Divided Saddle Theory: a new idea for rate constant calculation

János Daru[1,2], András Stirling[1]*

[1]*Institute of Organic Chemistry, Research Centre for Natural Sciences*
of the Hungarian Academy of Sciences, Budapest, Hungary and
[2]*Loránd Eötvös University, Budapest*
(Dated: January 13, 2014)

We present a theory of rare events and derive an algorithm to obtain rates from postprocessing the numerical data of a free energy calculation and the corresponding committor analysis. The formalism is based on the division of the saddle region of the free energy profile of the rare event into two adjacent segments called *Saddle Domains*. The method is built on sampling the dynamics within these regions: auxiliary rate constants are defined for the Saddle Domains and the absolute forward and backward rates are obtained by proper reweighting. We call our approach Divided Saddle Theory (DST). An important advantage of our approach is that it requires only standard computational techniques which are available in most molecular dynamics codes. We demonstrate the potential of DST numerically on two examples: rearrangement of alanine-dipeptide (CH_3CO-Ala-$NHCH_3$) conformers and the intramolecular Cope reaction of the fluxional barbaralane molecule.

I. INTRODUCTION

The mechanism of a chemical reaction is understood as a sequence of elementary steps. Chemical kinetics measurements have a crucial role in determining an unknown mechanism. Such measurements provide rate laws and rate constants which can then be used to derive a plausible reaction mechanism. Often the elementary steps cannot be identified unambiguously because they are not accessible directly from measurements. Over the last years several computational methods have been developed to calculate rate constants[1–11,15–17]. Theory can efficiently complement experiments, but the main challenge in computing rate constants is that a reaction is a rare event on the molecular time-scale.

The most often invoked framework to interpret rate constants is Transition State Theory (TST)[15–17]. TST connects the experimental rate constant with the concept of activation free energy ($\Delta F^{\ddagger}$):

$$k_{exp} = \frac{1}{\beta h} e^{-\beta \Delta F^{\ddagger}} \qquad (1)$$

where $\beta = 1/(k_B T)$, T is the temperature, k_B is the Boltzmann constant and h is the Planck constant. Although $\Delta F^{\ddagger}$ is not an experimentally observed quantity, it provides a convenient measure to interpret reaction rates[18]. TST has a couple of attractive features which boosted its widespread use in reaction mechanism studies. It offers an easy concept. $\Delta F^{\ddagger}$ can be estimated from the reactive potential energy surface. It can also be improved to incorporate the inefficiency of the reactive vibrations by introducing the κ transmission coefficient in Eq. 1. However TST suffers from limitations. It is based on the so-called quasi equilibrium hypothesis and the entropy of the initial and transition states is often difficult to calculate accurately.

Trajectory calculations on the underlying potential energy surface (PES) represent the straightforward way to obtain directly reaction rates[19]. Unfortunately these calculations require high computational costs and only for systems of few atoms can be done efficiently.

Although it has been originally developed for PES, TST is often applied directly on free energy surface (FES). $\Delta F^{\ddagger}$ is obtained by taking the difference between the saddle point and the reaction-side minimum of the FES determined as a function of a few reaction coordinates. The barriers calculated in this way are used to compare rates. This approach also has some drawbacks. In particular the arbitrary state definition brings ambiguity in the calculated $\Delta F^{\ddagger}$ and rate constants as the FES depends on the choice of the reaction coordinate[11]. Instead of applying Eq. 1 which is not strictly valid for FES, a more sophisticated TST-based approach is when the forward and backward rate constants are obtained from the mean residence times estimated from the full free energy surface[17,20]. All these methods overestimate the actual rates as recrossings are not taken into account and significant effort has already been invested to improve their efficiency and recover κ. One remedy is the variational TST[12–16] where the dynamical bottleneck of the reaction is varied to maximize κ. Another, computationally more demanding possibility is to explicitly calculate κ[2–5].

Equilibrium simulations can also be applied for direct rate constant calculations. However the separation of time scales prevents the sufficient sampling of the reactive events: their frequency is orders of magnitude lower than those of the typical atomic motions. Various rare event methods have been developed to circumvent this prob-

lem. Reactive flux methods[2–6] aim at improving the TST by calculating κ. Their efficiency depends on the quality of the reaction coordinate. Path-sampling methods, such as the original transition path sampling (TPS)[7], transition interface sampling[8], forward flux sampling[10] and related methods do not rely on a preliminary selection of reaction coordinate. All these methods can recover the transmission coefficient or directly yield rate constant. Still, their often complicated formalisms and the necessary implementation efforts prevented their widespread use. Instead, most mechanistic studies still employ the simple TST model and assume $\kappa = 1$ (see e.g Ref. 21).

