ |E_{\text{E,Schr}}^\pm - E_{\text{E,hop}}^\pm|/|E_{\text{E,Schr}}^\pm|.$$

Figure 5.4 shows these errors with respect to the lower level as a function of the

semiclassical parameter $\varepsilon$. For a clearer presentation, the errors with respect to the upper level, which are slightly smaller than those for the lower level, are not given in the plot. One observes, that the error decreases with decreasing $\varepsilon$ and that it is bounded by the function $\varepsilon \mapsto 2\sqrt{\varepsilon}$.

**5.5. Varying initial momentum.** Here, the experiments have the same potential as before, but fix the semiclassical parameter as $\varepsilon = 10^{-2}$. The initial Gaussian $\psi_0^+$ is centered in

$$q_0^\varepsilon = (5\sqrt{\varepsilon}, \tfrac{1}{2}\sqrt{\varepsilon}), \qquad p_0 = (-k, 0), \qquad k = 0, \tfrac{1}{2}, 1, 2, 3.$$

The length of the time interval $[t_i, t_f] = [0, t_f]$ is adapted to the strength of the different initial momenta, see table C.1 in appendix C. The sampling grid and tolerance are $(m_p, m_m) = (16, 16)$ and $\text{tol} = 10^{-3}$.
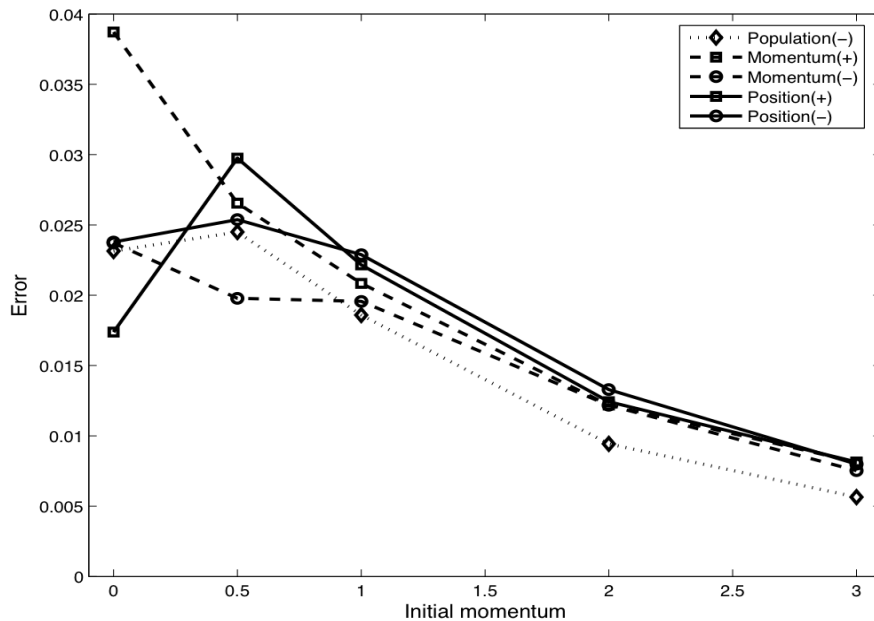


FIG. 5.5. *The plot shows the error of level populations and expectation values for position and momentum as the strength of the initial momentum varies within the set $k \in \{0, \tfrac{1}{2}, 1, 2, 3\}$. The dotted line with diamonds is the absolute error of populations. The solid lines with squares and circles are the relative error of position expectation for the plus and minus level, respectively. Similarly, the dashed lines refer to the relative errors of momentum expectation.*

In table 5.1, we observe increased level populations $P^-$ for smaller and larger values of initial momentum. The higher transitions for smaller momenta are related to the same effect as discussed above: for momenta of order $\sqrt{\varepsilon} \approx 0.3$, the position component of the flow stays $\varepsilon$-close to the position component of the free evolution, resulting in more non-adiabatic transfer. On the other hand, the higher transitions for larger momenta are explained by the growth of $T_*$ for increasing momentum.

The plot in figure 5.5 shows the absolute error of level populations as well as relative errors for position and momentum expectation values with respect to the lower and the upper level. Since the reference and the hopping algorithm both conserve the

