Semiclassical analysis of the quantum instanton approximation

Christophe L. Vaillant (iD), Manish J. Thapa (iD), Jiří Vaníček (iD), and Jeremy O. Richardson (iD)

View Online   Export Citation   CrossMark

Semiclassical analysis of the quantum instanton approximation

View Online   Export Citation   CrossMark

Christophe L. Vaillant,[1,a] (iD) Manish J. Thapa,[2] (iD) Jiří Vaníček,[1,b] (iD) and Jeremy O. Richardson[2,c] (iD)

AFFILIATIONS

[1] Laboratory of Theoretical Physical Chemistry, Institut des Sciences et Ingénierie Chimiques, École Polytechnique
  Fédérale de Lausanne (EPFL), CH-1015 Lausanne, Switzerland
[2] Laboratory of Physical Chemistry, ETH Zürich, 8093 Zürich, Switzerland

[a] Electronic mail: christophe.vaillant@epfl.ch
[b] Electronic mail: jiri.vanicek@epfl.ch
[c] Electronic mail: jeremy.richardson@phys.chem.ethz.ch

ABSTRACT

We explore the relation between the quantum and semiclassical instanton approximations for the reaction rate constant. From the quantum instanton expression, we analyze the contributions to the rate constant in terms of minimum-action paths and find that two such paths dominate the expression. For symmetric barriers, these two paths join together to describe the semiclassical instanton periodic orbit. However, for asymmetric barriers, one of the two paths takes an unphysically low energy and dominates the expression, leading to order-of-magnitude errors in the rate predictions. Nevertheless, semiclassical instanton theory remains accurate. We conclude that semiclassical instanton theory can be obtained directly from the semiclassical limit of the quantum instanton for symmetric systems. We suggest a modification of the quantum instanton approach which avoids sampling the spurious path and thus has a stronger connection to semiclassical instanton theory, giving numerically accurate predictions even for very asymmetric systems in the low temperature limit.

I. INTRODUCTION

The inclusion of nuclear quantum effects in molecular dynamics calculations is a challenging task. One of the main difficulties in calculating exact quantum dynamics is the need to solve the time-dependent Schrödinger equation with many degrees of freedom. For atomistic descriptions of most molecular systems, the dynamics are (currently) computationally impossible to solve exactly due to the exponentially increasing size of the basis sets required for these calculations. For this reason, approximate quantum dynamics methods based on path integrals are increasingly favored.

A rigorous formulation of reaction rate theory is provided by the flux-side or flux-flux time correlation function formalism.[1] These correlation functions can be evaluated with approximate quantum-dynamics methods, including linearized semiclassical initial-value representation (LSC-IVR),[2] centroid molecular dynamics,[3,4] ring-polymer molecular dynamics (RPMD),[5–7]

and the recently developed Matsubara dynamics[8,9] and its approximations.[10,11] The main approximation made by these methods is to neglect quantum coherences, which is not expected to be a cause of significant error for complex systems in thermodynamic equilibrium.[2] Thus, solving the system exactly to obtain the long-time dynamics is, in many cases, unnecessary as the relevant processes can be approximated to a good accuracy at short times. In fact, if the problem is well defined, no time propagation is necessary at all and one can turn to so-called quantum transition state theories (QTSTs), where the expressions involving time correlation functions are approximated using only their properties at zero time.[12,13]

Semiclassical instanton (SCI) theory provides one of the simplest formulations of a QTST for multidimensional tunneling.[14,15] It can be derived rigorously by taking asymptotic ($\hbar \to 0$) approximations to the flux-flux correlation function[16] and is known in a number of formulations,[17] including one obtained from the

"ImF" approximation,[18] which can all be shown to be equivalent.[19] SCI theory is defined in terms of the dominant tunneling pathway located on the full-dimensional potential-energy surface and employs a harmonic approximation for fluctuations around this pathway. The ring-polymer instanton approach[20–22] provides a computationally efficient algorithm for applying the theory in practice and can be automated for application to complex systems.[23]

Although SCI theory is very powerful, the harmonic expansion about a single dominant tunneling path can introduce errors, especially for low-frequency anharmonic systems. An alternative method, known as the quantum instanton (QI),[24] was developed as an attempt to correct the quantitative deficiencies of the semiclassical approximation. This method goes beyond the harmonic approximation employed by SCI theory and samples paths using efficient path-integral Monte Carlo or molecular dynamics methods.[25] The QI method has been used to calculate reaction rates in many multidimensional systems[26–28] and is a particularly efficient approach for the direct calculation of kinetic isotope effects.[29–32] However, for large asymmetries (meaning the reaction is highly exothermic or endothermic), the rate constant predicted by the QI method has been observed to give large errors, whereas SCI theory remains well-behaved.[17,33]

There are also a number of other path-integral-based QTSTs developed to go beyond the SCI approximation. Unlike in classical mechanics,[34] it is not possible to derive them directly from the flux-side correlation function, as this function tends to zero at zero time if used in its standard form. Instead, QTSTs have been proposed based on a connection to semiclassical instanton theory, including a free-energy version of instanton theory[35] and ring-polymer transition-state theory, which is itself related to the RPMD rate theory.[21] More recently, ring-polymer transition-state theory has been rederived from a generalized flux-side correlation function, which yields a good approximation to the rate at zero time.[36,37] This provides an extension of the centroid-based QTST of Voth, Chandler, and Miller,[38] a method which does not dominantly sample the instanton and thus fails for asymmetric barriers.[39]

In summary, it appears that the QTSTs which work best are those with the strongest connection to the SCI. In this article, we examine the relation between the QI and SCI theories in order to investigate the very plausible conjecture that the SCI is a semiclassical/steepest-descent approximation to the QI; put another way, the QI method is an improvement on the SCI with more accurate sampling of the paths. Surprisingly, we find that the two methods are not directly connected to each other except in the case of symmetric systems. We perform a semiclassical analysis of the quantities used in the QI method and find that the dominant paths that contribute are not the same as those which define the instanton periodic orbit. This leads to the observed error of the QI method for systems with large asymmetries and low temperatures. We then suggest an approach for modifying the QI expression to enforce sampling of the instanton orbit, which is seen to greatly improve the results.

This paper is organized as follows. We first give a derivation of the QI method from first principles in Sec. II, followed by a short investigation of the breakdown of the predicted rate for asymmetric systems. Motivated by this breakdown, we analyze the

contributions from semiclassical paths in Sec. III, and we show that the SCI expression can only be recovered if the spurious paths are removed. We suggest adding a projection operator to solve this problem and describe an improved approach, which we call the projected quantum instanton method, in Sec. IV. We follow this with a discussion of the numerical results of the various methods in Sec. V, including the semiclassical approximations, which justifies the analysis in terms of semiclassical paths, before concluding in Sec. VI.

II. QUANTUM INSTANTON APPROXIMATION

In Sec. II A, we re-examine the derivation of the QI method for calculating reaction rates. The derivation is based on a saddle-point approximation to the time integral of the flux-flux correlation function, and we show that the method breaks down for asymmetric barriers at low temperatures. The goal is to have a self-contained discussion of the QI method before undertaking a more detailed semiclassical analysis. There are two versions of the QI method: the original, which we shall continue to call QI, and a variant, which we call the second-order cumulant expansion (2OCE).

Although we limit ourselves to one dimension throughout the paper for simplicity, multidimensional extensions of the QI method exist.[24–26,29] Our conclusions remain valid for these multidimensional cases.

A. Derivation

We begin by rederiving the QI approximation from the Miller–Schwartz–Tromp expression for the reaction rate constant,[1]

$$kQ_r = \frac{1}{2} \int_{-\infty}^{\infty} dt \, C_{ff}(t). \tag{1}$$

Here, Q_r is the reactant partition function and C_{ff} is the symmetrized flux-flux time correlation function,

$$C_{ff}(t) = \text{Tr}\left(e^{-\beta\hat{H}/2} \hat{F}_1 e^{-\beta\hat{H}/2} e^{i\hat{H}t/\hbar} \hat{F}_2 e^{-i\hat{H}t/\hbar} \right), \tag{2}$$

generalized to two dividing surfaces. The Hamiltonian for a particle in a one-dimensional potential, $V(x)$, is $\hat{H} = \hat{p}^2/2m + V(\hat{x})$, and the flux operators are given by

$$\hat{F}_j = \frac{1}{2m}\left(\hat{\delta}_j \hat{p} + \hat{p} \hat{\delta}_j \right) \qquad (j \in \{1, 2\}), \tag{3}$$

where m is the mass and $\hat{\delta}_j = \delta(\hat{x} - x_j)$ indicates a Dirac delta function centered at the jth dividing surface located at $x = x_j$. Because both $\hat{H}$ and $\hat{F}_j$ are Hermitian as well as real operators, the correlation function [Eq. (2)] is a real and even function of t.

The main idea used in deriving the QI expression is to approximate the time integral in Eq. (1) using a steepest-descent (saddle-point) approximation taken around $t = 0$,

$$\int_{-\infty}^{\infty} dt \, e^{f(t)} \approx \int_{-\infty}^{\infty} dt \, e^{f(0) + \ddot{f}(0)t^2/2} = \sqrt{\frac{2\pi}{-\ddot{f}(0)}} e^{f(0)}, \tag{4}$$

where the dot implies differentiation with respect to time. In order to use this approximation, it is necessary that $\dot{f}(0) = 0$, which is guaranteed for even functions of t, and also that $\ddot{f}(0) < 0$. This steepest-descent approximation makes it possible to express the full reaction rate in terms of expressions defined at $t = 0$, which can then be evaluated exactly with imaginary-time path-integral methods. The original QI expression [see Eq. (14) below] has been derived with one of two approaches: one which uses the energy-integral formulation of the reaction rate (with complicated transformations and two separate steepest-descent integrations)[24] and another which involves multiplying and dividing by the delta-delta correlation function,

$$C_{dd}(t) = \text{Tr}\left(e^{-\beta \hat{H}/2} \hat{\delta}_1 e^{-\beta \hat{H}/2} e^{i\hat{H}t/\hbar} \hat{\delta}_2 e^{-i\hat{H}t/\hbar} \right), \quad (5)$$

as the first step.[29,40] Both derivations use the same prescription for choosing the dividing surfaces x_1 and x_2.

