

Monte Carlo sampling of Wigner functions and surface hopping quantum dynamics

Susanna Kube^{a,*}, Caroline Lasser^b, Marcus Weber^a

^a*Zuse Institute Berlin, Takustraße 7, D-14195 Berlin, Germany*

^b*Freie Universität Berlin, Arnimallee 6, D-14195 Berlin, Germany*

Abstract

The article addresses the achievable accuracy for a Monte Carlo sampling of Wigner functions in combination with a surface hopping algorithm for non-adiabatic quantum dynamics. The approximation of Wigner functions is realized by an adaption of the Metropolis algorithm for real-valued functions with disconnected support. The integration, which is necessary for computing values of the Wigner function, uses importance sampling with a Gaussian weight function. The numerical experiments agree with theoretical considerations and show an error of two to three percent.

Key words: Metropolis Monte Carlo, quantum dynamics, surface hopping

2000 MSC: 65D30, 81S30, 81Q20

*Corresponding author.

Email addresses: kube@zib.de (Susanna Kube), lasser@math.fu-berlin.de (Caroline Lasser), weber@zib.de (Marcus Weber)

1. Introduction

The fundamental equation of non-relativistic quantum nuclear dynamics is the time-dependent Schrödinger equation

$$i\hbar\partial_\tau\psi(\tau, q) = -\frac{\hbar^2}{2m}\Delta_q\psi(\tau, q) + V(q)\psi(\tau, q), \quad \psi(0, q) = \psi_0(q).$$

It results from the time-dependent Born-Oppenheimer approximation and describes nuclei of mass m moving in a potential $V : \mathbb{R}^d \rightarrow \mathbb{R}$, which is obtained from averaging out electronic degrees of freedom. Mostly, the dimension d of the configuration spaces \mathbb{R}^d is very large, since one might consider up to thirty nuclear degrees of freedom. On top of that, the solution $\psi(\tau, q)$ exhibits oscillations both in time and in space. For quantifying the oscillatory behavior, one switches to atomic units and introduces the crucial semiclassical parameter

$$\varepsilon \sim \sqrt{1/m}.$$

On the long time scale $t = \tau/\varepsilon$, on which non-trivial nuclear dynamics develop, the Schrödinger equation rewrites as

$$i\varepsilon\partial_t\psi(t, q) = -\frac{\varepsilon^2}{2}\Delta_q\psi(t, q) + V(q)\psi(t, q), \quad \psi(0, q) = \psi_0(q). \quad (1)$$

Then, all oscillations are roughly characterized by the frequency $1/\varepsilon$, which typically ranges between hundred and thousand.

The conventional interpretation of quantum mechanics does not assign any physical meaning to the wave function $\psi(t, q)$ itself, but to quadratic quantities of it. The probability for finding the quantum system at time t within the set $\Omega \subset \mathbb{R}^d$ is

$$\int_{\Omega} |\psi(t, q)|^2 dq.$$

The expectation value for the position and the momentum of the system at time t are for example

$$\langle \psi(t), q\psi(t) \rangle_{L^2}, \quad \langle \psi(t), -i\varepsilon \nabla_q \psi(t) \rangle_{L^2}.$$

More generally, with smooth functions $a : \mathbb{R}^{2d} \rightarrow \mathbb{C}$ on classical phase space \mathbb{R}^{2d} , one associates a Weyl quantized operator $\text{op}(a)$ by setting

$$\text{op}(a)\psi(q) = (2\pi\varepsilon)^{-d} \int_{\mathbb{R}^{2d}} a\left(\frac{1}{2}(q+x), p\right) e^{ip\cdot(q-x)/\varepsilon} \psi(x) \, dx \, dp.$$

The corresponding expectation values

$$\langle \psi(t), \text{op}(a)\psi(t) \rangle_{L^2}$$

then specialize to the components of the position and momentum expectation by choosing $a(q, p) = q_j$ and $a(q, p) = p_j$, respectively. Also the Schrödinger operator $-\frac{\varepsilon^2}{2}\Delta_q + V(q)$ can be seen as the Weyl operator $\text{op}(a)$ of the classical Hamiltonian function $a(q, p) = \frac{1}{2}|p|^2 + V(q)$.

In the semiclassical regime, the direct approach to quadratic quantities is advantageous, since their dynamics are less oscillatory than those of the wave function itself. The Egorov theorem provides the following approximation (e.g. Theorem IV.10 in (1)). Let

$$\dot{q} = p, \quad \dot{p} = -\nabla_q V(q) \tag{2}$$

be the Hamiltonian system associated to the Schrödinger operator, and let $\Phi^t : \mathbb{R}^{2d} \rightarrow \mathbb{R}^{2d}$ denote its flow. Then,

$$\langle \psi(t), \text{op}(a)\psi(t) \rangle_{L^2} = \langle \psi_0, \text{op}(a \circ \Phi^{-t})\psi_0 \rangle_{L^2} + O(\varepsilon^2),$$

where the constant of the error term depends on time t and bounds on derivatives of the function a and the potential V , which are greater or equal than order three. On the level of this general asymptotic approximation, oscillations in time do not show up any more, and space oscillations must only be resolved for the initial wave function ψ_0 .

Moreover, all expectation values associated with a wave function ψ can be expressed by its Wigner function $W(\psi) : \mathbb{R}^{2d} \rightarrow \mathbb{R}$, which is a function on phase space \mathbb{R}^{2d} . The definition

$$W(\psi)(q, p) = (2\pi)^{-d} \int_{\mathbb{R}^d} e^{ix \cdot p} \psi(q - \frac{\varepsilon}{2}x) \overline{\psi}(q + \frac{\varepsilon}{2}x) dx \quad (3)$$

grants

$$\langle \psi, \text{op}(a)\psi \rangle_{L^2} = \int_{\mathbb{R}^{2d}} W(\psi)(q, p) a(q, p) dq dp.$$

Hence, expectation values can be obtained by phase space integration. The Wigner function has first been proposed by Wigner in (2). Its main properties will be briefly discussed in §3. From the Wigner point of view, the Egorov theorem rephrases as

$$W(\psi(t)) = W(\psi_0) \circ \Phi^t + O(\varepsilon^2), \quad (4)$$

where the relation holds in a weak sense. One then deduces a simple particle method, which is built of the following steps.

Initial sampling. One samples the Wigner function $W(\psi_0)$ of the initial wave function ψ_0 to obtain phase space points $(q_1, p_1), \dots, (q_N, p_N)$.

Classical transport. The points are transported along the curves of the Hamiltonian system $\dot{q} = p, \dot{p} = -\nabla_q V(q)$ until the desired time t .

Final evaluation. The values of the initial Wigner function $W(\psi_0)$ in the points $(q_1, p_1), \dots, (q_N, p_N)$ approximate the values of $W(\psi(t))$ in the points $\Phi^t(q_1, p_1), \dots, \Phi^t(q_N, p_N)$. One computes expectation values.

It is our aim here to contribute to the initial sampling step, in particular to systematically evaluate the achievable accuracy of a Monte Carlo approach. To our best knowledge, neither the mathematical nor the chemical literature addresses this important point so far. In particular, we pursue the following two main objectives. First, the generation of phase space points according to the Wigner function $W(\psi)$ of a typical wave function ψ . Second, the computation of the Fourier integral in (3), which defines the value of the Wigner function in a given phase space point. The first task is related to the problem of approximation, while the second one concerns numerical quadrature.

Our article is organized as follows. The next section §2 reviews an extension of the simple particle method to Schrödinger systems with matrix-valued potential, called surface hopping algorithms. In §3, we discuss basic properties of Wigner functions and link them with quadratic quantities of interest like position and momentum densities. §4 presents Monte Carlo methods for the approximation and the quadrature problem at hand, while §5 contains the detailed set up for the numerical experiments. §6, §7, and §8 validate the proposed methods for initial wave functions, which are a single Gaussian wave packet, a superposition of two Gaussian wave packets, and a function resulting from a numerically computed laser excitation, respectively. Then, we offer an assessment of the obtained results in the final section §9.