$L^2$-norm, the difference of the absolute error of populations on the lower and the upper level is of the size of the sampling tolerance $\mathrm{tol} = 10^{-3}$ and hence not visible in the plot. One observes a systematic increase of errors, when momentum strength decreases from $k = 3$ to $k = 0$, while all errors stay below $0.05 \approx \frac{1}{2}\sqrt{\varepsilon}$. This tendency is expected, since for higher momenta, that is higher energies the semiclassical approximation is better. However, all errors have the same order of magnitude: from $k = 3$ to $k = 0$, they roughly differ by a factor five.
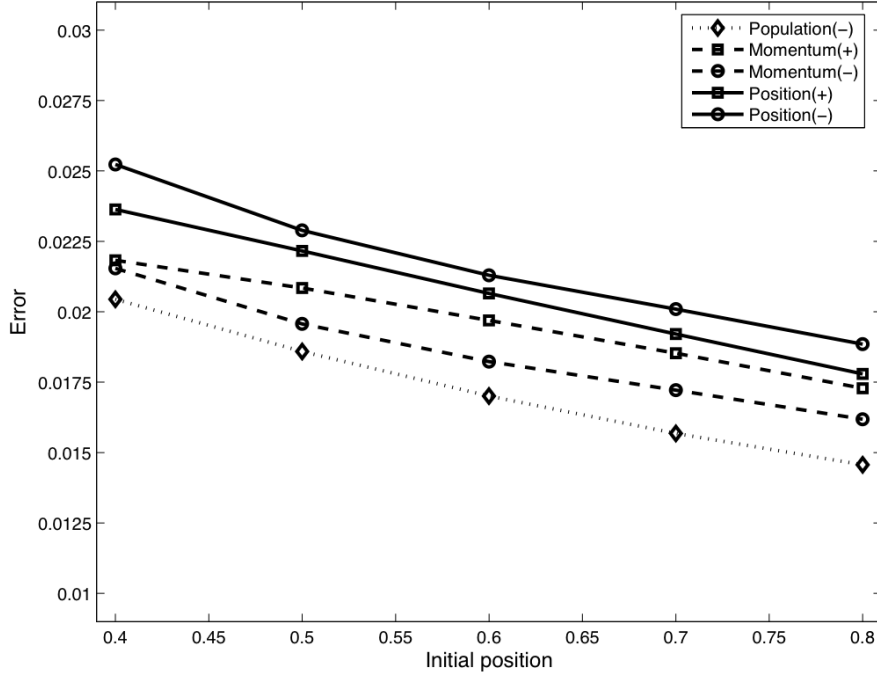


FIG. 5.6. *The error of level populations and relative errors for expectation values of position and momentum for varying parameter $\alpha = 4,5,6,7,8$, which gives the first component of the initial center in position space. The line style and symbol coding is the same as before in figure 5.5.*

**5.6. Varying position.**    The potential and the semiclassical parameter are the same as before, $\beta = 0$, $\gamma = 1$ and $\varepsilon = 10^{-2}$, while the initial Gaussian $\psi_0^+$ is centered in

$$q_0^\varepsilon = (\alpha\sqrt{\varepsilon}, \tfrac{1}{2}\sqrt{\varepsilon}), \qquad \alpha = 4,5,6,7,8, \qquad p_0 = (-1,0).$$

The time interval $[t_\mathrm{i}, t_\mathrm{f}] = [t_\mathrm{i}, 10\sqrt{\varepsilon}]$ is chosen such that the solution of

$$\tfrac{\mathrm{d}}{\mathrm{d}t}(x,\xi) = (\xi,-1), \qquad (x(t_\mathrm{i}),\xi(t_\mathrm{i})) = (\alpha\sqrt{\varepsilon},-1)$$

satisfies $x(0) = 5\sqrt{\varepsilon}$. That is, $t_\mathrm{i} = 1 - \sqrt{1 - 2(\alpha - 5)\sqrt{\varepsilon}}$. This auxiliary system is a rough approximation of the dynamics in $q_1$-direction. Depending on $\alpha$, the initial time $t_\mathrm{i}$ then varies between $-3\sqrt{\varepsilon}$ and $\sqrt{\varepsilon}$. The sampling grid and tolerance are $(m_\mathrm{p}, m_\mathrm{m}) = (16,16)$ and $\mathrm{tol} = 10^{-3} = \varepsilon/10$.

The farther away the initial wave function from the crossing $q = 0$, the more momentum is gained until the relevant non-adiabatic transitions are performed. The

transition rate $T_*$ increases with increasing momentum $p_*$. Hence, the increase of the final lower level population $P^-$ with increasing $\alpha$, see table 5.1, is qualitatively explained by $T_*$.

The initial data have negligible overlap with the crossing region for all values of $\alpha$. Still figure 5.6 shows that all errors increase with decreasing initial distance to the crossing. Although this behavior is heuristically expected, an appropriate explanation is somewhat subtle. For an improved approximation, the splitting of the initial data into upper and lower component should be done with reference to optimal super-adiabatic spaces, cf. [1, 18], not with respect to the eigenspaces of $V$. However, the difference is only noticeable close to the crossing point $q = 0$. We believe that our algorithm can be improved for initial data very close to the crossing region if we take into account the appropriate super-adiabatic splitting and will investigate this in a forthcoming publication.

**5.7. Varying confinement.**     In the chemical physics literature, Jahn-Teller Hamiltonians are a prominent class of theoretical model systems for non-adiabatic coupling in non-linear molecules, see for example chapter 10 in [3]. The most common one, the $E \otimes e$ Jahn-Teller Hamiltonian, describes the vibronic interaction of two electronic states via two nuclear coordinates, where both the electronic states and the nuclear coordinates are of $E$ symmetry. The simplest corresponding scenario concerns the displacement of a three atomic molecule from the equilateral triangle configuration.

