Binary contraction method for the construction of time-dependent dividing surfaces in driven chemical reactions

Robin Bardakcioglu,¹ Andrej Junginger,¹,* Matthias Feldmaier,¹ Jörg Main,¹ and Rigoberto Hernandez²,†

¹Institut für Theoretische Physik 1, Universität Stuttgart, 70550 Stuttgart, Germany
²Department of Chemistry, Johns Hopkins University, Baltimore, Maryland 21218, USA

(Received 7 May 2018; published 5 September 2018)

Transition state theory formally provides a simplifying approach for determining chemical reaction rates and pathways. Given an underlying potential energy surface for a reactive system, one can determine the dividing surface in phase space which separates reactant and product regions, and thereby also these regions. This is often a difficult task, and it is especially demanding for high-dimensional time-dependent systems or when a nonlocal dividing surface is required. Recently, approaches relying on Lagrangian descriptors have been successful at resolving the dividing surface in some of these challenging cases, but this method can also be computationally expensive due to the necessity of integrating the corresponding phase space function. In this paper, we present an alternative method by which time-dependent, locally recrossing-free, and global dividing surfaces can be constructed without the calculation of any auxiliary phase space function, but only from simple dynamical properties close to the energy barrier. Another benefit of this method is its exponential convergence which allows for the determination of high-accuracy dividing surfaces in higher-dimensional systems with relatively small computational effort.

DOI: 10.1103/PhysRevE.98.032204

I. INTRODUCTION

Predicting the rate of a chemical reaction is a central problem in the study of reaction dynamics. Chemical reaction dynamics are governed by classical equations of motion (EOM) driven by potentials whose primary nontrivial features can be approximated locally via rank-1 saddle points. Transition state theory (TST) [1–13] then provides a powerful framework for predicting both reaction rates and pathways. The theory is based on the identification of the dividing surface (DS) between reactant and product regions associated with a given rank-1 saddle which ipso facto defines these regions. In order to obtain exact—that is, not approximate—reaction rates, the DS must be free of recrossings. Otherwise, TST overestimates the reaction rate due to the incorrect attribution of nonreactive flux to the reaction rate. The central idea behind constructing nonrecrossing DSs is to resolve the good action-angle variables in phase space or the associated normally hyperbolic invariant manifold (NHIM) via its stable and unstable manifolds [14–17].

Local approximations to the saddle geometry can be applied in the case of low temperatures because passage over the rank-1 saddle is dominated by trajectories that cross near it. Harmonic approximations significantly reduce the computational effort because the reaction rate can then be directly formulated from the properties of the potential energy surface at the saddle. For higher-energy contributions which can feel the anharmonic potential in the vicinity of the saddle, normal form expansions [18–26] can be used to construct non-recrossing DSs to the desired accuracy.

A significant challenge arises if the DS is required for a potential energy surface whose reaction dynamics is determined by a nonlocal region, or one that is time-dependent such as when external forces are present. In recent years, this challenge has been successfully approached using Lagrangian descriptors (LDs) [27–30], which define a phase space function allowing for the identification of both stable and unstable manifolds. Despite their success in the application to nonautonomous systems [31–33], we recognize that LD-based methods can require a considerable amount of computational time, making it difficult for systems with large parameter sets or many degrees of freedom.

In this paper, we introduce a method that efficiently computes discrete points on the NHIM from very simple dynamical properties close to the energy barrier. A benefit of this algorithm lies in its avoidance of the calculation of any phase space function at the cost of the integration of a relatively small number of trajectories. It converges exponentially allowing for the determination of high-accuracy DSs with relatively small computational effort. The computational efficiency in the determination of these high-accuracy points can be extended through machine-learning algorithms—e.g., based on neural networks [34]. Specifically, a lower-resolution grid of points can be used as a training set for a neural net, thereby providing a noniterative and smooth interpolation of the NHIM. The recrossing-free dividing surface attached to this NHIM can then be used in TST to compute reaction rates by propagating trajectories from a suitably chosen initial distribution [33]. Note that application of TST requires the knowledge of which time each trajectory changes from reactant to product, and this is achieved with
the help of a recrossing-free time-dependent dividing surface [33–35].