In this study we present a theory of rare events to calculate rate constants of elementary reactions. The main motivation was to develop a method where the numerical results of a typical mechanistic study (a FES exploration with the subsequent committor analysis[22–24]) can be further exploited without additional simulations. The underlying theory has common motifs in general with the Bennett-Chandler formalism[2,3] and in particular with the effective positive flux variant (EPF)[4,25–28]. The method derived from the theory is a simple postprocessing of the available numerical data. It directly yields both the forward and backward rate constants without explicitly invoking the activation free energy concept.

In the following we first present the theory and the corresponding methodology. Then we illustrate its potential with two examples: rearrangements of alanine-dipeptide conformers and the intramolecular Cope reaction of barbaralane (Fig. 4). In both cases we compare the rate constants of the new method with those obtained from unbiased, long MD simulations employing the same computational setup. In this way the assessment of the performance of the new method is free from technical issues.

II. THEORY

The basic assumptions of the present kinetic model are the following: i) the elementary reaction step in question can be described with a single reaction coordinate (λ);[29] ii) reactant and product regions can be defined on the free energy profile, and the two regions are separated by the dividing surface (ds); iii) within these regions the system equilibrates much faster than the reaction time scale, therefore the assumption of Boltzmann distribution is valid within the regions.

Rate constant of an elementary process is generally described as the inverse of the mean lifetime (τ)[30]:

$$k = 1/\tau \qquad (2)$$

In a sufficiently long simulation the mean lifetime can be calculated as an average of the time intervals (t_{RS}) the system spends in the reactant state before a reaction takes place[4,31]. Hence k can be expressed as:

$$k = 1/\tau = \frac{N^t}{t_{RS}^t} = \frac{1}{\langle t_{RS} \rangle} \qquad (3)$$

where t_{RS}^t is the total time of the system spent in the reactant state (RS), and N^t is the total number of reactions in the actual observation.

Eq. 3 is a proper definition of the reaction rate constant. However in the present form it does not offer any useful strategy for an efficient rate estimation when the mean lifetime is much longer than the affordable simulation time. Note however that for fast reactions this is the simplest way to compute rate constant and for comparisons we used also this equation.

Eq. 3 can be further transformed:

$$k = \frac{N^t}{t_{RS}^t} \qquad (4)$$

$$k_{DST} = \frac{N^t}{t_{SD}^t} \cdot \frac{t_{SD}^t}{t_{RS}^t} \qquad (5)$$

$$= k_{SD} \cdot \alpha_{RS}^{SD} \qquad (6)$$

In Eq. 5 we have introduced the concept of Saddle Domain (SD) as depicted in Fig. 1. SD is defined as a region extending from the ds to arbitrary length within the reactant state. Another SD can be defined similarly for the product region. The two SD-s are adjacent at the ds and they constitute together the Divided Saddle (DS) region[32]. k_{DST} is the rate constant calculated by applying this concept. t_{SD} is the residence time the system spends in SD. α_{RS}^{SD} is the fractional concentration ie. the ratio of the probability to find the system in this interval and the probability to find the system in the reactant state:

$$\alpha_{RS}^{SD} = \frac{t_{SD}^t}{t_{RS}^t} = \frac{P_{SD}}{P_{RS}} = \frac{\int\limits_{\lambda \in SD} e^{-\beta F(\lambda)} d\lambda}{\int\limits_{\lambda \in RS} e^{-\beta F(\lambda)} d\lambda} \qquad (7)$$

It can be calculated easily from the probability distribution of λ.