In Cartesian coordinates $(q_1, q_2) \in \mathbb{R}^2$, the linear $E \otimes e$ Jahn-Teller Hamiltonian reads as

$$-\tfrac{\varepsilon^2}{2}\Delta_q + \beta|q|^2 + \gamma \begin{pmatrix} q_1 & q_2 \\ q_2 & -q_1 \end{pmatrix},$$

which is exactly a Schrödinger operator with isotropic potential of the form quadratic confinement plus linear trace-free matrix. In [7], electronic structure calculations for the triangular molecules silver, copper, sodium, and potassium ($Ag_3, Cu_3, Na_3$, and $K_3$) are used to estimate the strength of the confinement $\beta$ and the cone width $\gamma$, see also table 5.4. The corresponding semiclassical parameter $\varepsilon$ is then deduced from the relation

$$\varepsilon = \sqrt{\frac{\hbar^2}{m}} = \sqrt{\frac{\hbar^2 \omega^2}{2\beta}},$$

where $m$ is the nucleonic mass and $\omega$ the associated oscillator frequency.

TABLE 5.4. *Model parameters for a linear $E \otimes e$ Jahn-Teller Hamiltonian as suggested by the results of the electronic structure calculations for the triangular molecules silver, copper, sodium, and potassium as summarized in table III and IV of [7]. Energy is given in eV.*

| System | $\varepsilon$ | $\beta$ | $\gamma$ | $\hbar\omega$ |
|--------|------|------|------|------|
| $Ag_3$ | 0.005 | 2.613 | 0.871 | 0.012 |
| $Cu_3$ | 0.007 | 5.097 | 1.058 | 0.021 |
| $Na_3$ | 0.011 | 0.524 | 0.441 | 0.011 |
| $K_3$  | 0.008 | 0.260 | 0.276 | 0.006 |

The values in table 5.4 motivate a series of numerical experiments with fixed semi-classical parameter $\varepsilon = 10^{-2}$, initial Gaussian $\psi_0^+$ centered in position $q_0^\varepsilon = (5\sqrt{\varepsilon}, \tfrac{1}{2}\sqrt{\varepsilon})$

and momentum $p_0 = (-1,0)$, and cone width $\gamma = 1$. The quadratic confinement $\beta$ varies as

$$\beta = 0,1,2,3,4,5,$$

while the time interval is $[t_i, t_f] = [0, 6.4\sqrt{\varepsilon}]$. The initial sampling grid and tolerance are $(m_p, m_m) = (16,16)$ and $\mathrm{tol} = 10^{-3}$. The obtained final populations on the lower level given in table 5.1 increase with increasing $\beta$, which concurs with the fact that confinement directs the wave function towards the crossing as well as increases momentum favouring non-adiabatic transitions. Figure 5.7 shows the errors for level populations and expectation, which lie between 0.006 and 0.026 and decrease for increasing quadratic confinement.
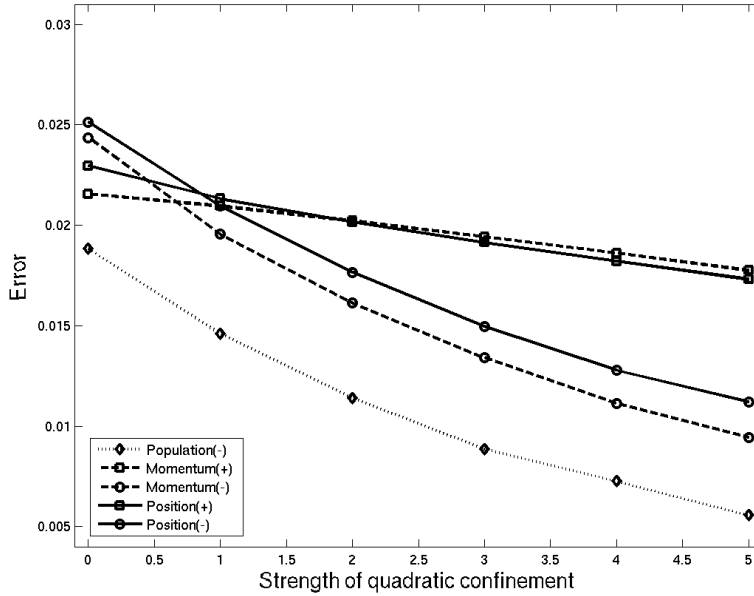


FIG. 5.7. *A plot of the absolute errors of energy level populations and the relative errors of position and momentum expectation values with respect to the strength of the quadratic confinement $\beta \in \{0,1,2,3,4,5\}$. All errors are bounded by 0.026 and decay with increasing strength of the confinement.*

**5.8. Varying cone width.** Again, the semiclassical parameter is $\varepsilon = 10^{-2}$, while $\beta = 0$ and the Gaussian $\psi_0^+$ is centered in $q_0^\varepsilon = (5\sqrt{\varepsilon}, \frac{1}{2}\sqrt{\varepsilon})$ and $p_0 = (-1,0)$. The cone width varies as

$$\gamma = \tfrac{1}{3}, \tfrac{1}{2}, 1, 2, 3.$$

The time interval is chosen as $[t_i, t_f] = [0, 10\sqrt{\varepsilon/\gamma}]$. The initial sampling grid and tolerance are $(m_p, m_m) = (16,16)$ and $\mathrm{tol} = 10^{-3}$. The results are shown in figure 5.8.