The paper is organized as follows: In Sec. II A we give a brief review of LDs and discuss a method based on nested iterations with LDs to obtain points on the NHIM. We then introduce in Sec. II B an efficient binary contraction method that does not rely on LDs, but which accomplishes the same goal. Important aspects of the contraction method are discussed in Secs. II B 1 and II B 2 in more detail. In Sec. III we compare the performance of the binary contraction method with the LD approach in a practical application. The accuracy of the NHIM is also verified by showing that points initially arbitrarily close to the NHIM deviate away from it in time without recrossing it. The limits and implications of this method in addressing higher-dimensional rank-1 saddles, multiple-ranked saddles, roaming reactions, and quantum effects are discussed in Sec. IV. Conclusions are drawn in Sec. V.

II. METHODS

We describe a driven chemical reaction as a dynamical system with d degrees of freedom of which only one is unstable and the remainder are stable. The unstable degree of freedom is used to classify the reaction coordinate, while we refer to the stable degrees of freedom that are coupled to the dynamics of the reaction coordinate as the bath coordinates. We model such a system by a Hamiltonian with a time-dependent rank-1 saddle driven by the underlying time-dependent potential.

When d = 1, the dynamics in the vicinity of a saddle point can be surmised by the hyperbolic fixed points in a moving frame. Stable and unstable manifolds—that is, continuous sets of points in phase space where trajectories either exponentially approach or leave the hyperbolic point, respectively—can be attached to the hyperbolic points. A critical observation lies in the fact that dynamical propagation contracts the stable manifold towards, and expands the unstable manifold away from, the hyperbolic fixed points.

In higher dimensions (d > 1), the notion of a hyperbolic fixed point is generalized by the (2d − 2)-dimensional NHIM mentioned above, and points on the stable and unstable manifolds exponentially approach or recede from the NHIM. Once the NHIM is known, a recrossing-free DS with increased dimension 2d − 1, which separates the (2d)-dimensional phase space into the product and reactant regions, can easily be attached to the NHIM [33].

A. Nested iterations with LDs

An LD is generally defined via the integral of a positive definite function along a trajectory within a given time interval [t − τ, t + τ]. Here, t describes the time coordinate and τ is some positive value large enough to cover the relevant dynamics. The phase space function used to define the LD is the velocity v of the particle, which makes the LD,

$$L(x, v, t) = \int_{t-\tau}^{t+\tau} ||v(t')|| \, dt'$$ (1)

a measure of the arc length of the corresponding trajectory [28,29,32,35]. An example of a two-dimensional LD plot in the x, v plane computed for the open, three-dimensional model system introduced later in Sec. III is shown in Fig. 1(a). The connection between the LD and the stable and unstable manifolds is surprisingly simple [29,30]. Since, the dynamics on these manifolds is extremal, such simplicity also holds for the LD. To be precise, a particle on such a manifold approaches the fixed point either in forward or in backward time. Consequently, the LD reveals minimal values on each of the manifolds and their intersection—that is, the hyperbolic fixed point—is a (local) minimum of the LD.

One can use this approach to construct an algorithm to find the intersection of the closure of the stable and unstable manifolds—i.e. the NHIM—as follows. First, at a given reaction coordinate x and bath coordinates, one can perform an optimization in the reaction velocity v such extremizing the LD so as to obtain the associated positions of stable and unstable manifolds [see Fig. 1(a)]. From this, one computes their difference Δv(x) in v space as a function of the reaction coordinate x. A one-dimensional root search for Δv(x0) = 0 [see Fig. 1(b)] yields the position x0 where the stable and
stable manifold intersect, i.e., a point on the NHIM. This can be repeated for a set of bath coordinates and time to obtain the time-dependent NHIM up to the desired resolution.

Although one can readily identify the intersection of the manifolds in Fig. 1(a), the construction of a high accuracy image requires the computation of tens of thousands of trajectories and is only useful for demonstration purposes or to determine a rough estimate for the location of the NHIM. Meanwhile, the computation of a single value of $\Delta v(x)$—i.e., a search of the manifolds for a fixed reaction coordinate—requires the computation of several LDs because extremal optimizations are in general iterative processes. This is especially taxing as one considers the fact that root finding methods, in and of themselves, are also iterative and require multiple values of $\Delta v(x)$ to provide accurate results. Doing so, provides a multiplicative factor to the required amount of computations. For multidimensional systems, these nested iterations have to be performed several times over large sets of bath coordinates, and this may easily result in the integration of several millions of trajectories.

