Quantum mechanical canonical rate theory: A new approach based on the reactive flux and numerical analytic continuation methods

Eran Rabani (a)
School of Chemistry, The Sackler Faculty of Science, Tel Aviv University, Tel Aviv 69978, Israel

Goran Krilov and B. J. Berne (b)
Department of Chemistry, Columbia University, 3000 Broadway, New York, New York 10027

(Received 4 October 1999; accepted 10 November 1999)

We present the reactive flux analytic continuation (RFAC) method, based on the quantum reactive flux formalism combined with a numerical analytic continuation approach to calculate quantum canonical rates in condensed phase systems. We express the imaginary time reactive-flux correlation function in terms of a frequency dependent rate constant, and use path integral formalism to derive a working expression suitable for Monte Carlo simulation techniques. The imaginary time data obtained by simulation is analytically continued to the real time using the maximum entropy method to obtain the reaction rate. Motivated by the success of the method to predict the rates for a simple one dimensional parabolic barrier model, we assess its accuracy for a condensed phase reaction modeled by a double-well coupled to a harmonic bath. We note that the method is applicable to a more general Hamiltonian as long as the reaction coordinate can be identified. The reaction rates computed in this fashion are in very good agreement with analytic and numerically exact results. We demonstrate the applicability of the method for a wide range of model parameters and temperatures. © 2000 American Institute of Physics. [S0021-9606(00)50606-0]

I. INTRODUCTION

The calculation of quantum mechanical reaction rates in condensed phase systems is a long standing problem. One of the ongoing challenges in rate theory is the proper formulation of a quantum thermodynamic theory which provides an estimate for the rate without any time propagation. Unlike the classical thermodynamic formulation, i.e., transition state theory (TST), the quantum mechanical version is not unique, and various authors have tried to formulate a quantum mechanical TST which would share the important properties of the classical version.

In many situations TST provides accurate estimates of the rate constant, however, it can also lead to significant deviation from the exact rate when one or more of its basic assumptions are broken. Namely, in the case when there are recrossings of the dividing surface, deviations from thermal equilibrium at the transition state, or when the reaction coordinate is not separable from the internal degrees of freedom at the saddle point. In such circumstances an exact estimate of the rate requires real time propagation. However, for reactions in condensed phases real time quantum propagation is extremely difficult. Thus, most theoretical work in this field has been based on approximate methods.

One approximate approach to estimate the quantum mechanical rate is based on the centroid density. This approach is very appealing since the centroid density can be estimated using path integral Monte Carlo (PIMC) techniques. Another class of approximate methods is based on replacing the exact quantum mechanical propagator with its semiclassical counterpart. While the semiclassical propagator provides a considerable reduction in the complexity of the real time calculation it does not eliminate completely the oscillatory nature of the quantum mechanical propagator and thus requires a large number of trajectories to obtain converged results. The semiclassical approach becomes practical for the condensed phase system when an additional approximation is introduced, namely when the quantum propagator is replaced by an ‘‘almost’’ classical one. This additional approximation can introduce larger deviations as compared to the fully semiclassical approach. It is still an open challenge to further reduce the computational complexity of the semiclassical method for condensed phase systems.

Another major class of approximate methods is based on mixed quantum classical propagation techniques. The need to propagate a small subset of degrees of freedom quantum mechanically along with a classical propagation of the remaining degrees of freedom makes this approach very attractive. Recently it was shown that the mixed quantum classical treatment of rate processes in a condensed phase environment is insufficient to provide accurate rates, and in some cases the relaxation rates deviate by several orders of magnitude compared to the exact quantum mechanical result.

All the aforementioned approximations are not free of limitations and often break down in various physical situations. Unfortunately, an accurate calculation of the quantum mechanical rate can only be performed for highly simplified models. Thus, it is very important to develop a method that can provide accurate quantum mechanical rates for gen-

A Most of this work was done while ER was at the Department of Chemistry, Columbia University.
B Electronic mail: berne@chem.columbia.edu
eral many-body anharmonic systems. In this paper we propose a new computational method, the reactive flux analytic continuation (RFAC), to calculate quantum reaction rates in condensed phase systems. Our approach combines the well known reactive flux formalism,\textsuperscript{36-40} which was generalized to the quantum mechanical case,\textsuperscript{41} with a numerical analytic continuation method.\textsuperscript{42} Since the flux–flux correlation function typically decays on a relatively fast time scale compared to the rate itself it is expected to provide a good starting point for obtaining the rate using analytic continuation methods.

This article is structured as follows: In Sec. II we provide a general formulation of the analytically continued rate equation. First, we introduce a frequency dependent rate which is related to the real time flux–flux correlation function. Next, we formulate this rate in terms of an imaginary time flux–flux correlation function. This time correlation function that is calculated along the imaginary time axis (also called the Euclidean time axis) can be uniquely analytically continued to the real time axis in order to obtain the rate constant. This mathematical transformation is well defined but rather unstable, and can result in an enormous amplification of the statistical and systematic errors present in the imaginary time data. In Sec. III we describe the maximum entropy (MaxEnt) method\textsuperscript{43,44} used to perform the analytic continuation. The choice of the best inversion method is still in debate,\textsuperscript{45,46} however, for the present problem we find that the MaxEnt method provides significantly better results than the singular value decomposition (SVD) method. In Sec. IV we provide a formulation of the imaginary time flux–flux correlation function which is suitable for PIMC techniques. We separate the calculation into a time dependent part which is efficiently sampled through the use of the reactive flux formalism, and a time independent part which is efficiently sampled through the use of the umbrella sampling method.\textsuperscript{47,48}