2. Surface hopping

For photoinduced molecular dynamics, scalar Schrödinger equations are not appropriate any more. Nuclear motion is governed by averages derived from at least two electronic eigenstates, and the basic equations are Schrödinger systems

$$i\varepsilon\partial_t\psi(t, q) = -\frac{\varepsilon^2}{2}\Delta_q\psi(t, q) + V(q)\psi(t, q), \quad \psi(0, q) = \psi_0(q)$$

with real symmetric potential matrix

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

The associated dynamics essentially differ from the scalar case when the eigenvalues of the potential matrix are not uniformly separated, as it happens for conical intersections (3). The potential matrix $V(q)$ has a conical intersection if its two eigenvalues $\lambda^+(q)$ and $\lambda^-(q)$ coincide on a smooth submanifold of codimension two. In this case, there is a suitable set of coordinates such that near the crossing set

$$\{q \in \mathbb{R}^d \mid \lambda^+(q) = \lambda^-(q)\}$$

the two eigenvalue surfaces $\mathbb{R}^d \rightarrow \mathbb{R}$, $q \mapsto \lambda^\pm(q)$ look like two cones touching each other in their end points. Conical intersections violate adiabatic decoupling in the following sense. If $\chi^\pm(q)$ denotes a normalized eigenvector of the matrix $V(q)$ and $\psi^\pm(t, q) = \langle \chi^\pm(q), \psi(t, q) \rangle_{\mathbb{C}^2}$ the solution's component in the corresponding eigenspace, then it may happen that

$$\psi^-(0) = 0 \quad \text{and} \quad \exists t : \psi^-(t) = O(1), \quad \varepsilon \rightarrow 0.$$

That is, the wave function performs a leading order non-adiabatic transition from one eigenspace to the other, from the plus space to the minus space or vice versa. For systems with conical intersections, the particle method has to be supplemented by a surface hopping step.

Initial sampling. One samples the Wigner functions $W(\psi^\pm(0))$ to obtain two families of phase space points $(q_1^\pm, p_1^\pm), \dots, (q_{N^\pm}^\pm, p_{N^\pm}^\pm)$ with associated real-valued weights $w_1^\pm, \dots, w_{N^\pm}^\pm$, which are the values of the Wigner function $W(\psi^\pm(0))$ in these points.

Classical transport. The phase space points are transported along the curves of the corresponding Hamiltonian system $\dot{q} = p, \dot{p} = -\nabla_q \lambda^\pm(q)$.

Surface hopping. Whenever a trajectory $t \mapsto (q_t, p_t)$ passes one of its minimal surface gaps at a point (q, p) , that is whenever the function

$$t \mapsto (\lambda^+(q_t) - \lambda^-(q_t))$$

attains a local minimum, then a branching occurs. The transition branch carries the old weight times the Landau-Zener factor

$$T(q, p) = \exp\left(-\frac{\pi}{\varepsilon} \frac{|v(q)|^2}{|dv(q)p|}\right),$$

where $dv(q)$ denotes the $2 \times d$ gradient matrix of $v(q) = (v_1(q), v_2(q))$, and starts a new trajectory in (q, p) , which is associated with the other eigenvalue. The remaining branch continues the old trajectory and carries the old weight times $1 - T(q, p)$.

Final evaluation. At the desired time t , one obtains two families of phase space points $(q_1^\pm, p_1^\pm), \dots, (q_{M^\pm}^\pm, p_{M^\pm}^\pm)$ and weights $w_1^\pm, \dots, w_{M^\pm}^\pm$, which

approximate the values of the Wigner function $W(\psi^\pm(t))$ in these points. One computes the final expectation values.

This particle method is called single switch surface hopping, since its constitutive branching condition allows for non-adiabatic switches just at minimal surface gaps along trajectories. It has been derived from a rigorous mathematical analysis of Schrödinger systems with generic crossings. If $W^\pm(t)$ denotes the phase space functions at time t generated by the single switch method, then a generalization of the Egorov theorem guarantees at least

$$W(\psi^\pm(t)) = W^\pm(t) + O(\varepsilon^{1/8}),$$

where the relation holds in a weak sense (Theorem 2.2 in (4)). However, all the numerical experiments so far have even shown a convergence rate of order $\varepsilon^{1/2}$, see (4; 5; 6) and §7.2 later on.

Surface hopping is one of the most popular approaches for simulating non-adiabatic quantum dynamics. The first algorithm of this type has been proposed by Tully and Preston in (7). The current chemical literature contains an overwhelming variety of surface hopping methods, which all differ in the way non-adiabatic transitions are performed. Most algorithms resemble the well established fewest switches method (8) and have random hops at every time step of the discretization. A comparative numerical study for a benchmark model of the internal conversion in pyrazine has shown that the single switch method reproduces the long-time oscillations of the system more accurately than the fewest switches algorithm (9).

3. Wigner functions

Since expositions of the main properties of Wigner functions have been given many times, see for example chapter 1.8 in (10) or chapter 4.3 in (11), we will only focus on those that are relevant for its intended Monte Carlo sampling to start an asymptotic particle method. Moreover, we compare the properties of Wigner functions with a different phase space representation, the Husimi function. In contrast to Wigner functions, Husimi functions are non-negative, which accounts for their popularity in computational chemistry.

3.1. Basic properties

We use the ε -scaled Fourier transform of a function $\psi \in L^2(\mathbb{R}^d)$,

$$(\mathcal{F}\psi)(p) = (2\pi\varepsilon)^{-d/2} \int_{\mathbb{R}^d} e^{-iq \cdot p/\varepsilon} \psi(q) \, dq.$$

Then, the Wigner function

$$W(\psi)(q, p) = (2\pi\varepsilon)^{-d} \int_{\mathbb{R}^d} e^{ix \cdot p/\varepsilon} \psi(q - \frac{1}{2}x) \bar{\psi}(q + \frac{1}{2}x) \, dx$$

is the inverse Fourier transform of the product $x \mapsto \psi(q - \frac{1}{2}x) \bar{\psi}(q + \frac{1}{2}x)$.

Hence,

$$W(\psi) : \mathbb{R}^{2d} \rightarrow \mathbb{R}$$

is a square integrable function on phase space, and one obtains for any $q_0 \in \mathbb{R}^d$ with $\psi(q_0) \neq 0$ the inversion formula

$$\psi(q) = \bar{\psi}(q_0)^{-1} \int_{\mathbb{R}^d} e^{i(q-q_0) \cdot p/\varepsilon} W(\psi)(\frac{1}{2}(q+q_0), p) \, dp.$$

Let $a : \mathbb{R}^{2d} \rightarrow \mathbb{C}$ be a Schwartz function on phase space and $\text{op}(a)$ the associated Weyl quantized pseudodifferential operator. Then, a short calculation yields

$$\langle \psi, \text{op}(a)\psi \rangle_{L^2} = \int_{\mathbb{R}^{2d}} W(\psi)(q, p) a(q, p) dq dp.$$

In addition to the relation with expectation values, the marginals are the position and momentum density,

$$\int_{\mathbb{R}^d} W(\psi)(q, p) dp = |\psi(q)|^2, \quad \int_{\mathbb{R}^d} W(\psi)(q, p) dq = |(\mathcal{F}\psi)(p)|^2,$$

if $\psi, \mathcal{F}\psi \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$. Consequently,

$$\int_{\mathbb{R}^{2d}} W(\psi)(q, p) dq dp = \|\psi\|_{L^2}^2.$$

We note that the condition $\psi, \mathcal{F}\psi \in L^1(\mathbb{R}^d) \cap L^2(\mathbb{R}^d)$ ensures absolute convergence of the preceding integrals. The wave functions considered in subsequent sections satisfy this integrability constraint. The balance between position and momentum is also observed in the identity

$$W(\psi)(q, p) = W(\mathcal{F}\psi)(p, -q).$$

Moreover, $W(\psi)(q, p) \neq 0$ implies that (q, p) lies in the convex hull of the supports

$$\text{supp}(\psi) \times \text{supp}(\mathcal{F}\psi).$$

The interpretation of the Wigner function as a phase space density has the defect that it might attain negative values. Indeed,

$$W(\psi)(0, 0) = -(\varepsilon\pi)^{-d} \|\psi\|_{L^2}^2$$

for odd functions $\psi(q) = -\psi(-q)$. However, in an averaged sense the negativity is rather mild. The sharp Gårding inequality (e.g. Chapter 2.10

in (12)) provides for non-negative Schwartz functions $a \geq 0$ a positive constant $C = C(a) > 0$ depending on derivative bounds of a such that

$$\int_{\mathbb{R}^{2d}} W(\psi)(q, p) a(q, p) dq dp \geq -C\varepsilon \|\psi\|_{L^2}^2.$$

3.2. Two examples

The simplest examples are Gaussian wave packets centered in a single phase space point $z_0 = (q_0, p_0)$, which are of the form

$$g_{z_0}(q) = (\pi\varepsilon)^{-d/4} \exp\left(-\frac{1}{2\varepsilon}|q - q_0|^2 + \frac{i}{\varepsilon}p_0 \cdot (q - q_0)\right).$$

Their Wigner function is a Gaussian

$$W(g_{z_0})(q, p) = (\pi\varepsilon)^{-d} \exp\left(-\frac{1}{\varepsilon}|q - q_0|^2 - \frac{1}{\varepsilon}|p - p_0|^2\right),$$

whose approximation is unproblematic, of course. However, if the Wigner function was not computable analytically, one would numerically solve a Gaussian Fourier integral of the type

$$F(p - p_0) = \int_{\mathbb{R}^d} e^{ix \cdot (p - p_0)} e^{-\frac{\varepsilon}{4}|x|^2} dx. \quad (5)$$

The relative condition number

$$\kappa_F(p - p_0) = \frac{|p - p_0|}{|F(p - p_0)|} |\nabla F(p - p_0)| = \frac{2}{\varepsilon} |p - p_0|^2$$

reflects the oscillatory behavior of the integrand for momenta with large distance from the center.