Evaluating the trace in Eq. (2) in the position representation, we find the well-known expression[1,40]

$$C_{ff}(t) = \left(\frac{\hbar}{2m} \right)^2 \left(\frac{\partial^2 \rho}{\partial x_1 \partial x_2} \rho^* - \frac{\partial \rho}{\partial x_1} \frac{\partial \rho^*}{\partial x_2} \right) + \text{c.c.}, \quad (6)$$

where "c.c." denotes the complex conjugate, and

$$\rho \equiv \rho(x_1, x_2, t) = \left\langle x_2 \left| e^{-\beta \hat{H}/2 - i\hat{H}t/\hbar} \right| x_1 \right\rangle. \quad (7)$$

The delta-delta correlation function [Eq. (6)] is then simply

$$C_{dd}(t) = |\rho(x_1, x_2, t)|^2. \quad (8)$$

For brevity, we will drop the explicit dependence of the time correlation functions on the dividing surfaces in our notation. To apply the steepest-descent approximation [Eq. (4)] to Eq. (1), we set

$$f_{ff}(t) = \ln C_{ff}(t), \quad (9)$$

which we note obeys the stationarity condition at time $t = 0$ because $\dot{C}_{ff}(0) = 0$. We are still free to choose the locations of the dividing surfaces, (x_1, x_2), and we should do this in such a way to ensure that $C_{ff}(t)$ has a maximum at $t = 0$. Many choices for the dividing surfaces have been suggested, but some choices lead to a local minimum at zero time for $C_{ff}(t)$, rather than a maximum.[40] The original and most common choice is that the dividing surfaces obey

$$\frac{\partial \rho(x_1, x_2, 0)}{\partial x_1} = \frac{\partial \rho(x_1, x_2, 0)}{\partial x_2} = 0, \quad (10)$$

which typically leads to $C_{ff}(t)$ being a maximum at $t = 0$. In our experience, there is always a solution to these equations with $x_1 = x_2$, known as merged dividing surfaces, which describe either a minimum or first-order saddle point of $\rho(x_1, x_2, 0)$. In the former case, another set of solutions exists as saddle points with $x_1 \neq x_2$, leading to what is known as split dividing surfaces. Adopting one of these choices of dividing surfaces such that x_1 and x_2 satisfy Eq. (10), we note that $C_{ff}(0) = C_{ff}^{sp}(0)$, where

$$C_{ff}^{sp}(t) = \left(\frac{\hbar}{2m} \right)^2 \frac{\partial^2}{\partial x_1 \partial x_2} C_{dd}(t), \quad (11)$$

and thus suggest the approximation $C_{ff}(t) \approx C_{ff}^{sp}(t)$ for $t \neq 0$.

We are now faced with a choice for performing the steepest-descent integration in time: we could either perform this approximate time integral on $C_{dd}(t)$ and then take the spatial derivatives, or we could perform the integral directly on $C_{ff}(t)$. Choosing first to perform the integral over time before evaluating the derivatives with respect to the positions of surfaces, the new function for the steepest-descent approximation and its second time derivative (in terms of the energy variance ΔH_{dd}^2) are

$$f_{dd}(t) = \ln C_{dd}(t), \quad (12a)$$

$$\left. \frac{d^2 f_{dd}}{dt^2} \right|_{t=0} = \frac{\ddot{C}_{dd}(0)}{C_{dd}(0)} = -\frac{2}{\hbar^2} \Delta H_{dd}^2, \quad (12b)$$

where $\dot{C}_{dd}(0) = 0$ and

$$\ddot{C}_{dd}(t) = 2|\dot{\rho}|^2 + \ddot{\rho}\rho^* + \rho\ddot{\rho}^*. \quad (13)$$

The QI rate is then defined as

$$k_{QI}Q_r = \frac{1}{2} \left(\frac{\hbar}{2m} \right)^2 \frac{\partial^2 C_{dd}(0)}{\partial x_1 \partial x_2} \int_{-\infty}^{\infty} dt\, e^{-\Delta H_{dd}^2 t^2 / \hbar^2}$$

$$= \frac{\hbar\sqrt{\pi}}{2\Delta H_{dd}} C_{ff}^{sp}(0), \quad (14)$$

where we have assumed that the spatial derivatives of ΔH_{dd} can be neglected as they are much smaller than the derivatives of $C_{dd}(0)$. Provided that the ratio $C_{ff}(t)/C_{dd}(t)$ is slowly varying in time, Eq. (14) can be generalized for dividing surfaces that do not obey the condition in Eq. (10) such that[40]

$$k_{QI}Q_r = \frac{\hbar\sqrt{\pi}}{2\Delta H_{dd}} C_{ff}(0), \quad (15)$$

where we have slightly abused the notation and kept the QI subscript in this final expression.

An alternative approach is derived by employing a saddle-point approximation directly to the time integral of $C_{ff}(t)$. In order to use this approximation, there is no particular requirement on the choice of dividing surfaces, provided $C_{ff}(t)$ is a maximum at $t = 0$ for that choice. The idea of this direct integration has previously been introduced in the context of creating a "cumulant expansion," where the expansion is truncated at the second order term[41] (although it could be extended to higher order terms and, in addition, the exact high-temperature behavior could be incorporated analytically).[41–43] The steepest-descent function is then

$$f_{ff}(t) = \ln C_{ff}(t), \quad (16a)$$

$$\left. \frac{d^2 f_{ff}}{dt^2} \right|_{t=0} = \frac{\ddot{C}_{ff}(0)}{C_{ff}(0)} = -\frac{2}{\hbar^2} \Delta H_{ff}^2, \quad (16b)$$

where

$$\ddot{C}_{\text{ff}}(t) = \left(\frac{\hbar}{2m}\right)^2 \left(\frac{\partial^2 \rho}{\partial x_1 \partial x_2}\ddot{\rho}^* + 2\frac{\partial^2 \dot{\rho}}{\partial x_1 \partial x_2}\dot{\rho}^* + \frac{\partial^2 \ddot{\rho}}{\partial x_1 \partial x_2}\rho^*\right.$$
$$\left. - \frac{\partial \ddot{\rho}}{\partial x_1}\frac{\partial \rho^*}{\partial x_2} - 2\frac{\partial \dot{\rho}}{\partial x_1}\frac{\partial \dot{\rho}^*}{\partial x_2} - \frac{\partial \rho}{\partial x_1}\frac{\partial \ddot{\rho}^*}{\partial x_2}\right) + \text{c.c.}, \quad (17)$$

which defines the energy variance ΔH_{ff}^2. The second-order cumulant expansion (2OCE) expression that results from this procedure is very similar to Eq. (15), namely,

$$k_{\text{2OCE}}Q_{\text{r}} = \frac{1}{2}C_{\text{ff}}(0)\int_{-\infty}^{\infty}dt\, e^{-\Delta H_{\text{ff}}^2 t^2/\hbar^2}$$
$$= \frac{\hbar\sqrt{\pi}}{2\Delta H_{\text{ff}}}C_{\text{ff}}(0). \quad (18)$$

We will show that both of these quantum instanton approximations have similar properties and break down for asymmetric systems for similar reasons.

B. Behavior of the quantum instanton method for asymmetric systems

To test the approximations, it is useful to apply the QI and 2OCE methods to a simple 1D Eckart barrier,

$$V(x) = \frac{(1-\alpha)V_0}{1+e^{-2ax}} + \frac{(1+\sqrt{\alpha})^2 V_0}{4\cosh^2(ax)}, \quad (19)$$

where, choosing the same parameters as in Ref. 24, $V_0 = 0.425$ eV, $a = 1.36\ a_0^{-1}$, and the mass is 1060 m_e (a_0 is the Bohr radius and m_e denotes the electron mass). α is the dimensionless asymmetry parameter controlling the degree of exothermicity between the reactants and products such that the barrier is symmetric for $\alpha = 1$. We investigated a variety of different temperatures and asymmetries, and found that the QI method shows pathological behavior for low temperatures and large asymmetries. We will demonstrate this behavior throughout this paper, unless otherwise stated, with a particular choice of $T = 100$ K and $\alpha = 1.425$, for which the QI prediction is an order of magnitude too large compared to the exact rate.

It would be reasonable to think that the QI method could be improved if a more suitable choice of dividing surfaces were found. We will therefore consider a range of choices for the dividing surfaces, as listed in Table I, with the above choice of parameters, showing examples of both split saddle points and a merged minimum of $C_{\text{dd}}(0)$ and a merged saddle point of $C_{\text{ff}}(0)$.[44] The positions of the dividing surfaces are also indicated in Fig. 1, showing the behavior of the propagator and the flux-flux correlation function as functions of the dividing surface locations. It is possible to choose dividing surfaces that are not at special points,[40] but we will show that it is not generally possible to find dividing surfaces that fix all the problems of the QI and 2OCE methods without needing to evaluate quantities at $t \neq 0$.

To summarize Sec. II A, the QI and 2OCE methods essentially approximate the time dependence of the relevant time correlation functions as Gaussians. The functions $C_{\text{dd}}(t)$ and $C_{\text{ff}}(t)$ are shown in Fig. 2, where the surfaces are chosen to be the split surfaces of ρ if they exist or the merged surfaces of ρ otherwise. For increasing asymmetry α, it is clear that the functions become less and less

TABLE I. Definition and positions of dividing surfaces for the asymmetric Eckart barrier with $\alpha = 1.425$ and $T = 100$ K. The remaining parameters are defined in the main text.