We consider two test cases to assess the accuracy of the RFAC method. The first is a simple parabolic barrier model where analytic results for the imaginary time flux–flux correlation function can be obtained. The other test case is a double-well coupled to a harmonic bath; this model was studied recently by Topaler and Makri\textsuperscript{34} and we compare our results with their exact numerical rates. However, unlike their work, we do not use an influence functional approach\textsuperscript{49} since we wish to illustrate that the method is applicable to a general many-body system. We cover a wide range of temperatures and coupling strength for the double-well model. The results for both models are shown and discussed in Sec. V. Conclusions and an outlook are given in Sec. VI.

II. ANALYTIC CONTINUATION OF RATES

In this section we generalize the reactive flux formalism and derive a working expression which is suitable for the analytic continuation method. In the reactive flux formalism, the quantum mechanical rate constant is given in terms of the time integral over the flux–flux correlation function,\textsuperscript{41}

\[ k = \int_0^\infty dt \ C_f(t), \]  

where the real time flux–flux correlation function is given by (from now on we set $\hbar = 1$)

\[ C_f(t) = \frac{1}{Z_r} \text{Tr}(Fe^{-\lambda H}e^{iHt}Fe^{-\lambda H}e^{-(\beta - \lambda)H}), \]  

and $Z_r$ is the partition function of the reactants. In the above equation $\beta = 1/k_B T$ is the inverse temperature and $\lambda$ is introduced for numerical convenience. For open reactive systems, the rate is independent of $\lambda$.\textsuperscript{41} However, the flux–flux correlation function does depend on it. This was the key point in the reduction of the computational complexity of the Yamamoto rate expression.\textsuperscript{50} Although the mathematical justification is not clear, the independence of the rate on $\lambda$ has been commonly used for bound systems as well.\textsuperscript{18,20,21,34,51,52} Nevertheless, our results indicate that numerically the reaction rate does appear to be independent of $\lambda$ even for a bound system like the double-well potential.

In Eq. (2), the symmetrized flux operator, $F$, is given by

\[ F = \frac{1}{2m}(p \delta(s) + \delta(s)p), \]  

where $s$ is the reaction coordinate with a conjugate momentum $p$ and mass $m$. Without loss of generality we assume that the transition state is located at $s = 0$.

It can easily be shown that the rate can be written as

\[ k = \frac{1}{2} \int_{-\infty}^{\infty} dt \ C_f(t), \]  

where $C_f(t)$ is given in Eq. (2). One can define a frequency dependent rate, $k(\omega)$,

\[ k(\omega) = \frac{1}{2} \int_{-\infty}^{\infty} dt \ e^{i\omega t} C_f(t), \]  

such that the rate in Eq. (4) is the zero frequency value of $k(\omega)$. The flux–flux correlation function can be obtained by inverting the Fourier relation in Eq. (5):

\[ C_f(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \ e^{-i\omega t} k(\omega). \]  

The frequency dependent rate is analogous to the spectral density used in the analytic continuation of spectral line shapes, and plays the same role. By performing the replacement $t \rightarrow -i\tau$ we obtain

\[ G_f(\tau) = \frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \ e^{-\omega^2 \tau} k(\omega), \]  

where $\tau, \tau \gg 0$, and

\[ G_f(\tau) = \frac{1}{Z_r} \text{Tr}(Fe^{-(\tau + \lambda)H}Fe^{-(\beta - (\tau + \lambda))H}). \]  

The reason that we have introduced this imaginary time flux–flux correlation function is that, unlike the real time flux–flux correlation function, it is straightforward to obtain it using Monte Carlo simulations even for a many-body system. In Sec. IV we provide a PIMC formulation that permits the calculation of $G_f(\tau)$ for condensed phase systems.
The imaginary time flux–flux correlation function is given by the two-sided Laplace transform of the frequency dependent rate, \( k(\omega) \). Analytic continuation to real time is performed by inverting the integral equation (7) to obtain \( k(\omega) \). This inversion is possible only if \( G_f(\tau) \) is analytic over the desired range of \( \tau \). By choosing a finite value for \( \lambda (0 < \lambda < \beta/2) \) it can be shown that \( G_f(\tau) \) is analytic for \( 0 < \tau < \beta - 2\lambda \). It turns out that for \( \lambda = 0 \) the imaginary time flux–flux correlation function is not analytic at \( \tau = 0 \), which is the reason behind introducing the parameter \( \lambda \) with a small positive value (recall that the rate is independent of this parameter, as will be discussed below in Sec. V).

Note that \( \lambda \) and \( \tau \) are interchangeable. A large value of \( \lambda \) implies a small range of imaginary time and vice versa. For numerical analytic continuation purposes it is desirable for \( G_f(\tau) \) to contain as much information as possible. For this reason it is important to maximize the imaginary time range while ensuring the analyticity of the imaginary time flux–flux correlation function.