Also the numerical approximation of the Wigner function suffers from oscillations as soon as the wave function microlocalizes around several phase

space points. An illustrative example is the superposition of two Gaussian wave packets with centers in $z_1, z_2 \in \mathbb{R}^{2d}$, whose Wigner function

$$W(g_{z_1} + g_{z_2})(q, p) = W(g_{z_1})(q, p) + W(g_{z_2})(q, p) + 2c(q, p)$$

contains a cross term

$$c(q, p) = (\pi\varepsilon)^{-d} e^{-|(q,p)-z_+|^2/\varepsilon} \cos\left(\frac{1}{\varepsilon}(p_+ \cdot q_- - ((q, p) - z_+) \wedge z_-)\right),$$

where $x \wedge y = p(x) \cdot q(y) - q(x) \cdot p(y)$ for vectors $x, y \in \mathbb{R}^{2d}$ (for detailed computations, see Appendix B). The cross term localizes around the arithmetic mean $z_+ = (z_1 + z_2)/2$ and oscillates with a frequency proportional to $z_- = z_1 - z_2$. However, an elaborate integration by parts, see Theorem 7.7.1 in (13), gives a constant $C > 0$ such that for all smooth compactly supported functions $a : \mathbb{R}^{2d} \rightarrow \mathbb{C}$ and all $k \in \mathbb{N}_0$

$$\left| \int_{\mathbb{R}^{2d}} c(q, p) a(q, p) \, dq \, dp \right| \leq C \varepsilon^k \sum_{|\alpha| \leq k} |z_-|^{\alpha/2 - k} \|D^\alpha a\|_\infty. \quad (6)$$

Thus, averages of the cross term are super-polynomially small with respect to the semiclassical parameter ε if the difference of the centers z_- is of order one.

The oscillations in the cross term can also be estimated by using another quadratic phase space representation, the ambiguity function

$$A(\psi) : \mathbb{R}^{2d} \rightarrow \mathbb{C}, \quad (x, \xi) \mapsto (2\pi\varepsilon)^{-d} \int_{\mathbb{R}^d} e^{-iq \cdot x/\varepsilon} \psi(q - \frac{1}{2}\xi) \overline{\psi}(q + \frac{1}{2}\xi) \, dq.$$

The ambiguity and the Wigner function are related via Fourier transformation, that is,

$$A(\psi)(x, \xi) = (\mathcal{F}W(\psi))(x, \xi).$$

The ambiguity function can be written as the convolution of the modulated wave function $q \mapsto e^{-iq \cdot x/\varepsilon} \psi(q)$ with its inflection $q \mapsto \overline{\psi}(-q)$,

$$A(\psi)(x, \xi) = (2\pi\varepsilon)^{-d} e^{\frac{i}{2\varepsilon} x \cdot \xi} \left((e^{-i \bullet \cdot x/\varepsilon} \psi(\bullet)) * \overline{\psi}(-\bullet) \right) (\xi).$$

The analogous expression holds on the Fourier level,

$$A(\psi)(x, \xi) = (2\pi\varepsilon)^{-d} e^{\frac{i}{2\varepsilon} x \cdot \xi} \left((e^{-i \bullet \cdot \xi/\varepsilon} (\mathcal{F}\psi)(\bullet)) * \overline{(\mathcal{F}\psi)}(-\bullet) \right) (x).$$

Hence, $A(\psi)(x, \xi) \neq 0$ implies

$$(x, \xi) \in (\text{supp}(\mathcal{F}\psi) - \text{supp}(\mathcal{F}\psi)) \times (\text{supp}(\psi) - \text{supp}(\psi)),$$

where the minus sign builds the set of pointwise differences. This estimate on the support of the ambiguity function yields an alternative explanation that the cross term c has oscillations of order z_- .

3.3. Husimi functions

An alternative quadratic phase space representation is the Husimi function (14). It can be defined as a Gauß transform of the Wigner function,

$$H(\psi)(q, p) = (\varepsilon\pi)^{-d} \int_{\mathbb{R}^{2d}} W(\psi)(x, \xi) e^{-(|q-x|^2 + |p-\xi|^2)/\varepsilon} dx d\xi.$$

A few lines of computation yield its non-negativity, since it can be expressed as the modulus squared of the Fourier-Bros-Iagolnitzer (FBI) transform

$$H(\psi)(q, p) = |T_{\text{FBI}}(\psi)(q, p)|^2$$

with

$$T_{\text{FBI}}(\psi)(q, p) = 2^{-d/2} (\pi\varepsilon)^{-3d/4} \int_{\mathbb{R}^d} e^{i(q-y) \cdot p/\varepsilon} e^{-|q-y|^2/(2\varepsilon)} \psi(y) dy,$$

see for example Chapters 3.1, 3.6 in (12).

The Husimi function of the Gaussian wave packet g_{z_0} is the Gaussian

$$H(g_{z_0})(q, p) = (2\pi\varepsilon)^{-d} \exp\left(-\frac{1}{2\varepsilon}|q - q_0|^2 - \frac{1}{2\varepsilon}|p - p_0|^2\right),$$

whose variance is larger than that of the corresponding Wigner function. Hence, in general the Husimi function's marginals are not position and momentum densities. For a superposition of two Gaussian wave packets with centers in $z_1, z_2 \in \mathbb{R}^{2d}$, one computes

$$H(g_{z_1} + g_{z_2})(q, p) = H(g_{z_1})(q, p) + H(g_{z_2})(q, p) + 2c(q, p),$$

where the cross term

$$c_H(q, p) = (2\pi\varepsilon)^{-d} e^{-\frac{1}{8\varepsilon}|z_-|^2} \exp\left(-\frac{1}{2\varepsilon}|(q, p) - z_+|^2\right) \cos\left(\frac{1}{2\varepsilon}(c_{1,2} - (q, p) \wedge z_-)\right)$$

expectedly localizes around the mean $z_+ = (z_1 + z_2)/2$. The cosine has a phase shift $c_{1,2} = q(z_1) \cdot p(z_1) - q(z_2) \cdot p(z_2)$ and oscillates with a frequency proportional to the difference $z_- = z_1 - z_2$. However, due to the damping term, which is exponentially small in $|z_-|^2$, the oscillations are absorbed by the tails of the two Gaussian functions $H(g_{z_1})$ and $H(g_{z_2})$.

The following proposition shows that averages of Husimi and Wigner functions differ by a term of order ε . Moreover, for the Husimi function the Egorov theorem only holds with a remainder of order ε , which is worse than the error of order ε^2 valid for Wigner functions.

Proposition 1. *Let $a : \mathbb{R}^{2d} \rightarrow \mathbb{C}$ be a Schwartz function. Then, there is a constant $C = C(a) > 0$ depending on derivatives of a of order greater or equal than two such that for all $\psi \in L^2(\mathbb{R}^d)$*

$$\left| \int_{\mathbb{R}^{2d}} (H(\psi) - W(\psi))(q, p) a(q, p) dq dp \right| \leq C\varepsilon \|\psi\|_{L^2}^2.$$

Let $\psi(t)$ solve the scalar Schrödinger equation (1) and let Φ^t be the flow of the associated Hamiltonian system (2). Then, there is a constant $C = C(a, \Phi^t) > 0$ depending on derivatives of a and Φ^t of order greater or equal than two such that for all $\psi_0 \in L^2(\mathbb{R}^d)$

$$\left| \int_{\mathbb{R}^{2d}} (H(\psi(t)) - H(\psi_0) \circ \Phi^t)(q, p) a(q, p) dq dp \right| \leq C\varepsilon \|\psi_0\|_{L^2}^2.$$

Proof. With the phase space Gaussian $G(q, p) = (\varepsilon\pi)^{-d} e^{-(|q|^2+|p|^2)/\varepsilon}$, the averaged Husimi function writes as the convolution

$$\int_{\mathbb{R}^{2d}} H(\psi)(q, p) a(q, p) dq dp = \int_{\mathbb{R}^{2d}} W(\psi)(q, p) (a * G)(q, p) dq dp.$$

Since the integral of G is one, while its mean and variance are zero and $\varepsilon/2$, respectively, a second order Taylor approximation of the function a gives

$$(a * G)(q, p) = a(q, p) + O(\varepsilon),$$

where the error depends on second order derivatives of a . Then, the first assertion follows from the Calderon-Vaillancourt Theorem (e.g. Chapter 2.8 in (12)). For proving the Egorov type result, one observes

$$\begin{aligned} ((a * G) \circ \Phi^{-t})(q, p) &= (a \circ \Phi^{-t})(q, p) + O(\varepsilon) \\ &= ((a \circ \Phi^{-t}) * G)(q, p) + O(\varepsilon), \end{aligned}$$

where the error depends on second order derivatives for a and Φ^t . Using (4), one obtains

$$\begin{aligned} &\int_{\mathbb{R}^{2d}} H(\psi(t))(q, p) a(q, p) dq dp \\ &= \int_{\mathbb{R}^{2d}} W(\psi_0)(q, p) ((a * G) \circ \Phi^{-t})(q, p) dq dp + O(\varepsilon^2) \\ &= \int_{\mathbb{R}^{2d}} H(\psi_0)(q, p) (a \circ \Phi^{-t})(q, p) dq dp + O(\varepsilon). \end{aligned}$$

□

4. Monte Carlo sampling

We investigate the achievable accuracy of Monte Carlo techniques for the approximation of Wigner functions and for the numerical quadrature to compute its function values. Our motivation for a Monte Carlo approach is twofold: First, the dimension of the configuration space \mathbb{R}^d is large, and the computational cost of randomized schemes need not increase exponentially in d . Second, the intended sampling accuracy is rather low, since the eigenvalue surfaces $q \mapsto \lambda^\pm(q)$ stem from difficult electronic structure calculations and typically have low resolution. Moreover, surface hopping is an asymptotic particle method whose accuracy depends on the size of the problem dependent parameter ε . In summary, the accuracy of a Monte Carlo approach is expected to be sufficient.

