Konrad-Zuse-Zentrum
für Informationstechnik Berlin

Takustraße 7
D-14195 Berlin-Dahlem
Germany

SUSANNA KUBE, CAROLINE LASSER, MARCUS WEBER

Monte Carlo sampling of Wigner functions and surface hopping quantum dynamics

ZIB-Report 07-17 (Dezember 2007)

Monte Carlo sampling of Wigner functions and surface hopping quantum dynamics

Susanna Kube* Caroline Lasser† Marcus Weber‡

Abstract

Wigner functions are functions on classical phase space, which are in one-to-one correspondence to square integrable functions on configuration space. For molecular quantum systems, classical transport of Wigner functions provides the basis of asymptotic approximation methods in the high energy regime. The article addresses the sampling of Wigner functions by Monte Carlo techniques. 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. The numerical experiments combine the sampling with a surface hopping algorithm for non-adiabatic quantum dynamics. In agreement with theoretical considerations, the obtained results show an accuracy of two to four percent.

AMS MSC 2000: 62D05, 65D30, 81S30, 81Q20

Keywords: Metropolis Monte Carlo, approximation, quadrature, oscillating functions, quantum dynamics, non-adiabatic, surface hopping

*Zuse Institute Berlin, Takustraße 7, D-14195 Berlin, Germany (kube@zib.de)
†FU Berlin, Arnimallee 6, D-14195 Berlin, Germany (lasser@math.fu-berlin.de)
‡Zuse Institute Berlin, Takustraße 7, D-14195 Berlin, Germany (weber@zib.de)

Contents

Introduction

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

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

It is a linear partial differential equation with a unique global solution provided by the spectral theorem. Indeed, under reasonable regularity and growth assumptions on the potential $V : \mathbb{R}^d \to \mathbb{R}$, the Hamiltonian $P = -\frac{\hbar^2}{2m}\Delta_q + V(q)$ is a self-adjoint operator in $L^2(\mathbb{R}^d)$, and the solution can be written as

$$\psi(\tau,q) = \mathrm{e}^{-iP\tau/\hbar}\psi_0(q).$$

For most molecules, the dimension d of the configuration spaces $\mathbb{R}^d$ is large. In typical applications, one might want to consider up to thirty degrees of freedom. On top of that, the solution $\psi(t,q)$ exhibits oscillations both in time and in space. For quantifying the oscillatory behavior, one switches to atomic units by setting $\hbar = 1$ and introduces the crucial semiclassical parameter

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

On the long time scale $t = \tau/\varepsilon$, on which the distinguished dynamical features develop, the Schrödinger equation writes as

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

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

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

Consequently, the initial data are always normalized, and the unitary time-evolution guarantees

$$\forall t \in \mathbb{R} : \ \|\psi(t)\|_{L^2} = 1.$$

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

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

More generally, one associates with smooth functions $a : \mathbb{R}^{2d} \to \mathbb{C}$ on classical phase space a Weyl quantized operator $\mathrm{op}(a)$ acting in $L^2(\mathbb{R}^d)$ by the appropriate interpretation of the definition

$$\mathrm{op}(a)\psi(q) = (2\pi\varepsilon)^{-d}\int_{\mathbb{R}^{2d}} a(\tfrac{1}{2}(q+x),p)\mathrm{e}^{ip\cdot(q-x)/\varepsilon}\psi(x)\,\mathrm{d}x\,\mathrm{d}p.$$

The corresponding expectation values

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

then specialize to components of the position and momentum expectation by choosing $a(q,p) = q_j$ and $a(q,p) = p_j$, respectively.

In the semiclassical regime with small parameter $0 < \varepsilon \ll 1$, the direct approach to quadratic quantities is advantageous, since their dynamics are less oscillatory than those of the wave function itself. The celebrated Egorov theorem provides the following approximation (Theorem IV.10 in [24]). Let

$$\dot{q} = p, \qquad \dot{p} = -\nabla_q V(q)$$

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

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

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. In particular, the Egorov theorem provides an exact description for the dynamics of harmonic oscillator, whose potential is a quadratic function. 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 $W(\psi) : \mathbb{R}^{2d} \to \mathbb{R}$, which is a function on classical phase space. The definition

$$W(\psi)(q,p) = (2\pi)^{-d} \int_{\mathbb{R}^d} \mathrm{e}^{ix\cdot p}\, \psi(q - \tfrac{\varepsilon}{2}x)\overline{\psi}(q + \tfrac{\varepsilon}{2}x)\, \mathrm{d}x \tag{1}$$

grants

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

for all square integrable functions ψ and smooth symbols a, where all integrals have to be interpreted carefully, of course. The Wigner function has first been proposed by Wigner, further developed by Moyal, and reintroduced in the context of signal analysis by Ville [31, 19, 29]. Its main properties will be briefly discussed in §1. From the Wigner point of view, the Egorov theorem rephrases as

$$W(\psi(t)) = W(\psi_0) \circ \Phi^t + O(\varepsilon^2), \tag{2}$$

where the relation has to be understood 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 a family of phase space points $(q_1,p_1), \ldots, (q_N,p_N)$.

Classical transport. The phase space 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. One regards the values of the initial Wigner function $W(\psi_0)$ in $(q_1,p_1), \ldots (q_N,p_N)$ as an approximation of the values of the Wigner function $W(\psi(t))$ in the points $\Phi^t(q_1,p_1), \ldots, \Phi^t(q_N,p_N)$ and computes expectation values.

It is our aim here to contribute to the initial sampling step. Hence, we pursue the following two main objectives. First, the generation of phase space points according to the Wigner function $W(\psi)$ of a typical square integrable function ψ. Second, the computation of the Fourier integral in equation (1), which defines the value of the Wigner function $W(\psi)$ in a given phase space point (q,p). The first task is related to the problem of approximation, while the second one concerns numerical quadrature.

The wave functions ψ we have in mind are laser-excited eigenstates of semiclassical Schrödinger operators

$$-\tfrac{\varepsilon^2}{2}\Delta_q + \widetilde{V}(q),$$

as they arise in the context of Born-Oppenheimer approximation. The configuration vector q contains the molecule's nuclear coordinates, the semiclassical parameter ε is of the same size as the inverse of the square root of the average nuclear mass, and the potential $\widetilde{V}$ gives the nth eigenvalue of the electronic Schrödinger equation for a given nuclear configuration q. Hence, we consider wave functions with the following properties: exponential decay, high order differentiability, microlocalization around a few phase space points $c_0, \ldots, c_s \in \mathbb{R}^{2d}$.

The current chemical literature mostly considers Gaussian wave packets centered in a single phase space point $z_0 = (q_0, p_0)$, which are of the form

$$g_{z_0}(q) = (\pi\varepsilon)^{-d/4} \exp\bigl(-\tfrac{1}{2\varepsilon}|q-q_0|^2 + \tfrac{i}{\varepsilon} p_0 \cdot (q-q_0)\bigr)\,.$$

Their Wigner function is a Gaussian

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

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

$$f(p-p_0) = \int_{\mathbb{R}^d} \mathrm{e}^{ix\cdot(p-p_0)} \mathrm{e}^{-\frac{\varepsilon}{4}|x|^2}\,\mathrm{d}x$$

numerically, whose relative condition number

$$\kappa_f(p-p_0) = \frac{|p-p_0|}{|f(p-p_0)|}|\nabla f(p-p_0)| = \tfrac{2}{\varepsilon}|p-p_0|^2$$

reflects the oscillatory behavior of the integrand for large distances $|p-p_0|$. Also the approximation problem suffers from oscillations as soon as the Wigner function localizes 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 the dangerous cross term [9]

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

which localizes around the arithmetic mean $z_+ = (z_1 + z_2)/2$ and oscillates with a frequency proportional to $z_- = z_1 - z_2$.

To summarize, one faces a highly oscillatory problem of approximation and numerical quadrature in a high dimensional setting. Our aim is the systematic exploration of how a Monte Carlo approach can deal with the situation. For the approximation, we propose an adaption of the Metropolis algorithm, which jumps between predefined phase space regions and accounts for the negative values of the Wigner function. The quadrature uses importance sampling with a Gaussian proposal distribution and is complemented by a convergence test for the oscillatory regime.

For the numerical validation we consider a variant of the above mentioned particle method, which belongs to the family of surface hopping algorithms. These algorithms are widely used for the simulation of non-adiabatic quantum dynamics as they are generated by Schrödinger systems with conical intersections. The Schrödinger system

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

with real symmetric potential matrix

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

has a conical intersection, if the eigenvalues $\lambda^+(q)$ and $\lambda^-(q)$ of $V(q)$ coincide on a smooth submanifold of codimension two of the configuration space. 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 $\mathbb{R}^d \to \mathbb{R}$, $q \mapsto \lambda^\pm(q)$ look like two cones touching each other in their end points. Conical intersections are ubiquitous for the description of radiationless decay and isomerization processes of polyatomic molecules [2]. The eigenvalues $\lambda^\pm(q)$ are also eigenvalues of an electronic Schrödinger operator, which parametrically depends on the nuclear position q. Their intersection violates the adiabatic Born-Oppenheimer approximation 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

$$\psi^-(0) = 0 \qquad \& \qquad \exists t : \psi^+(t) = O(1),\ \varepsilon \to 0.$$

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 $(q_1^\pm, p_1^\pm), \ldots, (q_{N^\pm}^\pm, p_{N^\pm}^\pm)$ with associated real-valued weights $w_1^\pm, \ldots, w_{N^\pm}^\pm$, 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

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

where $\mathrm{d}v(q)$ 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 $(q_1^\pm, p_1^\pm), \ldots, (q_{M^\pm}^\pm, p_{M^\pm}^\pm)$ and weights $w_1^\pm, \ldots, w_{M^\pm}^\pm$, which approximate the values of the Wigner function $W(\psi^\pm(t))$ in these points. One computes the final expectation values.

The chemical physics' literature contains an overwhelming variety of surface hopping algorithms, which all differ in the way the non-adiabatic transitions are performed. The method considered here is called *single switch surface hopping*, since its constitutive branching condition allows for non-adiabatic switches just at minimal surface gaps along trajectories, whereas most of the established algorithms have random jumps at every time step of the discretization [6]. Moreover, the single switch approach is the only way of surface hopping, which has been derived from a rigorous mathematical analysis of Schrödinger systems with generic crossings [16, 14]. The main observables for the evaluation of the initial sampling are the energy level populations $P^\pm(t) = \|\psi^\pm(t)\|_{L^2}^2$, which give the probability of the wave function to be in the plus or minus eigenspace at time t. In terms of Wigner functions they express as the phase space integral

$$P^\pm(t) = \int_{\mathbb{R}^{2d}} W(\psi^\pm(t))(q,p)\,\mathrm{d}q\,\mathrm{d}p.$$

If $W^\pm(t, q, p)$ denotes the value of the phase space functions at time t generated by the single switch method, then the generalization of the Egorov theorem guarantees

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

where the relation holds in a weak sense (Theorem 2.2 in [14]) . However, all the numerical experiments so far have even shown a convergence rate of order $\sqrt{\varepsilon}$, see [15, 14] and §5.3 later on.

Our article is organized as follows. In section §1, basic properties of Wigner functions are discussed. §2 contains the detailed set up for the numerical experiments, while §3 presents the Monte Carlo methods for the approximation and the quadrature problem at hand. §4, §5, and §6 validate the proposed method for initial wave functions, which are a single Gaussian wave packet, a superposition of two Gaussian wave packets, and an excited harmonic oscillator state, respectively. Then, we offer an assessment of the obtained results in the final section.