It is also useful to note that for a finite \( \lambda \) (specified in the above range) the frequency dependent rate obeys the following detailed balance relation:

\[
k(-\omega) = e^{-(\beta - 2\lambda)\omega}k(\omega). \tag{9}
\]

Using this relation, Eq. (7) can be rewritten as

\[
G_f(\tau) = \frac{1}{\pi} \int_0^\infty d\omega \left[ e^{-\omega\tau} + e^{(\tau - \beta + 2\lambda)\omega} \right] k(\omega). \tag{10}
\]

This form is more suitable for numerical analytic continuation methods since it eliminates the numerical instabilities arising from the negative frequency portion of the two-sided Laplace transform.

### III. MAXIMUM ENTROPY NUMERICAL ANALYTIC CONTINUATION

In this study, we seek to analytically continue the imaginary time flux–flux correlation function given in Eq. (10). Since \( G_f(\tau) \) is analytic for \( 0 < \tau < \beta - 2\lambda \), provided that it is known on this interval, the analytic continuation is accomplished by inverting the integral equation (10) in order to obtain a solution for \( k(\omega) \). The zero mode value \( k(0) \) would then correspond to the experimentally observable reaction rate. Due to the singular nature of the integration kernel the inversion of Eq. (10) is an ill-posed problem. As a consequence, a direct approach to the inversion would lead to an uncontrollable amplification of the statistical noise in the data for \( G_f(\tau) \), resulting in an infinite number of solutions that satisfy Eq. (10). Clearly, in this case, little can be said about real time flux–flux dynamics and the corresponding rates.

Recently, Bayesian ideas have been used to deal with the ill-posed nature of continuing the noisy imaginary time Monte Carlo data to real time. One of the most widely used approaches is the maximum entropy (MaxEnt) method. The method requires only that the transformation which relates the data and the solution be known. Furthermore, MaxEnt allows the inclusion of prior knowledge about the solution in a logically consistent fashion. As such, the method is well-suited for solving ill-posed mathematical problems.

The methods of MaxEnt have been successfully applied in the context of analytic continuation for a variety of quantum systems such as various quantum lattice models, and an excess electron solvated in water, helium, and xenon. More recently, vibrational relaxation and infrared line shapes have been studied as well.

For the purpose of the MaxEnt approach we rewrite the integral equation (10),

\[
D(\tau) = \int d\omega \, K(\tau, \omega)A(\omega). \tag{11}
\]

In this notation \( D(\tau) = G_f(\tau) \) is the data, in this case the imaginary time flux–flux correlation function, \( K(\tau, \omega) = e^{-\omega\tau} + e^{(\tau - \beta + 2\lambda)\omega} \) is the singular kernel, and \( A(\omega) \) is the solution, referred to as the map, corresponding to \( k(\omega) \). Maximum entropy principles provide a way to choose the most probable solution which is consistent with the data through the methods of Bayesian inference. Typically, the data is known only at a discrete set of points \( \{\tau_j\} \), and we likewise seek the solution at a discrete set of points \( \{\omega_k\} \).

The maximum entropy method selects the solution which maximizes the posterior probability, or the probability of the solution \( A \) given a data set \( D \). Using the Bayes theorem one can show that the posterior probability is given by

\[
\mathcal{P}(A|D) \propto \exp(\alpha S - \chi^2/2) = e^Q. \tag{12}
\]

Here \( \chi^2 \) is the standard mean squared deviation from the data

\[
\chi^2 = \sum_{j,k} \left( D_j - \sum_{l} K_{jl}A_l \right) [C^{-1}]_{jk} \left( D_k - \sum_{l} K_{kl}A_l \right), \tag{13}
\]

where \( C_{jk} \) is the covariance matrix,

\[
C_{jk} = \frac{1}{M(M-1)} \sum_{l=1}^{M} (\langle D_j \rangle - \langle D_l \rangle)(\langle D_k \rangle - \langle D_l \rangle), \tag{14}
\]

with \( M \) being the number of measurements. In this study, we find that the covariance matrix is diagonal within the noise level of our simulations, and therefore the expression for \( \chi^2 \) reduces to

\[
\chi^2 = \sum_{j,k} \frac{(D_j - \sum K_{jl}A_l)^2}{\sigma_j^2}, \tag{15}
\]

where \( \sigma_j \) are the standard deviations of each data point, i.e., the square root of the diagonal elements of the covariance matrix.

\( S \) is the information entropy, the form of which is axiomatically chosen to be

\[
S = \sum_k \Delta\omega \left( A_k - m_k - A_k \ln \frac{A_k}{m_k} \right). \tag{16}
\]

In this formulation the entropy is measured relative to a default model \( m(\omega) \) which can contain prior information about the solution and \( \alpha \) is a positive regularization parameter.

Obtaining the MaxEnt solution then involves finding a map \( A \) which maximizes the posterior probability and is therefore a maximization problem in \( N \) variables, where \( N \) is...
the number of points \( \{ \omega_k \} \) at which the solution is evaluated. The solution obtained in this way is still conditional on the arbitrary parameter \( \alpha \), which can be interpreted as a regularization parameter controlling the smoothness of the map. Large values of \( \alpha \) lead to a result primarily determined by the entropy function and hence the default model. Small \( \alpha \) in turn lead to a map determined mostly by the \( \chi^2 \) and thus to a closer fitting of the data. The principal drawback is that, along with the data, the errors would be fit as well.