4.1. Approximation by Metropolis Monte Carlo

The Metropolis algorithm (15) is one of the most popular and efficient Markov Chain Monte Carlo sampling schemes. We extend the standard approach to the sampling of real-valued functions $W : \mathbb{R}^{2d} \rightarrow \mathbb{R}$ with disconnected functional support. We assume a priori given phase space centers $\{c_i\}_{i=1}^s$ that define regions where W is non-negligible and has a local envelope of comparable variance. Then we use a jump method as in (16).

Choose a fixed jump rate $r_{\text{jump}} \in [0, 1]$. Select a point $(q_{\text{old}}, p_{\text{old}}) \in \mathbb{R}^{2d}$ in one of the sampling regions $i \in \{1, \dots, s\}$ and calculate $W_{\text{old}} = W(q_{\text{old}}, p_{\text{old}})$. Then start the following iteration.

1. Proposition step: Generate a random number r from the uniform distribution in the interval $[0, 1]$. If $r > r_{\text{jump}}$, generate a random vector Δ

from a normal distribution on \mathbb{R}^{2d} and give the old point a random displacement,

$$(q_{\text{new}}, p_{\text{new}}) = (q_{\text{old}}, p_{\text{old}}) + \Delta.$$

Otherwise, perform a jump step: Choose uniformly one of the other centers $j \in \{1, \dots, s\}$, $j \neq i$, and generate a new point via

$$(q_{\text{new}}, p_{\text{new}}) = (q_{\text{old}}, p_{\text{old}}) + c_j - c_i.$$

Always calculate $W_{\text{new}} = W(q_{\text{new}}, p_{\text{new}})$.

2. Acceptance step: Generate a random number ρ from a uniform distribution in the interval $[0, 1]$. Accept the trial move if

$$\rho < |W_{\text{new}}|/|W_{\text{old}}|$$

and set $(q_{\text{old}}, p_{\text{old}}) = (q_{\text{new}}, p_{\text{new}})$. Otherwise, reject the trial move and keep the old point $(q_{\text{old}}, p_{\text{old}})$.

The new point is located with respect to the new center as the old point with respect to the the old center. Hence, the proposal step keeps the symmetry of the standard Metropolis algorithm. The points $(q_k, p_k)_{k=1}^N$ shall form a Markov chain with stationary distribution $|W|/\|W\|_{L^1}$. If the chain is uniformly ergodic, then the central limit theorem holds (17), and the empirical means $\frac{1}{N} \sum_{k=1}^N a(q_k, p_k)$ approximate expectation values

$$\langle a \rangle_{|W|} = \int_{\mathbb{R}^{2d}} |W(q, p)| a(q, p) dq dp / \|W\|_{L^1}$$

in the following sense. If $\langle a^2 \rangle < \infty$, then there is a constant $\gamma_a > 0$ such that for all $c > 0$

$$\lim_{N \rightarrow \infty} \mathbb{P} \left(\left| \frac{1}{N} \sum_{k=1}^N a(q_k, p_k) - \langle a \rangle_{|W|} \right| \leq \frac{c\gamma_a}{\sqrt{N}} \right) = \frac{1}{\sqrt{2\pi}} \int_{-c}^c e^{-t^2/2} dt.$$

This well known convergence rate of order $N^{-1/2}$ does not depend on the dimension of the sampling space.

For us, W is the Wigner function of a normalized square integrable function. Hence, W is real-valued and satisfies $\int_{\mathbb{R}^{2d}} W(q, p) dq dp = 1$. Among the N sampling points are N_p and N_n points with positive and negative value of W , respectively. For the approximation of expectation values

$$\langle a \rangle_W = \int_{\mathbb{R}^{2d}} W(q, p) a(q, p) dq dp$$

by empirical means, we assign the weight $1/(N_p - N_n)$ and $-1/(N_p - N_n)$ for points with positive and negative function values, respectively.

4.2. Integration by importance sampling

Evaluating the Wigner function in a given phase space point, one has to solve a d -dimensional Fourier integral. For standard importance sampling, one rewrites the integral in the form

$$\begin{aligned} W(\psi)(q, p) &= (2\pi)^{-d} \int_{\mathbb{R}^d} e^{ix \cdot p} \psi(q - \frac{\varepsilon}{2}x) \bar{\psi}(q + \frac{\varepsilon}{2}x) dx \\ &= \int_{\mathbb{R}^d} \frac{f(x)}{w(x)} w(x) dx, \end{aligned}$$

where

$$f(x) = (2\pi)^{-d} e^{ix \cdot p} \psi(q - \frac{\varepsilon}{2}x) \bar{\psi}(q + \frac{\varepsilon}{2}x),$$

and $w(x) > 0$ is a normalized weight function, $\int_{\mathbb{R}^d} w(x) dx = 1$. If one generates sampling points $\{x_k\}_{k=1}^L$ according to w , which build a uniformly ergodic Markov chain, then the central limit theorem holds. One obtains an approximation of the integral

$$W(\psi)(q, p) \approx I := \frac{1}{L} \sum_{k=1}^L \frac{f(x_k)}{w(x_k)}$$

with a convergence rate of order $L^{-1/2}$ as $L \rightarrow \infty$. For our numerical experiments, we have chosen the Gaussian weight function

$$w(x) = \left(\frac{4\pi}{\varepsilon}\right)^{-d/2} \exp\left(-\frac{\varepsilon}{4}|x|^2\right), \quad (7)$$

since we expect integrands f with oscillations modulated by a Gaussian envelope whose variance is comparable to $2/\varepsilon$, see also the integral in (5).

4.2.1. Convergence test

The possible high-frequency oscillations of the integrand function f cannot be captured by any practicable weight function w , and convergence of the empirical means might be extremely slow due to a large error constant. We therefore use the following simple convergence test.

One defines a sequence of chain lengths $L_0 < \dots < L_{\max}$, sets a tolerance $\text{tol} > 0$, and starts by computing M different values $\{I_m\}_{m=1}^M$ of the integral based on independent sampling chains of length L_0 . We assume that these values are normally distributed with mean \bar{I} and variance σ_I^2 ,

$$\bar{I} = \frac{1}{M} \sum_{m=1}^M I_m, \quad \sigma_I^2 = \frac{1}{M-1} \sum_{m=1}^M (I_m - \bar{I})^2.$$

One computes a 95-percent confidence interval according to $K_I = \pm z \sigma_I / \sqrt{M}$ with $z = 1.96$ and checks whether

$$K_I < \bar{I} \text{ tol}. \quad (8)$$

If the tolerance is not met using sampling chains of length L_j , all chains are extended to length L_{j+1} until one reaches the maximal length L_{\max} . Either the last value of the mean \bar{I} or zero are used as an approximation to the

integral determining the function value $W(\psi)(q, p)$, depending on whether the convergence criterion has been satisfied or not.

In our experiments, we use the tolerance $\text{tol} = \varepsilon^{1/2}$, since the expected error of the single switch surface hopping algorithm is of order $\varepsilon^{1/2}$. We obtain satisfactory results with this simple test and therefore have not explored more sophisticated convergence criteria as for example the Gelman-Rubin criterion (18).

Setting the integral value to zero in the non-convergent case is motivated by the special class of wave functions ψ considered in our numerical experiments. Writing the integrand function in polar form $f(x) = e^{i\Phi(x)/\omega}|f(x)|$, its phase function $\Phi : \mathbb{R}^d \rightarrow \mathbb{R}$ does not have any stationary points, that is, there are no points $x \in \mathbb{R}^d$ with $\nabla_x \Phi(x) = 0$. Therefore, successive partial integration yields

$$\int_{\mathbb{R}^d} f(x) \, dx = \int_{\mathbb{R}^d} e^{i\Phi(x)/\omega} |f(x)| \, dx = O(\omega^\infty), \quad \omega \rightarrow 0,$$

and we regard integrals that do not meet the convergence criterion as being of this highly oscillatory type.