The errors for level population and for position and momentum on the lower level decrease for decreasing $\gamma$. This is related to the artificial energy shift introduced by

the non-adiabatic transitions. As the energetic gap becomes smaller for smaller $\gamma$, the position and momenta after the transition on the lower level are better approximated in this regime. We also mention, that the non-adiabatic transition rate $T_*$ increases when $\gamma$ tends to zero, which is reflected in the increasing level populations of the lower level given in table 5.1.
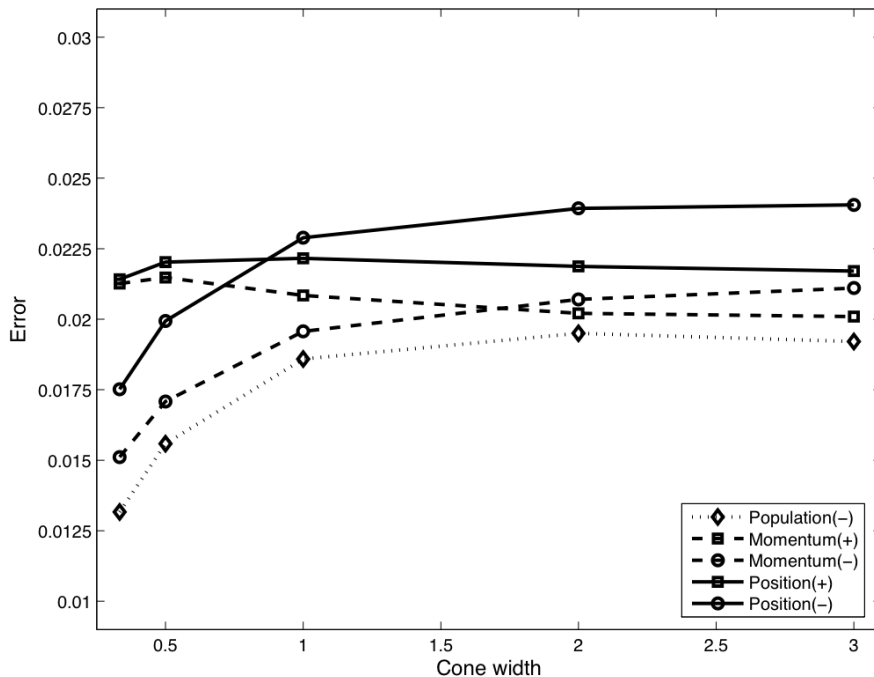


FIG. 5.8. *The error of level populations and relative errors for expectation values of position and momentum for varying cone width* $\gamma = \frac{1}{3}, \frac{1}{2}, 1, 2, 3$. *All errors lie between* $0.012$ *and* $0.025$. *They decrease for decreasing cone width.*

**Appendix A. Wigner functions & Weyl quantization.** Here, we collect some basic properties of Wigner functions and Weyl quantization in semiclassical scaling, which have been tacitly used before. For a detailed exposition of Wigner functions and pseudodifferential operators, one may consult [6].

DEFINITION A.1 (Wigner function). *Let* $\varepsilon > 0$. *The $\varepsilon$-scaled Wigner function* $W^\varepsilon(\psi)$ *of a square integrable function* $\psi \in L^2(\mathbb{R}^d, \mathbb{C}^N)$ *is defined as*

$$W^\varepsilon(\psi)(q,p) = (2\pi)^{-d} \int_{\mathbb{R}^d} \mathrm{e}^{\mathrm{i}x \cdot p} \psi(q - \tfrac{\varepsilon}{2}x) \otimes \overline{\psi}(q + \tfrac{\varepsilon}{2}x)\, \mathrm{d}x, \qquad (q,p) \in \mathbb{R}^{2d}.$$

The Wigner function $W^\varepsilon(\psi)$ of a wave function $\psi \in L^2(\mathbb{R}^d, \mathbb{C}^N)$ is the inverse Fourier transform of the $\varepsilon$-shifted product

$$\mathbb{R}^d \to L^2(\mathbb{R}^d, \mathbb{C}^{N \times N}), \qquad x \mapsto \psi(q - \tfrac{\varepsilon}{2}x) \otimes \overline{\psi}(q + \tfrac{\varepsilon}{2}x)$$

evaluated at $p \in \mathbb{R}^d$. Hence, it is a square integrable function on phase space. It takes values in the space of Hermitian $N$-by-$N$ matrices and consequently has real-valued

diagonal entries with respect to any $\mathbb{C}^N$-basis' decomposition. The marginals are position and momentum density,

$$\int_{\mathbb{R}^d} \mathrm{tr}\,(W^\varepsilon(\psi)(q,p))\,\mathrm{d}p = |\psi(q)|^2, \qquad \int_{\mathbb{R}^d} \mathrm{tr}\,(W^\varepsilon(\psi)(q,p))\,\mathrm{d}q = |(\mathcal{F}^\varepsilon\psi)(p)|^2,$$

where

$$(\mathcal{F}^\varepsilon\psi)(p) = (2\pi\varepsilon)^{-d/2} \int_{\mathbb{R}^d} \mathrm{e}^{-\mathrm{i}q\cdot p/\varepsilon}\psi(q)\,\mathrm{d}q, \qquad p \in \mathbb{R}^d,$$

is the $\varepsilon$-scaled Fourier transform of $\psi$. The support of $W^\varepsilon(\psi)$ is contained in the closed convex hull of