In this study \( \alpha \) is selected according to the L-curve method.\(^{36,57}\) In this context we regard \( \alpha \) as a regularization parameter controlling the degree of smoothness of the solution, and entropy as the regularizing function. The value of \( \alpha \) is selected by constructing a plot of \( \log[-S(A)] \) vs \( \log \chi^2 \). This curve has a characteristic L-shape, and the corner of the L, or the point of maximum curvature, corresponds to the value of \( \alpha \) which is the best compromise between fitting the data and obtaining a smooth solution.

We employ a maximization algorithm due to Bryan,\(^{58}\) which reduces the space in which the search for the solution is performed. The kernel is first factored using singular value decomposition \( K = \mathbf{V} \Sigma \mathbf{U}^T \). The singular nature of the kernel ensures only a small number of eigenvalues of \( \Sigma \) will be nonsingular. Since the space spanned by the rows of \( K \) is the same as that spanned by the columns of \( \mathbf{U} \) associated with nonsingular eigenvalues, the search for the solution can be performed in this singular space of dimensionality \( N_s \), where \( N_s \) is the number of nonsingular eigenvalues. The solution in singular space is expressed in terms of the vector \( \mathbf{u} \), which is related to the \( N \) dimensional map space via

\[
A_j = m_j \exp \left( \sum_{i=1}^{N_s} U_{ji} u_i \right).
\]

This exponential transformation is useful since it ensures the positivity of the solution.

In this study we use a flat default map, which satisfies a known sum rule, such as the integral over \( k(\omega) \). Other choices of \( m(\omega) \) and their effect on the quality of the results will be the subject of future investigation.

IV. PATH INTEGRAL MONTE-CARLO FORMULATION
OF THE IMAGINARY TIME FLUX–FLUX
CORRELATION FUNCTION

In this section we describe the path integral formulation of the imaginary time flux–flux correlation function. For reasons that will become clearer below, it is computationally more efficient to calculate the reactive-flux imaginary time correlation function given by

\[
G_{fs}(\tau) = \frac{1}{Z_r} \text{Tr}(F e^{-(\tau+i\lambda)H} h(s) e^{-(\beta-(\tau+i\lambda))H}),
\tag{18}
\]

where \( h(s) \) is the Heaviside function,

\[
h(s) = \begin{cases} 
1, & \text{if } s > 0, \\
0, & \text{if } s < 0.
\end{cases}
\tag{19}
\]

The flux–flux correlation function is obtained by taking the time derivative of the reactive-flux correlation function,

\[
G_f(\tau) = i \frac{\partial}{\partial \tau} G_{fs}(\tau).
\tag{20}
\]

Since PIMC methods provide the ratio between two integrals it is useful to multiply and divide \( G_{fs}(\tau) \) by the total partition function \( Z \):

\[
G_{fs}(\tau) = \frac{Z}{Z_r} \times \frac{1}{Z} \text{Tr}(F e^{-(\tau+i\lambda)H} h(s) e^{-(\beta-(\tau+i\lambda))H})
\]

\[
= \frac{Z}{Z_r} G(\tau).
\tag{21}
\]

To evaluate \( G_f(\tau) \) one needs to compute the ratio between the total and reactants partition functions given by

\[
\frac{Z}{Z_r} = \frac{\text{Tr} e^{-\beta H}}{\text{Tr} h(s) e^{-\beta H}},
\tag{22}
\]

which can be calculated using standard cyclic PIMC methods\(^{14,15}\) combined with umbrella sampling techniques.\(^{47,48}\) The other part of the calculation of \( G_f(\tau) \) is more involved (mainly due to the presence of the flux operator) and will be described in the remaining part of this section.

In the coordinate representation \( G(\tau) \) takes the following form:

\[
G(\tau) = \frac{1}{Z} \int ds \int dQ ds' Q^\dagger e^{-\tau H} h(s) e^{-(\beta-\tau)H} |Q s\rangle,
\tag{23}
\]

where \( \tau' = \tau + \lambda \) runs from \( \lambda \) to \( \beta - \lambda \), \( s \) is the reaction coordinate and all other degrees of freedom are labeled by \( Q \). Using the coordinate representation of the flux operator,

\[
\langle s Q | F | Q' s' \rangle = \frac{1}{2m_i} \left[ \delta'(s) \delta(s') - \delta(s) \delta'(s') \right] \delta(Q - Q'),
\tag{24}
\]

it is simple to show that

\[
G(\tau) = \frac{1}{2im\Delta} \int ds \int ds' \int ds'' dQ dQ' \left[ \delta''(s') \delta'(s) - \delta(s) \delta'(s') \right] \times \langle s' Q | e^{-\tau' H} | Q' s'' \rangle h(s'')
\]

\[
\times \langle s'' Q' | e^{-(\beta - \tau') H} | Q s \rangle.
\tag{25}
\]

The time interval can be discretized into \( P \) Trotter slices of size \( \epsilon = \beta/P \) such that \( \tau_j = j \epsilon \), where \( j \) is the slice number. Inserting complete sets of states between the short imaginary time propagators and integrating over the delta functions and their derivatives arising from the flux operator yields

\[
G(\tau) = \frac{1}{2im\Delta} \int dQ_0 \cdots dQ_p ds_1 \cdots ds_p
\]

\[
\times [h(s_j) - h(s_{p+1-j})]
\times \mathcal{P}(Q_0, Q_p | s_0 = \Delta, s_1, \ldots, s_p, s_{p+1} = 0).
\tag{26}
\]