5. Numerical set up

In this section, we present the general setup of our numerical experiments. We consider a two-dimensional Schrödinger system

$$i\varepsilon \partial_t \psi(t, q) = \left(-\frac{\varepsilon^2}{2} \Delta_q + V(q) \right) \psi(t, q), \quad \psi(0, q) = \psi_0(q) \quad (9)$$

with linear $E \otimes e$ Jahn-Teller potential matrix

$$V(q) = \gamma |q|^2 + \begin{pmatrix} q_1 & q_2 \\ q_2 & -q_1 \end{pmatrix}, \quad (10)$$

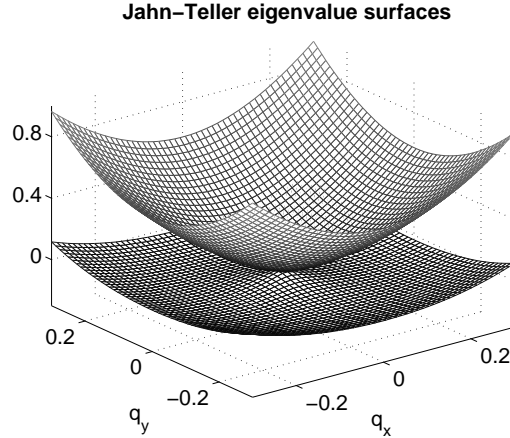


Figure 1: Conical intersection of eigenvalue surfaces of the Jahn-Teller potential (10).

which provides the most prominent example for a conical intersection of eigenvalues. Indeed, the potential matrix's eigenvalues

$$\lambda^\pm(q) = \gamma|q|^2 \pm \sqrt{q_1^2 + q_2^2} = \gamma|q|^2 \pm |q|$$

intersect in the point $q = 0$, which constitutes a codimension two submanifold of \mathbb{R}^2 , see Figure 1. Jahn-Teller Hamiltonians for silver, copper, sodium, and potassium fitted by recent electronic structure calculations (19) have associated semiclassical parameter $\varepsilon = 0.005, 0.007, 0.011, 0.008$ and quadratic confinement $\gamma = 2.613, 5.097, 0.524, 0.260$, respectively. These results motivate our default choice

$$\varepsilon = 0.01, \quad \gamma = 3.$$

The time interval is set to

$$[t_i, t_f] = [0, 10\sqrt{\varepsilon}].$$

For the initial data considered, the time interval allows the solution of the Schrödinger equation to pass the crossing point once and to generate leading order non-adiabatic transitions between the eigenspaces. The key quantity for the evaluation of the surface hopping algorithm with initial Monte Carlo sampling are the final level populations

$$P^\pm(t_f) = \|\psi^\pm(t_f)\|_{L^2}^2 = \int_{\mathbb{R}^d} |\psi^\pm(t_f, q)|^2 dq,$$

which give the probability of finding the system in one of the two eigenspaces at time t_f . The reference values for the final populations are computed by a highly resolved pseudospectral Strang splitting scheme, see Appendix A.

5.1. Analytic initial data

Our analytic initial data are the pointwise product of a scalar wave function $\psi_0^+ : \mathbb{R}^2 \rightarrow \mathbb{C}$ with an eigenvector $\chi^+(q)$ of the potential matrix associated with the eigenvalue $\lambda^+(q)$,

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

For the scalar wave function, we choose

$$\psi_0^+ \in \left\{ g_{z_0}, \frac{1}{\sqrt{2}} (g_{z_1} + g_{z_2}) \right\}$$

where $g_{z_0}, g_{z_1}, g_{z_2}$ are Gaussian wave packets centered in

$$z_0 = (5\sqrt{\varepsilon}, 0.5\sqrt{\varepsilon}, -1, 0), \quad z_1 = z_0, \quad z_2 = -z_1.$$

The supports of g_{z_1} and g_{z_2} have negligible overlap, such that the superposition $(g_{z_1} + g_{z_2})/\sqrt{2}$ can be regarded as a wave function of L^2 -norm one.

Eigenvectors of potentials with a conical crossing have non-removable discontinuities at crossing points, and smoothness away from the crossing is only possible if they are complex-valued. We consider the two cases

$$\chi^+(q) \in \{\tilde{\chi}(q), e^{\frac{i}{2}\vartheta_q}\tilde{\chi}(q)\}, \quad \tilde{\chi}(q) = (\cos(\frac{1}{2}\vartheta_q), \sin(\frac{1}{2}\vartheta_q))^T,$$

where $\vartheta_q \in (-\pi, \pi)$ is the polar angle of $q \in \mathbb{R}^2$. The complex-valued phase factor $\exp(\frac{i}{2}\vartheta_q)$ compensates the discontinuity of $\tilde{\chi}$ across the left half axis

$$\{q \in \mathbb{R}^2 \mid q_1 \leq 0, q_2 = 0\}.$$

Since z_0 is located in the upper right quadrant of position space, the overlap of the single Gaussian wave packet g_{z_0} with the left half axis is negligible. Therefore, we choose the real-valued eigenvector $\chi^+(q) = \tilde{\chi}(q)$ for g_{z_0} . On the other hand, g_{z_2} overlaps with the left half axis. Thus we use the complex-valued eigenvector $\chi^+(q) = e^{\frac{i}{2}\vartheta_q}\tilde{\chi}(q)$ for the superposition $(g_{z_1} + g_{z_2})/\sqrt{2}$.

5.2. Numerically computed laser excitation

Gaussian wave packets associated with the upper eigenvalue are a simple model for a molecule excited by light or a laser-pulse. As a third test case, we simulate laser excitation for the Jahn-Teller system numerically and use the resulting wave function as initial datum for the Schrödinger system (9).

The real-symmetric matrix

$$M(q) = \begin{pmatrix} 2 \cos \vartheta_q \sin \vartheta_q & \cos^2 \vartheta_q - \sin^2 \vartheta_q \\ \cos^2 \vartheta_q - \sin^2 \vartheta_q & -2 \cos \vartheta_q \sin \vartheta_q \end{pmatrix}$$

maps the eigenspaces of the Jahn-Teller matrix $V(q)$ onto each other. Hence, the Schrödinger system

$$i\varepsilon \partial_t \phi(t, q) = \left(-\frac{\varepsilon^2}{2} \Delta_q + V(q) + E(t)M(q) \right) \phi(t, q), \quad \phi(0, q) = \phi_0(q) \quad (11)$$

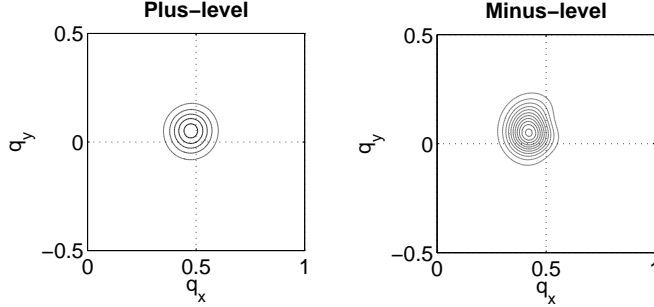


Figure 2: Contour plots of the position densities $|\psi_0^\pm(q)|^2$ of the numerically excited state.

with oscillatory electric field $E(t) = \frac{1}{\varepsilon} \exp(-t^2) \cos(\frac{t}{\varepsilon})$ models the excitation between the eigenspaces of the matrix $V(q)$, see also (20). We excite a Gaussian wave packet associated with the lower eigenvalue $\lambda^-(q)$,

$$\phi_0(q) = g_{z_0}(q) \left(-\sin(\frac{1}{2}\vartheta_q), \cos(\frac{1}{2}\vartheta_q) \right)^T,$$

and use the resulting wave function $\phi(t, q)$ at time $t = 0.032$ as initial datum $\psi_0(q)$ for the Jahn-Teller system (9). This initial state populates the upper level with a probability of 83 percent, and the scalar wave functions ψ_0^\pm have nearly Gaussian shape, compare Figure 2. However, $\|\psi_0 - g_{z_0}\chi^+\|_{L^2} = 0.65$.

6. Gaussian wave packet

The Wigner function of the Gaussian wave packet g_{z_0} is a Gaussian function on phase space. Hence, the sampling of $W(g_{z_0})$ can be realized by directly drawing from a normal distribution. However, ignoring the analytic knowledge on the explicit form of $W(g_{z_0})$, we also study the accuracy

Table 1: Gaussian wave packet. Statistics of $P^+(t_f)$ for Monte Carlo sampling of the initial Wigner distribution. The results are compared for different numbers of sampling points ($N = 100, \dots, 2000$), where $m = 10$ runs of the single switch algorithm are evaluated for each N . The reference value from Strang splitting amounts to $P_{\text{Strang}}^+(t_f) = 0.264$.

