

# Single switch surface hopping for a model of pyrazine

Caroline Lasser<sup>a)</sup> and Torben Swart<sup>b)</sup>*Fachbereich Mathematik, Freie Universität Berlin, 14195 Berlin, Germany*

(Received 17 April 2008; accepted 12 June 2008; published online 21 July 2008)

The single switch trajectory surface hopping algorithm is tested for numerical simulations of a two-state three-mode model for the internal conversion of pyrazine through a conical intersection of potential energy surfaces. The algorithm is compared to two other surface hopping approaches, namely, Tully's method of the fewest switches [J. Chem. Phys. **93**, 1061 (1990)] and the method by Voronin *et al.* [J. Phys. Chem. A **102**, 6057 (1998)]. The single switch algorithm achieves the most accurate results. Replacing its deterministic nonadiabatic branching condition by a probabilistic accept-reject criterion, one obtains the method of Voronin *et al.* without momentum adjustment. This probabilistic version of the single switch approach outperforms the considered algorithms in terms of accuracy, memory requirement, and runtime. © 2008 American Institute of Physics. [DOI: 10.1063/1.2954019]

## I. INTRODUCTION

The quantum dynamics of polyatomic molecules is often beyond the range of the Born–Oppenheimer approximation, since electronic potential energy surfaces, which are spanned by several nuclear coordinates, might approach each other or intersect. The induced nonadiabatic effects considerably extend the traditional picture, that nuclear motion mainly follows one relevant electronic surface.

There are numerous discretization methods for the numerical simulation of nonadiabatic phenomena. One class of approximation schemes directly solves the underlying Schrödinger equation, among them the multiconfiguration time-dependent Hartree algorithm, the coupled coherent states technique, and the matching-pursuit split-operator Fourier transform method.<sup>1–3</sup> Quantum-classical or semiclassical approaches, however, are motivated by the intrinsic high frequencies of the dynamics and suitably augment classical transport equations for the approximation of nonadiabatic transitions. Trajectory surface hopping constitutes a popular family of algorithms within this class. It combines classical transport along the electronic surfaces with some nonadiabatic hopping between them. The different kinds of hopping distinguish the many variants of the method.

Our aim here is to test the performance of the recently proposed single switch algorithm,<sup>4,5</sup> which has been derived from a rigorous mathematical analysis of Schrödinger systems with conical crossings of potential energy surfaces.<sup>6–8</sup> The single switch algorithm implements a deterministic branching scheme, which splits a classical trajectory whenever it attains a locally minimal surface gap and weights the branches according to a multidimensional Landau–Zener formula. For an assessment of the obtained numerical results, we compare with two other trajectory surface hopping methods, the well-established fewest switches algorithm<sup>9</sup> introduced by Tully in 1990 and the method of Voronin *et al.*,<sup>10</sup>

which can be viewed as an adoption of the original algorithm of Tully and Preston.<sup>11</sup> The three algorithms can be grouped in at least three ways. On the one hand, the fewest switches method and the approach of Voronin *et al.* are probabilistic, while the single switch algorithm implements deterministic dynamics. On the other hand, the fewest switches method allows for nonadiabatic hops at every time step of the discretization, while the other two algorithms perform transitions only when a classical trajectory attains a locally minimal surface gap. Moreover, the single switch approach does not adjust momentum, while the other two methods do.

Our numerical experiments use a two-state three-mode vibronic coupling model for the ultrafast internal conversion of pyrazine through a conical intersection of the two lowest excited singlet states. The model has been introduced by Schneider and Domcke<sup>12</sup> in 1988. It has been used as a demanding test for various approximations,<sup>13–16</sup> since it develops complex electronic and vibrational relaxation dynamics. Our main focus is on the accuracy, with which the three surface hopping algorithms compute level populations as well as position, momentum, and energy expectation values. However, issues of computational efficiency as memory requirement and runtime are also addressed.

The article is organized as follows. Section II discusses the three surface hopping algorithms in detail. Section III briefly introduces the considered pyrazine model as well as numerical implementation issues. Section IV presents the comparative numerical experiments, which are also concerned with momentum adjustment, efficiency, and Monte Carlo sampling of the initial data. Section V offers some concluding remarks.

## II. TRAJECTORY SURFACE HOPPING

We formulate a general scheme for trajectory surface hopping algorithms for the solution of time-dependent two-state Schrödinger systems,

<sup>a)</sup>Electronic mail: lasser@math.fu-berlin.de.<sup>b)</sup>Electronic mail: swart@math.fu-berlin.de.

$$i\partial_t\psi = -\sum_{j=1}^n \frac{1}{2m_j} \Delta_{q_j} \psi + V\psi, \quad \psi(0) = \psi_0,$$

whose potential is a real-symmetric two-by-two matrix,

$$V(q) = \begin{pmatrix} v_{11}(q) & v_{12}(q) \\ v_{12}(q) & v_{22}(q) \end{pmatrix}, \quad q \in \mathbb{R}^d = \mathbb{R}^{3n}.$$

Such systems approximately describe the dynamics of  $n$  nuclei with masses  $m_1, \dots, m_n$  under the influence of two electronic eigenstates. Note that the equation is cast in atomic units, such that  $\hbar=1$ . To transform the equation to dimensionless coordinates, one sets

$$h = \min(m_1, \dots, m_n)^{-1/2}$$

and rescales

$$q_j \rightarrow (m_j^{1/2}h)^{-1}q_j.$$

In this way, one reformulates the Schrödinger operator as

$$-\frac{h^2}{2}\Delta_q + V,$$

where  $h$  is typically ranging from 0.001 to 0.1. The timescales on which the associated dynamics develop their characteristic features are of order  $1/h$ . Therefore, one also rescales  $t \rightarrow t/h$  and obtains a Schrödinger system of the form

$$ih\partial_t\psi = -\frac{h^2}{2}\Delta_q\psi + V\psi, \quad \psi(0) = \psi_0. \quad (1)$$

We assume that the two eigenvalues  $\lambda^-(q) \leq \lambda^+(q)$  of the matrix have a conical intersection. That is, the crossing set  $\{q \in \mathbb{R}^d | \lambda^-(q) = \lambda^+(q)\}$  is a smooth codimension two sub-manifold of the configuration space  $\mathbb{R}^d$ . In the three-mode pyrazine model considered later on, the intersection set is a straight line in the plane spanned by the two tuning modes.