**B. Binary contraction method**

In this section, we provide an alternative method that does not explicitly search for the stable and unstable manifolds, and thereby avoids the nested iterations of the LD approach. To accomplish this, we assume that the saddle region is an appropriate interval in the reaction coordinate $[x_1, x_2]$ that covers the relevant dynamics of the time-dependently moving saddle. It can be found by sampling LDs in the region of interest, and has the structure shown in Fig. 1(a). In doing so, we are able to discriminate between four types of trajectories tied to one of four regions in phase space as shown in Fig. 2(a). Trajectories in regions (II) and (IV) are characterized as reactive, going from reactants to products and vice versa, respectively. Trajectories in regions (I) and (III) are nonreactive, remaining as products to products and reactants to reactants, respectively. Thus, every trajectory can be assigned to one of the four regions by integrating it forward and backward in time. While trajectories near the stable manifolds can spend a long time in the vicinity of the NHIM, they will eventually leave the barrier region when they are integrated over a sufficiently long time. For orbits extremely close to the stable manifolds, the classification to the regions (I)–(IV) may nevertheless depend on rounding errors or numerical errors of the integrator, which means that the stable and unstable manifolds and thus the crossing point on the NHIM can only be determined with finite numerical accuracy. For the benchmark example of this work, the accuracy of the integrator was sufficient for the determination of the four regions as illustrated in Sec. III A.

Our binary contraction method is based on the observation that, in the immediate vicinity of the hyperbolic point, the reactive and nonreactive regions are arbitrarily close to each other but still separated by stable and unstable manifolds. This is illustrated in Fig. 2(a). The manifolds associated with the NHIM determine the particle dynamics in the saddle region due to the asymptotic behavior of the related trajectories. More precisely, these regions determine which directions a particle enters and leaves the saddle region.

Similar to the LD method, we integrate the trajectories forward and backward in time. However, here the propagation of the trajectories does not occur for a fixed amount of time, rather each of the integrations in forward and backward time is stopped when the particle leaves the saddle region. This allows us to determine which of the four regions corresponds to the given initial phase space point. Using this information, we can construct a quadrangle in the $xv_x$ plane, with each vertex in one of the four regions. The algorithm is then divided into five steps [see Fig. 2(b)]:

1. Construct a quadrangle with each of its vertices in one of the four regions (I)–(IV).
2. Determine the midpoint between two adjacent vertices of the quadrangle.

FIG. 2. (a) View of the phase space in reaction coordinate and velocity. Stable and unstable manifolds intersect and divide the plane in four regions marked by (I)–(IV). The dynamics in a region is dictated by the dynamics on the manifolds that form the region’s boundary. The arrows indicate the general path of trajectories in a positive time direction. (b) The intersection of the manifolds is found by contracting a quadrangle with a binary search algorithm applied to all four edges consecutively. The highlighted quadrangle is the result of contracting all edges of the quadrangle with the vertices 1, 2, 3, and 4 once in a counterclockwise manner. The arrows with encircled numbers indicate the order and direction in which vertices are replaced. The procedure can be iterated to obtain the intersection point with the desired accuracy.
(3) Determine which region the midpoint corresponds to by integrating the dynamics in forward and backward time.
(4) Use the new vertex to replace the vertex in the same region. (If the region identified for the midpoint is not identical to either of the original points, perform an error correction as described in Sec. II B 2.)
(5) Repeat steps 2–4 for all edges, e.g., in a counterclockwise manner as in Fig. 2, until the longest edge of the quadrangle is below a desired error tolerance.

As implemented in our scheme, we walked through all the edges counterclockwise and successively in step 5. Alternative walks through the edges could provide faster or slower convergence, and the optimization of this scheme would be of possible interest to future work.