N	100	200	500	1000	1500	2000
Simple Monte Carlo						
mean	0.277	0.267	0.271	0.275	0.273	0.275
standard deviation	0.030	0.022	0.012	0.011	0.005	0.004
Metropolis approximation and importance sampling						
mean	0.238	0.269	0.243	0.256	0.275	0.271
standard deviation	0.088	0.025	0.041	0.032	0.016	0.015

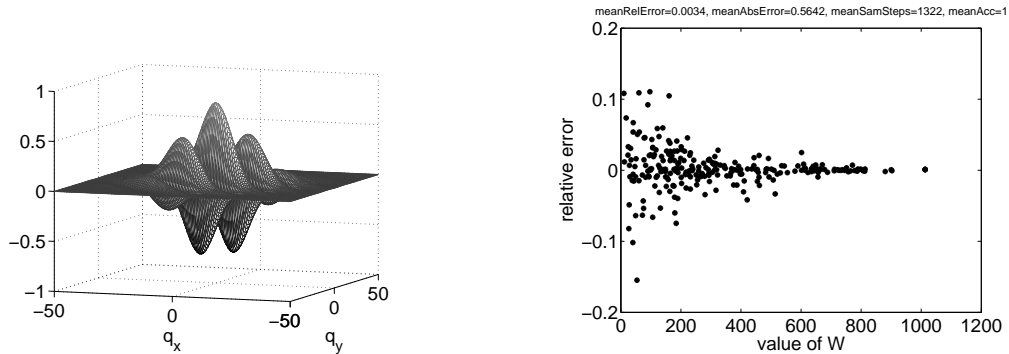
when using Metropolis Monte Carlo for the approximation and importance sampling for the integration.

6.1. Simple Monte Carlo

Since $W(g_{z_0})$ is a Gaussian function, one can simply generate approximation points $(q_k, p_k)_{k=1}^N$ by sampling from a multi-dimensional normal distribution and transforming it according to

$$(q, p) = (q_0, p_0) + \sigma y, \quad \sigma = \sqrt{\varepsilon/2}, \quad y \sim N_{2d}(0, 1).$$

We use these sampling points in combination with the surface hopping algorithm. As accuracy criterion, we take the deviation of the final population $P^+(t_f)$ from the reference value 0.264 stemming from the Strang splitting



(a) Typical integrand for computing the value of $W(g_{z_0})$.

(b) Distribution of relative errors when computing $W(g_{z_0})$.

Figure 3: Integration by importance sampling for computing the Wigner function of the Gaussian wave packet g_{z_0} . The left plot shows the integrand for computing the value of $W(g_{z_0})(q_0, p)$ with $p = p_0 + (0.3, 0.3)$. The right plot gives the distribution of relative errors with respect to the exact value of the Wigner function $W(g_{z_0})$ for 500 different phase space points.

scheme. The initial sampling is performed with $N = 100, \dots, 2000$ different numbers of sampling points. Then, for each fixed N there are $m = 10$ runs of the surface hopping algorithm. The results are listed in the upper lines of Table 1. The variance decreases monotonically as N increases. Hence, the results of a single run become more reliable for larger N . Though the surface hopping algorithm systematically overestimates the reference value, all the mean values differ at most by two percent.

6.2. Metropolis approximation and importance sampling

Approximating the Wigner function $W(g_{z_0})$, we use the Metropolis algorithm with normally distributed displacement

$$\Delta \sim N_{2d}(0, \varepsilon/2).$$

The starting point of the Markov chains is always the phase space center $z_0 = (q_0, p_0)$ of the Gaussian wave packet. Computing the value of $W(g_{z_0})$ in a point (q, p) , one has to solve the integral

$$\frac{1}{4\varepsilon\pi^3} e^{-\frac{1}{\varepsilon}|q-q_0|^2} \int_{\mathbb{R}^d} \cos(x \cdot (p - p_0)) e^{-\frac{\varepsilon}{4}|x|^2} dx. \quad (12)$$

Using that $\frac{\varepsilon}{4\pi} \int_{\mathbb{R}^2} e^{-\frac{\varepsilon}{4}|x|^2} dx = 1$, the integral can be approximated by importance sampling as

$$W(g_{z_0})(q, p) \approx \frac{1}{L\varepsilon^2\pi} e^{-\frac{1}{\varepsilon}|q-q_0|^2} \sum_{k=1}^L \cos(x_k \cdot (p - p_0)).$$

The sampling points $\{x_k\}_{k=1}^L$ are distributed according to $\exp(-\frac{\varepsilon}{4}|x|^2)$, which represents the two-dimensional normal distribution with mean zero and variance $2/\varepsilon$.

As the distance $|p - p_0|$ gets larger, the oscillation frequency of the integrand increases, which causes severe difficulties in Monte Carlo quadrature. Figure 3(a) shows the integrand for computing $W(g_{z_0})(q_0, p)$ with $p = p_0 + (0.3, 0.3)$. Even though the integral value is small in this case, the quadrature scheme can yield large errors due to numerical cancellation. We therefore use the convergence test from §4.2.1.

We set the number of chains $M = 5$ and the sequence of chain lengths $L_0 = 1000, L_1 = 2000, \dots, L_{\max} = 10000$. A typical distribution of relative

errors with respect to the exact value of the Wigner function is illustrated in Figure 3(b). The relative as well as the absolute error are centered around zero (mean relative/absolute error 0.003/0.564), supporting the expectation that oscillations with subsequent small integral value pose the main difficulty. On average, for 82 percent of the phase points (q, p) the computation of $W(g_{z_0})(q, p)$ does converge according to (8).

If the sampled phase points are propagated by the single switch algorithm, then the statistics for the upper level population at time $t = t_f$ are slightly worse than for the simple Monte Carlo approach. The lower lines in Table 1 contain the results for $N = 100, \dots, 2000$ sampling points and $m = 10$ surface hopping runs for each fixed N . The mean values differ from the reference at most by three percent.

7. Superposition of Gaussian wave packets

For the superposition of Gaussian wave packets $(g_{z_1} + g_{z_2})/\sqrt{2}$, the Wigner function is no longer positive and has three pronounced peaks with disconnected support. In the following, we examine the accuracy of the Metropolis type approximation proposed in §4.1. Then, we compare the accuracy for a sequence of surface hopping runs, which are either started with the sampling of the Wigner or the Husimi function of the initial wave function.

7.1. Approximation of the Wigner function

The Wigner function consists of the sum of the two phase space Gaussians

$$W(g_{z_j})(q, p) = (\pi\varepsilon)^{-d} \exp\left(-\frac{1}{\varepsilon}|(q, p) - z_j|^2\right), \quad j = 1, 2$$

Table 2: Superposition of Gaussian wave packets. Statistics of $P^+(t_f)$ for Metropolis approximation of the initial Wigner distribution. The results are for different numbers of sampling points ($N = 100, \dots, 2000$), where $m = 10$ runs of the single switch algorithm are evaluated for each N . The reference value amounts to $P_{\text{Strang}}^+(t_f) = 0.275$.

N	100	200	500	1000	1500	2000
Complete distribution						
mean	0.325	0.303	0.287	0.288	0.286	0.278
standard deviation	0.173	0.079	0.034	0.045	0.033	0.019
Without cross term						
mean	0.276	0.270	0.265	0.271	0.275	0.273
standard deviation	0.022	0.019	0.014	0.010	0.004	0.006

plus an oscillatory cross term localized around the middle point $z_+ = 0$,

$$c(q, p) = (\pi\varepsilon)^{-d} e^{-|(q,p)|^2/\varepsilon} \cos\left(\frac{1}{\varepsilon}(q, p) \wedge z_-\right),$$

which has a Gaussian envelope with the same variance and oscillates with a frequency proportional to the difference $z_- = (10\sqrt{\varepsilon}, \sqrt{\varepsilon}, -2, 0)$.

To sample from this distribution, we apply the Metropolis algorithm of §4.1 with jumps between the centers $\{c_1, c_2, c_3\} = \{z_1, z_2, z_+\}$ and jump rate $r_{\text{jump}} = 0.5$. The random displacement $\Delta \in \mathbb{R}^4$ in the proposition step is drawn from the normal distribution with standard deviation $\sqrt{\varepsilon/2}$. This choice of Δ and r_{jump} results in a mean acceptance ratio of about 56% (jump acceptance ratio about 80%). Examining the accuracy of the computed population $P^+(t_f)$, we consider different numbers of sampling points

$N = 100, \dots, 2000$ and perform the hopping algorithm $m = 10$ times for each fixed N . The results are listed in the upper lines of Table 2. Already with hundred sampling points, the mean value differs from the reference value by five percent. With more points the results improve towards an error of two percent and in the sense of variance reduction. Hence, fewer sampling points require several simulations, whereas for many sampling points, fewer simulations are sufficient to obtain reliable results.