In the above equation, \( \Delta \) arises from the use of a finite difference formula to express the integration over the derivatives of the delta functions,
\[
\int dx \, \delta'(x)f(x) = -f'(0) \approx \frac{f(0) - f(\Delta)}{\Delta},
\]
(27)
and \( \mathcal{P}(Q_0 \cdots Q_P,s_0 \cdots s_{P+1}) \) is given by
\[
\mathcal{P}(Q_0 \cdots Q_P,s_0 \cdots s_{P+1}) = \exp \left\{ \frac{-m}{2} \sum_{j=0}^{P} (s_{j+1} - s_j)^2 - e \sum_{j=0}^{P} V(s_j,Q_j) - \frac{e}{2} [V(s_0,Q_0) + V(s_{P+1},Q_0)] - \sum_{a=1}^{N} \frac{m_a}{2} \sum_{j=0}^{P} (Q_j^{(a)} - Q_{j+1}^{(a)})^2 \right\},
\]
(28)
such that \( Q_{P+1} = Q_0 \). \( N \) is the total number of modes other than the reaction coordinate, \( Q_a \) is the coordinate of the mode \( a \) with a mass \( m_a \) and \( V(s,Q) \) is the total potential energy surface. In the derivation of Eq. (28) we have used a symmetric factorization of the short time propagators. Note that in Eq. (26) \( s_0 \) and \( s_{P+1} \) are constrained to \( \Delta \) and 0, respectively. The total partition function can be expressed in terms of the kernel \( \mathcal{P}(Q_0 \cdots Q_P,s_0 \cdots s_{P+1}) \) as
\[
Z = \int dQ_0 \cdots dQ_P ds_0 \cdots ds_{P+1} \delta(s_0 - s_{P+1}) \mathcal{P}(Q_0 \cdots Q_P,s_0 \cdots s_{P+1}).
\]
(29)
In order to use importance sampling techniques it is useful to rewrite Eq. (26) in the following way:
\[
G(\tau_j) = \left \{ \frac{1}{2i\alpha \Delta} \gamma \delta_j, \right \}
\]
(30)
where \( \delta_j \) is the time dependent part given by
\[
\delta_j = \frac{\int dQ_0 \cdots dQ_P ds_1 \cdots ds_P (h(s_j) - h(s_{j+1})) \mathcal{P}(Q_0 \cdots Q_P,s_0 = \Delta,s_1 \cdots s_P,s_{P+1} = 0)}{\int dQ_0 \cdots dQ_P ds_1 \cdots ds_P \mathcal{P}(Q_0 \cdots Q_P,s_0 = \Delta,s_1 \cdots s_P,s_{P+1} = 0)}.
\]
(31)
and \( \gamma \) is a small time independent factor given by
\[
\gamma = \frac{\int dQ_0 \cdots dQ_P ds_0 \cdots ds_{P+1} \delta(s_0 - \Delta) \delta(s_{P+1}) \mathcal{P}(Q_0 \cdots Q_P,s_0 \cdots s_{P+1})}{\int dQ_0 \cdots dQ_P ds_0 \cdots ds_{P+1} \delta(s_0 - s_{P+1}) \mathcal{P}(Q_0 \cdots Q_P,s_0 \cdots s_{P+1})}.
\]
(32)
The computation of \( G(\tau_j) \) requires the evaluation of the multidimensional integrals in Eqs. (31) and (32). \( \delta_j \) can be computed by a PIMC simulation where the sampling function is given by \( \mathcal{P}(Q_0 \cdots Q_P,s_0 = \Delta,s_1 \cdots s_P,s_{P+1} = 0) \). Note that the initial and final time slices for the reaction coordinate are constrained to \( s_0 = \Delta \) and \( s_{P+1} = 0 \). The first is shifted slightly from the transition state while the other is fixed at the transition state location. This is the key point in the success of the reactive-flux method in the present formulation. During the simulation, the first and final beads are always near the transition state and thus the polymer chain can cross the barrier going from reactants to products, and vice versa, even for very high barriers. The constraints lead to an isomorphic classical system of an open polymer chain along the reaction coordinate, while the other degrees of freedom are cyclic, namely \( Q_{P+1} = Q_0 \). The time dependence of \( \delta_j \) can be computed with one PIMC simulation similar to standard PIMC methods. The time is related to the index \( j \) and different correlation times are generated by taking the average over different beads in a single simulation. This is not the case for the direct computation of the flux–flux correlation function due to the presence of an additional flux operator. We therefore compute the reactive-flux correlation function and take its time derivative numerically to obtain the flux–flux correlation function.