In Fig. 2(b), we illustrate these steps for the initial quadrangle with vertices \{(1, 2, 3, 4)\}. First, our algorithm contracts vertices 1 and 4, leading to the new quadrangle \{(5, 2, 3, 4)\}. It is then applied to vertices 5 and 2 to create \{(5, 6, 3, 4)\}. In the next step, the new vertex 6 (and not one of the old vertex points as in the two previous steps) is replaced with vertex 7 resulting in \{(5, 7, 3, 4)\}. After step 4, the highlighted quadrangle \{(5, 7, 3, 8)\} is created. Since this process is similar to a binary search algorithm, we call this method binary contraction.

Repeating this for all lines on the quadrangle ensures the reduction of mutual phase space distances between the vertices. Since reducing these distances in the \(x_v\) plane to an arbitrarily small size is only possible in the immediate vicinity of the hyperbolic point, we know that the quadrangle converges towards it. By computing the geometric center of the quadrangle in that plane, we have found the position of the hyperbolic point with an accuracy proportional to the size of the final quadrangle. By applying this method on an \(x_v\) plane of the reaction coordinates in a higher dimensional system, one can construct the NHIM \((x_{\text{NHIM}}(y, v_x, t_0), y_{\text{NHIM}}(y, v_y, t_0))\) for any suitable set of bath coordinates \((y, v_y)\) and initial time \(t_0\). As the contraction still takes place on a two-dimensional plane, the efficiency of the method is not impacted by additional dimensions in phase space.

An efficient initialization of the quadrangle in step 1, as well as the error correction in step 4, will now be discussed in detail in Secs. II B 1 and II B 2, respectively.

### 1. Exception handling in initialization

An important prerequisite of the binary contraction method is the identification of good initial conditions for the iteration. To construct a valid quadrangle, we rely on the observation that the shape of the cross which divides the regions in the \(x_v\) plane is robust in the vicinity of the hyperbolic fixed point. The construction of the NHIM can be visualized readily with the help of an LD plot such as that shown in Fig. 1(a). One first constructs a quadrangle by first guessing a phase space coordinate \((x_0, v_{x_0})\) in the vicinity of the NHIM. Then one constructs the vertices with constant shifts in the \(x_v\) plane \(\Delta x\) and \(\Delta v_x\), such that \((x_0 \pm \Delta x, v_{x_0})\) and \((x_0, v_{x_0} \pm \Delta v_x)\) form the vertices of the quadrangle.

For the initial set of selected vertices, we integrate the trajectories and compare their types to see if the construction of the required quadrangle was successful in the sense that there is a vertex in each of the 4 regions. If this is true, the iteration will start at this point. However some constructions may fail because they rely on both the accuracy of the guess and the size of the shifts. To mitigate this issue, one may increase shift sizes and try again. In addition, one can implement a bookkeeping method to combine valid vertices constructed from different shift sizes. Optimally, using a lightweight extrapolation method to determine an accurate guess from previous results would allow for a quick construction. This is especially efficient if we construct the NHIM in higher dimensions from a few already known points. In practice, a combination of both extrapolation and managed variation of shift sizes have proved to be highly effective for constructing the initial quadrangle.

One could also try to use an LD-based approach to determine points on the stable and unstable manifolds from which one can interpolate a valid quadrangle. However, because of the efficiency of the iteration in the binary contraction method, it would be ill-advised to use an initialization method that is computationally taxing in comparison, as the latter could dominate the runtime and overall cost of the computation.

### 2. Increasing convergence for edge cases

Depending on the choice of the underlying EOM and plane, one may run into situations where the borders of the regions form a heavily distorted cross. In situations like these, it is possible that the quadrangle contracts into an area where the intersection is not present; see Fig. 3. An iteration that is completely inclusive—that is, one that contracts the quadrangle into a subset of itself—may lead to convergence problems. This issue arises when the midpoint of a line is not contained in either of the regions of the original vertices, as this is the only situation where the quadrangle does not contract further. In that case, we allow the quadrangle to expand outward in the direction where the failure was discovered. The expansion is controlled such that the possible contraction of the iteration dominates.
Using this procedure, the quadrangle not only contracts around the intersection of the region, but also tumbles along the stable and unstable manifolds to actively center itself around the intersection. Although this procedure adds computational complexity to the iteration, we have observed that, on average, there is no noticeable increase in the computational time required. On the contrary, it even accelerates the computation for situations where convergence might not have been reached otherwise because it opens the quadrangle to more equally wrap around the targeted fixed point.