Let  $\chi^\pm(q)$  denote two normalized eigenvectors associated with the eigenvalues  $\lambda^\pm(q)$  and define the corresponding level functions

$$\psi^\pm(q, t) := \langle \psi(q, t) | \chi^\pm(q) \rangle_{\mathbb{C}^2} \quad (2)$$

of the solution of the Schrödinger Eq. (1). Here and in the following, the scalar product between vectors  $x$  and  $y$  in the Euclidian spaces  $\mathbb{C}^n$ ,  $n \in \mathbb{N}$ , is denoted by  $\langle x | y \rangle_{\mathbb{C}^n} = x^t \cdot y^*$   $= \sum_{j=1}^n x_j y_j^*$ .

One aims at the computation of expectation values

$$\langle A\psi^\pm(t) | \psi^\pm(t) \rangle_{L^2} = \int_{\mathbb{R}^d} A\psi^\pm(q, t) \psi^\pm(q, t)^* dq,$$

where the observables considered here are the identity operator  $\text{id}: \psi \mapsto \psi$ , which generates the adiabatic potential level populations, the position and momentum operators  $Q_j: \psi \mapsto q_j \psi$  and  $P_j: \psi \mapsto -ih\partial_j \psi$  for  $j=1, \dots, d$ , and the energy operators

$$E^\pm: \psi \mapsto -\frac{h^2}{2}\Delta_q\psi + \lambda^\pm\psi.$$

If the operator  $A$  is the Weyl quantization of a function  $a: \mathbb{R}^{2d} \rightarrow \mathbb{C}$  on classical phase space, then the expectation values can be expressed as phase space integrals,

$$\langle A\psi^\pm(t) | \psi^\pm(t) \rangle_{L^2} = \int_{\mathbb{R}^{2d}} a(q, p) W(\psi^\pm(t))(q, p) dq dp,$$

where

$$(W\psi)(q, p) = (2\pi h)^{-d} \int_{\mathbb{R}^d} e^{iy \cdot p/h} \psi(q - \frac{1}{2}y) \psi(q + \frac{1}{2}y)^* dy$$

is the Wigner function  $W(\psi): \mathbb{R}^{2d} \rightarrow \mathbb{R}$  of the wave function  $\psi$ . The level populations are generated by the function  $a(q, p) = 1$ , the position and momentum expectation values by  $a(q, p) = q_j$  and  $a(q, p) = p_j$ , respectively, and the energy expectations by  $a(q, p) = \frac{1}{2}|p|^2 + \lambda^\pm(q)$ . We adopt this phase space point of view for the following schematic exposition of a trajectory surface hopping algorithm.

*Phase space sampling.* Derive from the initial level functions  $\psi^\pm(0)$  two sets of phase space points  $(q_1^\pm, p_1^\pm), \dots, (q_N^\pm, p_N^\pm)$ . In our simulations, we perform either a grid-based or a Monte Carlo sampling of the two Wigner functions  $W(\psi^\pm(0))$ .

*Classical transport.* Propagate the phase space points along the trajectories of the associated Hamiltonian systems,

$$\dot{q} = p, \quad \dot{p} = -\nabla_q \lambda^\pm(q). \quad (3)$$

*Nonadiabatic transitions.* The way of incorporating nonadiabatic transitions is crucial and distinguishes the different surface hopping methods. The choices of the examined algorithms are thoroughly discussed in Secs. II A–II C.

*Computation of expectation values.* At some final time  $t$  one obtains two sets of phase space points  $(q_1^\pm, p_1^\pm), \dots, (q_M^\pm, p_M^\pm)$  with associated real-valued weights  $w_1^\pm, \dots, w_M^\pm$ . The size of the weights depends on the number of transitions the underlying trajectory has performed. For the probabilistic methods the number of initial and final phase space points coincides, while the deterministic approach generates a larger number of final points. The weighted sums

$$\sum_{k=1}^M a(q_k^\pm, p_k^\pm) w_k^\pm \delta_k$$

then approximate the desired expectation values. Depending on the initial sampling, the integration weight  $\delta_k > 0$  is the volume element of the grid or  $1/N$  for a Monte Carlo sampling with  $N$  initial phase space points per level.

## A. Single switch algorithm

For each individual classical trajectory  $t \mapsto (q_t^\pm, p_t^\pm)$ , the single switch algorithm<sup>4,5</sup> monitors the gap  $g(q) = \lambda^+(q) - \lambda^-(q)$  between the electronic surfaces. If the function  $t \mapsto g(q_t^\pm)$  passes a local minimum, a deterministic branching

occurs. The weights of the branches are given by a multidimensional Landau–Zener rate, which is derived from the diabatic potential.

One writes the potential matrix as the sum of half its trace and its trace-free part,

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

Let  $dv(q)$  denote the  $2 \times d$  gradient matrix of the vector  $v(q) = (v_1(q), v_2(q))$ . If  $t \mapsto g(q_t^\pm)$  attains a local minimum at time  $\tau$ , then

$$\langle dv(q_\tau^\pm) p_\tau^\pm | v(q_\tau^\pm) \rangle_{\mathbb{C}^2} = 0. \quad (4)$$

If  $(q, p)$  is the phase space point, at which the trajectory attains its local minimal gap, then the transition rate is

$$T(q, p) = \exp\left(-\frac{\pi}{h} \frac{|v(q)|^2}{|dv(q)p|}\right). \quad (5)$$

The branch switching to the other level takes the old weight times  $T(q, p)$ , while the one remaining on the same level is reweighted with the factor  $1 - T(q, p)$ .

If there are no trajectories on opposite levels, which perform simultaneously their branching at nearby phase space points, then the single switch algorithm is an asymptotically correct method. The mathematically proven convergence rate<sup>8</sup> is  $h^{1/8}$ , while all numerical experiments so far<sup>4,5,8</sup> have even shown a rate of order  $h^{1/2}$ .