In the present context it is important to define the reactive events contributing to N^t along the trajectory. This can be done in several ways which are compatible with the present theory. A simple and convenient way is to introduce indicator lines within the two free energy wells defining stable regions as shown in Fig. 1. The concept of the strict definition of the stable regions is also an important element of the EPF formalism[4,25–27]. A reactive event (forward or reverse) occurs when
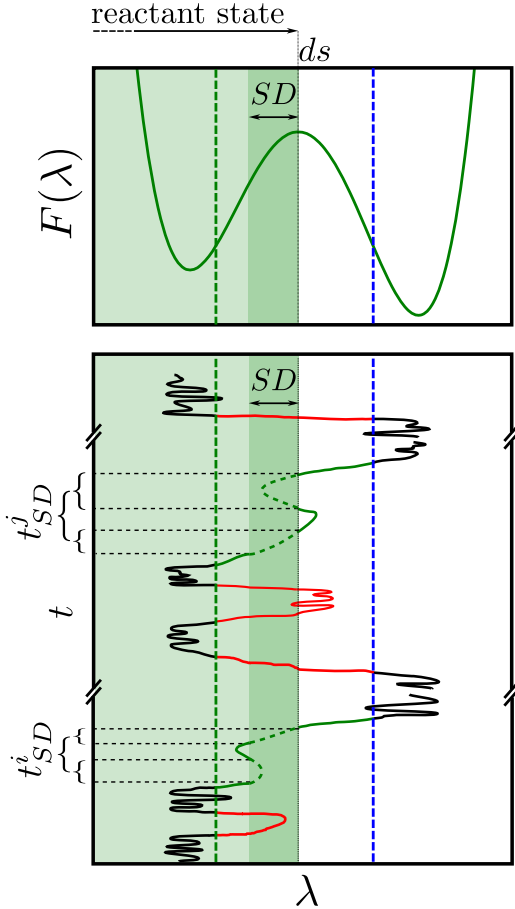
FIG. 1: Top: *Saddle Domain* (SD) defined on a typical free energy profile for the forward (*ie.* left-to-right) reaction. Green and blue dashed lines define stable reactant and product regions. The dividing surface is labelled by ds. Note that it does not necessarily coincide with the maximum on the profile. See text for details.
Bottom: Representation of a MD trajectory. Trajectory segments associated with forward reactive events are green, non-reactive segments are red. Note that some of these red segments can be reactive in backward direction. t^i_{SD} and t^j_{SD} are the sums of time intervals the i-th and j-th segments spent in the SD, respectively. For better visibility, these visits are indicated by dashed green lines.

the trajectory enters one of the selected regions after it came out from the other indicated region last. This definition is the same as employed in defining reactive trajectories in Transition Path Theory[33].

Led by chemical intuition we selected these ranges to include 90% of the configurations of a given well. Note that the calculated rate constants are not sensitive to small variations of the indicator lines due to the separation of timescales[34].

N^t is unaffected by the size of SD. Eq. 6

shows that α^{SD}_{RS} and k_{SD} are inversely proportional. Therefore extending SD affects only k_{SD} and α^{SD}_{RS}, but not k. In fact, in the limit where SD is extended over the whole reactant state ($SD \equiv RS$), α^{SD}_{RS} becomes unity and $k_{SD} = k$.

We stress that the present SD region differs fundamentally from the *transition state* used by TST. Indeed, SD is a part of the reactant state, it is in equilibrium with the rest of the reactant state by virtue of construction and does not extend over the ds.

III. METHOD

In the following we derive an expression to estimate k_{SD} in an efficient manner, where k_{SD} belongs to the forward reaction. Analogous expression can be derived for the backward rate. We start the derivation by analyzing a hypothetical, sufficiently long trajectory featuring many reactive transitions as depicted in Fig. 1. We select segments from this trajectory which connect an exit from one of the stable regions with the next entry to any of the stable regions after at least one visit of the SD. Some of these segments can be associated with a forward reactive event. In Fig. 1 we plotted these reactive segments in green while the others (reverse reactive events and nonreactive events) in red. We denote the total time spent by the ith segment in SD as t^i_{SD}. We assign a reactivity index N^i to each segment with $N^i = 1$ if it is forward reactive and $N^i = 0$ otherwise. Then k_{SD} can be written as follows.

$$k_{SD} = \frac{N^t}{t^t_{SD}} \tag{8}$$

$$= \frac{\sum N^i}{\sum t^i_{SD}} \tag{9}$$

$$= \frac{\langle N \rangle_{segments}}{\langle t_{SD} \rangle_{segments}} \tag{10}$$

In the last equation the averages are calculated on the ensemble of trajectory segments which is not straightforward to sample. In principle an infinitely long ergodic NVT trajectory visits all phase points of SD. In turn, all SD phase points are elements of these segments. Therefore we can generate these segments by propagating the equation of motions forward and backward from an NVT sample of the SD phase-phase points. However, the segments obtained in this way do not represent a proper sampling of the ensemble required in Eq. 10. This is because the probability of sampling segment i is proportional to the time spent within SD (t^i_{SD}). For example in Fig. 1 at the bottom panel the segment j would be sam-

pled with the factor t_{SD}^j / t_{SD}^i more often than segment i. Hence to calculate properly the average $\langle \ \rangle_{segments}$ we have to use an $1/t_{SD}^i$ weighting to obtain the final formula:

$$k_{SD} = \frac{\langle N \rangle_{segments}}{\langle t_{SD} \rangle_{segments}} \quad (11)$$

$$= \frac{\langle N/t_{SD} \rangle_{NVT}}{\langle t_{SD}/t_{SD} \rangle_{NVT}} \quad (12)$$

$$= \left\langle \frac{N}{t_{SD}} \right\rangle_{NVT} \quad (13)$$

Note, that reweighting by reciprocal path lengths also appears in TPS-based algorithms with variable path length[6].