1 Wigner functions

Considering dimension $d = 1$, the phase space $\mathbb{R}^2$ of classical and quantum mechanics can also be thought of as the time-frequency plane of signal analysis. In this context the Wigner function has been called the musical score [1]. Since expositions of its main properties have been given many times, see for example chapter 1.8 in [3] or chapter 4.3 in [7], we will only focus on those, which are relevant for its intended Monte Carlo sampling to start an asymptotic particle method.

1.1 Basic properties

By the Fourier transform of a square integrable function $\psi \in L^2(\mathbb{R}^d)$ we always mean the ε-scaled Fourier transform

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

Then, the Wigner function

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

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

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

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

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

Let $a : \mathbb{R}^{2d} \to \mathbb{C}$ be a smooth function on phase space and $\mathrm{op}(a)$ the associated Weyl quantized pseudodifferential operator,

$$\mathrm{op}(a)\psi(q) = (2\pi\varepsilon)^{-d} \int_{\mathbb{R}^{2d}} a(\tfrac{1}{2}(q+x), p)\, \mathrm{e}^{ip\cdot(q-x)/\varepsilon}\, \psi(x)\, \mathrm{d}x\, \mathrm{d}p.$$

Then,

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

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

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

and consequently

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

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

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

The interpretation of the Wigner function as a phase space density has the defect, that it might attain negatives 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 because of the sharp Gårding inequality (Chapter 2.10 in [17]). For smooth non-negative functions $a \geq 0$ there exists a positive constant $C = C(a) > 0$ depending on derivative bounds of a such that

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

1.2 Husimi functions

An alternative quadratic phase space representation of square integrable functions $\psi \in L^2(\mathbb{R}^d)$ is the Husimi function [11], which can be defined as a suitably scaled Gauß transform of the Wigner function,

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

A few lines of computation yield the non-negativity of the Husimi function, which can be expressed as the modulus squared of the FBI transform

$$H(\psi)(q,p) = |T(\psi)(q,p)|^2$$

with

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

The FBI transform is the inner product of the wave function with a Gaussian wave packet. The initials stand for Fourier, Bros and Iagolnitzer [12]. Averages of the Husimi and the Wigner function are rather close in the following asymptotic sense. For smooth phase space functions a one obtains

$$\begin{aligned}
&\int_{\mathbb{R}^{2d}} H(\psi)(q,p)\, a(q,p)\, \mathrm{d}q\, \mathrm{d}p \\
&\quad = (\varepsilon\pi)^{-d} \int_{\mathbb{R}^{4d}} W(\psi)(x,\xi)\, \mathrm{e}^{-(|q-x|^2+|p-\xi|^2)/\varepsilon}\, a(q,p)\, \mathrm{d}x\, \mathrm{d}\xi\, \mathrm{d}q\, \mathrm{d}p \\
&\quad = \int_{\mathbb{R}^{2d}} W(\psi)(x,\xi)\, a(x,\xi)\, \mathrm{d}x\, \mathrm{d}\xi + O(\varepsilon).
\end{aligned}$$

The above relation is due to the properties of the phase space Gaussian

$$G(q,p) = (\varepsilon\pi)^{-d}\, \mathrm{e}^{-(|q|^2+|p|^2)/\varepsilon},$$

whose integral is one, while its mean is zero, and its variance is $\varepsilon/2$. A second order Taylor approximation of the function a then gives for the convolution

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

where the error depends on second order derivatives of a. The Husimi function of the Gaussian wave packet g_{z_0} is the Gaussian

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

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

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

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 $H(g_{z_2})$. For the Husimi function, the Egorov theorem holds with a remainder of order ε, which is worse than the error of order ε^2 valid for Wigner functions. Indeed, if $\psi(t)$ solves the scalar Schrödinger equation

$$i\varepsilon\partial_t\psi = P\psi, \qquad \psi(0) = \psi_0,$$

and if Φ^t denotes the associated classical flow, then

$$H(\psi(t)) = H(\psi_0) \circ \Phi^t + O(\varepsilon) \tag{3}$$

holds in a weak sense. The remainder term of order ε is sharp, since

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

where the estimate of order ε^2 stems from the classical transport of Wigner functions and the order ε term from the relation

$$\left((a*G)\circ\Phi^{-t}\right)(q,p) = a(\Phi^{-t}(q,p)) + O(\varepsilon) = \left((a\circ\Phi^{-t})*G\right)(q,p) + O(\varepsilon).$$

Hence, the error in the Egorov theorem (3) depends on derivatives bounds for a and Φ^t, which are of order greater or equal than two.

1.3 Ambiguity functions

The Wigner function is also in close relation to another quadratic phase space representation, the ambiguity function

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

whose connection to radar theory is discussed in chapter 1.4 of [3] or chapter 4.2 of [7], for example. Indeed, Fourier transformation gives

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

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

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

The analogous expression holds on the Fourier level,

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

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

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

This estimate on the support of the ambiguity function $A(\psi) = \mathcal{F}W(\psi)$ agrees with the previous observation, that the Wigner function of the two superposed Gaussian wave packets with phase space centers z_1 and z_2 has oscillations of size $|z_1 - z_2|$.

2 Numerical set up

In this section, we present the general setup of our numerical experiments and give the results of the surface hopping algorithm with grid based sampling of the initial Wigner function. We consider the two-dimensional isotropic Schrödinger system

$$i\varepsilon\partial_t\psi(t,q) = -\tfrac{\varepsilon^2}{2}\Delta_q\psi(t,q) + \begin{pmatrix} q_1 & q_2 \\ q_2 & -q_1 \end{pmatrix}\psi(t,q), \qquad \psi(0,q) = \psi_0(q) \tag{4}$$

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

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

intersect in the point $q = 0$, which constitutes a codimension two submanifold of $\mathbb{R}^2$. Up to a quadratic diagonal term, our Hamiltonian is the linear $E \otimes e$ Jahn-Teller Hamiltonian, which models a large class of molecular systems including a triatomic molecule's displacement from the equilateral triangle configuration, see for example chapter 10 in [2]. The default choice for the semiclassical parameter is

$$\varepsilon = 0.01,$$

which corresponds to a real life's molecular system, where values between 0.001 and 0.01 have to be expected.

2.1 Initial data

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

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

Such initial data are typically chosen for the simulation of quantum molecular dynamics after excitation of the molecule by light or a laser-pulse. We have considered three different scalar functions ψ_0^+,

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

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

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

while

$$e(q) = 2\varepsilon^{-3/2}\pi^{-1/2}(q_1 - q_{0,1})(q_2 - q_{0,2})\, \mathrm{e}^{-\frac{1}{2\varepsilon}|q-q_0|^2}, \qquad q_0 = (5\sqrt{\varepsilon}, 0.5\sqrt{\varepsilon})$$

denotes one of the three excited states of the shifted two-dimensional harmonic oscillator with eigenvalue 3ε,

$$\left(-\tfrac{\varepsilon^2}{2}\Delta_q + \tfrac{1}{2}|q - q_0|^2 \right) e(q) = 3\varepsilon\, e(q).$$

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. The eigenvectors of a potential with conical crossings are discontinuous at the crossing points, and smoothness away from the crossing is only possible, if they are chosen complex-valued. We have considered the two cases

$$\chi^+(q) \in \left\{ \widetilde{\chi}(q), \mathrm{e}^{\frac{i}{2}\vartheta_q}\widetilde{\chi}(q) \right\}, \qquad \widetilde{\chi}(q) = \left(\cos(\tfrac{1}{2}\vartheta_q), \sin(\tfrac{1}{2}\vartheta_q) \right)^T,$$

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 $\widetilde{\chi}(q)$ across the left half axis $\{q \in \mathbb{R}^2 \mid q_1 \le 0,\, q_2 = 0\}$. Since the overlap of the single Gaussian wave packet and the excited oscillator state with the left half axis is negligible, we have chosen the real-valued eigenvector for them. The complex eigenvector is considered for the superposition. The time interval is set to

$$[t_\mathrm{i}, t_\mathrm{f}] = [0, 10\sqrt{\varepsilon}] \quad \text{or} \quad [t_\mathrm{i}, t_\mathrm{f}] = [0, 20\sqrt{\varepsilon}],$$

for the Gaussian wave packets and the excited state, respectively. It allows the solution of the Schrödinger equation to pass the crossing point once and to generate leading order non-adiabatic transitions to the eigenspace associated with the eigenvalue $\lambda^-(q)$.

2.2 Grid based surface hopping

Since the considered initial data are only associated to the eigenvalue $\lambda^+(q)$, the first step of the single switch algorithm requires the sampling of the Wigner

function $W(\psi_0^+)$ to obtain a family of phase space points $(q_k, p_k)_{k=1}^n$. Using the fact that

$$\int_{\mathbb{R}^4} W(\psi_0^+)(q,p)\,dq\,dp = 1,$$

a number of points $(q_k, p_k)_{k=1}^n$ is selected such that

$$\sum_{k=1}^{n} W(\psi_0^+)(q_k,p_k)\,\omega(q_k,p_k) = 1 - \mathrm{tol}_W$$

for a given tolerance tol_W. Here, ω denotes the weight of (q_k, p_k). For points distributed along a grid, ω is equal to the corresponding volume elements. If the location of the grid points does not depend on the value of $W(\psi_0^+)$, many points may be located in regions where the Wigner function nearly vanishes. In the following, we will demonstrate the need for an adaptive approach even in a low dimensional situation.

2.2.1 Single Gaussian wave packet

As in our previous work [15], the sampling domain $[z_0 - 5\sqrt{\varepsilon}, z_0 + 5\sqrt{\varepsilon}]$ is discretized by a uniform 16^4-grid, while the sampling tolerance is $\mathrm{tol}_W = 0.001$. We first sample from the marginal densities, which requires the evaluation of $|\psi_0^+|^2$ and $|\mathcal{F}\psi_0^+|^2$. The tolerances for position and momentum, $\mathrm{tol}_p = \mathrm{tol}_m = \mathrm{tol}_W/1000$, generate $108 \cdot 112 = 12096$ phase space points, at which the Wigner function must be evaluated. Thereof, 2188 points, that is 18 percent are subsequently selected to enter the hopping algorithm, which computes for the final level populations the values $P^+(t_\mathrm{f}) \approx 0.441$ and $P^-(t_\mathrm{f}) \approx 0.559$. As a reference solution of the Schrödinger system (4) we consider the outcome of a numerically converged pseudospectral Strang splitting scheme, see appendix A. For the level populations it computes the values 0.422 and 0.578. Hence, the surface hopping result differs by 0.02, which is well below the theoretically expected accuracy.

2.2.2 Superposition of Gaussian wave packets

The sampling domain $[z_2 - 5\sqrt{\varepsilon}, z_1 + 5\sqrt{\varepsilon}]$ is discretized by uniform m^4-grids with $m \in \mathbb{N}$. The sampling tolerance is $\mathrm{tol}_W = 0.001$. Due to the disconnected support of the position and momentum densities, a sampling of the marginals is not advantageous any more, since the Wigner function is supported in the convex hull of $\mathrm{supp}(\psi_0^+) \times \mathrm{supp}(\mathcal{F}\psi_0^+)$. If $m = 32$, then 5024 points are finally selected, which amounts to 0.48 percent of the grid. The minimal m for reaching the tolerance tol_W is 27, which results in 3089 points or a rate of 0.58 percent. The surface hopping starting from the 32^4-grid gives level populations of 0.405 and 0.594, respectively, in contrast to the values 0.436 and 0.564 obtained from the reference solution. Hence, the error is 0.03.