## B. Tully's fewest switches

In Tully's fewest switches approach,<sup>9</sup> one solves two additional ordinary differential systems. If the normalized eigenvectors are chosen real-valued,

$$\chi^+(q) = \begin{pmatrix} \cos \theta_q \\ \sin \theta_q \end{pmatrix}, \quad \chi^-(q) = \begin{pmatrix} -\sin \theta_q \\ \cos \theta_q \end{pmatrix}, \quad (6)$$

where the angle  $\theta_q = \frac{1}{2} \arctan(v_2(q)/v_1(q)) \in [-\pi/2, \pi/2]$  is half the four quadrant inverse tangent of  $v_1(q)$  and  $v_2(q)$ , then the so-called nonadiabatic coupling vector,

$$d(q) = \langle \chi^+(q) | \nabla_q \chi^-(q) \rangle_{\mathbb{C}^2},$$

is also real-valued. The additional systems have the form

$$ih \frac{d}{dt} c^\pm(t) = M^\pm(t) c^\pm(t), \quad (7)$$

with

$$M^\pm(t) = \begin{pmatrix} \lambda^\pm(q_t^\pm) & \mp ih \langle p_t^\pm | d(q_t^\pm) \rangle \\ \mp ih \langle p_t^\pm | d(q_t^\pm) \rangle & \lambda^\mp(q_t^\pm) \end{pmatrix}.$$

A trajectory switches from the upper to the lower level at time  $\tau$ , if

$$2\Delta\tau \Re(c_2^+(\tau) c_1^+(\tau)^* \langle p_\tau^+ | d(q_\tau^+) \rangle_{\mathbb{C}^d}) > \zeta |c_1^+(\tau)|^2,$$

where  $\Delta\tau$  is the length of the time interval since the last nonadiabatic hopping decision and  $\zeta$  a random number uniformly distributed in the interval  $[0, 1]$ . The corresponding condition for a switch from the lower to the upper level is

$$-2\Delta\tau \Re(c_2^-(\tau) c_1^-(\tau)^* \langle p_\tau^- | d(q_\tau^-) \rangle_{\mathbb{C}^d}) > \zeta |c_1^-(\tau)|^2.$$

If the switch is accepted, say in a point  $(q, p)$ , then the trajectory is assigned a new point before evolving on the new level. The new point has position  $q_* = q$  and its momentum  $p_*$  is chosen such that  $\frac{1}{2}|p|^2 + \lambda^\pm(q) = \frac{1}{2}|p_*|^2 + \lambda^\mp(q_*)$ . The precise prescription is the following. One defines two normalized nonadiabatic coupling vectors  $n^\pm(q) = \pm |d(q)|^{-1} d(q)$  and sets

$$\gamma^\pm = -\langle p | n^\pm(q) \rangle_{\mathbb{C}^d} + \sqrt{\langle p | n^\pm(q) \rangle_{\mathbb{C}^d}^2 \pm 2g(q)}.$$

For a switch from the upper to the lower level the new momentum is  $p_* = p + \gamma^+ n^+(q)$ . For an upward switch one chooses  $p_* = p + \gamma^- n^-(q)$ , if  $\gamma^-$  is real-valued, and suppresses the hop otherwise. The suppressed transitions are called classically forbidden hops.

## C. Method of Voronin *et al.*

The algorithm proposed by Voronin *et al.*<sup>10</sup> is also probabilistic. If  $t \mapsto g(q_t^\pm)$  reaches a local minimum at time  $\tau$  in a point  $(q, p)$ , then one uses the values of the gap for the last three points of time  $t$  to compute two parameters  $A$  and  $B$ , which are defined by

$$g(q_t^\pm) = \sqrt{B^2(t - \tau)^2 + 4A^2}.$$

These parameters constitute a Landau–Zener transition rate of the form

$$T_{\text{VMV}} = \exp\left(-\frac{\pi 2A^2}{h B|p|}\right).$$

If  $T_{\text{VMV}} \geq \zeta$  for a uniformly distributed random number  $\zeta \in [0, 1]$ , then the trajectory switches to the other surface, and one performs the same momentum adjustment as described in Sec. II B.

If nonadiabatic transitions are only performed when the energy gap  $g(q)$  is of order  $h^{1/2}$ , then an easy argument shows<sup>5</sup> that the Landau–Zener rate  $T_{\text{VMV}}$  rate agrees with the one of the single switch algorithm up to a term of order  $h^{1/2}$ . Moreover, for such small gaps the momentum adjustment is a correction of order  $h^{1/2}$  as well.

## III. NUMERICAL SETUP

We briefly provide the necessary details of the considered pyrazine model and discuss the numerical realization of the three surface hopping algorithms.

### A. The pyrazine model

The two-state three-mode Hamiltonian for the internal conversion of pyrazine is originally formulated in dimensionless normal coordinates.<sup>12</sup> It reads

$$\frac{1}{2} \sum_{j=1}^3 \omega_j (-\partial_j^2 + q_j^2) + \begin{pmatrix} E_1 + \langle \kappa | q \rangle_{\mathbb{C}^3} & \lambda q_3 \\ \lambda q_3 & E_2 + \langle k | q \rangle_{\mathbb{C}^3} \end{pmatrix},$$

where the parameters are given in Table I. To obtain a Schrödinger equation of the semiclassical form (1), we set

TABLE I. The parameters of the original (Ref. 12) and the rescaled pyrazine model.