Choice of surfaces	Shorthand	x_1	x_2
Split saddle points of C_{dd}	A	-0.968	0.063
Merged minimum of C_{dd}	B	-0.379	-0.379
Merged saddle point of C_{ff}	C	-0.623	-0.623
Semiclassical turning points	D	-1.847	0.828

Gaussian (a similar situation to that discussed in Refs. 45 and 46 for electron-transfer rates). This will lead to a significant error in the prediction of the rate constant. In fact, we find that none of the choices considered for the dividing surfaces make much difference in the QI prediction for the rate, as the QI prediction still deviates significantly from the exact value. In order to gain insight into this behavior, in Sec. III, we analyze the correlation functions in terms of semiclassical paths and discover the cause of the breakdown of the approximation.
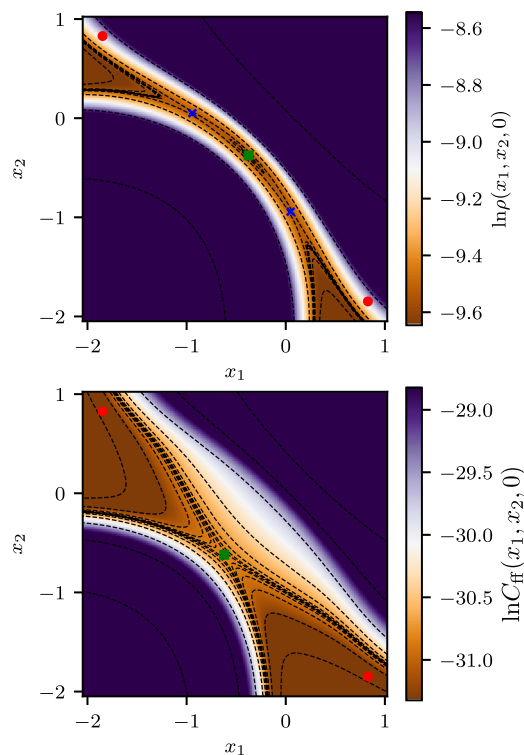

FIG. 1. Contour plots of $\rho(x_1, x_2, 0)$ and $C_{\text{ff}}(0)$ for the asymmetric barrier at 100 K and $\alpha = 1.425$, plotted as functions of the dividing surface locations. The saddle points of ρ [and therefore $C_{\text{dd}}(0)$, which is simply the square of ρ] that define the split surfaces are shown as blue crosses, the semiclassical instanton turning points are shown as red circles, and the merged surfaces locations of either $C_{\text{dd}}(0)$ or $C_{\text{ff}}(0)$ are shown by the green squares. These coordinates are given in Table I. All quantities are in atomic units.
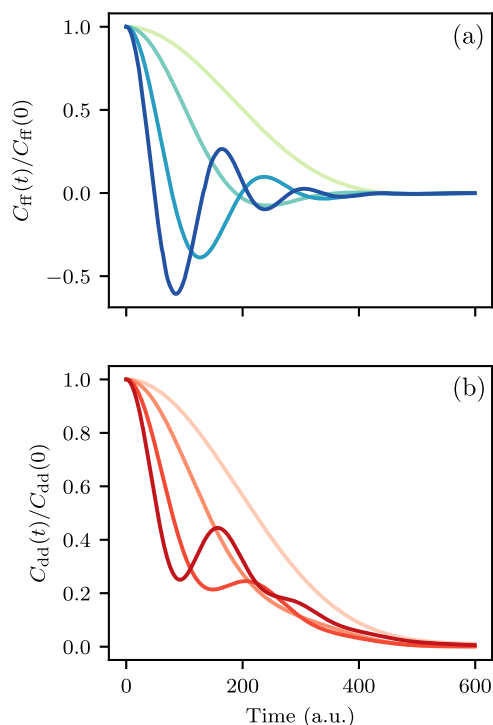
FIG. 2. (a) Flux-flux and (b) delta-delta time correlation functions as functions of time for several values of the asymmetry parameter α. Darker lines indicate larger values of α, with α = 1, 2, 3, and 4. For each value of α, the dividing surfaces are reoptimized with the split saddle points on C_{dd} chosen if they exist, corresponding to A-type surfaces in Table I. In this way, the same surfaces are used for both C_{ff} and C_{dd}, although similar behavior is also seen for other choices of dividing surfaces.

III. SEMICLASSICAL ANALYSIS

In this section, we will examine the various terms in the QI and 2OCE approximations using a semiclassical analysis, which gives the asymptotic behavior in the $\hbar \to 0$ limit of a quantum-mechanical expression in terms of minimum-action paths. We will identify the dominant contributions that cause the observed deviation from the semiclassical instanton expression and determine what improvements would be necessary to fix these problems.

A. Semiclassical contributions to quantum instanton quantities

All terms in the QI and 2OCE rate expressions are defined at $t = 0$, and thus, all of our analysis can be performed using the imaginary-time quantum propagator, $\rho(x_1, x_2, 0)$, and its derivatives. Its semiclassical limit is the imaginary-time van Vleck propagator, which involves a sum over all minimum-action paths which travel from x_1 to x_2 in imaginary time $\beta\hbar/2$. These are clearly the paths which dominate in a path-integral Monte Carlo evaluation of the quantities, in which paths are weighted by the exponential of the negative action. The minimum-action paths follow imaginary-time

classical trajectories, which are equivalent to Newtonian trajectories on the upside-down potential-energy surface.[47] In fact, there are two such minimum-action paths, one which bounces to the left and one to the right. Note that the imaginary-time classical trajectories which travel directly from one dividing surface to the other or which bounce more than once are saddle points of the action[48] and do not contribute to the semiclassical limit. The semiclassical limit of the quantum propagator is thus given by

$$\rho = \left\langle x_2 | e^{-\beta\hat{H}/2} | x_1 \right\rangle \underset{\hbar \to 0}{\sim} K_\ell + K_r, \tag{20}$$

where $\underset{\hbar \to 0}{\sim}$ denotes the asymptotic behavior in the limit $\hbar \to 0$,[49] while K_ℓ and K_r are the semiclassical contributions from the left and right paths, given by

$$K_\gamma = (2\pi\hbar)^{-\frac{1}{2}} \left(-\frac{\partial^2 S_\gamma}{\partial x_1 \partial x_2} \right)^{\frac{1}{2}} e^{-S_\gamma/\hbar}, \tag{21}$$

where $\gamma \in \{\ell, r\}$. These propagators are functions of x_1, x_2 and τ_γ, and in this case, both have the same imaginary-time length $\tau_\gamma = \beta\hbar/2$. The paths are written in terms of the classical Euclidean action (the action in imaginary time),

$$S_\gamma = \int_0^{\tau_\gamma} d\tau' \left\{ \frac{1}{2} m \left| \frac{dX_\gamma(\tau')}{d\tau'} \right|^2 + V[X_\gamma(\tau')] \right\}, \tag{22}$$

where $X_\gamma(\tau')$ are the imaginary-time-dependent left and right paths from x_1 to x_2 via a "bounce."[50] Each action is related to an eikonal W_γ (also known as the reduced action) by the Legendre transform,[17,48]

$$S_\gamma \equiv S_\gamma(x_1, x_2, \tau_\gamma) = W_\gamma(x_1, x_2, E_\gamma) + E_\gamma \tau_\gamma, \tag{23}$$

and

$$W_\gamma \equiv W_\gamma(x_1, x_2, E_\gamma) = \int_{X_\gamma} dx \sqrt{2m[V(x) - E_\gamma]}, \tag{24}$$

where x is integrated along the relevant path X_γ. The energy, E_γ, is chosen to solve $\frac{\partial W_\gamma}{\partial E_\gamma} = -\tau_\gamma$ and the path bounces at the point where $E_\gamma = V(x)$.

The two minimum-action paths are shown for the A, B, and D dividing surfaces in Fig. 3. It is important to notice that the two paths do not join together to form the semiclassical instanton solution, which is an imaginary-time periodic orbit with constant energy. Because the reaction is exothermic, for all choices of dividing surfaces, the left-bouncing path (green) follows half of the instanton periodic orbit, but the right-bouncing path (red) is spurious and has a much lower energy, E_r. This remains true even when using dividing surfaces placed at the turning points of the instanton trajectory (D), as advocated in the appendix of Ref. 24. In all these cases, the second half of the instanton periodic orbit cut at the location of the dividing surfaces is a first-order saddle point of the action, as it passes through a conjugate point,[48] which is why the minimum-action path must follow a different classical trajectory. In fact, we found that for this system below about 142 K there is no way to split the instanton into two trajectories of equal imaginary time without encountering a conjugate point. The picture which we show is inconsistent with
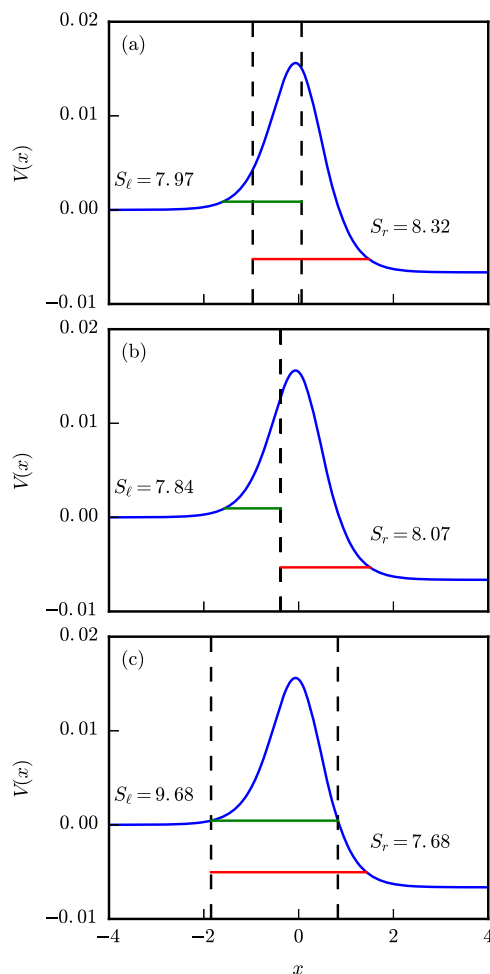
FIG. 3. The two semiclassical minimum-action trajectories are shown plotted at their values of the energy for three different choices of dividing surfaces (indicated by dashed lines). The green trajectory starts at one of the dividing surfaces with initial momentum in the negative direction, bounces against the barrier on the left, and returns to the other dividing surface with positive momentum. The red trajectory travels in the opposite direction and bounces against the right-hand side of the barrier. The dividing surfaces are chosen at (a) the saddle points of the quantum-mechanical $\rho(x_1, x_2, 0)$, (b) the merged-surface minimum of $\rho(x_1, x_2, 0)$, and (c) the semiclassical turning points. These choices correspond to A, B, and D in Table I, respectively. The choice C gives qualitatively the same picture as B. The actions of each semiclassical path are also indicated. In (c), the green line follows one half of the instanton periodic orbit. All quantities are given in atomic units.

the qualitative picture suggested in the appendix of the original QI paper[24] and, as we will show, the spurious path is the crux of the observed error of the QI method in asymmetric systems.