2.2.3 Excited harmonic oscillator state

The excited oscillator state has the position and momentum expectation value $(5\sqrt{\varepsilon}, 0.5\sqrt{\varepsilon})$ and $(0,0)$, respectively. Therefore, the sampling domain is set to $[z_3 - 5\sqrt{\varepsilon}, z_3 + 5\sqrt{\varepsilon}]$ with $z_3 = (5\sqrt{\varepsilon}, 0.5\sqrt{\varepsilon}, 0, 0)$. With 16 grid points per direction and the tolerance $\mathrm{tol}_W = 0.001$, we end up with 5068 sampling points,

that is a rate of 7.73 percent. The level populations amount to 0.616 and 0.383 on the upper and lower level, respectively, in contrast to the values 0.571 and 0.429 from the Strang splitting scheme. The minimum number of grid points per direction for achieving the desired tolerance tol_W is 12, which results in 1136 sampling points, that is a rate of 5.48 percent, and level populations of 0.604 and 0.395. Hence, the two different grids give errors of roughly 0.04.

3 Monte Carlo sampling

We propose Monte Carlo sampling techniques for the generation of approximation points for the Wigner function. Furthermore, for the evaluation of the Wigner function at these points, Monte Carlo quadrature methods will be used as well. The reason is the following: deterministic algorithms for approximation and integration problems imply computational costs, which increase exponentially with the dimension d of the problem. Randomized approaches like Markov chain Monte Carlo (MCMC) can break this curse of dimensionality [23, 26, 20].

3.1 Metropolis Monte Carlo

MCMC methods construct a random walk through the region in sampling space where a non-negative function W is non-negligible. In this random walk, a trial move is rejected if W becomes too small and is accepted otherwise. The rule for this decision must satisfy the constraint that the probability of finding the system in a point (q, p) is proportional to $W(q, p)$.

3.1.1 Standard approach

The Metropolis-Hastings algorithm [18, 8, 22, 4] is one of the most popular sampling schemes. Select a point $(q_{\mathrm{old}}, p_{\mathrm{old}})$ in sampling space and calculate $W_{\mathrm{old}} = W(q_{\mathrm{old}}, p_{\mathrm{old}})$. Then start the following iteration.

1. Proposition step: Give the point a random displacement,

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

   and calculate $W_{\mathrm{new}} = W(q_{\mathrm{new}}, p_{\mathrm{new}})$.

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

$$r < W_{\mathrm{new}}/W_{\mathrm{old}}$$

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

In our examples, the random displacement will be chosen from the normal distribution which corresponds to a symmetric proposal density as in the original work of Metropolis et al. [18]. Therefore, we will always speak of Metropolis Monte Carlo when referring to this algorithm. The Metropolis points $(q_k, p_k)_{k=1}^N$ form a Markov chain, which has W as equilibrium distribution. If the chain is

uniformly ergodic, then the central limit theorem holds [28], and the empirical means $\frac{1}{N}\sum_{k=1}^{N} a(q_k, p_k)$ approximate expectation values

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

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

$$\lim_{N\to\infty} \mathbb{P}\left(\left| \frac{1}{N} \sum_{k=1}^{N} a(q_k, p_k) - \langle a \rangle_W \right| \leq \frac{c\gamma_a}{\sqrt{N}} \right) = \Theta(c) \tag{5}$$

converges with

$$\Theta(c) = \frac{1}{\sqrt{2\pi}} \int_{-c}^{c} \exp(-t^2/2)\, dt = \mathrm{erf}(c/\sqrt{2}).$$

In particular, $\Theta(2.2414) = 0.975$ corresponds to a probability of 2.5 percent that the difference between the empirical mean and the expectation value is larger than $2.2414\gamma_a/\sqrt{N}$.

3.1.2 Sampling from a real-valued function

The Metropolis Monte Carlo algorithm has been developed for the sampling of probability distributions. Taking into account different signs of the Wigner function, we do not sample from W but from $|W|$. The acceptance criterion changes to

$$r < |W_{\mathrm{new}}|/|W_{\mathrm{old}}|.$$

Afterwards, in the summation for the integral value the algebraic sign is reassigned.

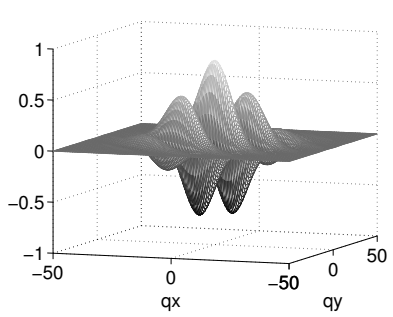
3.1.3 Disconnected support

If W decomposes into several peaks with virtually disjoint support, as it is the case for $W = W(g_{z_1} + g_{z_2})$, then the Metropolis chain cannot switch between the peaks, because intermediate points with small value of W are rejected, and the chain is not ergodic. Then one must generate a proposition step jumping to another peak from time to time. For the approximation of initial Wigner functions, we assume a priori given phase space centers $\{c_i\}_{i=1}^{s}$, which define regions on which W is non-negligible and has a local envelope of comparable variance. Given a jump rate $r_{\mathrm{jump}} \in [0,1]$ and the current sampling region $i \in \{1,\ldots,s\}$, the proposal step is modified as follows [30]:

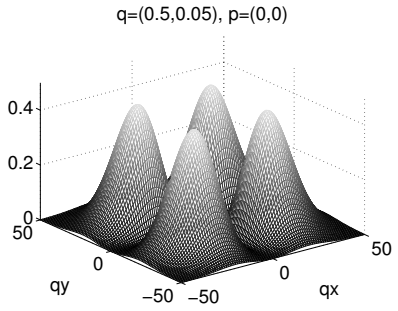
1.* Generate a random number r from the uniform distribution in the interval $[0,1]$. If $r > r_{\mathrm{jump}}$, generate a new point by random displacement. Else, perform a jump step: Choose uniformly one of the other centers $j \in \{1,\ldots,s\}$, $j \neq i$, and generate a new point via

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

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 is still symmetric.



(a) Integrand for Gaussian wave packet.



(b) Integrand for excited oscillator state.

Figure 1: Typical integrands in the computation of the Wigner function by Monte Carlo quadrature. The left and right plot show the integrands for the single Gaussian wave packet at (q_0, p) with $p = p_0 + (0.3, 0.3)$ and for the excited oscillator state at (q_0, p_0), respectively.

3.2 Importance Sampling

For evaluating the Wigner function, one has to solve the d-dimensional Fourier integral $W(\psi)(q,p) = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x$ with

$$f(x) = (2\pi)^{-d}\,\mathrm{e}^{ix\cdot p}\,\psi(q - \tfrac{\varepsilon}{2}x)\,\overline{\psi}(q + \tfrac{\varepsilon}{2}x).$$

The real part of two typical integrands $f(x)$ is plotted in Figure 1. With conventional quadrature the integrand is evaluated on a grid, which does not only require a very fine resolution, but also leads to the curse of dimensionality. Most of the computing time is spent on points, where the integrand is negligible. To sample many points in regions where the integrand is large and few elsewhere, is the basic idea behind importance sampling.

3.2.1 Standard approach

One rewrites the integral in the form

$$W(\psi)(q,p) = \int_{\mathbb{R}^d} \frac{f(x)}{w(x)}\,w(x)\,\mathrm{d}x,$$

where $w(x) \geq 0$ is a positive normalized weight function, $\int_{\mathbb{R}^d} w(x)\,\mathrm{d}x = 1$. If one generates sampling points according to the function $w(x)$, then the integral can be approximated by

$$W(\psi)(q,p) \approx I := \frac{1}{L}\sum_{k=1}^{L} \frac{f(x_k)}{w(x_k)}. \tag{6}$$

If the sampling points are independent and identically distributed or a uniformly ergodic Markov chain, then the central limit theorem guarantees convergence of the empirical mean as $L \to \infty$. The quality of the approximation crucially depends on the choice of the weight function w. If the sampling points are

independent and identically distributed, then I has mean $\mu = \int_{\mathbb{R}^d} f(x)\,\mathrm{d}x$ and variance

$$\left\langle (I-\mu)^2 \right\rangle_w = \frac{1}{L} \int_{\mathbb{R}^d} \left(\tfrac{f(x)}{w(x)} - \mu \right)^2 w(x)\,\mathrm{d}x.$$

Hence, small variances are achieved for weight functions w, which closely resemble the integrand f. On the other hand, w should allow for an efficient sampling, and we have chosen a Gaussian function

$$w(x) = (2\pi)^{-d/2} \sigma_w^{-d} \exp\left(-\frac{|x-\mu|^2}{2\sigma_w^2} \right), \tag{7}$$

whose mean $\mu \in \mathbb{R}^d$ and standard deviation $\sigma_w > 0$ depend on the integrand under consideration.

Due to the possible oscillations of the integrand, convergence might be extremely slow, and we have used the following simple convergence test. Assume we have computed M different values $\{I_m\}_{m=1}^M$ of the integral determining $W(\psi)(q,p)$. Moreover, we assume that these values are normally distributed with mean $\bar{I}$ and variance σ_I^2,

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

We compute a 95-percent confidence interval according to $K_I = \pm z\,\sigma_I/\sqrt{M}$ with $z = 1.96$. The sampling is continued until

$$K_I < \bar{I}\sqrt{\varepsilon}. \tag{8}$$

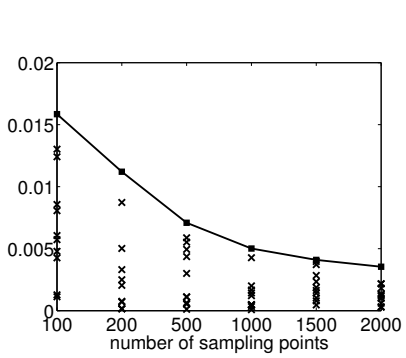
If the tolerance is not reached within a maximum number $N_{\max}$ of sampling steps, the point (q,p) will be rejected. There are more sophisticated convergence criteria as for example the Gelman-Rubin criterion [5], which is not only based on information between different sequences but also on within-sequence information. Its output is a number $R > 1$ which indicates how much the distributional estimate might improve if the simulations run longer. However, since the simple test yields satisfactory results, we have not explored this possibility.

3.2.2 Embedded importance sampling

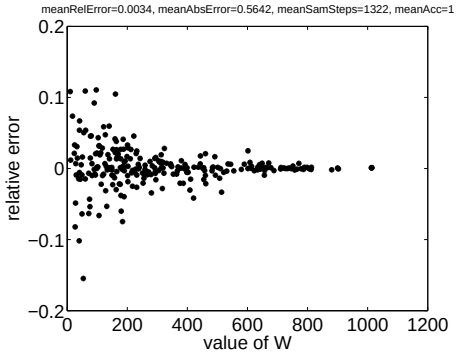
If the integrand is the product of two functions, $f(x) = g(x)w(x)$, where one of them has a known integral $I_w = \int_{\mathbb{R}^d} w(x)\,\mathrm{d}x$ and can be used as a weight function, then the approximation (6) simplifies to

$$W(\psi)(q,p) \approx \frac{1}{L} \frac{1}{I_w} \sum_{k=1}^{L} g(x_k),$$

where the points x_k are distributed according to $w(x)$. This approach uses more information on the integrand, and the numerical experiments show that it expectedly outperforms the standard approach.