III. RESULTS

A. Tracing the motion of the NHIM for fixed bath coordinates

In this section, we will compare the performance of the LD-based method with nested iterations to the binary contraction method by computing several points on the NHIM and comparing the average number of necessary trajectories to compute a single point as a function of the algorithm’s tolerance. This is accomplished by applying both methods to a three-dimensional model system with a time-dependently moving rank-1 saddle. To this end, we use a three-dimensional extension of the two-dimensional time-dependent potential introduced in Ref. [33],

\[
V(x, y, z, t) = E_b \exp \left( -a(x - \hat{x} \sin(\omega_x t))^2 \right) \\
+ \frac{\omega^2}{2} \left( y - \frac{2}{\pi} \arctan(2x) \right)^2 \\
+ \frac{\omega^2}{2} \left( z - \frac{2}{\pi} \arctan(2x) \right)^2, \tag{2}
\]

and we use parameters in simulation units \( E_b = 2, a = 1, \hat{x} = 0.4, \omega_x = \pi, \omega_y = 2, \) and \( \omega_z = 1. \) With these parameters, the potential has a periodicity of \( T = 2. \) The algorithms are set to find a point on the NHIM for fixed bath coordinates \( y = z = v_y = 0 \) and \( v_z = 0.5, \) for 200 equidistant, initial time coordinates during one oscillation period, thereby tracing the NHIM’s time-dependent motion. Note that this tracing of the NHIM is not, in general, a trajectory of the system, since the results of the algorithms are not propagated by the dynamics of the system to find further points on the NHIM. Each point is computed separately. This allows us to control the accuracy of the data, as well as the specific bath coordinate for which the point is computed.

To compare the performance, we measure the computational complexity in terms of the average number of integrated trajectories per hyperbolic fixed point found as a function of the tolerance. Note that such a point can be found for any set of bath coordinates and time. The results in Fig. 4 show that the binary contraction method requires far fewer trajectories to reach a given tolerance. Especially for very low tolerances, i.e., for high accuracy, it is highly superior in terms of performance. The difference in scaling is also quite evident, as the contraction exhibits an exponential convergence to the hyperbolic fixed point, while the performance of the nested iterations is clearly worse, even for large values of the tolerance. The fast convergence of the binary contraction method is evident from the fact that it divides the lengths of the quadrangle’s sides by roughly two for every successful iteration, therefore contracting the quadrangle exponentially fast.

This is further illustrated in Fig. 5, where we apply the method to the model system with the time-dependent potential \( (2) \) for a specific choice of the bath coordinates and time. [Note that in this figure we use a different set of bath coordinates and time compared to Fig. 1(a).] The quadrangles of the first and second iteration are marked by numbers 1–8, and the dashed lines complete intermediate quadrangles. The quadrangles converge rapidly towards the intersection of the stable and unstable manifolds as marked by the largest values of the LD shown in the darkest shade of gray.

![FIG. 5. The binary contraction in action for the model system with the time-dependent potential given in Eq. (2). One iteration corresponds to one counterclockwise revolution around the quadrangle, contracting each side once, starting from the vertex line \( (1, 2) \) of the quadrangle \( \{1, 2, 3, 4\}. \) In the next iteration the quadrangle \( \{2, 7, 4, 8\} \) is obtained. The LD values are shown through shading in the background, where one can clearly discern a cross which corresponds to the stable and unstable manifolds in the darkest shade. The quadrangles rapidly contract towards their intersection.](image-url)
in Eq. (2) as it is sufficient to illustrate the behavior even the results of an investigation for the model system defined they are started in its close vicinity. In this section, we present we investigate how trajectories deviate from the NHIM when points on the NHIM for any suitable bath coordinate and time, stated assumptions are satisfied. The binary contraction was for systems with higher dimensionality as long as our earlier achieved as long as sufficient numerical stability is provided. Consequently, the binary contraction method can indeed be achieved as long as sufficient numerical stability is provided. This accuracy only applies to the discrete dynamics of the numerically integrated system. The analytical accuracy of the equations of motion that may increase in higher dimensional systems will increase the computational cost of computing these trajectories. Additionally, one needs to keep in mind that even though the complexity of the contraction for a single point on the NHIM scales modestly with the dimensionality of the equations of motion, the number of points needed to characterize the NHIM for a given system may scale exponentially with its dimensionality, depending on how one chooses to construct the DS.