We use this semiclassical analysis in terms of minimum-action paths to find the dominant contributions to the various quantities used in the QI and 2OCE approximations. For instance, the semiclassical limit of $C_{dd}(0)$ is simply

$$C_{dd}(0) \underset{\hbar \to 0}{\sim} (K_\ell + K_r)^2 = K_\ell^2 + 2K_\ell K_r + K_r^2. \tag{25}$$

The semiclassical limits of the spatial first derivatives and mixed second derivative of the imaginary-time propagator are

$$\frac{\partial \rho}{\partial x_1} \underset{\hbar \to 0}{\sim} \frac{1}{\hbar}\left(-p_1^\ell K_\ell + p_1^r K_r\right), \tag{26a}$$

$$\frac{\partial \rho}{\partial x_2} \underset{\hbar \to 0}{\sim} \frac{1}{\hbar}\left(-p_2^\ell K_\ell + p_2^r K_r\right), \tag{26b}$$

$$\frac{\partial^2 \rho}{\partial x_1 \partial x_2} \underset{\hbar \to 0}{\sim} \frac{1}{\hbar^2}(p_1^\ell p_2^\ell K_\ell + p_1^r p_2^r K_r), \tag{26c}$$

where $p_j^\gamma = +\sqrt{2m[V(x_j) - E_\gamma]}$ for $j \in \{1, 2\}$. Note that we take the positive root such that p_j^γ is always a positive scalar and is therefore the magnitude of the momentum (without the direction). It is also clear from Fig. 3 that, contrary to what was previously thought,[24,29] the dividing surfaces according to choice A in Table I are not close to the turning points of the semiclassical instanton, and under a semiclassical analysis, there is no relation between the two choices.

Using Eqs. (6), (20), and (26), we find the semiclassical limit of $C_{ff}(0)$ to be

$$C_{ff}(0) \underset{\hbar \to 0}{\sim} \frac{1}{2m^2}(p_1^\ell + p_1^r)(p_2^\ell + p_2^r)K_\ell K_r. \tag{27}$$

It is important to note that the terms proportional to K_ℓ^2 and K_r^2, which appear in $C_{dd}(0)$, completely cancel out in the semiclassical limit of $C_{ff}(0)$ as a consequence of the flux operators.

The real time derivatives (denoted by dots above quantities, as above) of the semiclassical propagator can be written using the Cauchy-Riemann equations, $\dot{K}_\gamma = i\frac{\partial K_\gamma}{\partial \tau_\gamma}$ and $\dot{E}_\gamma = i\frac{\partial E_\gamma}{\partial \tau_\gamma}$, as

$$\dot{\rho} \underset{\hbar \to 0}{\sim} -\frac{i}{\hbar}(E_\ell K_\ell + E_r K_r) \tag{28}$$

and

$$\ddot{\rho} \underset{\hbar \to 0}{\sim} \frac{1}{\hbar^2}\left(\hbar\frac{\partial E_\ell}{\partial \tau_\ell}K_\ell + \hbar\frac{\partial E_r}{\partial \tau_r}K_r - E_\ell^2 K_\ell - E_r^2 K_r\right). \tag{29}$$

The combination of these two equations and Eq. (13) gives the energy variance of C_{dd} as

$$\Delta H_{dd}^2 \underset{\hbar \to 0}{\sim} \frac{K_\ell K_r (E_\ell - E_r)^2}{(K_\ell + K_r)^2}$$
$$- \frac{\hbar}{K_\ell + K_r}\left(\frac{\partial E_\ell}{\partial \tau_\ell}K_\ell + \frac{\partial E_r}{\partial \tau_r}K_r\right). \tag{30}$$

A similar analysis gives ΔH_{ff} in terms of the semiclassical quantities

$$\Delta H_{ff}^2 \underset{\hbar \to 0}{\sim} \frac{1}{2}\left[(E_\ell - E_r)^2 - \hbar\left(\frac{\partial E_\ell}{\partial \tau_\ell} + \frac{\partial E_r}{\partial \tau_r}\right)\right]. \tag{31}$$

The final expressions for the semiclassical limits of the QI and 2OCE rates are then simply obtained by inserting either Eqs. (27) and (30) into Eq. (15) for the QI result or Eqs. (27) and (31) into Eq. (18) for the 2OCE result.

Note that for less asymmetric systems or at higher temperatures (above 142 K in this case), it is possible to find a new definition for the dividing surface for which the two minimum-action paths have the same energy and thus describe the instanton orbit, as shown in Fig. 4. However, despite the fact that these two paths together describe the correct periodic orbit, we find that the QI approximation nonetheless overpredicts the rate by many orders of magnitude. This is because the left-bouncing path has a much smaller action than the right-bouncing path. As a result, the semiclassical analysis of $C_{\text{ff}}(0)$ given in Eq. (27) is no longer a good estimate of the quantum-mechanical value. The breakdown of the semiclassical analysis is due to the fact that the term proportional to K_ℓ^2 only cancels to first order in $\hbar$ and cannot be neglected when compared with the much smaller $K_\ell K_r$ term. This problem is avoided by choosing dividing surfaces according to the standard prescriptions in Table I, for which K_ℓ and K_r are of a similar order of magnitude, for which the semiclassical analysis is valid.

For a completely symmetric system at any temperature, it is always possible to split the instanton in the middle, thereby generating two minimum-action paths of equal energy with $K_\ell = K_r$. In the following, we shall use this fact to explain the success of the QI method for symmetric systems.

B. Connection to the semiclassical instanton theory

The semiclassical limits to the QI and 2OCE approximations derived above show that the predicted rates do not just depend on the instanton path, but instead have contributions from the spurious path, X_r. However, the SCI result for the rate depends only on the instanton periodic orbit.[17] The implication is that, contrary to the original conjecture, the SCI method is not a simple semiclassical approximation to the QI or 2OCE methods. In fact, it can be seen from both the above expressions and Fig. 3 that the contributions from the unphysical right-hand path will dominate in many situations and lead to erroneous rates. The only exception, when using the standard dividing surfaces satisfying Eq. (10), is for perfectly symmetric barriers where the two paths combine into the instanton periodic orbit.

These problems would go away if ρ were dominated by the semiclassical paths which together make up the instanton periodic
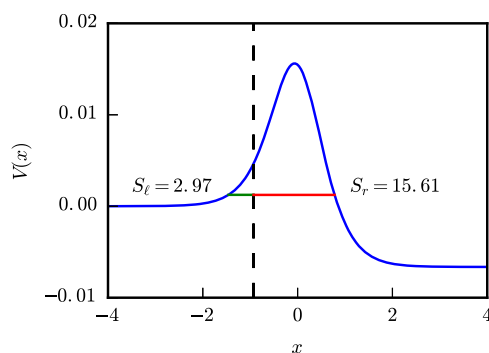
FIG. 4. Two semiclassical minimum-action trajectories at 150 K, with $\alpha = 1.425$, and with dividing surfaces chosen at $x_1 = x_2 = -0.930$ such that two paths combine to form an instanton.

orbit. This is of course already the case for a symmetric system. Here, we will show that if this were generally the case, then the SCI rate formula would be recovered as the asymptotic limit of the QI expression.

If the two paths together describe a periodic orbit, then we can use $E_\ell = E_r$ and hence $p_1^\ell = p_1^r = p_1$ and $p_2^\ell = p_2^r = p_2$. The flux-flux correlation function [Eq. (27)] thus simplifies to

$$C_{\text{ff}}(0) \underset{\hbar \to 0}{\sim} \frac{2p_1 p_2}{m^2} K_\ell K_r \tag{32}$$

and

$$\Delta H_{\text{ff}}^2 \underset{\hbar \to 0}{\sim} -\frac{\hbar}{2}\left(\frac{\partial E_\ell}{\partial \tau_\ell} + \frac{\partial E_r}{\partial \tau_r}\right). \tag{33}$$

In order to study the QI approximation based on ΔH_{dd}, we define dividing surfaces which obey Eq. (10) in the semiclassical limit [see also Eq. (26)], which implies that $K_\ell = K_r$. The semiclassical limit of the energy variance ΔH_{dd} [Eq. (30)] becomes identical to that of ΔH_{ff} given in Eq. (33).

Equations (32) and (33) can thus be used to obtain the semiclassical limit of the QI and 2OCE approximations. Using the known expressions (for the 1D case)[48]

$$-\frac{\partial^2 S_\gamma}{\partial x_1 \partial x_2} = \frac{m^2}{p_1 p_2}\left(\frac{\partial^2 W_\gamma}{\partial E_\gamma^2}\right)^{-1} \tag{34}$$

and

$$\left(\frac{\partial^2 W_\gamma}{\partial E_\gamma^2}\right)^{-1} = -\frac{\partial E_\gamma}{\partial \tau_\gamma}, \tag{35}$$

Eq. (32) becomes

$$C_{\text{ff}}(0) \underset{\hbar \to 0}{\sim} \frac{1}{\pi\hbar}\sqrt{\frac{\partial E_\ell}{\partial \tau_\ell}\frac{\partial E_r}{\partial \tau_r}}\, e^{-(S_\ell + S_r)/\hbar}. \tag{36}$$

Using[16]

$$\hbar\left(\frac{dE}{d\beta}\right)^{-1} = \left(\frac{\partial E_\ell}{\partial \tau_\ell}\right)^{-1} + \left(\frac{\partial E_r}{\partial \tau_r}\right)^{-1}, \tag{37}$$

where E is the energy of the instanton orbit, we obtain

$$k_{\text{sc}} Q_r = \frac{1}{\sqrt{2\pi\hbar^2}}\left(-\frac{dE}{d\beta}\right)^{1/2} e^{-(S_\ell + S_r)/\hbar}, \tag{38}$$

which matches Miller's original SCI expression[14] in the one-dimensional case, bearing in mind that the total action of the instanton periodic orbit is given by $S_\ell + S_r$. This equation is equal to the ImF instanton expression,[18] which can be written in a number of equivalent ways[17,19] including the ring-polymer formulation.[21] Thus, if the right-hand and left-hand paths combine to form the instanton and the dividing surfaces are chosen to satisfy Eq. (10), the QI and 2OCE expressions reduce to the SCI rate in the semiclassical limit. In general, however, the two paths only join exactly into the instanton periodic orbit for a symmetric system; hence, the QI method is a particularly accurate method for symmetric barriers. For asymmetric barriers, the QI rate cannot be directly connected to the SCI expression.