$$\mathrm{supp}(\psi) \times \mathrm{supp}(\mathcal{F}^\varepsilon\psi).$$

DEFINITION A.2 (Weyl quantized operator). *Let $\varepsilon > 0$. The Weyl quantized operator $a(q, -\mathrm{i}\varepsilon\nabla_q)$ of a square integrable function on phase space $a \in L^2(\mathbb{R}^{2d}, \mathbb{C}^{N\times N})$ is defined by its action on square integrable functions $\psi \in L^2(\mathbb{R}^d, \mathbb{C}^N)$ as*

$$a(q, -\mathrm{i}\varepsilon\nabla_q)\psi(q) = (2\pi)^{-d} \int_{\mathbb{R}^{2d}} a\left(\tfrac{1}{2}(q+y), \varepsilon p\right) \mathrm{e}^{\mathrm{i}(q-y)\cdot p}\psi(y)\,\mathrm{d}y\,\mathrm{d}p. \qquad (1.1)$$

For symbols $a(q,p)$, which are not square integrable, the above definition applies to wave functions $\psi$ in a suitably restricted subspace of $L^2(\mathbb{R}^d, \mathbb{C}^N)$. If $a(q,p) = \frac{1}{2}|p|^2$, for example, then equation (1.1) defines the $\varepsilon$-scaled Laplace operator,

$$a(q, -\mathrm{i}\varepsilon\nabla_q) = -\tfrac{\varepsilon^2}{2}\Delta_q : H^2(\mathbb{R}^d, \mathbb{C}^N) \to L^2(\mathbb{R}^d, \mathbb{C}^N).$$

Similarly, the Weyl quantized operator of $a(q,p) = q_j$ is the position operator

$$a(q, -\mathrm{i}\varepsilon\nabla_q) : \psi(q) \mapsto q_j\psi(q),$$

and the symbols $a(q,p) = \frac{1}{2}|p|^2 + \lambda^\pm(q)$ give the level operators

$$a(q, -\mathrm{i}\varepsilon\nabla_q) : \psi(q) \mapsto -\tfrac{\varepsilon^2}{2}\Delta_q\psi(q) + \lambda^\pm(q)\psi(q).$$

Expectation values of Weyl quantized observables can be computed by phase space integration of Wigner functions, that is

$$\langle\psi, a(q, -i\varepsilon\nabla_q)\psi\rangle_{L^2} = \int_{\mathbb{R}^{2d}} \mathrm{tr}\,(W^\varepsilon(\psi)(q,p)a(q,p))\,\mathrm{d}q\,\mathrm{d}p.$$

This relation together with the approximation

$$\chi^\pm(q)a(q, -\mathrm{i}\varepsilon\nabla_q)\chi^\pm(q)^t = (\Pi^\pm a\Pi^\pm)(q, -\mathrm{i}\varepsilon\nabla_q) + \mathcal{O}(\varepsilon), \qquad \varepsilon \to 0,$$

gives

$$\int_{\mathbb{R}^{2d}} \mathrm{tr}\,(\Pi^\pm(q)W^\varepsilon(\psi)(q,p)\Pi^\pm(q)\,a(q,p))\,\mathrm{d}q\,\mathrm{d}p =$$

$$\int_{\mathbb{R}^{2d}} W^\varepsilon(\psi^\pm)(q,p)\,a(q,p)\,\mathrm{d}q\,\mathrm{d}p + \mathcal{O}(\varepsilon), \qquad \varepsilon \to 0.$$

for scalar observables $a \in C_0^\infty(\mathbb{R}^{2d}, \mathbb{C})$ with support away from the crossing set, $\text{supp}(a) \cap (\Gamma \times \mathbb{R}^d) = \emptyset$. In this weak* sense, the diagonal components $w^\pm$ of the Wigner matrix $W^\varepsilon(\psi)$ coincide approximately with the Wigner function of the projected wave function $W^\varepsilon(\psi^\pm)$,

$$W^\varepsilon(\psi^\pm) \stackrel{*}{=} w^\pm + \mathcal{O}(\varepsilon), \qquad \varepsilon \to 0.$$

**Appendix B. Discretizations.** Our reference solver for the time-dependent Schrödinger equation (5.1) is based on a Strang splitting scheme: the unitary group

$$U^\varepsilon(t) = \exp\left(-\tfrac{i}{\varepsilon} t \left(-\tfrac{\varepsilon^2}{2} \Delta_q + V(q)\right)\right)$$

is approximated by $M \in \mathbb{N}$ powers of the Strang split operator

$$S^\varepsilon(\Delta_t) = \exp\left(-\tfrac{i}{\varepsilon} \tfrac{\Delta_t}{2} \left(-\tfrac{\varepsilon^2}{2} \Delta_q\right)\right) \exp\left(-\tfrac{i}{\varepsilon} \Delta_t V(q)\right) \exp\left(-\tfrac{i}{\varepsilon} \tfrac{\Delta_t}{2} \left(-\tfrac{\varepsilon^2}{2} \Delta_q\right)\right)$$

with timestep $\Delta_t = t/M$. The local error in $L^2$-norm is

$$\|U^\varepsilon(\Delta_t)\psi - S^\varepsilon(\Delta_t)\psi\|_{L^2} \leq C \left(\Delta_t/\varepsilon\right)^3,$$

where the constant $C = C(\psi, V) > 0$ depends on regularity and decay properties of the wave function $\psi$ as well as the potential $V(q)$, see [10].