$E_1=3.94$	$\kappa=(0.037, -0.105, 0)$	$\omega=(0.126, 0.074, 0.118)$
$E_2=4.84$	$k=(-0.254, 0.149, 0)$	$\lambda=0.262$
$C_1=4.39$	$a=(-0.521, 0.081, 0)$	$\alpha=(1.703, 1, 1.595)$
$C_2=-0.45$	$b=(0.698, -0.467, 0)$	$\beta=1.216$

the smallest of the three oscillator frequencies as  $h=0.074$  and rescale the coordinates according to

$$q_j \rightarrow \frac{h}{\sqrt{\omega_j}} q_j \quad (j = 1, 2, 3).$$

The resulting Schrödinger equation is

$$ih\partial_t\psi = -\frac{h^2}{2}\Delta_q\psi + V\psi, \quad \psi(0) = \psi_0,$$

and has a potential matrix defined by

$$\frac{1}{2}\text{tr}(V(q)) = C_1 + \langle a|q\rangle_{C^3} + \frac{1}{2}\sum_{j=1}^3 \alpha_j^2 q_j^2,$$

$$v(q) = (C_2 + \langle b|q\rangle_{C^3}, \beta q_3).$$

The new parameters are also collected in Table I. The eigenvalues  $\lambda^\pm(q)$  intersect on a straight line in the plane spanned by the two tuning modes, that is, for  $q_2 = -(C_2 + b_1 q_1)/b_2$  and  $q_3 = 0$ . As in the previous benchmark computations for the model,<sup>13–16</sup> the initial wave function is chosen as  $\psi_0 = \psi_0^+ \chi^+$ , where

$$\psi_0^+(q) = \prod_{j=1}^3 \left(\frac{\alpha_j}{\pi h}\right)^{1/4} \exp\left(-\frac{\alpha_j}{2h} q_j^2\right)$$

is the ground state of the quantum mechanical harmonic oscillator in semiclassical scaling  $\frac{1}{2}\sum_{j=1}^3 (-h^2 \partial_j^2 + \alpha_j^2 q_j^2)$  and  $\chi^+$  is an eigenvector associated with the  $S_2$ -state of pyrazine. There is some freedom in the choice of the phase of the eigenvectors, which can be used to alleviate the discontinuity of the real-valued eigenvectors (6) on the slit  $v_1(q) < 0$ ,  $v_2(q) = 0$ . Since the functional support of the initial oscillator state is rather close to the discontinuity, we have used the complex-valued eigenvectors,

$$\chi^+(q) = e^{i\theta_q} \begin{pmatrix} \cos \theta_q \\ \sin \theta_q \end{pmatrix}. \quad (8)$$

The initial level functions  $\psi^\pm(q, 0)$  defined by Eq. (2) are insensitive to any choice of phase, while the quantum solution  $\psi(q, t)$  slightly changes in momentum. An easy argument shows<sup>5</sup> that level expectation values are altered by terms of order  $h$ , which is below the accuracy of surface hopping algorithms.

For a reference solution, the Schrödinger Eq. (1) is solved by a Strang splitting scheme, which discretizes the Laplacian using the fast Fourier transform. The time interval  $[0, 500 \text{ fs}]$  is divided into 200 000 time steps, while the computational domain  $[-2, 2]^3$  is discretized by 128 uniformly spaced grid points per direction. Comparing the solution with

TABLE II. Differences of the solution computed by the pseudospectral Strang splitting scheme with full and half resolutions. Each entry shows the maximum error of the recorded points of time. The error of the energy is a relative error, all other numbers are absolute errors.

$L^2$ -norm	Position exp.	Momentum exp.	Energy exp.
0.022	0.0013	0.0011	0.000 04

the one obtained by half the number of time steps and half the number of grid points per direction, one observes a norm difference of roughly 2% and deviations in the expectation values around and below one permille, see Table II. Hence, the solution is sufficiently accurate for the validation of surface hopping algorithms and can be regarded as a reference.

## B. Numerical implementation of surface hopping

The numerical implementation of the three surface hopping algorithms has to realize the initial sampling and the solution of ordinary differential equations. For the present test case a grid based initial sampling is still feasible, while the complementary numerical experiments with a simple Monte Carlo approach presented in Sec. IV D achieve comparable accuracy. The initial Wigner functions are

$$W(\psi_0^+)(q, p) = (\pi h)^{-3} \prod_{j=1}^3 \exp\left(-\frac{\alpha_j}{h} q_j^2 - \frac{1}{\alpha_j h} p_j^2\right) \quad (9)$$

and  $W(\psi_0^-)(q, p) = 0$ . Hence, for the lower level there are no sampling points. The upper level function is sampled by the following grid-based approach. One discretizes the six-dimensional rectangle  $[-A, A]^3 \times [-B, B]^3$  with  $A = 3.5h^{1/2}$  and  $B = 5h^{1/2}$  using  $m = 12$  uniformly spaced grid points per direction. Then one chooses the minimal number of points  $(q_1, p_1), \dots, (q_N, p_N)$  such that

$$\sum_{j=k}^N W(\psi_0^+)(q_k, p_k) \delta_k \geq 1 - s_{\text{tol}}, \quad (10)$$

where  $\delta_k = (4AB/m^2)^3$  is the volume element of the sampling grid and  $s_{\text{tol}}$  a predefined sampling tolerance. Our default choice  $s_{\text{tol}} = 0.05$  is smaller than the expected accuracy of the surface hopping algorithms and results in 6656 sampling points.

For the ordinary differential systems (3) and (7), we use the algorithm of Dormand and Prince DOPRI5(4), a Runge Kutta method of order of 5 with automatic step-size control.<sup>17</sup> For the detection of the minimal gap, the regula falsi method is combined with the interpolation routine of DOPRI5(4), which does not require any further evaluations of the right-hand side of Eq. (3). For the single switch algorithm as well as the approach of Voronin *et al.*, the rigorously derived Landau–Zener rate (5) is employed.

## IV. NUMERICAL RESULTS

The numerical experiments are concerned with momentum adjustment, runtime and memory requirements, the achieved accuracy of the computations, and the performance of a simple Monte Carlo sampling of the initial data.