PIMC techniques can be used to evaluate \( \gamma \) as well. For this purpose, we rewrite Eq. (32) as \( \gamma = \gamma_n/\gamma_d \) where
\[
\gamma_n = \frac{\int dQ_0 \cdots dQ_P ds_0 \cdots ds_{P+1} \delta(s_0 - \Delta) \delta(s_{P+1}) \mathcal{P}(Q_0 \cdots Q_P,s_0 \cdots s_{P+1})}{\int dQ_0 \cdots dQ_P ds_0 \cdots ds_{P+1} \delta(s_0 - s_{P+1}) \mathcal{P}(Q_0 \cdots Q_P,s_0 \cdots s_{P+1})},
\]
(33)
and
\[
\gamma_d = \frac{\int dQ_0 \cdots dQ_P ds_0 \cdots ds_{P+1} \delta(s_0 - s_{P+1}) \mathcal{P}(Q_0 \cdots Q_P,s_0 \cdots s_{P+1})}{\int dQ_0 \cdots dQ_P ds_0 \cdots ds_{P+1} \mathcal{P}(Q_0 \cdots Q_P,s_0 \cdots s_{P+1})}.
\]
(34)
Both \( \gamma_n \) and \( \gamma_d \) can be computed simultaneously from a single open chain PIMC simulation with the sampling function given by \( \mathcal{P}(Q_0 \cdots Q_P,s_0 \cdots s_{P+1}) \). However, in this case the initial and final reaction coordinate time slices are not constrained. Since both \( \gamma_n \) and \( \gamma_d \) are time independent we use umbrella sampling techniques to efficiently sample \( \mathcal{P}(Q_0 \cdots Q_P,s_0 \cdots s_{P+1}) \). Note that the evaluation of \( \gamma_n \) and \( \gamma_d \) involves the calculation of the expectation values of the delta functions. In this work, these are approximated by computing the averages of normalized narrow Gaussians.

In summary, to compute the reactive-flux imaginary time correlation function we perform the following PIMC simulations.
The ratio between $Z$ and $Z_r$ is obtained from a standard cyclic PIMC simulation using the umbrella sampling method.

(2) $\gamma$ is computed using an open chain PIMC with a sampling function given by $\mathcal{P}(Q_0 \cdots Q_p, s_0 \cdots s_{p+1})$, and since it is time independent, umbrella sampling methods can be used as well.

(3) $\xi_t$ is the only time dependent factor. It is obtained from an open chain PIMC simulation where the initial and final time slices of the reaction coordinate are fixed, and the sampling function is $\mathcal{P}(Q_0 \cdots Q_p, s_0 = \Delta, s_1 \cdots s_p, s_{p+1} = 0)$.

The flux–flux imaginary time correlation function needed to obtain the reaction rate is generated from the reactive-flux imaginary time correlation function using the relation specified in Eq. (20).

V. RESULTS AND DISCUSSIONS

A. Parabolic barrier

The first test case of our RFAC method is a simple one dimensional parabolic barrier model. The real time flux–flux correlation function for this model is known analytically,\(^{41}\) and the frequency dependent rate can be obtained from a Fourier transform of $C_f(t)$. On the other hand, it is straightforward to obtain analytically the imaginary time flux–flux correlation function and to use it with the MaxEnt inversion method. The real time flux–flux correlation function is given by

$$C_f(t) = -\frac{\omega_b}{8\pi Z_r \sqrt{\sinh(-\omega_b \tau_1) \sinh(\omega_b \tau_2)}} \left(\frac{1}{\sinh(-\omega_b \tau_1)} + \frac{1}{\sinh(\omega_b \tau_2)}\right),$$

where $\omega_b$ is the barrier frequency, $\tau_1 = t - i\lambda$ and $\tau_2 = t + i(\beta - \lambda)$. The imaginary time flux–flux correlation function can be obtained by performing the replacement $t \rightarrow -i\tau$ in Eq. (35),

$$G_f(\tau) = -\frac{\omega_b}{8\pi Z_r \sqrt{\sin(-\omega_b \tau_1) \sin(\omega_b \tau_2)}} \left(\frac{1}{\sin(-\omega_b \tau_1)} + \frac{1}{\sin(\omega_b \tau_2)}\right),$$

where $\tau_1 = \tau + \lambda$ and $\tau_2 = \beta - (\tau + \lambda)$.

In Fig. 1 we plot $k(\omega) Z_r$ for the parabolic barrier at four different temperatures. The exact results were computed from the Fourier transform of $C_f(t)$ given in Eq. (35) with a barrier frequency of $\omega_b = 1$ a.u. and $\lambda = 0.05\beta$. Maximum entropy numerical analytic continuation was performed on mock data generated by adding 1% Gaussian noise [with respect to the maximum value of $G_f(\tau)$] to the analytic results of $G_f(\tau)$ given by Eq. (36). The number of time points was set to 256 for the entire time range. Since $G_f(\tau)$ is symmetric, all useful information is contained in the time range 0 to $\beta/2 - \lambda$. Hence only this portion was used in the analytic continuation procedure. The number of frequency points ranged from 1000 to 1200 and the frequency range was adjusted to ensure converged results. Our default model was a flat distribution with an integral that equals $\pi G_f(0)$.

The analytically continued frequency dependent rates are in excellent agreement with the exact results over the entire frequency range. Moreover, the best agreement is achieved at low frequencies which are expected to have the largest influence on the reaction rates. The latter are in very good agreement with the exact analytic results and the average deviation does not exceed 10%. We also verified that the rates are independent of the value of $\lambda$. Since $k(\omega \neq 0)$ does depend on $\lambda$, it follows that $C_f(t)$ depends on it too, which is consistent with the prediction of Yamamoto.\(^{50}\) The high quality of these results was a motivation to treat real simulation data of a more realistic model which we describe in the next subsection.