The Laplacian on $\mathbb{R}^2$ is realized by Fourier differencing: the wave function is regarded as a $2K$-periodic function in $q_1$- and $q_2$-direction, $K > 0$, and is approximated by its trigonometric interpolant

$$\psi_N(q) = \sum_{\|p\|_\infty \leq N} \widehat{\psi}_p \exp\left(i \tfrac{\pi}{K} p \cdot q\right), \qquad N \in \mathbb{N}.$$

(For notational simplicity, we have suppressed the fractional weights required at the end points of the sum.) For the computation of the interpolant's Laplacian one just has to multiply the interpolation coefficients by $|p|^2$. The $L^2$-error reads as

$$\left\|\varepsilon^2 \Delta_q \psi - \varepsilon^2 \Delta_q \psi_N\right\|_{L^2([-K,K]^2)} \leq C (\varepsilon N)^{2-s} K^{s-1} \|\psi\|_{\varepsilon,s},$$

with $\|\psi\|_{\varepsilon,s}^2 = \sum_{p \in \mathbb{Z}^2} (1 + \varepsilon|p|)^{2s} |\widehat{\psi}(p)|^2$. The constant $C = C(s) > 0$ depends on the Sobolev regularity $s \geq 2$ of the wave function $\psi$, see [17]. In *Matlab 7.0*, this approximation of $-\varepsilon^2 \Delta_q \psi(q)$ can be efficiently realized by calling the two-dimensional fast Fourier transform `fft2`, multiplication with the weight vector

$$-\varepsilon^2 \left(\tfrac{\pi}{K}\right)^2 \left(0, 1, 4, \ldots, (N-1)^2, -N^2, -(N-1)^2, \ldots, -4, -1\right),$$

and calling the inverse fast Fourier transform `ifft2`.

The classical flows $\Phi_\pm^t$ are discretized by the embedded Runge-Kutta method `DOPRI45` of Dormand and Prince [2], which is realized by the *Matlab 7.0* built-in function `ode45`. Writing the Hamiltonian systems (1.3) as

$$\dot{x} = f(x),$$

one computes two approximate flows

$$\widehat{\Phi}^{t+\tau} x = x + \tau \widehat{b} \cdot k, \qquad \widetilde{\Phi}^{t+\tau} x = x + \tau \widetilde{b} \cdot k$$

with

$$k = (k_i) \in \mathbb{R}^7, \qquad k_i = f\Big(x + \tau \sum_{j=1}^{i-1} a_{ij} k_j\Big), \qquad i = 1,\dots,7.$$

The matrix $A = (a_{ij}) \in \mathbb{R}^{7 \times 7}$ and the vectors $\widehat{b}, \widetilde{b} \in \mathbb{R}^7$ are chosen such that $\widehat{\Phi}$ and $\widetilde{\Phi}$ give effectively six-stage methods of order five and four, respectively. The difference

$$\widehat{\Phi}^{t+\tau} x - \widetilde{\Phi}^{t+\tau} x$$

between the two discrete evolutions is used as a local error estimator for the adaptive step-size control. The event function

$$\kappa(t) = v(q(t)) \cdot Dv(q(t)) p(t)$$

is evaluated after each time step $t_n \in [t_{\mathrm{i}}, t_{\mathrm{f}}]$, $n = 1,\dots,N$. If a sign change is detected, the time $t_*$ at which the trajectory hits the submanifold $S$ is estimated by the Regula-Falsi formula:

$$t_* \approx t_n - \kappa(t_n) \frac{t_{n+1} - t_n}{\kappa(t_{n+1}) - \kappa(t_n)} =: \tilde{t}_*.$$

Next, $\kappa(\tilde{t}_*)$ is evaluated, and the procedure is iterated on one of the subintervals $[t_n, \tilde{t}_*]$ and $[\tilde{t}_*, t_{n+1}]$ until a given tolerance is reached. The value $\kappa(\tilde{t}_*)$ is computed by interpolation, and hence without further evaluations of the right hand side $f(x)$. To interpolate the solution for some time $t_n + \theta\tau$, $0 < \theta < 1$, one determines a vector $\bar{b}(\theta)$, whose entries depend polynomially on $\theta$, such that the formula

$$\overline{\Phi}^{t+\theta\tau} x = x + \theta\tau\, \bar{b}(\theta) \cdot k$$

is of comparable accuracy as the Runge-Kutta method. Note that the stage vector $k$, which represents the main computational cost, is reused. For the precise definition of the vector $\bar{b}(\theta)$, we refer to Chapter II.6 in [9].