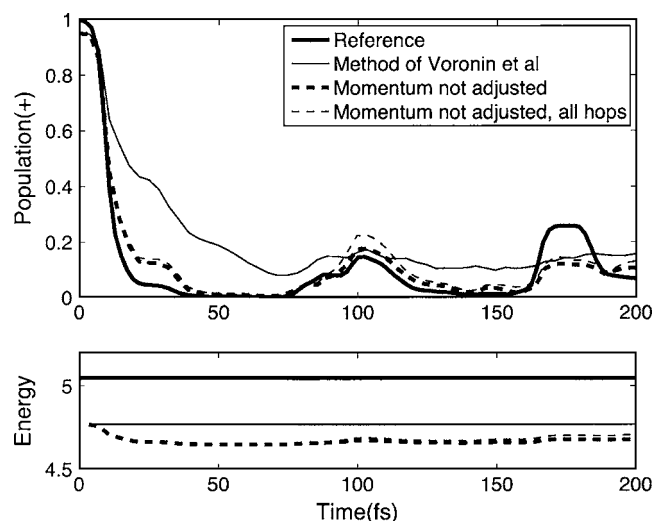


FIG. 1. The population of the  $S_2$ -state (top) and the total energy of the system (bottom) computed for the first 200 fs by variants of the method proposed by Voronin *et al.* The population improves considerably without momentum adjustment, whereas the loss in total energy is of the order of magnitude of the error induced by the initial sampling error.

### A. Momentum adjustment

We numerically investigate the effects of the momentum adjustment contained in the fewest switches algorithm and the method of Voronin *et al.*

Figure 1 shows the upper level population for the first 200 fs and the total energy of the system as computed by the Voronin method using the rigorously derived Landau-Zener rate (5). The thick solid line is the outcome of the grid-based reference solution. The thin solid line represents the result of the original suggestion, that is, with momentum adjustment. For the dashed lines, momentum is not adjusted, while classical forbidden jumps are either suppressed (thin dashed) or allowed (thick dashed). Since the initial sampling according to Eq. (10) achieves an initial population of  $1 - s_{\text{tol}}$ , all surface hopping computations do not start with unity.

The outcome of the two computations without momentum adjustment deviates only slightly, namely, at most 0.05 and 0.02 in average. However, an adjustment of the momentum strongly deteriorates the resulting population. The computations with and without momentum adjustment show errors around 0.12/0.39 and 0.04/0.14 (mean error/maximal error), respectively. On the other hand, the improvement in total energy, which is shown in the lower plot of Fig. 1, is negligible compared to the error already induced by the initial sampling. Thus, for the following comparison in Sec. IV C, we use the Voronin approach without momentum adjustment and allow for forbidden hops. Since this version differs from the single switch algorithm only with respect to the probabilistic nonadiabatic transfer, we refer to it as the *probabilistic single switch method*.

For the fewest switches algorithm the situation is different. Figure 2 shows again the population of the upper state and the total energy of the system, using the same line styles as before. The three variants of the fewest switches method differ less drastically as in the previous experiments. More precisely, the average errors are between 0.04 and 0.08,

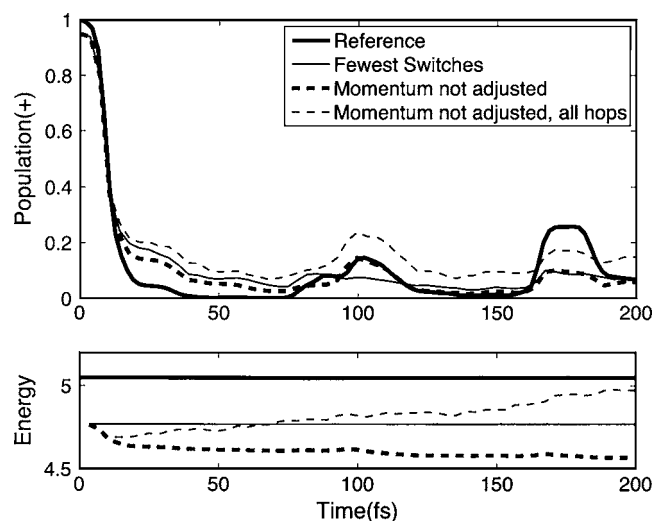


FIG. 2. The population of the  $S_2$ -state (top) and the total energy (bottom) for 200 fs computed by variants of the fewest switches method. Momentum adjustment does not alter the outcome as drastically as in the method by Voronin *et al.*

whereas the maximal errors range between 0.16 and 0.17. The two revivals around  $t=100$  fs and  $t=175$  fs are best reproduced by two different versions without momentum adjustment, namely, the one with suppressed forbidden jumps on the one hand and the one allowing all hops on the other hand. Since a good deal of the time, the original method seems to interpolate between the other two, we do not alter the fewest switches approach for the comparison in Sec. IV C.

### B. Efficiency

A crucial problem for the deterministic single switch algorithm is the branching of trajectories at points of minimal gap, which leads to an unmanageable growth of particle numbers as time evolves. However, most particles carry negligible weight and can be removed without loss of accuracy. Thus, every time a trajectory with weight  $w$  attains a local minimal gap at a point  $(q, p)$ , a new particle is added if and only if its weight  $T(q, p)w$  exceeds a given tolerance  $w_{\text{tol}}$ . If the remaining particle's weight  $(1 - T(q, p))w$  is smaller than this tolerance, the weight is transferred completely.

Setting the sampling tolerance  $s_{\text{tol}}=0.05$ , one starts with 6656 initial sampling points, and the unfiltered single switch computation breaks down at approximately 220 fs, when 44 682 174 particles occupy the available 2 GB RAM. Introducing a filtering with  $w_{\text{tol}}=0.1$ , produces 21 484 particles after 200 fs and results in a population of the  $S_2$ -state, which differs at most by 0.003 from the unfiltered computation. However, too generous filtering tolerances deteriorate the accuracy. Figure 3 shows the population of the  $S_2$ -state for two different choices of  $w_{\text{tol}}$ . At the final time 500 fs, the tolerances  $w_{\text{tol}}=0.1$  and  $w_{\text{tol}}=0.5$  produce 46 532 and 14 352 particles, respectively. However, the level populations of both computations differ by 0.055 on average with a maximal deviation of 0.072.