IV. PROJECTED QI

In this section, we suggest a modification of the quantum instanton approach to avoid sampling the spurious paths, which we will call the Projected Quantum Instanton (PQI) method. This defines a novel method in the spirit of the quantum instanton but which should give accurate rate predictions even for strongly asymmetric systems in the low-temperature limit.

In order to fix the QI method for asymmetric systems, it will be necessary to require that the two minimum-action paths have matching energies and thus combine into the instanton periodic orbit. Yet, for very asymmetric systems, it is impossible to choose dividing surfaces such that both minimum-action paths of imaginary-time $\beta\hbar/2$ have the same energy. Therefore, we will need to relax the requirement that the left- and right-bouncing paths have equal imaginary-time lengths[17] and will therefore also need a projection scheme which can categorize any general path into left and right sets. This can be achieved by defining the projected propagator

$$\hat{U}_\gamma(t - i\tau_\gamma) = e^{-i\hat{H}(t-i\tau_\gamma)/2\hbar}\hat{P}_\gamma e^{-i\hat{H}(t-i\tau_\gamma)/2\hbar}, \quad (39)$$

where the projection operators are written in terms of the Heaviside step function, θ, as

$$\hat{P}_\ell = \theta(x_0 - \hat{x}) = \int_{-\infty}^{x_0} dx' |x'\rangle\langle x'| \quad (40a)$$

$$\hat{P}_r = \theta(\hat{x} - x_0) = \int_{x_0}^{\infty} dx' |x'\rangle\langle x'| \quad (40b)$$

such that

$$\hat{U}_\ell(t - i\tau) + \hat{U}_r(t - i\tau) = e^{-i\hat{H}(t-i\tau)/\hbar}. \quad (41)$$

Paths projected in this way can thus be categorized depending on whether their central point is to the right or left of the point x_0. The location of x_0 is somewhat arbitrary for the following arguments as long as it appears between the instanton turning points. This way, $\langle x_2|\hat{U}_\ell(-i\tau_\ell)|x_1\rangle$ will be dominated by the left-bouncing minimum-action path and $\langle x_2|\hat{U}_r(-i\tau_r)|x_1\rangle$ by the right-bouncing minimum-action path. The projection operator we have used is the simplest choice to implement in our simulations. However, there are many other possible definitions which also pick out the correct semiclassical pathways. Note that the multidimensional extension of the approach follows directly by projecting along a reaction coordinate.

By inserting projection operators into the Miller–Schwartz–Tromp formula for the rate [Eq. (1)], we will effectively neglect contributions from pairs of paths which both bounce either to the left or to the right. This is expected to be a good approximation because within a semiclassical analysis these paths would give a zero contribution to the rate.[16,17] We therefore propose to approximate the rate as

$$kQ_r \approx \frac{1}{2}\int_{-\infty}^{\infty} dt\, C_{ff}^{\mathcal{P}}(t), \quad (42)$$

where

$$C_{ff}^{\mathcal{P}}(t) = 2\operatorname{Tr}\left(\hat{U}_\ell(t - i\tau_\ell)\hat{F}_1\hat{U}_r(-t - i\tau_r)\hat{F}_2\right), \quad (43)$$

which is a general (nonsymmetric and complex) correlation function whose form permits the imaginary time to be different for

the two propagators, although their sum, $\tau_\ell + \tau_r = \beta\hbar$, is fixed. This extra flexibility allows us to choose an appropriate dividing surface for which both the left- and right-bouncing paths are minimum-action paths and also join together to describe the instanton. This idea follows a similar approach to that used in a first-principles derivation of the SCI method.[16,17] It is absolutely necessary to project onto left and right paths when using a nonsymmetric split of the imaginary time to avoid finding the spurious minimum-action paths of a right-bouncing path in time τ_ℓ and vice versa. Note that the factor of 2 in Eq. (43) accounts for the alternative ordering of the projection operators, which integrates to the same result.[17]

Equation (42) is an approximation in general, but remarkably it reproduces the exact rate of a free particle as shown in Appendix A. In this important limit, despite the fact that the projected and unprojected correlation functions are not equivalent, their integrals over time are identical.

The final expression for the PQI rate is the second-order cumulant expansion of

$$C_{ff}^{\mathcal{P}}(t) = \frac{\hbar^2}{2m^2}\left(\frac{\partial^2 \rho_\ell}{\partial x_1 \partial x_2}\rho_r^* - \frac{\partial \rho_\ell}{\partial x_1}\frac{\partial \rho_r^*}{\partial x_2} - \frac{\partial \rho_\ell}{\partial x_2}\frac{\partial \rho_r^*}{\partial x_1} + \rho_\ell\frac{\partial^2 \rho_r^*}{\partial x_1 \partial x_2}\right), \quad (44)$$

where the projected matrix elements are given by

$$\rho_\gamma \equiv \rho_\gamma(x_1, x_2, t) = \langle x_2|\hat{U}_\gamma(t - i\tau_\gamma)|x_1\rangle, \quad (45)$$

and their time derivatives calculated using $\dot{\rho}_\gamma = i\frac{\partial \rho_\gamma}{\partial \tau_\gamma}$ and $\ddot{\rho}_\gamma = -\frac{\partial^2 \rho_\gamma}{\partial \tau_\gamma^2}$. The PQI rate expression is then defined identically to the 2OCE approximation [Eq. (18)], except that C_{ff} is replaced by $C_{ff}^{\mathcal{P}}$ throughout.

By following a similar approach as in Sec. III B we obtain the semiclassical limit of $C_{ff}^{\mathcal{P}}(0)$ equal to that of Eq. (32), and of ΔH_{ff}^2 given by Eq. (33). The only difference with the previous analysis is that in PQI the minimum-action paths really do describe the instanton periodic orbit and no extra assumptions need be made. One can therefore show that the semiclassical limit of the rate expression reduces to SCI using the results of Sec. III B. This result is actually unsurprising, as the derivation of SCI follows similar lines of reasoning in Ref. 17.

In principle, one can use any choice of dividing surfaces for which the two halves of the instanton periodic orbit are both minimum-action paths. We chose to place both the dividing surfaces as well as x_0 at the location of the barrier maximum and found that this choice obeyed the rule in each case tested. One should then optimize τ_ℓ (keeping $\tau_\ell + \tau_r = \beta\hbar$ fixed) until $\dot{C}_{ff}^{\mathcal{P}}(0) = 0$. However, the value of τ_ℓ obtained directly from SCI was found to be an excellent approximation to this optimal value.

We note that the steps used to derive the PQI method share some similarity to Wolynes' nonadiabatic quantum-transition theory.[51] In this approach, the two paths are forced to bounce left or right depending on a projection onto the two electronic states and for asymmetric systems may have different imaginary times. The idea for deriving new rate theories by ensuring that the instanton is the dominant path contributing to the rate has also been used

TABLE II. Numerical values (in a.u.) of various quantities used in the QI expressions are calculated using quantum-mechanical (QM) and semiclassical (SC) methods for the specified dividing surfaces. The parameters of the potential are the same as in Table I, and the temperature is 100 K. In each case, the semiclassical version is within a factor of two of the quantum result, confirming that our semiclassical analysis of the quantum instanton is valid.

Div. surf.	$C_{dd}(0)$		$C_{ff}(0)$		ΔH_{dd}		ΔH_{ff}	
	QM	SC	QM	SC	QM	SC	QM	SC
A	5.40×10^{-9}	4.16×10^{-9}	4.60×10^{-14}	3.88×10^{-14}	0.003 23	0.003 23	0.004 37	0.004 48
B	5.12×10^{-9}	4.03×10^{-9}	6.60×10^{-14}	5.59×10^{-14}	0.003 37	0.003 34	0.004 54	0.004 60
C	1.44×10^{-7}	1.31×10^{-7}	4.11×10^{-14}	3.07×10^{-14}	0.001 29	0.001 06	0.004 09	0.004 41
D	1.58×10^{-8}	2.27×10^{-8}	1.61×10^{-14}	3.00×10^{-14}	0.002 31	0.002 95	0.003 51	0.004 03

in previous work on QTST[21,35,36] and most recently in Ref. 52 for nonadiabatic rates.

The PQI method is also applicable to multidimensional problems and can be efficiently computed using the path-integral Monte Carlo approach with a simple extension to the standard methodology. The central bead of each path plays the role of the x' variable in Eq. (40) and, for example, only paths for which $x' < x_0$ should contribute to ρ_ℓ. In this work, however, we implement the projection by integrating numerically over the allowed range of x' and evaluate the imaginary-time propagator, as described in Appendix B.

V. RESULTS AND DISCUSSION

In order to validate our semiclassical analysis of the quantum instanton and cumulant expansion, it is useful to compute the numerical values of the various terms making up these approximations and to compare the values obtained from quantum mechanics (using the eigenfunctions for the Eckart barrier discussed in Appendix B) and from the semiclassical approximation. The results of the quantum and semiclassical calculations for the same system and temperature as those used in Figs. 1 and 3 are compared in Table II. The semiclassical limit is seen to be very close to the quantum values for each of these choices of dividing surfaces, being at most a factor of two different. The discrepancy is of course larger for the dividing surface choice made in Fig. 4, which leads to $K_\ell \gg K_r$. The good agreement is also seen in Table III, where the QI and 2OCE rate predictions are compared using both quantum and semiclassical calculations. However, all these approximations are an order of magnitude larger than the exact rate constant, showing that the source of the error is in the QI or 2OCE rate formulas themselves. (The system and temperature were in fact specifically chosen to demonstrate this order of magnitude error.) This justifies our use of a semiclassical evaluation of the relevant quantities for analyzing the QI results. The PQI method predicts a rate of 4.34×10^{-12} a.u. without optimization of the imaginary-time split (i.e., using the semiclassical ratio of $\tau_r/\tau_\ell \approx 0.258$), but when it is optimized to $\tau_r/\tau_\ell \approx 0.298$, the rate prediction is 4.85×10^{-12} a.u., which is even closer to the exact result (cf. the caption of Table III).