**Appendix C. Accuracy of Schrödinger solutions.** The initial data for the Schrödinger systems (5.1) are centered at points $(p_0, q_0^\varepsilon)$ in phase space, where the initial position center's second component $q_0^\varepsilon = (q_{0,1}^\varepsilon, \frac{1}{2}\sqrt{\varepsilon})$ has a fixed relation to the semiclassical parameter and the one of the momentum center $p_0 = (p_{0,1}, 0)$ is vanishing. In this case, one expects a solution with higher momentum in $q_1$-direction, and to prevent aliasing, the number of grid points in $q_1$-direction is twice the number in $q_2$-direction. Most of the calculations are performed on uniform $1024 \times 512$ grids, except for the case $\varepsilon = 10^{-4}$, where a $2048 \times 1024$ grid is used. The number of time steps is fixed as $10^5$. The other parameters, that is the semiclassical parameter $\varepsilon$, the cone width $\gamma$, and the time-interval $[t_{\mathrm{i}}, t_{\mathrm{f}}]$ are given in table C.1. The domain of computation is chosen as

$$[2x(t_{\mathrm{f}}), 2x(t_{\mathrm{i}})] \times ([-1,1] \cdot |x(t_{\mathrm{f}}) - x(t_{\mathrm{i}})|)$$

where $x(t)$ is the solution of the auxiliary system

$$\tfrac{\mathrm{d}}{\mathrm{d}t}(x, \xi) = (\xi, -\gamma), \qquad (x(t_{\mathrm{i}}), \xi(t_{\mathrm{i}})) = (q_{0,1}^\varepsilon, p_{0,1}).$$

TABLE C.1. *The parameters of the five different sets of numerical experiments and the accuracy of the Schrödinger solution measured in $L^2$-norm. For the choice of $t_i$ see section 5.6.*

| $\varepsilon$ | $q_{0,1}^\varepsilon$ | $p_{0,1}$ | $\beta$ | $\gamma$ | Time | Accuracy |
|---|---|---|---|---|---|---|
| $10^{-1}$ | $5\sqrt{\varepsilon}$ | $-1$ | $0$ | $1$ | $[0,7\sqrt{\varepsilon}]$ | $7.37\cdot10^{-6}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-1$ | $0$ | $1$ | $[0,10\sqrt{\varepsilon}]$ | $4.68\cdot10^{-7}$ |
| $10^{-3}$ | $5\sqrt{\varepsilon}$ | $-1$ | $0$ | $1$ | $[0,10\sqrt{\varepsilon}]$ | $3.91\cdot10^{-7}$ |
| $10^{-4}$ | $5\sqrt{\varepsilon}$ | $-1$ | $0$ | $1$ | $[0,10\sqrt{\varepsilon}]$ | $2.67\cdot10^{-7}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $0$ | $0$ | $1$ | $[0,20\sqrt{\varepsilon}]$ | $1.07\cdot10^{-4}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-0.5$ | $0$ | $1$ | $[0,20\sqrt{\varepsilon}]$ | $2.95\cdot10^{-5}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-1$ | $0$ | $1$ | $[0,10\sqrt{\varepsilon}]$ | $4.68\cdot10^{-7}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-2$ | $0$ | $1$ | $[0,5\sqrt{\varepsilon}]$ | $4.90\cdot10^{-7}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-3$ | $0$ | $1$ | $[0,10/3\sqrt{\varepsilon}]$ | $6.51\cdot10^{-7}$ |
| $10^{-2}$ | $4\sqrt{\varepsilon}$ | $-1$ | $0$ | $1$ | $[t_i,10\sqrt{\varepsilon}]$ | $3.62\cdot10^{-5}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-1$ | $0$ | $1$ | $[0,10\sqrt{\varepsilon}]$ | $4.68\cdot10^{-7}$ |
| $10^{-2}$ | $6\sqrt{\varepsilon}$ | $-1$ | $0$ | $1$ | $[t_i,10\sqrt{\varepsilon}]$ | $4.92\cdot10^{-9}$ |
| $10^{-2}$ | $7\sqrt{\varepsilon}$ | $-1$ | $0$ | $1$ | $[t_i,10\sqrt{\varepsilon}]$ | $5.33\cdot10^{-9}$ |
| $10^{-2}$ | $8\sqrt{\varepsilon}$ | $-1$ | $0$ | $1$ | $[t_i,10\sqrt{\varepsilon}]$ | $6.30\cdot10^{-9}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-1$ | $0$ | $1/3$ | $[0,10\sqrt{3\varepsilon}]$ | $5.88\cdot10^{-7}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-1$ | $0$ | $0.5$ | $[0,10\sqrt{2\varepsilon}]$ | $5.40\cdot10^{-7}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-1$ | $0$ | $1$ | $[0,10\sqrt{\varepsilon}]$ | $4.68\cdot10^{-7}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-1$ | $0$ | $2$ | $[0,10\sqrt{\varepsilon/2}]$ | $4.20\cdot10^{-7}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-1$ | $0$ | $3$ | $[0,10\sqrt{\varepsilon/3}]$ | $4.86\cdot10^{-7}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-1$ | $0$ | $1$ | $[0,6.4\sqrt{\varepsilon}]$ | $3.57\cdot10^{-7}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-1$ | $1$ | $1$ | $[0,6.4\sqrt{\varepsilon}]$ | $3.40\cdot10^{-7}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-1$ | $2$ | $1$ | $[0,6.4\sqrt{\varepsilon}]$ | $3.20\cdot10^{-7}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-1$ | $3$ | $1$ | $[0,6.4\sqrt{\varepsilon}]$ | $3.26\cdot10^{-7}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-1$ | $4$ | $1$ | $[0,6.4\sqrt{\varepsilon}]$ | $3.15\cdot10^{-7}$ |
| $10^{-2}$ | $5\sqrt{\varepsilon}$ | $-1$ | $5$ | $1$ | $[0,6.4\sqrt{\varepsilon}]$ | $3.21\cdot10^{-7}$ |