For the sampling and filtering tolerances  $s_{\text{tol}}=0.01$  and  $w_{\text{tol}}=0.05$ , one starts with 15 911 and ends with 89 190 par-

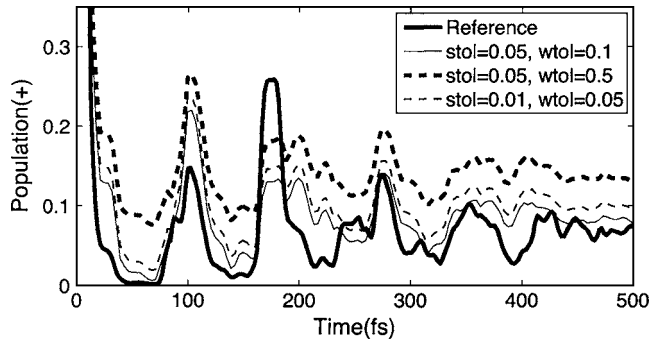


FIG. 3. The population of the  $S_2$ -state as computed by the single switch method with different sampling and filtering tolerances,  $s_{\text{tol}} \in \{0.01, 0.05\}$  and  $w_{\text{tol}} \in \{0.01, 0.1, 0.5\}$ . The choice of the tolerances strongly influences the final particle numbers as well as the quality of the approximation.

ticles. Table III illustrates that the deterministic single switch algorithm even deteriorates by 1%, since the finer sampling seems to amplify the tendency to overestimate the population. The other two probabilistic methods are rather robust with respect to the initial sampling. For the following comparison of the algorithms, we therefore choose the sampling and filtering tolerances  $s_{\text{tol}}=0.05$  and  $w_{\text{tol}}=0.1$ .

The runtime of each of the three surface algorithms on a 3 GHz Pentium IV computer with an initial grid-based sampling with  $s_{\text{tol}}=0.05$  and 6656 points is shown in Table IV. The probabilistic single switch and the fewest switches method are the fastest and the slowest method, respectively. The long runtime of the fewest switches method despite the constant number of particles is due to the additional ordinary differential systems (7), which are more oscillatory than the pure classical transport, and the fact that the frequent checking of the hopping rate prohibits long time steps for the ordinary differential equation solver.

### C. Comparison of the algorithms

We turn to the central comparison of the three surface hopping algorithms, the deterministic and probabilistic single switch, and the fewest switches method. The first quantity we study is the population of the  $S_2$ -state again. Figure 4 shows the population produced by the three algorithms for the first 200 fs. Both single switch results are more accurate than the one of the fewest switches method, whereas the difference between the deterministic single switch and its probabilistic counterpart is on average of 0.012 and does not exceed 0.044.

Figure 5 show the corresponding long-time behavior. The bottom plot illustrates that the fewest switches method produces an almost constant population for times larger than

TABLE III. Accuracy of the algorithms depending on the initial sampling tolerance. The table shows the average and maximal errors of the upper level population for the time interval  $[0, 500 \text{ fs}]$ . The single switch method uses the filtering tolerances  $w_{\text{tol}}=0.1$  and  $w_{\text{tol}}=0.05$ , respectively.

$s_{\text{tol}}$	Single switch	Fewest switches	Prob. single switch
0.05	0.032/0.128	0.033/0.170	0.023/0.140
0.01	0.044/0.140	0.032/0.164	0.026/0.136

TABLE IV. Runtime of the algorithms.

Single switch	Fewest switches	Probabilistic single switch
5 m 10 s	11 m 31 s	1 m 23 s

200 fs. Both single switch approaches, however, also capture the long-time oscillations. Even the two population increases around  $t=400 \text{ fs}$  are qualitatively reproduced. Again, the difference between both approaches is negligible.

The upper plot in Fig. 6 shows the position expectation of the  $q_1$ -mode for the  $S_2$ -state as computed by the deterministic single switch algorithm. The average error is 0.07 with a maximum of 0.45 at 80 fs. As before, the difference between both single switch approaches is small, namely, at most 0.12 and 0.02 on average. The fewest switches method's error is comparable for the first 20 fs and increases up to 0.32 in the long-time regime. This behavior is qualitatively met for all expectation values of position and momentum of the two levels.

The third direction shows an interesting symmetry. Both the position and the momentum expectation value of the single switch method are constantly equal zero, see the momentum expectation in Fig. 6. Since the eigenvalues obey  $\partial_3 \lambda^\pm(q_1, q_2, q_3) = -\partial_3 \lambda^\pm(q_1, q_2, -q_3)$ , the classical trajectories are axially symmetric with respect to the third coordinate. The related property of the vector  $v(q)$ , which builds the diabatic potential matrix,

$$v_1(q_1, q_2, q_3) = v_1(q_1, q_2, -q_3),$$

$$v_2(q_1, q_2, q_3) = -v_2(q_1, q_2, -q_3),$$

guarantees that the minimal gap condition (4) and the Landau-Zener rate (5) are symmetric in the third direction, too. Moreover, the grid-based initial sampling preserves the symmetry of the initial wave function  $\psi_0^\pm(q)$ . Hence, the single switch algorithm produces vanishing position and momentum expectations for the third coordinate, whereas the probabilistic elements of the other two methods do not preserve this symmetry. The oscillations in the reference solution are induced by the complex phase of the eigenvector (8), which causes a small initial momentum of the wave function of order  $\hbar$  not being resolved by the surface hopping approximations. Analytic computations show that an initial da-

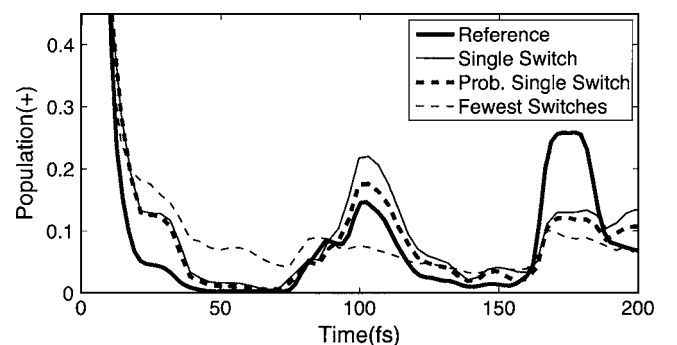


FIG. 4. Population of the  $S_2$ -state during the first 200 fs as computed by the three surface hopping algorithms. The single switch results are more accurate than the one of the fewest switches approach.