(a) Convergence behavior of simple Monte Carlo sampling for the observable $q_1(0)$.



(b) Distribution of errors for the integration of the Wigner function by importance sampling.

Figure 2: Gaussian wave packet. The left plot illustrates the convergence behavior of the simple Monte Carlo sampling for the observable $q_1(t = 0)$. Single points show the absolute difference to the true value 0.5, while the solid curve indicates the values of $c\gamma_{q_1}/\sqrt{N}$ with $\gamma_{q_1} = \sqrt{\varepsilon/2} \approx 0.07$ and $c = 2.2414$. The right plot gives the distribution of relative errors with respect to the exact value of the Wigner function W for quadrature by importance sampling for $N = 500$ phase space sampling points.

4 Gaussian wave packet

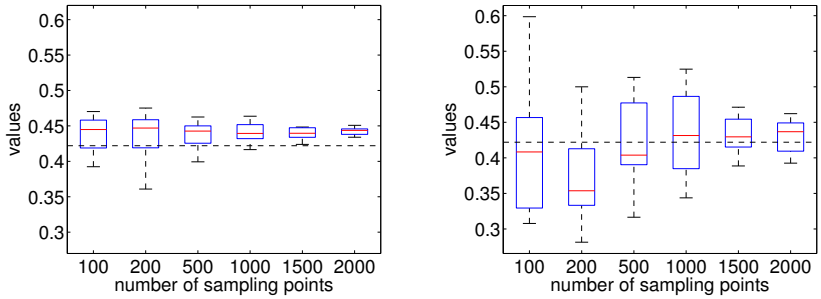
The Gaussian wave packet g_{z_0} is one of the simplest examples for an initial wave function. Since its Wigner function is also Gaussian, the initial phase space sampling can be simplified a lot compared to the proposed strategy. Nevertheless, ignoring the analytic knowledge on the explicit form of $W(g_{z_0})$, we also study the performance of the Metropolis approach.

4.1 Approximation

Since $W(g_{z_0})$ is a Gaussian, one can simply generate approximation points by sampling from a multi-dimensional normal distribution. This is much faster than Metropolis Monte Carlo because the summation of the displacement vector is omitted and the acceptance ratio is equal to 1. We refer to this method as simple Monte Carlo in the following and analyze its performance in combination with the surface hopping algorithm. Afterwards, we forget about the Gaussian shape and examine the choice of the displacement Δ in the Metropolis Monte Carlo method.

4.1.1 Simple Monte Carlo

We first sample the Gaussian Wigner function by generating points from a four-dimensional standard normal distribution and multiplying them with the standard deviation $\sigma = \sqrt{\varepsilon/2}$. Let us briefly verify the convergence rate of order $1/\sqrt{N}$ for the observable $a(q, p) = q_1$, as it is given by the central limit theorem for independent identically distributed random variables. In this case,



(a) Analytic evaluation of W. (b) Monte Carlo computation of W.

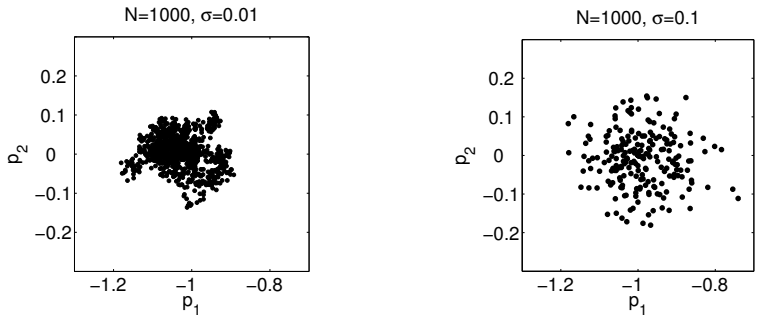
Figure 3: Gaussian wave packet. Statistics of $P^+(t_\mathrm{f})$ for simple Monte Carlo approximation of the initial distribution. For the left and right plot the values of the Wigner function are determined analytically and by importance sampling, respectively. The results are compared for different numbers of sampling points ($N = 100, \ldots, 2000$), where $m = 10$ runs of the single switch algorithm are evaluated for each N. The dashed horizontal line indicates the reference value 0.422 from Strang splitting.

the constant $\gamma_{q_1} = \sigma$ is the standard deviation of the random variables. Figure 2(a) plots the values of the difference $|\frac{1}{N}\sum_{k=1}^{N} q_{1,k} - 0.5|$ for ten repetitions of the sampling with a fixed number of points $N \in \{100, \ldots, 2000\}$ as well as the reference function $N \mapsto c\gamma_{q_1}/\sqrt{N}$ with $c = 2.2414$. Nearly all values are below this boundary, which should indeed be satisfied with probability $\Theta(c) = 0.975$ in the limit $N \to \infty$.

The next goal is to examine the results from the single switch algorithm for simple Monte Carlo sampling of the initial distribution. As accuracy criterion, we take the deviation of the final population $P^+(t_\mathrm{f})$ from the reference value 0.422 stemming from the Strang splitting scheme. The sampling of the initial distribution is performed with $N = 100, \ldots, 2000$ different numbers of sampling points. Then, for each fixed N there are $m = 10$ runs of the surface hopping algorithm. The results are illustrated as boxplots in Figure 3(a). The boxes have lines at the lower quartile, median, and upper quartile values. The dashed lines extending from each end of a box show the extent of the rest of the data. They extend out to the most extreme data value within 1.5 times the interquartile range of the sample. Data values beyond these lines are marked as outliers. The variances become smaller 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 only differ by two to three percent.

4.1.2 Metropolis Monte Carlo

Investigating sampling strategies applicable for arbitrary initial distributions, we repeat the experiments from the previous paragraph with Metropolis Monte Carlo sampling. We use the algorithm explained in §3.1 with normally distributed displacement $\Delta \sim \sigma_\mathrm{appr} N_4(0, 1)$. The starting point of the Markov



(a) $\sigma_{\text{appr}} = 0.01$ (b) $\sigma_{\text{appr}} = 0.1$

Figure 4: Gaussian wave packet. Distribution of momenta $p(0)$ resulting from Metropolis Monte Carlo approximation of the Wigner function with different values of the standard deviation σ_{appr} of the random displacement.

Table 1: Gaussian wave packet. Mean acceptance ratios over 10 runs of Metropolis Monte Carlo sampling with different values of the standard deviation σ_{appr} of the random displacement.

σ_{appr}	0.01	0.05	0.1	0.2
acceptance ratio	0.9059	0.5258	0.2211	0.0033

chains is always the center $z_0 = (q_0, p_0)$ of the Gaussian $W(g_{z_0})$. In the literature, see for example [25, 27], more elaborate proposal densities are discussed. However, in the following experiments the simple normal distribution yields quite large acceptance ratios.

The standard deviation σ_{appr} of the displacement should be comparable to the one of the distribution to be sampled, which is $\sigma = \sqrt{\varepsilon/2}$. We perform the algorithm with a fixed number of sampling points $N = 1000$ and different values of $\sigma_{\text{appr}} \in \{0.1\sqrt{\varepsilon}, 0.5\sqrt{\varepsilon}, \sqrt{\varepsilon}, 2\sqrt{\varepsilon}\}$. We repeat each run ten times and analyze the corresponding mean values. As expected, the mean acceptance ratio, that is the number of proposed points over the number of accepted points, increases with smaller displacement, see Table 1. Though a high acceptance ratio is desirable, small displacements prevent the Markov chain from quickly exploring the complete distribution. Indeed, Figure 4 illustrates that the distribution of $p(t = 0)$ is more dense for $\sigma_{\text{appr}} = 0.01$ than for $\sigma_{\text{appr}} = 0.1$. Hence, we use $\sigma_{\text{appr}} = 0.5\sqrt{\varepsilon}$ throughout the following experiments as a compromise between acceptance ratio and exploration of sampling space.

4.2 Integration

Standard importance sampling introduces an importance sampling function w and samples from this distribution. In general, this sampling is performed via Metropolis Monte Carlo. However, if w is Gaussian, simple Monte Carlo is faster. Further simplifications are possible if embedded importance sampling

can be applied. The embedded importance sampling function can be sampled via Metropolis Monte Carlo or, if it is Gaussian, via simple Monte Carlo. In the following, we compare these different approaches and use simple Monte Carlo sampling for the approximation of the Wigner function.

4.2.1 Standard importance sampling

Evaluation of the Wigner function in a point (q, p) requires solving the integral

$$\frac{1}{4\varepsilon\pi^3}\,\mathrm{e}^{-\frac{1}{\varepsilon}|q-q_0|^2}\int_{\mathbb{R}^2}\cos(x\cdot(p-p_0))\,\mathrm{e}^{-\frac{\varepsilon}{4}|x|^2}\,\mathrm{d}x. \tag{9}$$

There are oscillations with respect to x which are modulated by a Gaussian envelope. As the distance $|p-p_0|$ gets larger the oscillation frequency increases, which causes severe difficulties in Monte Carlo quadrature. Figure 1(a) shows the integrand for (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 test two strategies to circumvent these difficulties. The first method introduces a cut-off for p, meaning that during the sampling of the Wigner function all points with $|p-p_0| > \mathrm{tol}_p$ are rejected. For the Gaussian with variance $\sigma = \sqrt{\varepsilon/2}$, the cut-off tolerance is set to $\mathrm{tol}_p = 3\sqrt{\varepsilon/2} \approx 0.21$, which includes 99.9% of the probability density. However, this requires some a priori knowledge about the distribution function, which is not always accessible. The second method uses the convergence criterion (8).

We apply importance sampling with Gaussian importance function w as defined in (7) with $\mu = 0$ and $\sigma_w = \sqrt{2/\varepsilon}$, according to the variance of the Gaussian part of the integrand, and draw from w by simple Monte Carlo sampling. We set the number of chains $M = 5$ and the maximum number of steps $N_{\max} = 10^4$. A typical distribution of relative errors with respect to the exact value of the Wigner function is illustrated in Figure 2(b). The relative as well as the absolute error are centered around zero (mean relative error = 0.0034, mean absolute error = 0.5642), supporting the expectation that oscillations with subsequent small integral value pose the main difficulty. If the sampled phase points are propagated by the single switch algorithm, then the statistics for the upper level population at time $t = t_\mathrm{f}$ are worse than before, when the analytic values of the Wigner function have been used. Figure 3(b) shows the results for $N = 100, \ldots, 2000$ sampling points and $m = 10$ surface hopping runs for each fixed N. The variances are larger, and the mean values differs from the reference by a few percent only for $N \geq 500$.

Finally, we compare the cut-off and the multiple chain strategy by applying them to two points (q_0, p), whose momentum component is inside and outside the cut-off range, that is $p = p_0 + (0.1, 0.1)$ and $p = p_0 + (0.3, 0.3)$. The importance function w is sampled via Metropolis Monte Carlo with random displacement chosen from the normal distribution $N_2(0, \sigma^2)$ with $\sigma = \frac{1}{2}\sigma_w$. In the first case, the exact value is 137.12. The sampling converges after 4000 steps with mean $\bar{I} = 148.71$ and confidence bound $K_I = 14.5$. For the second point, the analytic value of the integral is $1.5 \cdot 10^{-5}$. We obtain $\bar{I} = -0.5$ with $K_I = 10.65$. The sampling does not converge within the maximum number of steps, and the second point is rejected according to both strategies.