Since the difference vector z_- has length of order one with respect to the semiclassical parameter ε , averages of the cross term c are super-polynomially small in ε , see the upper bound (6). Therefore, in a second set of experiments we ignore the oscillatory middle peak and approximate the two remaining Gaussians $W(g_{z_1})$ and $W(g_{z_2})$ independently by a simple Monte Carlo approach. Each Gaussian is approximated with the same number of sampling points. The lower lines of Table 2 illustrate no loss in accuracy but a reduction of variance.

7.2. Approximation of the Husimi function

The chemical literature often treats the Wigner and the Husimi function on the same footing for the initial sampling of semiclassical particle methods. In the following, we compare the accuracy of the single switch algorithm when the phase space sampling of the initial wave function $(g_{z_1} + g_{z_2})/\sqrt{2}$ is performed for the Husimi function, the Wigner function, and the Wigner function without cross term.

The quadratic confinement γ in the Jahn-Teller potential (10) is set to zero, and the semiclassical parameter ε varies as $\{0.0005, 0.001, 0.01, 0.1\}$. All three initial distributions are sampled by the Metropolis type algorithm

Table 3: Superposition of Gaussian wave packets. Absolute error of $P^+(t_f)$ for different values of ε and different phase space representations (Wigner or Husimi). The reference value is calculated with a highly resolved Strang splitting. The surface hopping values are mean values over $m = 10$ runs with $N = 5000$ particles each.

ε	0.1	0.01	0.001	0.0005
$\varepsilon^{1/2}$	0.316	0.1	0.032	0.0224
Husimi	0.125	0.102	0.096	0.096
Wigner	0.018	0.003	0.043	0.052
Wigner(no cross term)	0.010	0.002	0.014	0.015

used in the previous section §7.1.

The surface hopping results in Table 3 are mean values over $m = 10$ runs with $N = 5000$ sampling points each. The Wigner functions' errors are smaller than those for the Husimi function by at least a factor of two. This difference can be explained by the better asymptotic properties of the Wigner function with respect to classical transport. As discussed in section §3.3 before, the Egorov theorem for scalar Schrödinger equations gives an approximation error of order ε^2 for the Wigner function, whereas the Husimi function only yields an error of order ε . The rougher approximation of the single switch algorithm does not completely bury this difference. Moreover, the smaller the semiclassical parameter the higher is the oscillation frequency, and neglecting the oscillatory cross term of the Wigner function remarkably improves accuracy. The errors obtained using Wigner functions are well below $\varepsilon^{1/2}$, confirming earlier numerical evidence (4; 5; 6).

Table 4: Laser excited state. Statistics of $P^+(t_f)$ for Metropolis approximation and integration by importance sampling for the initial Wigner function. The results are for different numbers of phase space sampling points ($N = 100, \dots, 2000$), where $m = 10$ runs of the single switch algorithm are evaluated for each N . The reference value from Strang splitting amounts to $P_{\text{Strang}}^+(t_f) = 0.312$.

N	100	200	500	1000	1500	2000
mean	0.369	0.321	0.311	0.324	0.306	0.310
standard deviation	0.100	0.055	0.026	0.029	0.026	0.020

8. Laser excited state

For the numerically computed laser excitation, the input for the surface hopping algorithm is a vector-valued function $\psi_0 : \mathbb{R}^2 \rightarrow \mathbb{C}^2$, $q \mapsto \phi(t = 0.032, q)$ whose values are only known at grid points in position space. Since this function results from the excitation of a Gaussian wave packet g_{z_0} , we assume that the two level functions $\psi_0^\pm : \mathbb{R}^2 \rightarrow \mathbb{C}$ have a comparable localization and frequency range as g_{z_0} , see also Figure 2.

For the approximation of the real-valued Wigner functions $W(\psi_0^\pm)$, we therefore apply the Metropolis type algorithm proposed in §4.1 with starting point z_0 and displacement vector $\Delta \sim N_{2d}(0, \varepsilon/2)$. For the computation of the function values

$$W(\psi_0^\pm)(q, p) = (2\pi)^{-2} \int_{\mathbb{R}^d} e^{ix \cdot p} \psi_0^\pm(q - \frac{\varepsilon}{2}x) \overline{\psi_0^\pm}(q + \frac{\varepsilon}{2}x) dx,$$

we again solve the integrals by importance sampling with sampling points $\{x_k\}_{k=1}^L$ distributed according to the Gaussian weight function w from (7) and use the convergence test with $M = 5$ chains of maximal length $L_{\max} = 10000$.

Similar to the single Gaussian wave packet, the importance sampling converged within the maximum number of sampling steps for about 82% of the phase points. The single difference to the preceding experiments is that the functions ψ_0^\pm are only known at grid points. Therefore, each evaluation of the integrand in a sampling point x_k requires an additional linear interpolation with respect to the grid points. Nevertheless, the overall computing time for the approximation of $W(\psi_0^\pm)$ is comparable to the Metropolis approximation of the single Gaussian wave packet (about 460 seconds CPU-time for 1000 phase points with MATLAB 7.1 on a 2200 MHz Dual Core Processor).

Table 4 summarizes the numerical results for $N = 100, \dots, 2000$ phase space sampling points and $m = 10$ surface hopping runs for each fixed N . Similarly to the analytic initial data, the mean level population achieves an accuracy of roughly one percent, and the variance decreases with increasing number of sampling points.

9. Conclusion

We have studied the accuracy of Monte Carlo sampling for Wigner functions in the context of particle methods for non-adiabatic Schrödinger systems. The approximation step is realized by an adaption of the Metropolis algorithm for real-valued functions with disconnected support. The quadrature, which computes values of the Wigner function, uses importance sampling with a Gaussian weight function.

Three sets of numerical experiments have shown that the proposed Monte Carlo schemes combined with the single switch surface hopping algorithm give results with an error of two to three percent, which is in very good

agreement with the methodological accuracy of surface hopping algorithms. We have also considered the Husimi function as an alternative phase space representation. However, the subsequent surface hopping results are systematically less accurate than for the Wigner function.

The presented numerical experiments are low-dimensional, since an accuracy study requires the validation against a convergent solution of the underlying time-dependent Schrödinger systems, which is only feasible for few degrees of freedom. High-dimensional experiments as well as the Monte Carlo integration of oscillatory functions with stationary points have to be addressed in future work.

A. Reference solutions

For evaluating the different initial sampling strategies in combination with the single switch algorithm, we directly solve the Jahn-Teller system (9) with a pseudo-spectral Strang splitting scheme. For this two-dimensional problem a space discretization based on the fast Fourier transform and an operator splitting with third order local convergence in time (21) provides accurate reference solutions. The number of time steps is set to 5000 for all experiments.

Table 5 contains the computational domains, the grid sizes, the final population $P^+(t_f)$, and the achieved accuracy. The accuracy of the solution refers to the difference $\|\psi(t_f) - \psi_c(t_f)\|_{L^2}$ of the final reference solution $\psi(t_f)$ and a coarser solution $\psi_c(t_f)$, which is computed with fourth the number of grid points and half the number of time steps.

In section §7.2, we have varied the semiclassical parameter ε to compare

Table 5: Input parameters and results for the grid-based solution of the Jahn-Teller system (9) in the top lines and for the Schrödinger system with electric field (11) in the bottom line. The final times t_f are 1 and 0.032 for system (9) and (11), respectively. In all cases, $t_i = 0$, $\varepsilon = 0.01$ and $\gamma = 3$.

	domain	fine grid size	$P^+(t_f)$	accuracy
$\psi_0^+ = g_{z_0}$	$[-2, 2] \times [-1, 1]$	1024×512	0.264	$2.3 \cdot 10^{-6}$
$\psi_0^+ = \frac{1}{\sqrt{2}}(g_{z_1} + g_{z_2})$	$[-\frac{3}{2}, \frac{3}{2}] \times [-\frac{3}{4}, \frac{3}{4}]$	2048×1024	0.275	$2.3 \cdot 10^{-6}$
$\psi_0 = \phi(t = 0.032)$	$[-2, 2] \times [-1, 1]$	1024×512	0.312	$2.3 \cdot 10^{-6}$
Laser excitation	$[-2, 2] \times [-1, 1]$	1024×512	0.833	$2.1 \cdot 10^{-7}$

Wigner and Husimi functions. The input parameters as well as the accuracy of the corresponding reference solutions are listed in Table 6. The achieved errors are all sufficient for the validation of the single switch algorithm, whose accuracy for the computation of quadratic quantities of the wave function typically varies between two and three percent.

For the computation of the laser excited initial state, the Schrödinger system with time-dependent electric field (11) is solved by a pseudo-spectral Trotter splitting with second order local convergence in time. Again, the two-dimensional Laplacian is realized by the fast Fourier transform. The time interval $[0, 0.04]$ is discretized by 500 time steps.