The example system used in Fig. 3 shows a significant error only at low enough temperatures. However, we stress that the error seen in the QI method depends on both the asymmetry and the temperature such that for more asymmetric systems, even high temperatures can result in very large errors. For a more general outlook, Fig. 5

shows the relative error between the exact rate and the rates calculated using the four methods (QI, 2OCE, SCI, and PQI) as a function of asymmetry, α, and for higher temperatures than that considered in Tables II and III.[53] To generate these results, we have generally favored A-type dividing surfaces, although it would have been possible to choose surfaces for the 2OCE calculation that minimize the contributions of higher order terms.[41] Nonetheless, our semiclassical analysis implies that no dividing surface choice will significantly improve these results.

The temperature dependence of the exact and approximate rate constants is shown in Fig. 6(a), along with the corresponding relative errors in Fig. 6(b). While the SCI rates tend to a small constant relative error in the deep tunneling regime,[17,54] the error in the QI rates increases exponentially with asymmetry. We attribute this growing error in the QI rate to the increasing contribution of the right-hand path in Fig. 3. In principle, 2OCE could be improved by going to a higher order expansion;[41] however, in practice, very high orders will typically be required to cancel the spurious dominant contribution at zero time. In contrast, the PQI results in Figs. 5 and 6 show a substantial improvement over both the QI and 2OCE results and are at least as accurate as SCI. For these PQI results, the value of τ_ℓ is not optimized but simply chosen to match the imaginary time for the left bounce in the instanton trajectory.

TABLE III. Prediction of rates (in a.u.) using different methods, where k_{2OCE} is the rate from the cumulant expansion and k_{QI} is the rate from the QI method. The parameters of the potential are the same as in Table I, and the temperature is 100 K. The exact rate is 4.63×10^{-12} a.u., and the semiclassical instanton rate is 3.93×10^{-12} a.u. For comparison, the classical TST rate is 8.31×10^{-26} a.u., which confirms that quantum tunneling effects are of significant importance. Rate constants were evaluated using either the quantum-mechanical (QM) or semiclassical (SC) components from Table II.

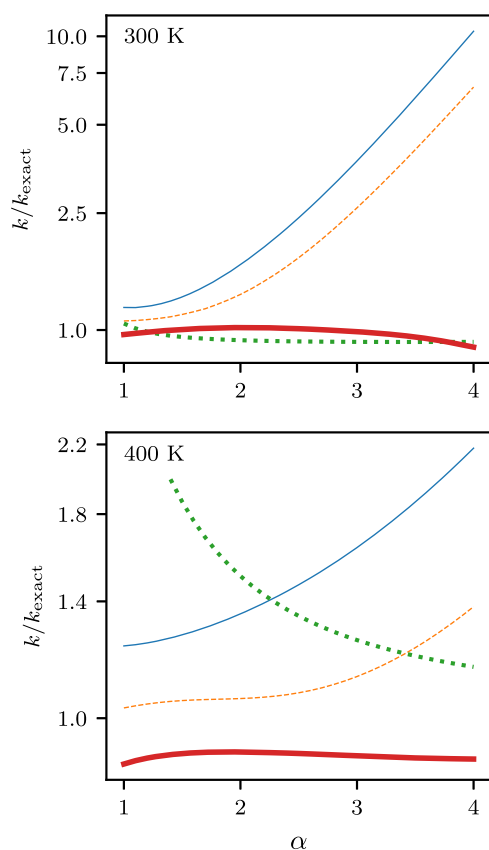| Div. surf.	$k_{QI} \times 10^{12}$		$k_{2OCE} \times 10^{12}$	
	QM	SC	QM	SC
A	54.6	46.0	40.3	33.2
B	75.1	64.1	55.7	46.6
C	123	111	38.6	26.7
D	26.7	39.0	17.5	28.5

FIG. 5. Dependence of relative errors of various approximations for the rate constant on the asymmetry parameter α of the Eckart barrier for two different temperatures. The error is measured with the ratio k/k_{exact} (on a logarithmic scale for clarity), where k_{exact} and k are, respectively, the exact and approximate rate constants. The different approximations shown are the quantum instanton with A-type dividing surfaces (thin blue line), the second-order cumulant expansion with A-type surfaces (thin orange dashed line), the semiclassical instanton (green dashed line), and the projected quantum instanton evaluated with dividing surfaces at the barrier maximum (thick red line). The SCI result is only plotted below the crossover temperature.

FIG. 6. Arrhenius plots of the temperature dependence of (a) the rate constant and (b) the relative error of the rate constant evaluated with various approximations for a fixed asymmetry $\alpha = 1.425$. The error and various line types are defined as in Fig. 5. The crossover temperature for this system is 407 K (β = 776 a.u.).

Close to the crossover temperature, $T_c = \hbar\omega_b/2\pi k_B$ (where ω_b is the absolute value of the imaginary barrier frequency),[17] the SCI approximation breaks down. The accuracy of the PQI rate is also decreased at higher temperatures, although it is not much worse than the original QI and 2OCE methods, a statement which we can quantify by studying the rate of a free particle (and hence the high-temperature limit for arbitrary barriers). In Appendix A, we calculate the PQI rate for this case and our analytical results explain why the PQI prediction slightly underestimates the exact rate in the high-temperature regime. However, we also show that direct integration of $C_{ff}^{\mathcal{P}}(t)$ yields the exact rate for the free particle, suggesting that the only error lies in the steepest-descent approximation.

Overall, these promising results show that the new PQI method is a substantial improvement over the traditional QI and 2OCE
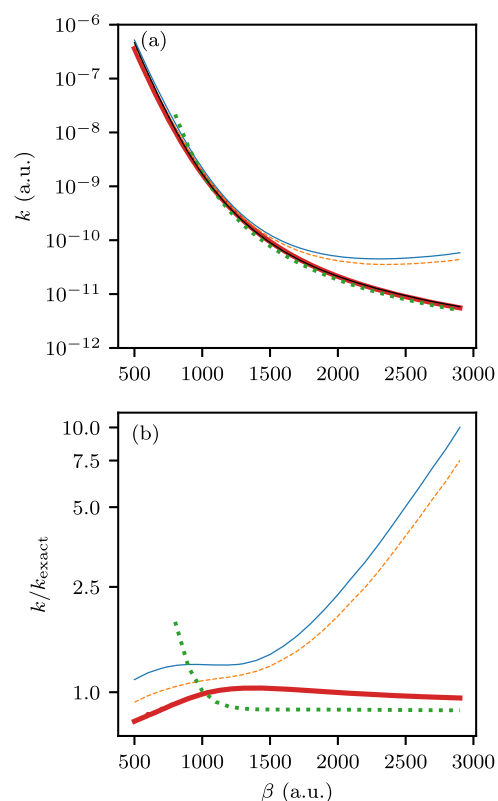
methods. Future extensions of PQI for use with path-integral methods are therefore likely to fulfill the initial promise of QI as an extension to SCI, suitable for accurately predicting deep-tunneling reaction rates in highly anharmonic and asymmetric systems.

VI. CONCLUSIONS

In conclusion, we have presented a semiclassical analysis of the quantum instanton method. We have shown that the dominant contribution to the QI expression can arise from spurious paths, which leads to very large errors for very asymmetric barriers, especially at low temperatures. Consequently, despite conventional wisdom suggesting otherwise, no choice of dividing surface can remove this problem. We justify our analysis by showing that semiclassical evaluation of the relevant quantities yields very similar numerical values to those from exact quantum mechanics for the asymmetric Eckart barrier. The major discrepancy from the exact rate therefore lies with the QI and 2OCE approximations themselves.

From our analysis, we conclude that the SCI rate is the semiclassical limit of the QI rate *only* when the barrier is perfectly symmetric. Therefore, the QI method provides an accurate prediction of the quantum rate for symmetric barriers. However, we find that SCI is much more accurate than QI at describing tunneling in asymmetric

systems, although the SCI clearly cannot describe anharmonicity as well as the QI does.

Our proposed new method, PQI, fixes the inherent path-sampling problems with the QI and 2OCE methods mentioned above, and, at the same time, takes into account anharmonicities that SCI neglects. There exists some similarity between the PQI and theories employed to simulate electron-transfer reactions,[51,52] and we hope that this study will inspire future development of nonadiabatic rate theories. It remains to be seen whether a path-integral application of PQI will become a competitive method for reaction rate calculations of complex systems. Although the theory can be easily generalized to multidimensional systems, it may not be so easy to apply to atomistic simulations of reactions in solution where SCI theory is not well defined.[15] RPMD rate theory[7] also dominantly samples instanton paths even for asymmetric systems[21] and, in addition, includes classical recrossing effects neglected by SCI, QI, and PQI. It thus remains the method of choice for atomistic simulations of reactions in solution.[55]

APPENDIX A: PQI, 2OCE, AND QI RATES FOR THE FREE PARTICLE

It is informative to examine the behavior of PQI for the case of the free particle, which can be treated analytically. The exact propagator for the free particle is well known,[56]

$$\langle x | e^{-iHt/\hbar} | x' \rangle = \sqrt{\frac{m}{2\pi i \hbar t}} \exp\left[\frac{im(x-x')^2}{2\hbar t}\right]. \quad (A1)$$

If we substitute $t \to -i\tau_\gamma$ in Eq. (A1) and then insert Eq. (A1) into Eq. (45), we find

$$\rho_\ell = \frac{m}{\pi \hbar \tau_\ell} \exp\left[-\frac{m}{\hbar \tau_\ell}(x_1^2 + x_2^2)\right]$$
$$\times \int_{-\infty}^{x_0} dx' \exp\left\{-\frac{2m}{\hbar \tau_\ell}\left[(x')^2 - x'(x_1 + x_2)\right]\right\} \quad (A2a)$$

$$\rho_r = \frac{m}{\pi \hbar \tau_r} \exp\left[-\frac{m}{\hbar \tau_r}(x_1^2 + x_2^2)\right]$$
$$\times \int_{x_0}^{\infty} dx' \exp\left\{-\frac{2m}{\hbar \tau_r}\left[(x')^2 - x'(x_1 + x_2)\right]\right\}. \quad (A2b)$$

The integrals over x' can be done analytically with standard results[57] to obtain

$$\rho_\ell = \frac{\rho}{2}[1 + \Phi(z_\ell)], \quad (A3a)$$

$$\rho_r = \frac{\rho}{2}[1 - \Phi(z_r)], \quad (A3b)$$

where $\Phi(z)$ is the Gauss error function and

$$z_\gamma = \sqrt{\frac{m}{2\hbar \tau_\gamma}}(2x_0 - x_1 - x_2). \quad (A4)$$

Note that, as required from Eq. (41), $\rho = \rho_\ell + \rho_r$.