4.2.2 Embedded importance sampling

In general, an analytic calculation of the Wigner function is impossible but the integrand might contain some fast decaying function whose integral is known. Using that

$$\frac{\varepsilon}{4\pi}\int_{\mathbb{R}^2} \mathrm{e}^{-\frac{\varepsilon}{4}|x|^2}\,\mathrm{d}x = 1,$$

the integral (9) can be evaluated via Monte Carlo quadrature as

$$W(g_{z_0})(q,p) \approx \frac{1}{L\varepsilon^2\pi}\,\mathrm{e}^{-\frac{1}{\varepsilon}|q-q_0|^2}\sum_{k=1}^{L}\cos(x_k\cdot(p-p_0)). \tag{10}$$

The sampling points $\{x_k\}_{k=1}^L$ are distributed according to $\exp(-\frac{\varepsilon}{4}|x|^2)$, which represents the 2-dimensional normal distribution $N_2(0, 2/\varepsilon)$.

Comparing accuracy and complexity of the four different ways of importance sampling, we calculate the Wigner function at the 2188 grid points of §2.2.1. The first method uses the importance function w from (7) with $\mu = 0$, $\sigma_w = \sqrt{2/\varepsilon}$ and samples it via Metropolis Monte Carlo with random displacement from the normal distribution $N_2(0,\sigma^2)$ with $\sigma = \frac{1}{2}\sigma_w$. The second method replaces Metropolis Monte Carlo by directly drawing from w. The third method applies formula (10) and generates the quadrature points by Metropolis Monte Carlo with random displacement from $N_2(0,\sigma^2)$. The fourth approach directly samples from $\exp(-\frac{\varepsilon}{4}|x|^2)$.

The results are listed in Table 2. The relative and absolute errors refer to the quadrature error in the computation of $W(g_{z_0})(q_k,p_k)$ and are mean values over all points $(q_k,p_k)_{k=1}^{2188}$. All the relative errors are around one percent, while the absolute errors improve for simple Monte Carlo. The sampling error is the deviation from the vector of expectation values in position and momentum z_0 measured in the supremum norm. It lies in the permille range. The error of the final population on the upper level $P^+(t_\mathrm{f})$ is computed from this single run of the surface hopping algorithm. It is roughly two percent. Time refers to the total amount of time to compute the Wigner function for all 2188 points. It reduces significantly for embedded importance sampling with simple Monte Carlo. The Monte Carlo steps give the number of quadrature points required to compute the Wigner function averaged over all grid points (q_k,p_k). The acceptance ratio is the mean value over all 2188 importance sampling runs. The mass ratio is the ratio of the sum of the Wigner function at phase space points, for which the quadrature achieves the convergence criterion (8), over the sum of the Wigner function at all grid points.

5 Superposition of Gaussian wave packets

For the superposition of Gaussian wave packets the Wigner function is no longer positive and has several peaks with disconnected support. In the following, we examine the performance of the strategies which have been proposed in §3.

5.1 Approximation

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

$$W(g_{z_j})(q,p) = (\pi\varepsilon)^{-d}\exp\bigl(-\tfrac{1}{\varepsilon}|(q,p)-z_j|^2\bigr), \qquad j=1,2$$

Table 2: Single Gaussian wave packet. Accuracy and complexity in the computation of the Wigner function via different importance sampling schemes.

	standard import. sampling		embedded import. sampling	
	Metropolis MC	Simple MC	Metropolis MC	Simple MC
relative error	0.012	0.009	0.013	0.009
absolute error	0.473	0.031	0.438	0.153
sampling error	0.0052	0.0004	0.0052	0.0011
$P^+(t_\mathrm{f})$ error	0.019	0.018	0.021	0.019
time (sec)	7940	6031	3349	1960
MC steps	3163	2351	3118	2370
acceptance ratio	0.76	1	0.76	1
mass ratio	0.88	0.94	0.86	0.95

plus an oscillatory cross term $c(q,p)$ localized around the middle point $z_+ = 0$,

$$c(q,p) = (\pi\varepsilon)^{-d} \mathrm{e}^{-|(q,p)|^2/\varepsilon} \cos\bigl(\tfrac{1}{\varepsilon}(q,p) \wedge z_-\bigr)\,,$$

which has a Gaussian envelope with the same variance and oscillates with a frequency proportional to the difference $z_- = (10\sqrt{\varepsilon}, \sqrt{\varepsilon}, -2, 0)$. Consequently, we choose the random displacement as before, namely from the normal distribution with standard deviation $\sigma_\mathrm{appr} = 0.5\sqrt{\varepsilon}$. Moreover, an elaborate integration by parts, see Theorem 7.7.1 in [10], gives a positive constant $C > 0$ such that for all smooth compactly supported functions $a : \mathbb{R}^{2d} \to \mathbb{C}$ and all $k \in \mathbb{N}_0$

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

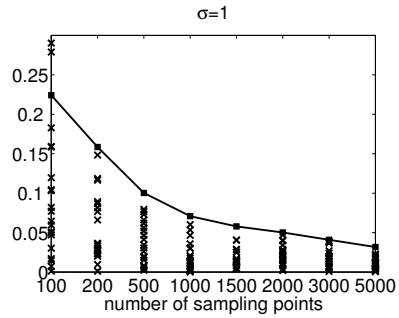
Thus, averages of the cross term are super-polynomially small with respect to the semiclassical parameter, which indicates that the cross term might be neglected without affecting the overall accuracy of the surface hopping algorithm.

During the sampling, jumps between the centers z_1 and z_2 of the two Gaussians and the middle point z_+ are performed, if a random number r uniformly distributed in $[0,1]$ is below the jump rate r_jump. As expected, the larger jump rates increase the overall acceptance ratio, see Table 3. However, the final population $P^+(t_\mathrm{f})$ of the surface hopping is rather insensitive to the rate. Figure 5(b) presents the statistics for three choices of $r_\mathrm{jump} \in \{0.2, 0.5, 0.8\}$. For the two smaller values $r_\mathrm{jump} = 0.2, 0.5$ there is more mixing within the components and consequently no outliers, while for $r_\mathrm{jump} = 0.8$ there are three of them. Hence, we have fixed $r_\mathrm{jump} = 0.2$ for the following experiments.

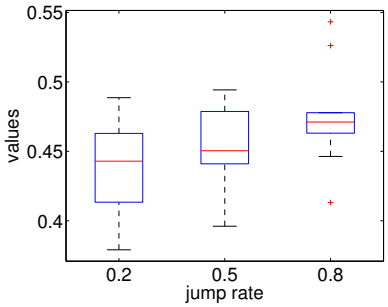
As for the single Gaussian wave packet before, one observes a convergence rate of order $1/\sqrt{N}$ for the approximation of observables via Monte Carlo sampling, see Figure 5(a). Examining the influence of the initial sampling on the accuracy of the surface hopping population $P^+(t_\mathrm{f})$, we have considered different numbers of sampling points $N = 100, \ldots, 3000$ and have performed the hopping algorithm $m = 20$ times for each fixed number of sampling points. The results are illustrated in Figure 6(a). With a few hundred sampling points the mean value already differs from the reference value 0.436 by a few percent.

Table 3: Superposition of Gaussian wave packets. The acceptance ratio increases with growing jump rate. The ratios are mean values over 10 runs with 2000 sampling points each.

r_{jump}	0.2	0.5	0.8
acceptance ratio	0.55	0.66	0.75

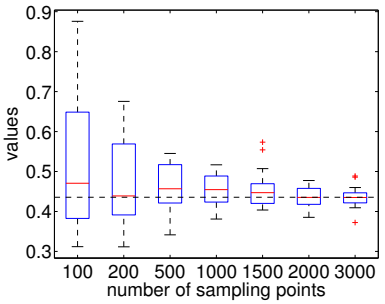


(a) Convergence behavior of Monte Carlo sampling for the observable $q_1(0)$.

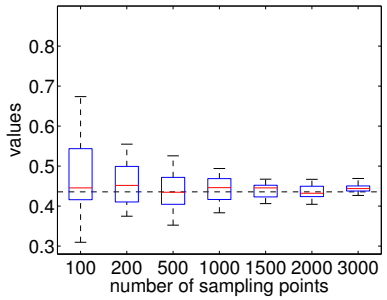


(b) Boxplots for the distribution of $P^+(t_f)$ around the exact value 0.436.

Figure 5: Superposition of Gaussian wave packets. The left plot illustrates the convergence behavior of the sampling for the observable $q_1(0)$. Single points show the absolute difference to the true value 0, while the solid curve indicates the values of $c\gamma_{q_1}/\sqrt{N}$ with $\gamma_{q_1} = 1$ and $c = 2.2414$. The right plot shows the final population of the upper level for three different values of the jump rate. For each jump rate, 10 runs with 2000 points each have been performed.

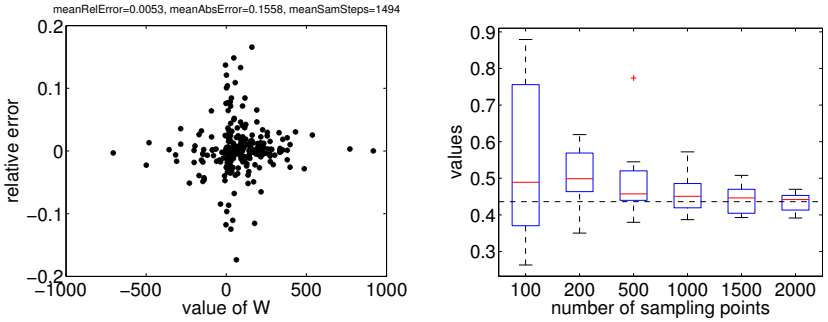


(a) Complete distribution.



(b) Distribution without oscillatory part.

Figure 6: Superposition of Gaussian wave packets. Statistics of $P^+(t_f)$ for Metropolis Monte Carlo sampling of the complete initial distribution (left hand side) and without the oscillatory middle peak (right hand side). The results are for different numbers of sampling points ($N = 100, \ldots, 3000$), where $m = 20$ runs of the single switch algorithm were evaluated for each N. The dashed horizontal line indicates the reference value 0.436 from Strang splitting.



(a) Distribution of errors for the integration of W by importance sampling.

(b) Boxplots for the distribution of $P^+(t_\mathrm{f})$ around the exact value 0.436.

Figure 7: Superposition of Gaussian wave packets. The left plot shows the distribution of relative errors w.r.t. the exact value of the Wigner function for 500 sampling points. The right plot shows the statistics of $P^+(t_\mathrm{f})$ for initial Monte Carlo sampling and evaluation of the Wigner function via importance sampling. The results are compared for different numbers of sampling points ($N = 100, \ldots, 2000$), where $m = 10$ runs of the single switch algorithm were evaluated for each N.

Furthermore, with more points the results improve 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. A reduction of variance is also achieved by ignoring the oscillatory middle peak of the initial distribution, see Figure 6(b). This supports the previous prediction that the cross term can be neglected because its contribution to integration errors is much more significant than its portion w.r.t. the phase space density.