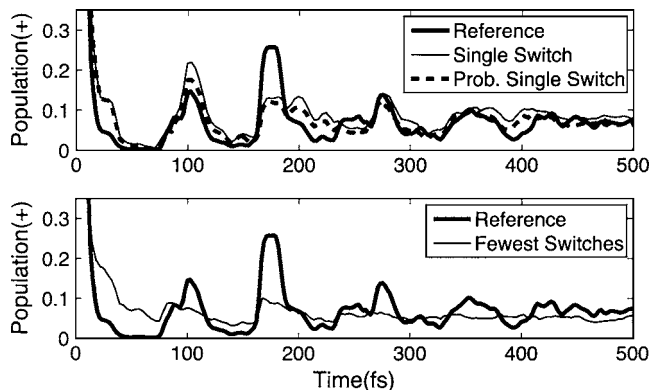


FIG. 5. Long-time behavior of the population of the  $S_2$ -state as computed by the single switch algorithms (top) and the fewest switches approach (bottom). For times larger than 200 fs, fewest switches produce an almost constant population.

tum, which uses a real-valued eigenvector (6), yields a solution, whose position and momentum expectation values for the third direction vanish identically, compare the Appendix.

Finally, we investigate the energy expectation values with respect to both the  $S_1$ - and the  $S_2$ -state. Figure 7 illustrates that all algorithms approximate the  $S_2$ -energy with comparable accuracy, while the fewest switches algorithm benefits from the momentum adjustment and reproduces the  $S_1$ -energy best. Indeed, for the  $S_2$ -state, the single switch algorithms show average relative errors around 0.022 compared to 0.035 for fewest switches. For the  $S_1$ -state, the fewest switches approach has an average relative error of 0.008 compared to 0.03 for the single switch approaches. The maximal relative  $S_2$ -errors are around 0.14 for all three algorithms and occur at times with low population.

#### D. Monte Carlo sampling

A grid-based sampling of the initial data is clearly bound to low-dimensional problems. However, a Monte Carlo sampling of the simple Gaussian function of the test case at hand also allows for a drastic reduction of sampling points. One

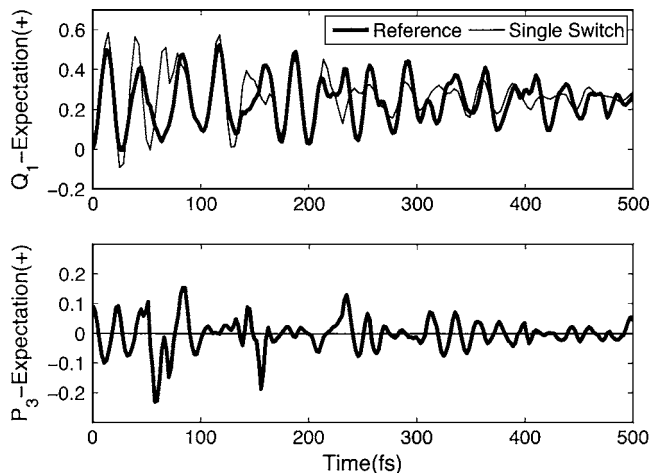


FIG. 6. Position expectation in  $q_1$ -direction (top) and momentum expectation in  $p_3$ -direction (bottom) with respect to the  $S_2$ -state as computed by the single switch algorithm.

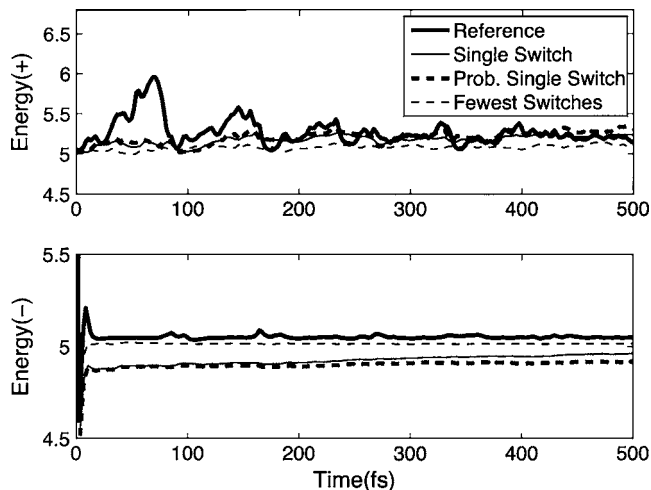


FIG. 7. Energy expectation with respect to the  $S_2$ -state (top) and the  $S_1$ -state (bottom). The  $S_2$ -approximations are of comparable accuracy, while the fewest switches algorithm reproduces the  $S_1$ -energy best.

views the Wigner function (9) as a six-dimensional normal distribution with mean zero and diagonal covariance matrix generated by the vector

$$\frac{\hbar}{2} \left( \frac{1}{\alpha_1}, \frac{1}{\alpha_2}, \frac{1}{\alpha_3}, \alpha_1, \alpha_2, \alpha_3 \right)$$

and draws  $n_{MC}$  sampling points from it. For our experiments we consider  $n_{MC}=100, 250, 500, 1000, 2500,$  and  $5000$ .

Figure 8 shows boxplots of the mean error in the short-time regime  $[0, 250 \text{ fs}]$  and the maximal error in the long-time regime  $[250, 500 \text{ fs}]$  of the deterministic single switch algorithm for 20 runs with initial Monte Carlo sampling. The plots contain the median, the box bounded by the first and third quartile, whiskers extending the box out to the most extreme error within three halves of the interquartile range, and plus marks for outliers. Even small particle numbers result in a mean error comparable to the grid-based approach, which is indicated by the horizontal line in the plot. As expected, the variance between different runs is reduced with

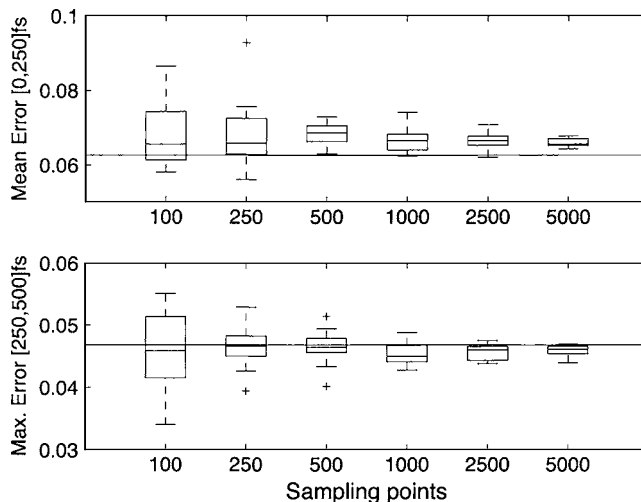


FIG. 8. Boxplots of the error in the population of the  $S_2$ -state for the single switch algorithm. The upper plot shows the mean error on  $[0, 250 \text{ fs}]$ , the lower the maximal error on  $[250, 500 \text{ fs}]$ .

higher initial particle numbers. The other two surface hopping methods show the same behavior with a larger overall variance, which is due to their probabilistic treatment of nonadiabatic transfer.