The free-particle delta-delta correlation function is

$$C_{dd}(t) = \frac{m}{\pi \hbar \sqrt{4t^2 + \hbar^2 \beta^2}} \exp\left[\frac{-2m\beta(x_1 - x_2)^2}{4t^2 + \hbar^2 \beta^2}\right], \quad (A5)$$

and the stationarity condition [Eq. (10)] for the dividing surfaces, equivalent to setting $\partial C_{dd}(0)/\partial x_\gamma = 0$, implies $x_1 = x_2$. Taking $x_0 = x_1 = x_2$, we find $\rho_\ell = \rho_r = \rho/2$.

For $x_1 = x_2$, the exact quantum flux-flux correlation function for the free particle is[1]

$$C_{ff}(t) = \frac{1}{2\pi \hbar \beta} \frac{2\hbar^2 \beta^2}{[4t^2 + \hbar^2 \beta^2]^{3/2}}. \quad (A6)$$

Quantities required for the QI rate [Eq. (14)] are thus $\Delta H_{dd} = \sqrt{2}/\beta$ and $C_{ff}(0) = (\pi \hbar^2 \beta^2)^{-1}$. The resulting QI rate,

$$k_{QI} Q_r = \frac{1}{2\pi \hbar \beta} \sqrt{\frac{\pi}{2}}, \quad (A7)$$

is a factor of $\sqrt{\pi/2}$ from the correct free-particle result, $kQ_r = (2\pi \hbar \beta)^{-1}$, leading to an overestimation of the rate by 25%.[24]

For the 2OCE rate (18), the relevant quantity is instead $\Delta H_{ff} = \sqrt{6}/\beta$, giving the rate

$$k_{2OCE} Q_r = \frac{1}{2\pi \hbar \beta} \sqrt{\frac{\pi}{6}}, \quad (A8)$$

which differs from the exact free particle rate by a factor $\sqrt{\pi/6}$ (an underestimation by about 28%).

Likewise, we can calculate the projected correlation function for the free particle using Eqs. (44) and (A3) with the substitution $\tau_\gamma \to \tau_\gamma + it$ to obtain

$$C_{ff}^{\mathcal{P}}(t) = \frac{1}{2\pi \hbar \beta} \frac{\hbar^2 \beta^2}{[4t^2 + \hbar^2 \beta^2]^{3/2}} + \frac{1}{\pi^2(4t^2 + \hbar^2 \beta^2)}. \quad (A9)$$

The relevant quantities needed to calculate the PQI rate for the free particle are therefore

$$C_{ff}^{\mathcal{P}}(0) = \frac{2 + \pi}{2\pi^2 \hbar^2 \beta^2} \quad (A10)$$

and

$$\Delta H_{ff}^{\mathcal{P}} = \frac{\sqrt{2}}{\beta} \sqrt{\frac{4 + 3\pi}{2 + \pi}}. \quad (A11)$$

Therefore, the PQI rate for the free particle is given as

$$k_{PQI} Q_r = \frac{1}{2\pi \hbar \beta} \frac{\sqrt{(2 + \pi)^3}}{\sqrt{8\pi(4 + 3\pi)}}, \quad (A12)$$

which underestimates the exact rate by about 37%.

It is possible to directly integrate $C_{ff}^{\mathcal{P}}(t)$ over t to yield the exact rate for this system. Then, using Eq. (42), we get $kQ_r = (2\pi\hbar\beta)^{-1}$, which is the exact rate for this system.[24] Therefore, for the special case of the free particle, no approximation is made by adding projections in Eq. (42). However, in their simplest forms, QI, 2OCE, and PQI each give a different result for the rate of a free particle, none of which is correct. The errors in these approximate methods stem from the fact that neither the standard nor the projected flux-flux correlation function is a Gaussian function of t, and thus, the steepest-descent approximation introduces an error. This is well understood from the original QI work, and simple corrections have been suggested[24,41,58] which could also be applied to improve the PQI rate formula.

APPENDIX B: ECKART BARRIER EIGENFUNCTIONS

In this appendix, we derive closed-form expressions for the wavefunctions of the one-dimensional Eckart barrier. Eckart[59] presented a derivation for only one wavefunction, which was enough to obtain the reflection and transmission coefficients. For our purpose, however, we will need two orthogonal wavefunctions to serve as a complete basis. In nearly all other respects, we follow his approach.

The Schrödinger equation for this problem can be written as

$$\xi^2 \frac{\partial^2 u}{\partial \xi^2} + \xi \frac{\partial u}{\partial \xi} + \frac{ml^2}{2\pi^2\hbar^2}\left[\frac{A\xi}{1-\xi} + \frac{B\xi}{(1-\xi)^2} + E\right]u = 0, \quad \text{(B1)}$$

where comparison with Eq. (19) gives $\xi = -e^{2\pi x/l}$, $A = V_0(1-\alpha)$, $B = V_0(1+\sqrt{\alpha})^2$, and $l = \pi/a$. Following Eckart, we define

$$\bar{\alpha} = \sqrt{\frac{Eml^2}{2\pi^2\hbar^2}}, \quad \text{(B2a)}$$

$$\bar{\beta} = \sqrt{\frac{(E-A)ml^2}{2\pi^2\hbar^2}}, \quad \text{(B2b)}$$

$$\bar{\delta} = \sqrt{\frac{2Bml^2 - \pi^2\hbar^2}{4\pi^2\hbar^2}}, \quad \text{(B2c)}$$

and then rearrange into the form

$$\xi^2(1-\xi)^2 \frac{\partial^2 u}{\partial \xi^2} + \xi(1-\xi)^2 \frac{\partial u}{\partial \xi}$$
$$+ \left[\bar{\alpha}^2 + (-\bar{\beta}^2 - \bar{\alpha}^2 + \bar{\delta}^2 + 1/4)\xi + \bar{\beta}^2\xi^2\right]u = 0. \quad \text{(B3)}$$

Note that if $E < 0$, the root in Eq. (B2a) is taken such that $\text{Im}\,\bar{\alpha} < 0$. We assume that $A < 0$, i.e., the asymptote on the right-hand side is lower than that on the left.

It is known that the hypergeometric function solves Eq. (B3). Klein[60] showed that solutions to this differential equation can be obtained using the ansatz,

$$u(\xi) = \xi^{-i\bar{\alpha}}(1-\xi)^{1/2+i\bar{\delta}} F[a_0, b_0, c_0, \xi], \quad \text{(B4)}$$

with

$$a_0 = \tfrac{1}{2} + i(-\bar{\alpha} + \bar{\beta} + \bar{\delta}), \quad \text{(B5a)}$$
$$b_0 = \tfrac{1}{2} - i(\bar{\alpha} + \bar{\beta} - \bar{\delta}), \quad \text{(B5b)}$$
$$c_0 = 1 - 2i\bar{\alpha}. \quad \text{(B5c)}$$

However, $F[a_0, b_0, c_0, \xi]$ is only one possible solution to the hypergeometric differential equation, and instead we choose to use Forsyth's solutions[61] called III and XIII to give

$$u_1(\xi) = (-1)^{-i\bar{\alpha}}\xi^{i\bar{\alpha}}(1-\xi)^{1/2+i\bar{\delta}} F\left[\tfrac{1}{2} + i(\bar{\alpha} - \bar{\beta} + \bar{\delta}),\right.$$
$$\left. \tfrac{1}{2} + i(\bar{\alpha} + \bar{\beta} + \bar{\delta}), 1 + 2i\bar{\alpha}, \xi\right], \quad \text{(B6a)}$$

$$u_2(\xi) = (-1)^{i\bar{\alpha}}\xi^{-i\bar{\alpha}}(1-\xi)^{i(\bar{\alpha}-\bar{\beta})} F\left[\tfrac{1}{2} + i(-\bar{\alpha} + \bar{\beta} + \bar{\delta}),\right.$$
$$\left. -i/2(i + 2\bar{\alpha} - 2\bar{\beta} + 2\bar{\delta}), 1 + 2i\bar{\beta}, 1/(1-\xi)\right]. \quad \text{(B6b)}$$

The former is only a solution if $E > A$ and the second if $E > 0$. In order to find the $\xi \to -\infty$ limit, we first apply Abramowitz and Stegun's formula 15.3.7,[62] and after premultiplying by a constant, we find the desired asymptotic limits for u_1 to be

$$\lim_{\xi \to 0} u_1(\xi) = e^{+ik_{\bar{\alpha}}x}, \quad \text{(B7a)}$$

$$\lim_{\xi \to -\infty} u_1(\xi) = C_1 e^{-ik_{\bar{\beta}}x} + D_1 e^{+ik_{\bar{\beta}}x}, \quad \text{(B7b)}$$

where

$$C_1 = \frac{\Gamma(1 + 2i\bar{\alpha})\Gamma(-2i\bar{\beta})}{\Gamma(\tfrac{1}{2} + i(\bar{\alpha} - \bar{\beta} - \bar{\delta}))\Gamma(\tfrac{1}{2} + i(\bar{\alpha} - \bar{\beta} + \bar{\delta}))}, \quad \text{(B8a)}$$

$$D_1 = \frac{\Gamma(1 + 2i\bar{\alpha})\Gamma(2i\bar{\beta})}{\Gamma(\tfrac{1}{2} + i(\bar{\alpha} + \bar{\beta} - \bar{\delta}))\Gamma(\tfrac{1}{2} + i(\bar{\alpha} + \bar{\beta} + \bar{\delta}))}, \quad \text{(B8b)}$$

and $k_{\bar{\alpha}} = 2\pi\bar{\alpha}/l$ and $k_{\bar{\beta}} = 2\pi\bar{\beta}/l$. The asymptotic limits of the second wavefunction are (using Abramowitz and Stegun 15.3.9)[62]

$$\lim_{\xi \to -\infty} u_2(\xi) = e^{-ik_{\bar{\beta}}x}, \quad \text{(B9a)}$$

$$\lim_{\xi \to 0} u_2(\xi) = A_2 e^{-ik_{\bar{\alpha}}x} + B_2 e^{+ik_{\bar{\alpha}}x}, \quad \text{(B9b)}$$

where

$$A_2 = \frac{\Gamma(2i\bar{\alpha})\Gamma(1 + 2i\bar{\beta})}{\Gamma(\tfrac{1}{2} + i(\bar{\alpha} + \bar{\beta} - \bar{\delta}))\Gamma(\tfrac{1}{2} + i(\bar{\alpha} + \bar{\beta} + \bar{\delta}))}, \quad \text{(B10a)}$$

$$B_2 = \frac{\Gamma(-2i\bar{\alpha})\Gamma(1 + 2i\bar{\beta})}{\Gamma(\tfrac{1}{2} + i(-\bar{\alpha} + \bar{\beta} - \bar{\delta}))\Gamma(\tfrac{1}{2} + i(-\bar{\alpha} + \bar{\beta} + \bar{\delta}))}. \quad \text{(B10b)}$$

Note that Eckart's results for reflection and transmission are given by $|B_2/A_2|^2$ and $|\bar{\beta}/\bar{\alpha}|/|A_2|^2$ given a particle incident on the left.