5.2 Integration

Now, we also compute the Wigner function via importance sampling. The two integrals defining the Wigner functions $W(g_{z_1})$ and $W(g_{z_2})$ have already been dealt with in the previous section §4.2. The integrals occurring for the oscillatory cross term $c(q,p)$

$$\int_{\mathbb{R}^2} \cos(x \cdot (p - p_+))\, \mathrm{e}^{-\frac{1}{2}x\cdot q_-}\, \mathrm{e}^{-\frac{\varepsilon}{4}|x|^2}\, \mathrm{d}x$$

are only slightly different due to the additional term $\mathrm{e}^{-\frac{1}{2}x\cdot q_-}$. We apply standard importance sampling with a Gaussian proposal distribution w of mean $\mu = 0$ and standard deviation $\sigma_w = \sqrt{2/\varepsilon}$, where the quadrature points are directly drawn from w. For each phase space point (q,p) we use $M = 5$ different chains of maximal length $N_\mathrm{max} = 10^4$.

There are the same two regimes observed for the single Gaussian wave packet. If the distance of (q,p) to the middle center $z_+ = 0$ is sufficiently large, then the integrand is dominated by oscillations, and the point is rejected if the convergence criterion (8) is not met. If (q,p) is close enough to z_+, then the integrand is a regular function with bell-shaped envelope, and the quadrature achieves

Table 4: Superposition of Gaussian wave packets. Final level population for different values of ε, different phase space representations (Wigner or Husimi), and different sampling strategies (grid based or MCMC). The reference value is calculated with a highly resolved Strang splitting. The particle numbers from the grid based sampling are given in parentheses, while the Monte Carlo sampling always uses 5000 points.

ε	10^{-1}	10^{-2}	10^{-3}	$5 \cdot 10^{-4}$
1. reference value	0.37798	0.43565	0.52582	0.54363
2. Husimi (grid)	0.5095	0.5371	0.6037	0.6149
	(4754)	(6170)	(5548)	(5394)
3. Wigner (coarse grid)	0.3372	0.4054	0.5299	0.5079
	(5108)	(5024)	(2920)	(4811)
4. Wigner (fine grid)	0.3856	0.4397	0.5268	0.4613
	(31163)	(55799)	(6302)	(11743)
5. Husimi (MCMC)	0.5031	0.5379	0.6221	0.6396
6. Wigner (MCMC)	0.3963	0.4385	0.5692	0.5956
7. Wigner (MCMC, no middle part)	0.3876	0.4375	0.5404	0.5585

the expected accuracy. Indeed, for 500 phase points the quadrature produces a mean relative and mean absolute error with respect to the exact value of the Wigner function, which is of size 0.0053 and 0.1558, respectively. The mean chain length to achieve convergence is 1494, and the mean acceptance ratio for phase space points is 0.9. Figure 7(a) illustrates the corresponding distribution of relative errors, which again are largest for small values of the Wigner function.

Also the results of the single switch algorithm have the same tendencies as before. Sampling the initial Wigner function with different numbers of sampling points $N = 100, \ldots, 2000$ and computing the function value via importance sampling, Figure 7(b) shows larger variances than for the experiments with analytic function values, but also the expected variance reduction for a growing number of initial phase space points.

5.3 Approximation of the Husimi function

Now we explore the performance of the single switch algorithm, when the Wigner function of the initial wave function is replaced by the Husimi function. While the semiclassical parameter is set to $\varepsilon \in \{0.0005, 0.001, 0.01, 0.1\}$, the two different phase functions are sampled in a grid based way or by the previous Monte Carlo method. As for the Wigner function, the Metropolis Monte Carlo sampling of the Husimi function is rather insensitive to the variance of the random displacement, and we fix for both functions $\sigma_{\text{appr}} = 0.5\sqrt{\varepsilon}$.

Let us first consider the grid based results, see Table 4, lines two to four. For the Husimi function, the number of grid points and hence the number of particles is increased until the final population $P^+(t_\text{f})$ converges up to an error of roughly one percent. Such behavior already occurs for moderate particle numbers around $N = 5000$. For the Wigner function, the same grids do not yield convergence, and one has to refine significantly. For small semiclassical

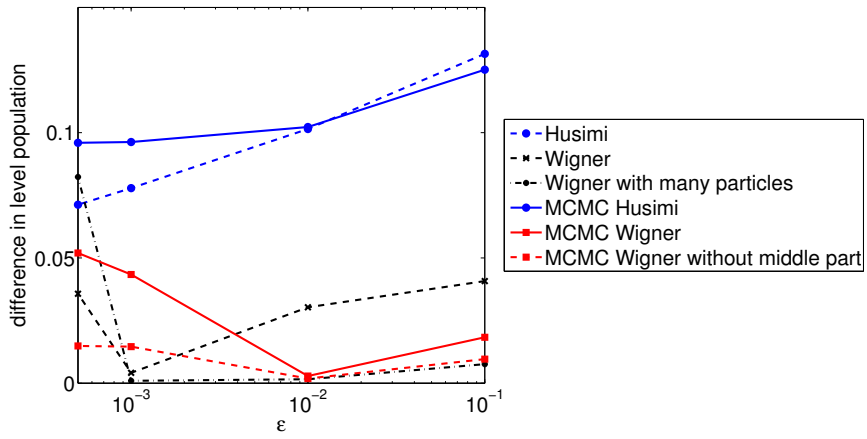


Figure 8: Superposition of Gaussian wave packets. The semilogarithmic plot shows the error in the final level population $P^+(t_\mathrm{f})$ for the Husimi and Wigner function with varying value of ε. The first three curves are based on grid based sampling, while the last three ones result from MCMC sampling.

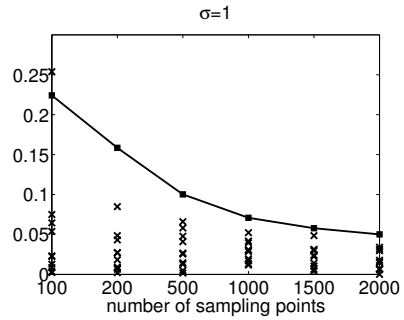
parameter $\varepsilon = 0.0005$ convergence is not obtained, since a reasonably sized uniform grid cannot resolve the rapid oscillations of the middle peak. However, the population errors for the converged sampling of the Wigner function are roughly three times smaller than those for the Husimi function, see also Figure 8. This difference can be explained by the better asymptotic properties of the Wigner function with respect to classical transport. As discussed in section §1.2, 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 bury this difference.

The Monte Carlo results all use $N = 5000$ sampling points and are the mean values of ten different runs, see Table 4, lines five to seven. The outcome for the Husimi function changes only slightly. The results are a bit better for larger semiclassical parameters and worse for small values. The Monte Carlo sampled Wigner function without oscillatory middle peak achieves the smallest error, which differs from those for the Husimi function by roughly a factor three.

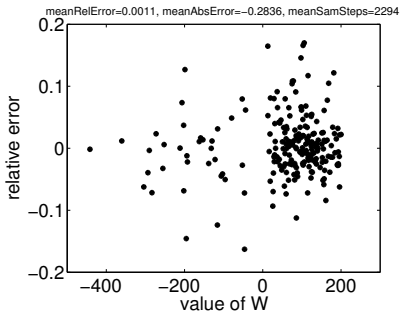
6 Excited harmonic oscillator state

The Wigner function of the excited harmonic oscillator state e is a Gaussian of variance $\varepsilon/2$ multiplied with an even quadratic polynomial attaining negative values for small momenta,

$$W(e)(q,p) = \frac{4\,\mathrm{e}^{-\frac{1}{\varepsilon}|(q-q_0,p)|^2}}{\varepsilon^4\pi^2} \tag{11}$$
$$\Big(\prod_{j=1,2}(q_j-q_{0,j})^2 + \sum_{(j,k)=(1,2),(2,1)}(q_j-q_{0,j})^2(p_k^2-\tfrac{\varepsilon}{2}) + \prod_{j=1,2}(p_j^2-\tfrac{\varepsilon}{2})\Big).$$



(a) Convergence behavior of Monte Carlo sampling for the observable $q_1(0)$.



(b) Distribution of errors for the integration of W by importance sampling.

Figure 9: Excited oscillator state. The left plot illustrates the convergence behavior of Monte Carlo sampling for the observable $q_1(0)$, when the analytic expression of the Wigner function W is used. Single points show the absolute difference to the true value 0.5, while the solid curve indicates the values of $c\gamma_{q_1}/\sqrt{N}$ with $\gamma_{q_1} = 1$ and $c = 2.2414$. The right plot shows the distribution of relative errors w.r.t. the exact value of W for 500 sampling points.

At first, we use the analytic expression for the Wigner function and sample from this distribution via Metropolis Monte Carlo with random displacement

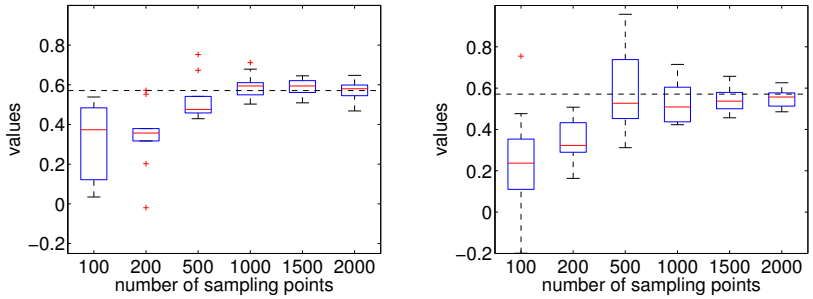
$$\Delta \sim N_4(0, \sigma_{\text{appr}}^2), \qquad \sigma_{\text{appr}} = 0.5\sqrt{\varepsilon}.$$

The error in the computation of the position expectation value in the first component $q_1(0)$ again decreases as $1/\sqrt{N}$, see Figure 9(a). The results for the final population on the upper level obtained by the surface hopping algorithm are illustrated in Figure 10(a). As in the previous examples, the error decreases down to a few percent, if the number of sampling points is large enough. Hence, the approximation is not affected by the negative values of the Wigner function.

The integrals for the computation of the Wigner function $W(e)(q,p)$ are

$$\int_{\mathbb{R}^2} \cos(x \cdot p) \left(q_1^2 - \tfrac{\varepsilon^2}{4} x_1^2\right) \left(q_2^2 - \tfrac{\varepsilon^2}{4} x_2^2\right) \mathrm{e}^{-\frac{\varepsilon}{4}|x|^2} \, \mathrm{d}x.$$

Their integrand differs from those for the Gaussian wave packets by a polynomial term. A typical integrand is illustrated in Figure 1(b). There are oscillations, but they are moderate as long as the momentum p is small and the magnitude of $W(e)(q,p)$ large enough. We apply embedded importance sampling with respect to $x \mapsto \exp(-\frac{\varepsilon}{4}|x|^2)$ and use simple Monte Carlo to sample from this proposal function. For each integral, there are $M = 5$ chains of maximal length $N_{\max} = 10^4$ stopped by the convergence criterion (8). Figure 9(b) shows the relative errors for 500 phase space points. As for the Gaussian wave packets, the errors are bounded by 0.2. The mean relative and absolute errors are 0.0011 and -0.2836, respectively. The mean chain length is 2294, hence somewhat larger than the length 1494 obtained in the Gaussian case. The mean acceptance ratio is about 0.7 and smaller than the previously observed 0.9. In clear contrast to the Gaussian wave packets, there is no direct relationship between the errors



(a) Analytic evaluation of W. (b) Monte Carlo computation of W.