B. Double-well coupled to a harmonic bath

In this subsection we use the reactive flux analytic continuation method to calculate the canonical rate constants for a double-well bilinearly coupled to a harmonic bath. The Hamiltonian of our system is given by

$$H = \frac{p^2}{2m} + V(s) + \sum_{\alpha=1}^{N} \left[ \frac{p_{\alpha}^2}{2m_{\alpha}} + \frac{1}{2} m_{\alpha} \omega_{\alpha}^2 \left( Q_{\alpha} - \frac{g_{\alpha} s}{m_{\alpha} \omega_{\alpha}^2} \right)^2 \right].$$

(37)
In the above equation, \( s \) is the reaction coordinate with conjugate momentum \( p \) and mass \( m \); \( Q_{0} \) is the harmonic mode \( \alpha \) of the discretized bath with conjugate momentum \( P_{\alpha} \), mass \( m_{\alpha} \) and frequency \( \omega_{\alpha} \). The potential along the reaction coordinate is a symmetric double-well given by

\[
V(s) = -a_{1}s^{2} + a_{2}s^{4}.
\] (38)

The \( g_{\alpha} \) are the coupling coefficients between the reactive system and the bath that are determined by the Ohmic spectral density function

\[
J(\omega) = \frac{\pi}{2} \sum_{\alpha} \frac{g_{\alpha}^{2}}{m_{\alpha}\omega_{\alpha}} \delta(\omega - \omega_{\alpha}) = \eta \omega e^{-\frac{\omega}{\omega_{a}}}. \tag{39}
\]

\( J(\omega) \) was discretized evenly with an increment \( \Delta \omega_{a} \), and the coupling coefficients were calculated according to

\[
g_{\alpha}^{2} = \frac{2}{\pi} m_{\alpha} \omega_{\alpha}^{2} \eta e^{-\omega_{a}/\omega_{\alpha}} \Delta \omega. \tag{40}
\]

In order to compare the escape rates to the work of Topaler and Makri,\textsuperscript{34} we have studied the same double-well systems as those described in their work. These include a high barrier model with a wide spectral density labeled DW1, and a lower barrier model with a narrower spectral density labeled DW2. The parameters were chosen to model a proton transfer reaction described by the above generic Hamiltonian, and are summarized in Table I.

We performed PIMC simulations described in Sec. IV in order to calculate the imaginary time reactive-flux correlation functions for the above two systems. In both cases we investigated a range of coupling strengths and temperatures studied by Topaler and Makri.\textsuperscript{34} For the symmetric double-well, the ratio between the full and reactant partition function is known analytically to be 2, and hence did not need to be determined by simulation. However, our results confirm that it is possible to estimate this ratio from standard umbrella sampling PIMC simulations. In performing the simulations the number of bath modes was chosen to be 100 for DW1 and 50 for DW2. This number of modes was sufficient to obtain converged rates in both cases.

The quantities \( \gamma_{a} \) and \( \gamma_{d} \), defined in the previous section, were computed from open chain staging PIMC\textsuperscript{59–61} simulations. The width of the Gaussians used to approximate the delta functions in Eqs. (33) and (34) was chosen to be 0.1 a.u. This should be compared to the natural distance \( \sqrt{a_{1}/2a_{2}} \), which is 2 a.u. for DW1 and \( \sqrt{2} \) a.u. for DW2. In both cases, the systems were equilibrated for 200,000 MC passes and data was collected for 10 million passes. Each pass consisted of attempting moves of all the beads for each degree of freedom. In most cases this was sufficient to obtain converged results for \( \gamma \). The number of Trotter slices was \( P = 50 \) to \( P = 128 \) depending on the temperature. In all the results shown below we set \( \lambda = \beta/P \). In some of the cases we verified that the rates are independent of \( \lambda \) within the statistical error.

The time dependent part of the reactive-flux correlation function, \( \xi_{j} \), was computed by an open chain staging PIMC simulation. In this case the first and last beads of the reaction coordinate, \( s_{0} \) and \( s_{T+1} \), were fixed at \( \Delta = 0 \) and \( 0 \), respectively, where \( \Delta \) arises from a finite difference approximation in Eq. (27). In both cases, \( \Delta \) was chosen to be 0.1 a.u. The lengths of the simulations and the number of Trotter slices were the same as those for \( \gamma \) described above.

The imaginary time flux–flux correlation functions were computed from the reactive-flux ones by taking a time derivative. This was done numerically using a simple two point forward difference. Our findings indicate that this procedure significantly amplifies the statistical noise present in \( G_{jk}(\tau) \). However, it should be noted that higher order finite difference methods were not observed to improve the results.

The correlation functions evaluated by the above procedure were then used as input data for the MaxEnt numerical analytic continuation procedure. The covariance matrices required by the MaxEnt procedure were computed by block averaging the MC data. We find that, within the statistical noise, the matrices are diagonal, with the diagonal elements giving the estimates of the variances for each of the imaginary time data point \( \tau_{j} \). The MaxEnt procedure was then used to determine the frequency dependent rate constant corresponding to each correlation function, by inverting Eq. (10). The rates computed from the zero mode values of \( k(\omega) \) were found to be only weakly dependent on the regularization parameter estimated from the L-curve. In each case the spacing and the range of the frequency grid \( \{\omega_{k}\} \) on which the solution was represented was adjusted to obtain converged results. Similar to the parabolic barrier case, a flat default model (prior distribution) with the integral given by \( \pi G(0) \) was used.

A sample input imaginary time flux–flux correlation function for DW2 is shown in the upper panel of Fig. 2 along with the MaxEnt fit. The lower panel shows the corresponding \( k(\omega) \) computed by the numerical analytic continuation procedure. Note the structured shape of \( k(\omega) \) which was also observed with the SVD inversion method (not shown). This structure indicates deviations from quantum transition state theory, since it implies recrossings of the dividing surface which will give rise to a structured reactive flux correlation function. The reaction rate is given by the zero value of \( k(\omega) \).