The overall rate constant can be expressed as:

$$k_{DST} = \left\langle \frac{N}{t_{SD}} \right\rangle_{NVT} \frac{\int\limits_{\lambda \in SD} e^{-\beta F(\lambda)} d\lambda}{\int\limits_{\lambda \in RS} e^{-\beta F(\lambda)} d\lambda} \quad (14)$$

In practice we can employ the following procedure. First a series of equilibrium simulations is done with appropriate umbrella potentials to generate the FES as a function of λ. To identify the ds, committor analysis is performed, ie. from an NVT ensemble of configurations trajectories are initiated forward with momenta drawn isotropically from a Maxwell-Boltzmann distribution and backward with reversed momenta. The configurations are selected from the saddle region which is defined to include the ds and the SD-s for both the forward and reverse reactions. The committor function $p(\lambda)$ is determined as the average committor value measured at λ. After identifying ds where $p(\lambda) = 0.5$, the SD ranges are defined. Then for both the forward and reverse reactions the rate constants can be calculated by post-processing the trajectory segments generated for the committor analysis using Eqs. 7, 13, and then 14. We stress that $\Delta F^\ddagger$ does not enter our equations and our formalism does not require that the dynamical bottleneck be at the FES maximum; thus the method provides a larger flexibility in λ selection. The width of SD is important and our experience shows that it should cover a region where the free energy change is of order of kT.

In the following we show how our formalism is related to the Bennett-Chandler-type formalisms. To this end we select the EPF method[4,25–27], which employs an analogous definition for stable reactant and product states. We first note that the conceptual difference between BC-type methods and DST is that the formers measure the flux through the ds, whereas DST counts reactive and nonreactive events. In case of proper sampling,

both formalisms give the same, exact rate constant. Comparison of the two methods can be done only at the zero-width SD limit because of the definition of EPF. We now demonstrate that both formalisms lead to the same rate constant value although for the individual trajectories the evaluations proceed differently. At the zero-width SD limit Eq. 14 transforms to:

$$k_{DST} = \left\langle \frac{N}{t_{SD}} \right\rangle_{NVT} \cdot d\lambda \cdot \frac{e^{-\beta F(\lambda = ds)}}{\int\limits_{\lambda \in RS} e^{-\beta F(\lambda)} d\lambda} \quad (15)$$

One can recognize the so-called unnormalized transmission coefficient[4] in our equation:

$$R_{DST} = \left\langle \frac{N}{t_{SD}} \right\rangle_{NVT} \cdot d\lambda \quad (16)$$

We can compare this expression with that of the EPF method[4,25–27]:

$$R_{EPF} = \left\langle \dot{\lambda}_{ds} \cdot \Theta(\dot{\lambda}_{ds}) \chi^{EPF} \right\rangle_{NVT}, \quad (17)$$

where $\chi^{EPF} = 1$ if integrating backward the trajectory it reaches the stable reactant region before crossing ds and integrating forward reaches the stable product region before reaching the stable reactant region. Otherwise $\chi^{EPF} = 0$. The dot notation indicates time derivative. The Heaviside function Θ selects the positive velocities. We can rewrite this expression by taking into account only the non-zero contributions to the average:

$$R_{EPF} = \left\langle \frac{\chi^{EPF}}{t_{SD}^f} \right\rangle_{NVT} \cdot d\lambda \quad (18)$$

Here t_{SD}^f denotes the infinitesimal time period spent by a reactive trajectory with velocity $d\lambda/t_{SD}^f$ during the *first* crossing of the SD of $d\lambda$ width. Comparing Eqs. 16 and 18 we find that the terms of the fraction within the averages represent different quantities and it is not evident that they yield the same NVT average. To show this we notice that in sampling the phase-space points belonging to SD, a given reactive trajectory yields a single $1/t_{SD}^f$ term within the EPF framework, whereas in the DST it gives t_{SD}/t_{SD}^f times more non-zero $1/t_{SD}$ terms. For this trajectory this factor exactly compensates the difference between the lengths of the time periods in the denominator. Therefore the two expressions yield the same rate.