Figure 10: Excited oscillator state. Statistics of $P^+(t_\mathrm{f})$ for Metropolis Monte Carlo sampling of the initial distribution with analytic evaluation and Monte Carlo computation of the Wigner function. The results are compared for different numbers of sampling points ($N = 100, \ldots, 2000$), where $m = 10$ runs of the single switch algorithm were evaluated for each N. The dashed horizontal line indicates the reference value 0.571 from Strang splitting.

and the magnitude of the Wigner function, since the polynomial factor of the integrand causes a modulation of the envelope function, which in turn prolongs the mixing time of the Markov chain and generates errors dominating those due to high frequency oscillations. However, as before, the results for the upper level population obtained by a subsequent surface hopping nicely converge for larger particle numbers; see Figure 10(b).

Conclusion

We have addressed the Monte Carlo sampling of Wigner functions, as it arises in the context of particle methods for Schrödinger equations in the semiclassical regime. The sampling poses an approximation and a quadrature problem. For both we have considered a Monte Carlo approach, motivated by the following observations. Quantum molecular dynamics is typically formulated on high-dimensional configuration spaces. The potential of Schrödinger equations for nuclear propagation is mostly determined with coarse resolution, since its value for each nuclear configuration is based on the solution of a high-dimensional electronic structure problem. Moreover, also particle methods for Wigner functions in the spirit of an Egorov theorem are asymptotic approximations with rather coarse accuracy as well.

For the approximation problem we have proposed and tested an adaption of the Metropolis Monte Carlo algorithm to real-valued functions with disconnected support. The algorithm requires a priori physical knowledge on the functional support of the Wigner function, in the sense that phase space points, which specify the different components, are used as an input. The method enforces random jumps between the regions, within which the same local proposal steps generate the chain. While the construction of jump proposal steps with high acceptance ratio is a challenging task in classical molecular dynamics

simulations, the similarities between the jump regions ensured the success of this approach in our examples. In the numerical experiments we have observed convergence of computed observables with an error of order $1/\sqrt{N}$, where N is the number of sampling points, as it is implied by the central limit theorem for ergodic Markov chains.

For the integration problem, which arises when evaluating the Wigner function, we have tested the performance of importance sampling with Gaussian weight function. For certain phase space points, the quadrature problem is ill-conditioned due to a highly oscillatory integrand. For the considered test cases, however, these points are characterized by small values of the Wigner function and can thus be neglected. The proposed simple convergence criterion, which is based on the variance between different sampling chains, is able to detect such points.

We have demonstrated in three sets of numerical experiments that the proposed Monte Carlo approach to the sampling of Wigner functions yields surface hopping results, which reach an accuracy comparable to the one obtained by a converged grid-based sampling with analytic evaluation of the Wigner function. We have also considered the Husimi function as an alternative phase space representation, since it is non-negative and less oscillatory. However, the subsequent surface hopping results are systematically less accurate than for the Wigner function, which can be explained by the different order of approximation error in the Egorov theorem for scalar Schrödinger equations.

The presented numerical experiments are low-dimensional, since we have aimed at a combination of initial Monte Carlo sampling with surface hopping and its validation against a reliable solution of the underlying Schrödinger systems. Such a comparison is naturally bound to a 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 Schrödinger system 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 [13] provides accurate reference solutions. The number of time steps is set to 5000 for all experiments. The length of the time interval allows the wave function to pass the crossing point once. The final time is $t_\mathrm{f} = 10\sqrt{\varepsilon}$ for the Gaussians and $t_\mathrm{f} = 20\sqrt{\varepsilon}$ for the excited harmonic oscillator state.

Table 5 contains the computational domains, the grid sizes, the final population $P^+(t_\mathrm{f})$ and the achieved accuracy. The accuracy of the solution refers to the difference $\|\psi(t_\mathrm{f}) - \psi_c(t_\mathrm{f})\|_{L^2}$ of the final reference solution $\psi(t_\mathrm{f})$ and a coarser solution $\psi_c(t_\mathrm{f})$, which is computed with fourth the number of grid points and half the number of time steps. In section §5.3, we varied the semiclassical parameter ε to compare 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

Table 5: Input parameters and results for the reference solution if $\varepsilon = 0.01$.

	domain	fine grid size	$P^+(t_\mathrm{f})$	accuracy
Gaussian wave packet	$[-2,2]\times[-1,1]$	1024×512	0.422	$4.7\cdot10^{-7}$
Superposition	$[-\frac{3}{2},\frac{3}{2}]\times[-\frac{3}{4},\frac{3}{4}]$	2048×1024	0.436	$1.9\cdot10^{-4}$
Excited oscillator	$[-4,4]\times[-2,2]$	2048×1024	0.571	$5.2\cdot10^{-4}$

Table 6: Input parameters and results for the reference solution in case of the superposition of two Gaussians in dependence on the semiclassical parameter ε.

ε	domain	fine grid size	$P^+(t_\mathrm{f})$	accuracy
10^{-1}	$[-10,10]\times[-5,5]$	2048×1024	0.378	$1.0\cdot10^{-3}$
10^{-2}	$[-\frac{3}{2},\frac{3}{2}]\times[-\frac{3}{4},\frac{3}{4}]$	2048×1024	0.436	$1.9\cdot10^{-4}$
10^{-3}	$[-1,1]\times[-\frac{1}{2},\frac{1}{2}]$	2048×1024	0.526	$1.6\cdot10^{-7}$
$5\cdot10^{-4}$	$[-\frac{1}{2},\frac{1}{2}]\times[-\frac{1}{4},\frac{1}{4}]$	2048×1024	0.544	$1.0\cdot10^{-7}$

of the wave function typically varied around three percent.

B Analytical Wigner transformation

For the cross term of the Wigner function $W(g_{z_1}+g_{z_2})$, one solves the integral

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

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

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

one has

$$\begin{aligned}W(g_{z_1},g_{z_2})(q,p) &= (2\pi)^{-d}(\pi\varepsilon)^{-d/2}\,\mathrm{e}^{-\frac{1}{2\varepsilon}|q-q_1|^2-\frac{1}{2\varepsilon}|q-q_2|^2}\\ &\mathrm{e}^{\frac{i}{\varepsilon}(p_1\cdot(q-q_1)-p_2\cdot(q-q_2))}\int_{\mathbb{R}^d}\mathrm{e}^{ix\cdot(p-p_+-\frac{i}{2}q_-)}\,\mathrm{e}^{-\frac{\varepsilon}{4}|x|^2}\,\mathrm{d}x.\end{aligned}\tag{12}$$

One uses the value of the Gaussian integral

$$\int_{\mathbb{R}}\mathrm{e}^{iyw}\mathrm{e}^{-\alpha y^2}\mathrm{d}y = \sqrt{\tfrac{\pi}{\alpha}}\,\mathrm{e}^{-w^2/(4\alpha)},\qquad w\in\mathbb{C},\ \alpha>0$$

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

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

one gets

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

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

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

and

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

Finally we compute the Wigner function for the two-dimensional excited oscillator state $e(q)$ for the case $q_0 = (0,0)$. The defining integral

$$W(e)(q,p) = (\varepsilon\pi)^{-3} \int_{\mathbb{R}^2} \mathrm{e}^{ix\cdot p}(q_1^2 - \tfrac{\varepsilon^2}{4}x_1^2)(q_2^2 - \tfrac{\varepsilon^2}{4}x_2^2)\, \mathrm{e}^{-\frac{1}{\varepsilon}|q|^2 - \frac{\varepsilon}{4}|x|^2}\, \mathrm{d}x$$

falls into four parts, whose value can be determined due to

$$\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}yw} y^2\, \mathrm{e}^{-\alpha y^2} \mathrm{d}y = -\tfrac{\sqrt{\pi}}{4}\alpha^{-5/2}(w^2 - 2\alpha)\, \mathrm{e}^{-\frac{w^2}{4\alpha}}, \qquad w\in\mathbb{C},\ \alpha > 0.$$

With $\alpha = \frac{\varepsilon}{4}$ and $w = p_j$ for $j = 1,2$, one obtains

$$\begin{aligned} \int_{\mathbb{R}^2} \mathrm{e}^{\mathrm{i}x\cdot p}\, \mathrm{e}^{-\frac{\varepsilon}{4}|x|^2}\, \mathrm{d}x &= \tfrac{4\pi}{\varepsilon}\, \mathrm{e}^{-\frac{1}{\varepsilon}|p|^2}, \\ \int_{\mathbb{R}^2} \mathrm{e}^{\mathrm{i}x\cdot p} x_1^2\, \mathrm{e}^{-\frac{\varepsilon}{4}|x|^2}\, \mathrm{d}x &= \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}x_1 p_1} x_1^2\, \mathrm{e}^{-\frac{\varepsilon}{4}x_1^2}\, \mathrm{d}x_1 \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}x_2 p_2}\, \mathrm{e}^{-\frac{\varepsilon}{4}x_2^2}\, \mathrm{d}x_2, \\ &= -8\sqrt{\pi}\varepsilon^{-5/2}(p_1^2 - \tfrac{\varepsilon}{2})\, \mathrm{e}^{-\frac{1}{\varepsilon}p_1^2}\sqrt{\tfrac{4\pi}{\varepsilon}}\, \mathrm{e}^{-\frac{1}{\varepsilon}p_2^2} \\ &= -\tfrac{16\pi}{\varepsilon^3}(p_1^2 - \tfrac{\varepsilon}{2})\, \mathrm{e}^{-\frac{1}{\varepsilon}|p|^2}, \\ \int_{\mathbb{R}^2} \mathrm{e}^{\mathrm{i}x\cdot p} x_2^2\, \mathrm{e}^{-\frac{\varepsilon}{4}|x|^2}\, \mathrm{d}x &= -\tfrac{16\pi}{\varepsilon^3}(p_2^2 - \tfrac{\varepsilon}{2})\, \mathrm{e}^{-\frac{1}{\varepsilon}|p|^2}, \\ \int_{\mathbb{R}^2} \mathrm{e}^{\mathrm{i}x\cdot p} x_1^2 x_2^2\, \mathrm{e}^{-\frac{\varepsilon}{4}|x|^2}\, \mathrm{d}x &= \textstyle\prod_{j=1,2} \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}x_j p_j} x_j^2\, \mathrm{e}^{-\frac{\varepsilon}{4}x_j^2}\, \mathrm{d}x_j \\ &= \textstyle\prod_{j=1,2} -8\sqrt{\pi}\varepsilon^{-5/2}(p_j^2 - \frac{\varepsilon}{2})\, \mathrm{e}^{-\frac{1}{\varepsilon}p_j^2} \\ &= \tfrac{64\pi}{\varepsilon^5}(p_1^2 - \tfrac{\varepsilon}{2})(p_2^2 - \tfrac{\varepsilon}{2})\, \mathrm{e}^{-\frac{1}{\varepsilon}|p|^2}. \end{aligned}$$