In Fig. 3 we show the dependence of the transmission coefficient, \( \kappa = k/k_{\text{TST}} \), on the damping parameter \( \eta \) for the DW1 system. The rate was the classical transition state theory rate given by

\[
k_{\text{TST}} = \frac{\omega_{b}}{2\pi} e^{-\beta E_{b}}, \tag{41}
\]

where \( \omega_{b} = 2\sqrt{a_{1}/m} \). We compare the results with the calculations by Topaler and Makri using the quasi adiabatic propagator path integral (QUAPI) method and with the quantum Grote–Hynes theory (QGH).\textsuperscript{62} Our results are in very good agreement with the QUAPI numerical results over the
entire range of frictions. Near the turnover friction, our method slightly underestimates the transmission coefficient, however, it does capture the turnover from energy to spatial diffusion. For this high barrier case the imaginary time correlation functions were more noisy than the results for DW2 shown below. As a consequence, the error bars (not calculable) on the transmission coefficients obtained using the MaxEnt procedure were larger compared to the results for DW2. We note that we have also carried out simulations for DW1 with a larger number of Trotter slices. The level of noise in those simulation was much larger compared to the results shown, leading to slightly larger deviation from the QUAPI results.

It turns out that the classical results for the parameters chosen for DW1 provide reasonable estimates of $\kappa$. In view of that, it is interesting to assess the accuracy of our method for a more ‘‘quantum’’ systems, such as DW2. The dependence of the transmission coefficient on the damping parameter for DW2 is shown in Fig. 4 with the QUAPI and the QGH results. The classical transmission coefficient for this case differ by almost an order of magnitude compared to their quantum mechanical counterparts. Our results are in good agreement with the QUAPI results over the entire friction range. They capture the turnover in the transmission coefficient signifying the crossover from energy to spatial diffusion.

It is also interesting to study the temperature dependence of the rate calculated using the RFAC method. An Arrhenius plot of the escape rate for DW1 is shown in Fig. 5 where we compare the RFAC and QUAPI results. Our results are in excellent agreement with the QUAPI rates over the range of temperatures characterized by the thermally activated regime. In this range the rates change by 7 orders of magnitude. This suggests the wide range of applicability of the RFAC method. In the low temperature tunneling regime (i.e., below the crossover temperature), the statistical noise was too large to obtain converged PIMC results for the imaginary time reactive-flux correlation function, and thus we could not compute the tunneling rates.

VI. CONCLUSIONS

In this paper we presented a method to compute canonical quantum reaction rate constants in the condensed phase. The method is based on the reactive-flux formalism com-
combined with the numerical analytic continuation approach. We expressed the imaginary time reactive-flux correlation function in terms of a frequency dependent rate constant, and used the path integral formalism to derive a working expression suitable for Monte Carlo simulation techniques. The maximum entropy method was then used to analytically continue the imaginary time data to real time to obtain the frequency dependent rate and the reaction rate.

Unlike the existing methods based on semiclassical and mixed quantum-classical approaches, the reactive flux analytic continuation formalism treats the whole system in a fully quantum mechanical fashion. In addition, compared to other fully quantum mechanical methods such as the quasi adiabatic propagator path integral method, the present approach is not limited to the treatment of harmonic baths. Furthermore, the method does not require that the Hamiltonian be separable into a “system” part and a “bath” part consisting of oscillators whether harmonic or anharmonic. Rather, the only requirement is that one is able to identify the reaction coordinate. To our knowledge this is the only method that may be capable of determining reasonably accurate quantum mechanical rates for a general anharmonic many-body system.

The accuracy of the method was tested for a simple parabolic barrier model, and for a double-well coupled to a harmonic bath. It is remarkable to note that the Arrhenius behavior is captured, as well as the turnover from energy to spatial diffusion. The method fails when the statistical noise exceeds the limits set by the numerical inversion method or when long time dynamics govern the reactive-flux correlation function. This may happen at very low temperatures when tunneling dominates the escape rate, or at low frictions where energy diffusion becomes important.

There are numerous possible improvements for our method. On one hand, more efficient sampling techniques are needed to converge the results for PIMC at very low temperatures in order to obtain tunneling rates. On the other hand, the numerical analytic continuation method used in this work is a very basic implementation of the maximum entropy method. As such, it is quite surprising that the results for the rates we obtained are in such a good agreement with the numerically exact ones. We believe that considerable improvement can be achieved by using a more sophisticated maximum entropy procedure. For example, rather than using a flat default model, one could use a more informative one. Such a model could be obtained from approximate methods, such as a multi dimensional parabolic approximation. In addition, recently it was shown that combining short real-time dynamical information with the imaginary time data can significantly improve the quality of the analytically continued results. Real time PIMC simulations could be used to compute the real-time flux–flux correlation function for short times to be used in combination with imaginary time data as input for the MaxEnt procedure. All of these approaches will be the subject of future investigation.

ACKNOWLEDGMENTS

E.R. would like to acknowledge financial support from The Israeli Council of Higher Education Alon fellowship. This work was supported by a grant to BJB from the National Science Foundation.

1 P. Hänggi, P. Talkner, and M. Borkovec, Rev. Mod. Phys. 62, 251 (1990), and references therein.