Finally, we normalize the wavefunctions

$$\psi_1(x;E) = \frac{u_1(\xi(x))}{\sqrt{\pi N_1}}, \quad \text{(B11a)}$$

$$\psi_2(x;E) = \frac{u_2(\xi(x))}{\sqrt{\pi N_2}} \quad \text{(B11b)}$$

such that

$$\int dx\, \psi_1(x;E)^* \psi_1(x;E') = \delta(E - E'). \quad \text{(B12)}$$

Using

$$\frac{1}{\pi} \int_0^\infty dx\, e^{i[f(E)-f(E')]x} = \left|\frac{\partial f}{\partial E}\right|^{-1} \delta(E - E'), \quad \text{(B13)}$$

we find the normalization constants should be chosen as

$$N_1 = \frac{1}{k'_{\tilde{\alpha}}} + \frac{|C_1|^2}{k'_{\tilde{\beta}}} + \frac{|D_1|^2}{k'_{\tilde{\beta}}}, \tag{B14a}$$

$$N_2 = \frac{1}{k'_{\tilde{\beta}}} + \frac{|A_2|^2}{k'_{\tilde{\alpha}}} + \frac{|B_2|^2}{k'_{\tilde{\alpha}}}, \tag{B14b}$$

where $k'_{\tilde{\alpha}} = \frac{\partial k_{\tilde{\alpha}}}{\partial E} = k_{\tilde{\alpha}}/(2E)$ and $k'_{\tilde{\beta}} = \frac{\partial k_{\tilde{\beta}}}{\partial E} = k_{\tilde{\beta}}/(2(E - A))$. The two wavefunctions must be orthogonal because one corresponds to an incoming particle from the left and the other from the right. Mathematically, it occurs because $B_2/k'_{\tilde{\alpha}} + C_1^*/k'_{\tilde{\beta}} = 0$.

The propagator can be evaluated numerically as an integral over energies using

$$\langle x_2 | e^{-i\hat{H}t/\hbar} | x_1 \rangle = \int_A^\infty dE\, \psi_1(x_2; E) \psi_1(x_1; E)^* e^{-iEt/\hbar}$$
$$+ \int_0^\infty dE\, \psi_2(x_2; E) \psi_2(x_1; E)^* e^{-iEt/\hbar}, \tag{B15}$$

where t can also be complex in order to obtain the density matrix. Derivatives of the propagator can be evaluated by explicit differentiation of the wavefunctions, although in practice we found that it was simpler and sufficiently accurate to evaluate them with finite differences.

REFERENCES

[1] W. H. Miller, S. D. Schwartz, and J. W. Tromp, J. Chem. Phys. **79**, 4889 (1983).

[2] W. H. Miller, J. Phys. Chem. A **105**, 2942 (2001).

[3] J. Cao and G. A. Voth, J. Chem. Phys. **100**, 5106 (1994).

[4] S. Jang and G. A. Voth, J. Chem. Phys. **111**, 2371 (1999).

[5] I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **121**, 3368 (2004).

[6] I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **122**, 084106 (2005).

[7] I. R. Craig and D. E. Manolopoulos, J. Chem. Phys. **123**, 034102 (2005).

[8] T. J. H. Hele, M. J. Willatt, A. Muolo, and S. C. Althorpe, J. Chem. Phys. **142**, 134103 (2015).

[9] T. J. H. Hele, M. J. Willatt, A. Muolo, and S. C. Althorpe, J. Chem. Phys. **142**, 191101 (2015).

[10] M. J. Willatt, M. Ceriotti, and S. C. Althorpe, J. Chem. Phys. **148**, 102336 (2018).

[11] G. Trenins and S. C. Althorpe, J. Chem. Phys. **149**, 014102 (2018).

[12] W. H. Miller, J. Chem. Phys. **61**, 1823 (1974).

[13] E. Pollak and J. L. Liao, J. Chem. Phys. **108**, 2733 (1998).

[14] W. H. Miller, J. Chem. Phys. **62**, 1899 (1975).

[15] J. O. Richardson, J. Chem. Phys. **148**, 200901 (2018).

[16] J. O. Richardson, J. Chem. Phys. **144**, 114106 (2016).

[17] J. O. Richardson, Int. Rev. Phys. Chem. **37**, 171 (2018).

[18] I. Affleck, Phys. Rev. Lett. **46**, 388 (1981).

[19] S. C. Althorpe, J. Chem. Phys. **134**, 114104 (2011).

[20] S. Andersson, G. Nyman, A. Arnaldsson, U. Manthe, and H. Jónsson, J. Phys. Chem. A **113**, 4468 (2009).

[21] J. O. Richardson and S. C. Althorpe, J. Chem. Phys. **131**, 214106 (2009).

[22] J. B. Rommel, T. P. M. Goumans, and J. Kästner, J. Chem. Theory Comput. **7**, 690 (2011).

[23] C. L. Vaillant, S. C. Althorpe, and D. J. Wales, J. Chem. Theory Comput. **15**, 33 (2019).

[24] W. H. Miller, Y. Zhao, M. Ceotto, and S. Yang, J. Chem. Phys. **119**, 1329 (2003).

[25] T. Yamamoto and W. H. Miller, J. Chem. Phys. **120**, 3086 (2004).

[26] Y. Zhao, T. Yamamoto, and W. H. Miller, J. Chem. Phys. **120**, 3100 (2004).

[27] W. Wang and Y. Zhao, J. Chem. Phys. **130**, 114708 (2009).

[28] T. Yamamoto and W. H. Miller, J. Chem. Phys. **122**, 044106 (2005).

[29] J. Vaníček, W. H. Miller, J. F. Castillo, and F. J. Aoiz, J. Chem. Phys. **123**, 054108 (2005).

[30] K. Karandashev, Z. H. Xu, M. Meuwly, J. Vaníček, and J. O. Richardson, Struct. Dyn. **4**, 061501 (2017).

[31] K. Karandashev and J. Vaníček, J. Chem. Phys. **143**, 194104 (2015).

[32] K. Karandashev and J. Vaníček, J. Chem. Phys. **146**, 184102 (2017).

[33] K. Karandashev, "Accelerating path integral evaluation of equilibrium and kinetic isotope effects," Ph.D. thesis, École Polytechnique Fédérale de Lausanne, 2017.

[34] D. Chandler, J. Chem. Phys. **68**, 2959 (1978).

[35] G. Mills, G. K. Schenter, D. E. Makarov, and H. Jónsson, Chem. Phys. Lett. **278**, 91 (1997).

[36] T. J. H. Hele and S. C. Althorpe, J. Chem. Phys. **138**, 084108 (2013).

[37] S. C. Althorpe and T. J. H. Hele, J. Chem. Phys. **139**, 084115 (2013).

[38] G. A. Voth, D. Chandler, and W. H. Miller, J. Chem. Phys. **91**, 7749 (1989).

[39] D. E. Makarov and M. Topaler, Phys. Rev. E **52**, 178 (1995).

[40] C. Aieta and M. Ceotto, J. Chem. Phys. **146**, 214115 (2017).

[41] M. Ceotto, S. Yang, and W. H. Miller, J. Chem. Phys. **122**, 044109 (2005).

[42] S. Yang, T. Yamamoto, and W. H. Miller, J. Chem. Phys. **124**, 084102 (2006).

[43] C. Predescu, Phys. Rev. E **70**, 066705 (2004).

[44] We were unable to locate an asymmetric system with split saddle points of $C_{ff}(0)$.

[45] P. Huo, T. F. Miller III, and D. F. Coker, J. Chem. Phys. **139**, 151103 (2013).

[46] J. O. Richardson and M. Thoss, J. Chem. Phys. **141**, 074106 (2014).

[47] W. H. Miller, J. Chem. Phys. **55**, 3146 (1971).

[48] M. C. Gutzwiller, *Chaos in Classical and Quantum Mechanics* (Springer-Verlag, New York, 1990).

[49] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1978).

[50] J. O. Richardson, R. Bauer, and M. Thoss, J. Chem. Phys. **143**, 134115 (2015).

[51] P. G. Wolynes, J. Chem. Phys. **87**, 6559 (1987).

[52] M. J. Thapa, W. Fang, and J. O. Richardson, J. Chem. Phys. **150**, 104107 (2019).

[53] The error in the QI rates is slightly larger than those reported in Ref. 25 because they used an *ad hoc* correction to the energy variance for their calculations, but this does not affect our conclusions.

[54] J. O. Richardson, Faraday Discuss. **195**, 49 (2016).

[55] S. Habershon, D. E. Manolopoulos, T. E. Markland, and T. F. Miller III, Annu. Rev. Phys. Chem. **64**, 387 (2013).

[56] R. P. Feynman and A. R. Hibbs, *Quantum Mechanics and Path Integrals* (McGraw-Hill, New York, 1965).

[57] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series and Products*, 6th ed. (Academic Press, San Diego, 2000).

[58] N. F. Hansen and H. C. Andersen, J. Phys. Chem. **100**, 1137 (1996).

[59] C. Eckart, Phys. Rev. **35**, 1303 (1930).

[60] F. Klein, *Vorlesungen über die Hypergeometrische Funktion* (Springer-Verlag, 1933), pp. 3–4.

[61] A. R. Forsyth, *A Treatise on Differential Equations*, 6th ed. (Macmillan, 1948), pp. 214–215.

[62] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions* (U.S. Government Printing Office, 1972).