Hence,

$$W(e)(q,p) = \frac{4\, \mathrm{e}^{-\frac{1}{\varepsilon}|(q,p)|^2}}{\varepsilon^4\pi^2}\Big(q_1^2 q_2^2 + q_2^2(p_1^2 - \tfrac{\varepsilon}{2}) + q_1^2(p_2^2 - \tfrac{\varepsilon}{2}) + \textstyle\prod_{j=1,2}(p_j^2 - \frac{\varepsilon}{2})\Big).$$

C Position densities by density estimation

At some final time $t = t_\mathrm{f} > 0$ the single switch algorithm produces two sets of irregularly spaced points

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

with associated weights $w_k^\pm$, which approximate the value of the Wigner functions $W(\psi^\pm(t_\mathrm{f}))$ in these points. A simple approximation in arbitrary phase space points $(q^\star, p^\star)$ can be deduced by kernel density estimation [21]

$$W(\psi^\pm(t_\mathrm{f}))(q^\star, p^\star) \approx \frac{1}{M^\pm} \sum_{k=1}^{M^\pm} w_k^\pm \, K((q^\star, p^\star) - (q_k, p_k))$$

with Gaussian kernel

$$K : \mathbb{R}^{2d} \to \mathbb{R}, \quad x \mapsto (\pi/c)^{-d} \exp(-c|x|^2).$$

Let $\{p_1, \ldots, p_N\} \subset \mathbb{R}^d$ be a grid in momentum space with uniform patch size $\omega(p_1) = \ldots = \omega(p_N)$. Then, the position densities in a point $q^\star \in \mathbb{R}^d$ are approximately

$$\begin{aligned} |\psi^\pm(t_\mathrm{f}, q^\star)|^2 &= \int_{\mathbb{R}^d} W(\psi^\pm(t_\mathrm{f}))(q^\star, p) \, \mathrm{d}p \\ &\approx \frac{1}{M^\pm} \sum_{j=1}^{N} \sum_{k=1}^{M^\pm} w_k^\pm K((q^\star, p_j) - (q_k, p_k)) \omega(p_j). \end{aligned}$$

We use this approach for a visual comparison of the position densities obtained from Strang splitting with the approximation resulting from surface hopping with Monte Carlo quadrature. The obtained plots have been insensitive with respect to the variation of c in the interval $[10, 50]$. Choosing $c = 20$, Figures 11, 12, and 13, collect the densities for the single Gaussian wave packet, the superposition of two Gaussians, and the excited oscillator state, respectively. One observes a good agreement between the mean positions, however, as expected, no pointwise agreement, since the surface hopping approximation of the Wigner function only holds in a weak sense.

References

[1] N. de Bruijn. Uncertainty principles in Fourier analysis. In O. Shisha, editor, *Inequalities*, pages 57–71. Academic Press, 1967.

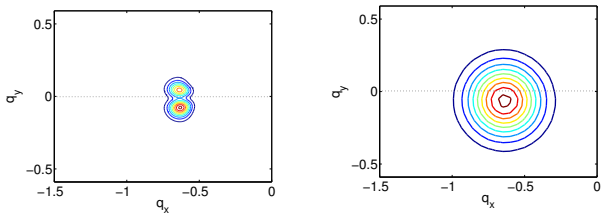
[2] W. Domcke, D. Yarkony, and H. Köppel, editors. *Conical intersections*. World Scientific, Singapore, 2004.

[3] G. Folland. *Harmonic Analysis in Phase Space*. Princeton University Press, Princeton, 1989.

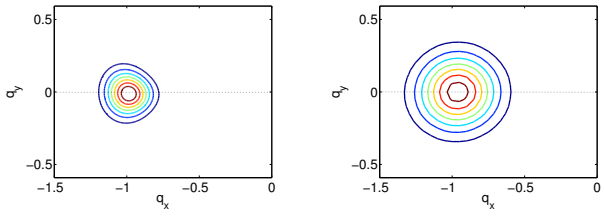
[4] D. Frenkel and B. Smit. *Understanding Molecular Simulation – From Algorithms to Applications*, volume 1 of *Computational Science Series*. Academic Press, 2002.

[5] A. Gelman and D. B. Rubin. Inference from iterative simulation using multiple sequences. *Statistical Sience*, 7:457–511, 1992.

[6] G. Granucci and M. Persico. Critical appraisal of the fewest switches algorithm for surface hopping. *J. Chem Phys.*, 126:134114, 2007.

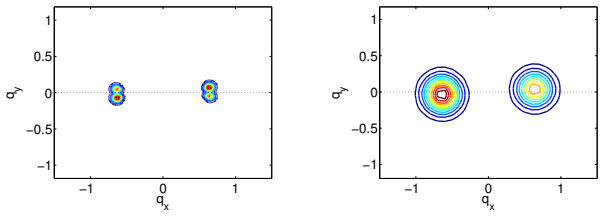


(a) Strang: $|\psi^+(q,t_\mathrm{f})|^2$ (b) SH: $|\psi^+(q,t_\mathrm{f})|^2$

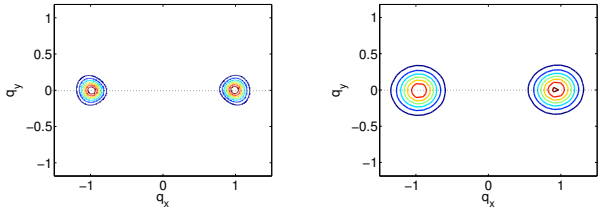


(c) Strang: $|\psi^-(q,t_\mathrm{f})|^2$ (d) SH: $|\psi^-(q,t_\mathrm{f})|^2$

Figure 11: Gaussian wave packet. Position density of the upper and lower level functions $|\psi^+(q,t)|^2$ and $|\psi^-(q,t)|^2$ at time $t = t_\mathrm{f}$. The figures on the left show the results from Strang splitting, the figures on the right the results from a single surface hopping run with 1000 sampling points.

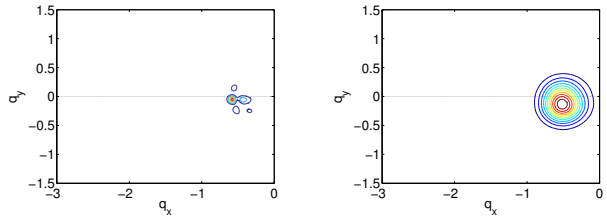


(a) Strang: $|\psi^+(q,t_\mathrm{f})|^2$ (b) SH: $|\psi^+(q,t_\mathrm{f})|^2$

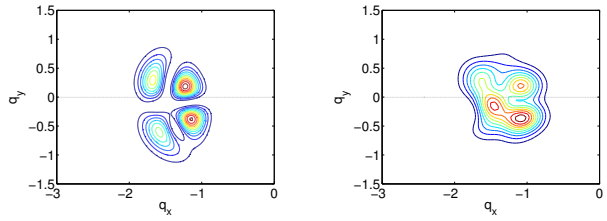


(c) Strang: $|\psi^-(q,t_\mathrm{f})|^2$ (d) SH: $|\psi^-(q,t_\mathrm{f})|^2$

Figure 12: Superposition of Gaussian wave packets. Position density of the upper and lower level function $|\psi^+(q,t)|^2$ and $|\psi^-(q,t)|^2$ at time $t = t_\mathrm{f}$. The figures on the left show the results from Strang splitting, the figures on the right the results from a single surface hopping run with 1000 sampling points, which does not properly reflect the symmetries between the two peaks.



(a) Strang: $|\psi^+(q,t_\mathrm{f})|^2$ (b) SH: $|\psi^+(q,t_\mathrm{f})|^2$



(c) Strang: $|\psi^-(q,t_\mathrm{f})|^2$ (d) SH: $|\psi^-(q,t_\mathrm{f})|^2$

Figure 13: Excited oscillator state. Position density of the upper and lower level function $|\psi^+(q,t)|^2$ and $|\psi^-(q,t)|^2$ at time $t = t_\mathrm{f}$. The figures on the left show the results from Strang splitting, the figures on the right the results from a single surface hopping run with 1000 sampling points.

[7] K. Gröchenig. *Foundations of Time-Frequency Analysis*. Birkhäuser, Boston, 2001.

[8] W. K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57:97–109, 1970.

[9] E. Heller. Wigner phase space method: analysis for semiclassical applications. *J. Chem Phys.*, 65:1289–1298, 1976.

[10] L. Hörmander. *The Analysis of Linear Partial Differential Operators I*. Springer-Verlag, Berlin, 1983.

[11] K. Husimi. Some formal properties of the density matrix. *Proc. Phys. Math. Soc. Japan*, 22:264–314, 1940.

[12] D. Iagolnitzer. Microlocal essential support of a distribution and decomposition theorems - an introduction. In F. Pham, editor, *Hyperfunctions and Theoretical Physics*, pages 121–132. Springer, New York, 1975.

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

[14] C. Fermanian Kammerer and C. Lasser. Propagation through generic level crossings: a surface hopping semigroup. *SIAM J. Math. An.*, 2008.

[15] C. Lasser, T. Swart, and S. Teufel. Construction and validation of a rigorous surface hopping algorithm for conical crossings. *Comm. Math. Sci.*, 2007.

[16] C. Lasser and S. Teufel. Propagation through conical crossings: an asymptotic semigroup. *Comm. Pure Appl. Math*, 58:1188–1230, 2005.

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

[18] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1092, 1953.

[19] J. Moyal. Quantum mechanics as a statistical theory. *Proc. Cambridge Philos. Soc.*, 45:99–124, 1949.

[20] E. Novak. Is there a curse of dimension for integration? In C.-H. Lai, P. E. Bjorstad, M. Crss, and O. B. Widlund, editors, *Eleventh Int. Conf. on Domain Decomposition Methods*, pages 88–95. Domain Decomposition Press, Bergen, 1999.

[21] E. Parzen. On estimation of a probability density function and mode. *Ann. Math. Stat.*, 33:1065–1076, 1962.

[22] P. H. Peskun. Optimum Monte Carlo sampling using Markov chains. *Biometrika*, 60:607–612, 1973.

[23] L. Plaskota, K. Ritter, and G. W. Wasilkowski. Average case complexity of weighted approximation and integration over $\mathbb{R}^d$ with isotropic weight. In K.T. Fang, F.J. Hickernell, and H. Niederreiter, editors, *Monte Carlo and Quasi Monte Carlo Methods 2000*, pages 479–522. Springer, 2002.

[24] D. Robert. *Autour de l'Approximation Semi-Classique*. Birkhäuser, Boston, 1987.

[25] C. P. Roberts and G. Casella. *Monte Carlo Statistical Methods*. Springer, New York, 1999.

[26] J. Rust. Using randomization to break the curse of dimensionality. *Econometrica*, 65(3):487–516, 1997.

[27] A. M. Stuart, J. Voss, and P. Wiberg. Conditional path sampling of SDEs and the Langevin MCMC method. *Comm. Math. Sci.*, 2:685–697, 2004.

[28] L. Tierney. Markov chains for exploring posterior distributions. *The Annals of Statistics*, 22:1701–1728, 1994.

[29] J. Ville. Theory and applications of the notion of complex signal. Rept. T-92, The Rand Corp., Santa Monica, California, 1958.

[30] L. Walter and M. Weber. ConfJump: a fast biomolecular sampling method which drills tunnels through high mountains. ZIB-Report 06-26, Zuse Institute Berlin, 2006. `http://www.zib.de/Publications/abstracts/ZR-06-26/`.

[31] E. Wigner. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.*, 40:749–759, 1932.