B. Analytical Wigner transformation

The Wigner function of Gaussian wave packets can be calculated explicitly. Here, we give some details for the analytic computation of the cross

Table 6: Input parameters and results for the reference solution of the Jahn-Teller system (9) with $\gamma = 0$, $\psi_0^+ = (g_{z_1} + g_{z_2})/\sqrt{2}$, and $t_f = 10\sqrt{\varepsilon}$ in dependence on the semiclassical parameter ε .

ε	domain	fine grid size	$P^+(t_f)$	accuracy
10^{-1}	$[-10, 10] \times [-5, 5]$	2048×1024	0.378	$1.0 \cdot 10^{-3}$
10^{-2}	$[-\frac{3}{2}, \frac{3}{2}] \times [-\frac{3}{4}, \frac{3}{4}]$	2048×1024	0.436	$1.9 \cdot 10^{-4}$
10^{-3}	$[-1, 1] \times [-\frac{1}{2}, \frac{1}{2}]$	2048×1024	0.526	$1.6 \cdot 10^{-7}$
$5 \cdot 10^{-4}$	$[-\frac{1}{2}, \frac{1}{2}] \times [-\frac{1}{4}, \frac{1}{4}]$	2048×1024	0.544	$1.0 \cdot 10^{-7}$

term of the Wigner function $W(g_{z_1} + g_{z_2})$. One solves the integral

$$\begin{aligned} W(g_{z_1}, g_{z_2})(q, p) &= (2\pi)^{-d} \int_{\mathbb{R}^d} e^{ix \cdot p} g_{z_1}(q - \frac{\varepsilon}{2}x) \bar{g}_{z_2}(q + \frac{\varepsilon}{2}x) dx \\ &= (2\pi)^{-d} (\pi\varepsilon)^{-d/2} \\ &\quad \int_{\mathbb{R}^d} e^{ix \cdot p} e^{-\frac{1}{2\varepsilon}(|q - \frac{\varepsilon}{2}x - q_1|^2 + |q + \frac{\varepsilon}{2}x - q_2|^2)} e^{\frac{i}{\varepsilon}(p_1 \cdot (q - \frac{\varepsilon}{2}x - q_1) - p_2 \cdot (q + \frac{\varepsilon}{2}x - q_2))} dx, \end{aligned}$$

where $q_{1,2}$ and $p_{1,2}$ denote the position and momentum component of the phase space points $z_{1,2}$ for the rest of the calculation. Rewriting the quadratic part as

$$|q - \frac{\varepsilon}{2}x - q_1|^2 + |q + \frac{\varepsilon}{2}x - q_2|^2 = |q - q_1|^2 + |q - q_2|^2 + \frac{\varepsilon^2}{2}|x|^2 + \varepsilon x \cdot q_-,$$

one has

$$\begin{aligned} W(g_{z_1}, g_{z_2})(q, p) &= (2\pi)^{-d} (\pi\varepsilon)^{-d/2} e^{-\frac{1}{2\varepsilon}|q - q_1|^2 - \frac{1}{2\varepsilon}|q - q_2|^2} \\ &\quad e^{\frac{i}{\varepsilon}(p_1 \cdot (q - q_1) - p_2 \cdot (q - q_2))} \int_{\mathbb{R}^d} e^{ix \cdot (p - p_+ + \frac{i}{2}q_-)} e^{-\frac{\varepsilon}{4}|x|^2} dx. \end{aligned}$$

One uses the value of the Gaussian integral

$$\int_{\mathbb{R}} e^{iyw} e^{-\alpha y^2} dy = \sqrt{\frac{\pi}{\alpha}} e^{-w^2/(4\alpha)}, \quad w \in \mathbb{C}, \alpha > 0$$

for $w^j = p^j - p_+^j + \frac{i}{2}q_-^j$ with $j = 1, \dots, d$ and $\alpha = \varepsilon/4$. Since

$$\sum_{j=1}^d (w^j)^2 = |p - p_+|^2 - \frac{1}{4}|q_-|^2 + i(p - p_+) \cdot q_-,$$

one gets

$$\begin{aligned} W(g_{z_1}, g_{z_2})(q, p) &= (\pi\varepsilon)^{-d} e^{-\frac{1}{2\varepsilon}|q-q_1|^2 - \frac{1}{2\varepsilon}|q-q_2|^2} e^{\frac{i}{\varepsilon}(p_1 \cdot (q-q_1) - p_2 \cdot (q-q_2))} \\ &\quad e^{-\frac{1}{\varepsilon}|p-p_+|^2} e^{\frac{1}{4\varepsilon}|q_-|^2} e^{-\frac{i}{\varepsilon}(p-p_+) \cdot q_-}. \end{aligned}$$

Observing that $-\frac{1}{2}|q - q_1|^2 - \frac{1}{2}|q - q_2|^2 + \frac{1}{4}|q_-|^2 = -|q - q_+|^2$ and

$$p_1 \cdot (q - q_1) - p_2 \cdot (q - q_2) - (p - p_+) \cdot q_- = ((q, p) - z_+) \wedge z_- - p_+ \cdot q_-$$

one arrives at

$$W(g_{z_1}, g_{z_2})(q, p) = (\pi\varepsilon)^{-d} e^{-\frac{i}{\varepsilon}p_+ \cdot q_-} e^{-\frac{1}{\varepsilon}|(q,p) - z_+|^2} e^{\frac{i}{\varepsilon}((q,p) - z_+) \wedge z_-}.$$

and

$$\begin{aligned} c(q, p) &= \frac{1}{2}(W(g_{z_1}, g_{z_2})(q, p) + W(g_{z_2}, g_{z_1})(q, p)) \\ &= (\pi\varepsilon)^{-d} e^{-\frac{1}{\varepsilon}|(q,p) - z_+|^2} \cos\left(\frac{1}{\varepsilon}(p_+ \cdot q_- - ((q, p) - z_+) \wedge z_-)\right). \end{aligned}$$

References

- [1] D. Robert, *Autour de l'Approximation Semi-Classique*, Birkhäuser, Boston, 1987.
- [2] E. Wigner, On the quantum correction for thermodynamic equilibrium, *Phys. Rev.* 40 (1932) 749–759.
- [3] W. Domcke, D. Yarkony, H. Köppel (Eds.), *Conical intersections*, World Scientific, Singapore, 2004.

- [4] C. Fermanian-Kammerer, C. Lasser, Propagation through generic level crossings: a surface hopping semigroup, *SIAM J. Math. An.* 40 (2008) 103–133.
- [5] C. Lasser, T. Swart, S. Teufel, Construction and validation of a rigorous surface hopping algorithm for conical crossings, *Comm. Math. Sci.* 5 (2007) 789–814.
- [6] C. Fermanian-Kammerer, C. Lasser, Single switch surface hopping for molecular dynamics with transitions, *J. Chem. Phys.* 128 (2008) 144102–1–9.
- [7] J. Tully, R. Preston, Trajectory surface hopping approach to nonadiabatic molecular collisions: the reaction of H^+ with D_2 , *J. Chem. Phys.* 55 (1971) 562–572.
- [8] J. Tully, Molecular dynamics with electronic transitions, *J. Chem. Phys.* 93 (1990) 1061–1071.
- [9] C. Lasser, T. Swart, Single switch surface hopping for a model of pyrazine, *J. Chem. Phys.* 129 (2008) 034302–1–8.
- [10] G. Folland, *Harmonic Analysis in Phase Space*, Princeton University Press, Princeton, 1989.
- [11] K. Gröchenig, *Foundations of Time-Frequency Analysis*, Birkhäuser, Boston, 2001.
- [12] A. Martinez, *An Introduction to Semiclassical and Microlocal Analysis*, Springer, New York, 2002.

- [13] L. Hörmander, *The Analysis of Linear Partial Differential Operators I*, Springer-Verlag, Berlin, 1983.
- [14] K. Husimi, Some formal properties of the density matrix, *Proc. Phys. Math. Soc. Japan* 22 (1940) 264–314.
- [15] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, E. Teller, Equation of state calculations by fast computing machines, *J. Chem. Phys.* 21 (1953) 1087–1092.
- [16] L. Walter, M. Weber, ConfJump: a fast biomolecular sampling method which drills tunnels through high mountains, ZIB-Report 06-26, Zuse Institute Berlin, <http://opus.kobv.de/zib/volltexte/2006/920/> (2006).
- [17] L. Tierney, Markov chains for exploring posterior distributions, *The Annals of Statistics* 22 (1994) 1701–1728.
- [18] A. Gelman, D. B. Rubin, Inference from iterative simulation using multiple sequences, *Statistical Science* 7 (1992) 457–511.
- [19] P. Garcia-Fernandez, I. Bersuker, M. 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 (2005) 184117–1–10.
- [20] G. Hagedorn, S. Jilcott, V. Rousse, The AC Stark effect, time-dependent Born-Oppenheimer approximation, and Franck-Condon factors, *Ann. I.H.P.* 7 (2006) 1065–1083.

- [21] T. Jahnke, C. Lubich, Error bounds for exponential operator splitting,
BIT 40 (4) (2000) 735–744.