## V. CONCLUSION

We have tested the single switch surface hopping algorithm for a two-state three-mode model for the internal conversion of pyrazine. The computed level populations and expectation values differ from their reference values on average by 3%. Even the long-time oscillations of the system are well reproduced.

The related approach of Voronin *et al.*, which also combines nonadiabatic transitions at points of locally minimal surface gap with a multidimensional Landau–Zener formula, computes level populations only with an average error of 12%. However, if one omits the method's momentum adjustment, one reaches an accuracy comparable to the single switch approach. Without momentum adjustment the Voronin method can be viewed as a probabilistic counterpart to the single switch algorithm, since in the regime of small surface gap their Landau–Zener rates only differ by a term of order  $h^{1/2}$ . Therefore, we call the Voronin algorithm with the rigorously derived transition rate (5) and without momentum adjustment the probabilistic single switch method. It outperforms the deterministic algorithm with respect to runtime and memory requirement.

The well-established algorithm of the fewest switches is comparable to the single switch results in the short-time regime, but does not reproduce the long-time oscillations of the population. However, it approximates the energy of the  $S_1$ -state with an average relative error of 1%, which has to be compared with an average relative error of 3% for the single switch approaches.

In summary, the probabilistic single switch approach obtains the best results among the considered algorithms with respect to accuracy and efficiency. A detailed study of the algorithm's performance for other systems than the pyrazine model has to be the subject of future research.

## ACKNOWLEDGMENTS

We thank B. Schmidt for providing his implementation of the fewest switches algorithm.

## APPENDIX: SYMMETRY OF THE REFERENCE SOLUTION

Assume an initial datum of the form  $\psi_0 = \psi_0^+ \chi^+$ , where  $\chi^+$  denotes the real-valued eigenvector (6). We define the operator  $J$ , which acts as

$$J \begin{pmatrix} \psi_1(q_1, q_2, q_3) \\ \psi_2(q_1, q_2, q_3) \end{pmatrix} := \begin{pmatrix} \psi_1(q_1, q_2, -q_3) \\ -\psi_2(q_1, q_2, -q_3) \end{pmatrix}.$$

Using  $\theta_{(q_1, q_2, -q_3)} = -\theta_{(q_1, q_2, q_3)}$ , it is easily seen that  $J\psi_0 = \psi_0$ . Moreover,  $J$  commutes with the Schrödinger operator  $H = -(h^2/2)\Delta_q + V$ , that is,  $JH = HJ$ . As a consequence we have

$$J e^{-iHt/h} = e^{-iHt/h} J$$

and  $J\psi(t) = \psi(t)$ . Thus,

$$\begin{pmatrix} \psi_1(q_1, q_2, q_3, t) \\ \psi_2(q_1, q_2, q_3, t) \end{pmatrix} = \begin{pmatrix} \psi_1(q_1, q_2, -q_3, t) \\ -\psi_2(q_1, q_2, -q_3, t) \end{pmatrix}.$$

Combining this with the explicit form of the eigenvectors, one obtains that the level functions fulfill  $\psi^\pm(q_1, q_2, -q_3, t) = \pm \psi^\pm(q_1, q_2, q_3, t)$ . Hence, the expectation values of position and momentum in the third direction are zero for all times.

<sup>1</sup>G. Worth, H. Meyer, and L. Cederbaum, in *Conical Intersections*, edited by W. Domcke, D. Yarkony, and H. Köppel (World Scientific, Singapore, 2004).

<sup>2</sup>D. Shalashilin and M. Child, *Chem. Phys.* **304**, 103 (2004).

<sup>3</sup>X. Chen and V. Batista, *Photochem. Photobiol.* **190**, 274 (2007).

<sup>4</sup>C. Lasser, T. Swart, and S. Teufel, *Commun. Math. Sci.* **5**, 789 (2007).

<sup>5</sup>C. Fermanian-Kammerer and C. Lasser, *J. Chem. Phys.* **128**, 144102 (2008).

<sup>6</sup>C. Fermanian-Kammerer and P. Gérard, *Ann. Henri Poincaré* **4**, 513 (2003).

<sup>7</sup>C. Lasser and S. Teufel, *Commun. Pure Appl. Math.* **224**, 113 (2005).

<sup>8</sup>C. Fermanian-Kammerer and C. Lasser, *SIAM J. Math. Anal.* **40**, 103 (2008).

<sup>9</sup>J. Tully, *J. Chem. Phys.* **93**, 1061 (1990).

<sup>10</sup>A. Voronin, J. Marques, and A. Varandas, *J. Phys. Chem. A* **102**, 6057 (1998).

<sup>11</sup>J. Tully and R. Preston, *J. Chem. Phys.* **55**, 562 (1971).

<sup>12</sup>R. Schneider and W. Domcke, *Chem. Phys. Lett.* **150**, 235 (1988).

<sup>13</sup>M. Prasad, *Chem. Phys. Lett.* **194**, 27 (1992).

<sup>14</sup>G. Stock, *J. Chem. Phys.* **103**, 2888 (1995).

<sup>15</sup>J. Fang and S. Hammes-Schiffer, *J. Phys. Chem. A* **103**, 9399 (1999).

<sup>16</sup>G. Stock and M. Thoss, *Adv. Chem. Phys.* **131**, 243 (2005).

<sup>17</sup>J. Dormand and P. Prince, *J. Comput. Appl. Math.* **6**, 19 (1980).