Although we have been successful in applying the binary contraction method to a system with a rank-1 saddle potential, we have yet to see how well it can be applied to systems with higher-rank saddle potentials, such as the systems investigated in Refs. [39–41], where the classification of trajectories is not as straightforward as for rank-1 saddle potentials because the characterization of the dynamics does not necessarily contract to the requisite quadrangle. Thus it remains to be seen if this approach can be extended to higher-rank saddles.

Another challenge to this and related approaches lies in its possible application to roaming pathways for which no clear saddle exists [42–46]. However, if a NHIM exists for such a reaction, then that NHIM would by definition not be associated with a rank-1 saddle potential. This makes the identification of trajectories for the binary contraction even more of a challenge because the requisite quadrangle does not exist. Thus the future use of a binary contraction method to address roaming reactions would require a classification scheme to identify which trajectories—on a contracted two-dimensional surface—are reactive or not.

Finally, an additional concern for this and other related approaches addressing chemical reactions using transition state theory is whether quantum effects provide a significant correction [47,48]. Indeed, one of the constructions of semiclassical transition state theory has relied on the construction of the good-action variables that underlie the NHIM [20,21,49–52]. Thus, one possible new direction beyond this work would involve the use of the binary contraction method to construct not only the NHIM but also an underlying coordinate space.
that represents the nonseparable but integrable coordinates needed to construct a semiclassical rate formula. However, this would only be necessary if the chosen system exhibits significant quantum effects. The present work is restricted to fully classical treatments of driven chemical reactions. Such dynamics arises frequently, particularly when heavy molecular motions occur in solution.

V. CONCLUSION

We have introduced an exponentially converging algorithm that finds hyperbolic fixed points, i.e., the NHIM for nonautonomous systems with rank-1 saddles. The binary contraction method achieves the high efficiency by avoiding the explicit search for stable and unstable manifolds in nested iterations with LDs. The method relies on the identification of the saddle region, as well as the ability to describe the system in coordinates where the stable and unstable manifold intersect within two-dimensional planes in phase space.

By independently performing the algorithm on several two-dimensional planes one can map the NHIM via the bath coordinates of the system. We have discussed ideas to make this method reliable for difficult phase space structures, and methods for efficient initialization even in higher dimensions.

We have confirmed the relative stability of the NHIM computed numerically from the binary contraction method. Specifically, trajectories initiated on the time-dependent NHIM remain in its vicinity over time, subject to the accuracy of the initial computation of the NHIM. The binary contraction method presented here is one important component for a useful advance to the existing set of methods for the comprehensive handling of driven chemical reactions that can be described by time-dependent rank-1 saddles. Using TST, fluxes and rates in such systems can be obtained by propagating a large ensemble of trajectories and determining the time when each trajectory crosses a time-dependent DS [33]. The binary contraction method provides, in a first step, a large set of points on the NHIM, which can then be used, in a second step, to construct a time-dependent and locally recrossing-free DS by interpolating the discrete points on the NHIM. As with the other approaches, listed in the discussion in Sec. IV, it remains challenging to address high dimensionality in the degrees of freedom orthogonal to the reaction coordinate. This could perhaps be resolved using machine learning methods, such as neural networks, as demonstrated in Ref. [34].

ACKNOWLEDGMENTS

The German portion of this collaborative work was partially supported by Deutsche Forschungsgemeinschaft (DFG) through Grant No. MA1639/14-1. The U.S. portion was partially supported by the National Science Foundation through Grant No. CHE 1700749. A.J. acknowledges the Alexander von Humboldt Foundation, Germany, for support through a Feodor Lynen Fellowship. M.F. is grateful for support from the Landesgraduiertenförderung of the Land Baden-Württemberg. This collaboration has also benefited from support by the European Union’s Horizon 2020 Research and Innovation Program under the Marie Skłodowska-Curie Grant Agreement No. 734557.