The accuracy of solutions, given in the last column of table C.1, refers to the $L^2$-norm difference $\|\psi(t_f)-\widetilde{\psi}(t_f)\|_{L^2}$ of the final reference solution $\psi(t_f)$ and a coarser solution $\widetilde{\psi}(t_f)$, which is computed with a fourth of the number of grid points and half the number of time steps. The actually evaluated quantity $\|\psi(t_f)-\widetilde{\psi}(t_f)\|_{L^2}^2$ is the square of the difference in norm. Hence, the accuracy for most of the calculations is close to machine precision. The four reference solutions with an error larger than $10^{-6}$ are still accurate enough for our purpose, the validation of our surface hopping algorithm, whose experimentally observed errors vary around $10^{-2}$. Those four outliers are due to the less localized nature of the solution in the cases $\varepsilon=10^{-1}$, $p_{0,1}\in\{0,-0.5\}$, and an initial datum closer to the singularity of the crossing in the case $q_{0,1}=4\sqrt{\varepsilon}$.

TABLE C.2. *The accuracy of less resolved Strang splitting solutions for the standard test case. The reference solution corresponding to the second row of table C.1 is compared in $L^2$-norm to coarser solutions computed on a $256\times128$ grid with varying timesteps.*

| Timesteps | $10^4$ | $10^3$ | 100 | 10 |
|---|---|---|---|---|
| Accuracy | $9.54\cdot10^{-7}$ | $1.19\cdot10^{-5}$ | 0.001 | 0.166 |

In table C.2 the accuracy of less resolved Strang splitting solutions is summarized. The one with $256\times128$ grid and one hundred time steps takes less than half a minute computing time. Its accuracy is of the order reached by the surface hopping algorithm.

## REFERENCES

[1] V. Betz and S. Teufel, *Precise coupling terms in adiabatic quantum evolution: The generic case*, Commun. Math. Phys., 260(2), 481–509, 2005.

[2] J. R. Dormand and P. J. Prince, *A family of embedded Runge-Kutta formulae*, J. Comput. Appl. Math., 6, 19–26, 1980.

[3] W. Domcke, D. Yarkony, and H. Köppel, *Conical intersections*, World Scientific Publishing, 2004.

[4] C. Fermanian-Kammerer and P. Gérard, *Mesures semiclassiques et croisements de modes*, Bull. Soc. Math. Fr., 130(1), 123–168, 2002.

[5] C. Fermanian-Kammerer and C. Lasser, *Propagation through generic level crossings: a surface hopping semigroup*, in preparation.

[6] G. Folland, *Harmonic analysis in phase space*, Annals of Mathematics Studies, 122, Princeton University Press, 1989.

[7] P. Garcia-Fernández, I. Bersuker, A. Aramburu, M. Barriuso, M. Moreno, *Origin of warping in the $E \otimes e$ Jahn-Teller problem: Quadratic vibronic coupling versus anharmonicity and application to $NaCl\!:\!Rh^{2+}$ and triangular molecules*, Phys. Rev. B, 71, 184117-1–184117-10, 2005.

[8] G. Hagedorn, *Molecular propagation through electron energy level crossings*, Mem. Am. Math. Soc., 536, 1994.

[9] E. Hairer, S. P. Nørsett, and G. Wanner, *Solving ordinary differential equations I: Nonstiff problems*, 2. rev. ed., Springer Series in Computational Mathematics, 8, Springer-Verlag, 1993.

[10] T. Jahnke and C. Lubich, *Error bounds for exponential operator splittings*, BIT, 40(4), 735–744, 2000.

[11] T. Kato, *Perturbation theory for linear operators*, Grundlehren der mathematischen Wissenschaften, 132, Springer-Verlag, 1980.

[12] H. Köppel, *Regularized diabatic states and quantum dynamics on intersecting potential energy surfaces*, Faraday Discuss., 127, 35–47, 2004.

[13] C. Lasser and S. Teufel, *Propagation through conical crossings: an asymptotic semigroup*, Commun. Pure Appl. Math., 58(9), 1188–1230, 2005.

[14] R. Schneider, W. Domcke, and H. Köppel, *Aspects of dissipative electronic and vibrational dynamics of strongly vibronically coupled systems*, J. Chem. Phys., 92(2), 1045–1061, 1990.

[15] H. Spohn and S. Teufel, *Adiabatic decoupling and time-dependent Born-Oppenheimer theory*, Commun. Math. Phys., 224(1), 113–132, 2001.

[16] G. Stock and M. Thoss, *Classical description of nonadiabatic quantum dynamics*, Adv. Chem. Phys., 131, 243–375, 2005.

[17] E. TADMOR, *The exponential accuracy of Fourier and Chebyshev differencing methods*, SIAM J. Numer. Anal., 23, 1–10, 1986.

[18] S. Teufel, *Adiabatic perturbation theory in quantum dynamics*, Lecture Notes in Mathematics 1821, Springer-Verlag, 2003.

[19] J. Tully and R. Preston, *Trajectory surface hopping approach to nonadiabatic molecular collisions: the reaction of $H^+$ with $D_2$*, J. Chem. Phys., 55(2), 562–572, 1971.