In case of postprocessing available trajectories, at the zero-width SD limit our method is more efficient than EPF. Indeed, DST exploits the full data set available. In contrast, EPF utilizes only those trajectories which contribute to the effective

positive flux and throws out the others including several reactive ones. This implies that for the same statistical accuracy EPF requires more trajectories than DST. In contrast EPF is more advantageous when trajectories are not available and have to be generated for the rate constant calculations. We compare numerically the two methods for alanine-dipeptide in the next section.

In summary, the main advantage of the present method is practical: to obtain forward and backward rate constants no additional calculations are required, the necessary ingredients are available from the simulations carried out for the exploration and analysis of the FES along the putative λ. Indeed, the NVT ensemble of the SD points are available from the sampling of the probability distribution of λ, whereas the trajectories are available from the analysis of the committor function $p(\lambda)$ employed to identify the true dynamical bottleneck along the λ[35]. At variance with the Bennett-Chandler-type methods, where only trajectories initiated at exactly the dividing surface (in practice within a very thin λ region) can be utilized, here all the committor trajectories initiated from the preselected finite λ region are employed. This can significantly improves the statistics and can be very useful in case of transformations where the FES is very flat around the ds. In addition, we can see that by varying the width of the SD within the reactant state region, DST unifies the direct, brute-force rate constant calculations (when SD covers the full reactant basin) with the Bennett-Chandler type methods at the zero-width limit.

IV. COMPUTATIONAL DETAILS

The CP2K program package has been used for the calculations[36]. For alanine-dipeptide we have employed the Amber force field[37] in a nonperiodic model. Electrostatic interactions were handled analytically. The integration of the equations of motion has been carried out with the Verlet-algorithm, applying 0.3 fs time steps in the simulations. The NVT conditions at 300 K were maintained by coupling the system to a Nose-Hoover global thermostat-chain. The free energy surface as a function of Φ, and Ψ has been obtained by the metadynamics method[38] and plotted in Fig. 2. In the simulations for obtaining rate constants we have used a 1D projection of Φ, and Ψ: $\lambda = a \cdot \Phi + b \cdot \Phi + c$ for describing the progress of the reaction. The a, b and c were obtained by projecting Φ, and Ψ-s onto the line connecting the two FES minima and rescaling the new coordinate by setting the positions of the two minima to 0 and 1 ($a = 0.30, b = -0.36, c = 1.79$).

The saddle region has been sampled by a 15 kcal/λ^2 umbrella deposited at the maximum of the 1D FES at $\lambda = 0.57$. The NVT ensemble of the configurations has been recovered by reweighting. For each sampled configuration a trajectory was initialized with velocity taken from a 300 K Maxwell-Boltzmann distribution and the equations of motion were integrated for both the forward and reverse directions till the first crossing of an indicator line.

For barbaralene we have employed the BLYP exchange correlation functional[40]. We have used the double-ζ basis sets with a set of polarization functions in conjunction with the Goedecker-Teter-Hutter[41] pseudopotentials. The edge length of the periodic cubic simulation box was 20.8 bohr. A plane-wave basis set with an energy cutoff of 300 Ry has been used to expand the electron densities. The Martyna-Tuckermann method[42] has been employed to decouple the electrostatics of the periodic images. The time step for integrating the Verlet-equations was 0.3 fs. NVT sampling at 1000 K has been achieved by using Nose-Hoover global thermostat. The λ for the system was chosen as the difference of the actual bond lengths of the breaking and forming C-C bonds ($d_1 - d_2$), as depicted in Fig. 4. The free energy profile has been calculated directly from the probability density histogram. SD was sampled according to the 1000 K probability distribution with a 15.0 kcal/bohr^2 umbrella and trajectories were initialized with velocity taken from a 1000 K Maxwell-Boltzmann distribution. We followed again the trajectories in both forward and backward directions till the first indicator line crossing. Note that the intramolecular reaction of barbaralane is degenerate: the forward and backward reactions are the same transformation, hence the rate constants are equal.

V. RESULTS

As our first test case we have selected alanine-dipeptide in vacuum. Alanine-dipeptide is a suitable model to study protein backbone conformation equilibria which play important role in the motifs of protein folding such as α-helix and β-sheet . In the present case we studied the transition equilibrium between the extended C5 and the C7 equatorial conformations. To this end we have first explored the FES as a function of the Ramachandran Φ and Ψ torsional angles[44] and identified the corresponding free energy minima (Fig. 2) by metadynamics. We have sampled 2000 configurations in the saddle region which were used for the shootings. The total simulation time of these trajectories was 1.06 ns. The
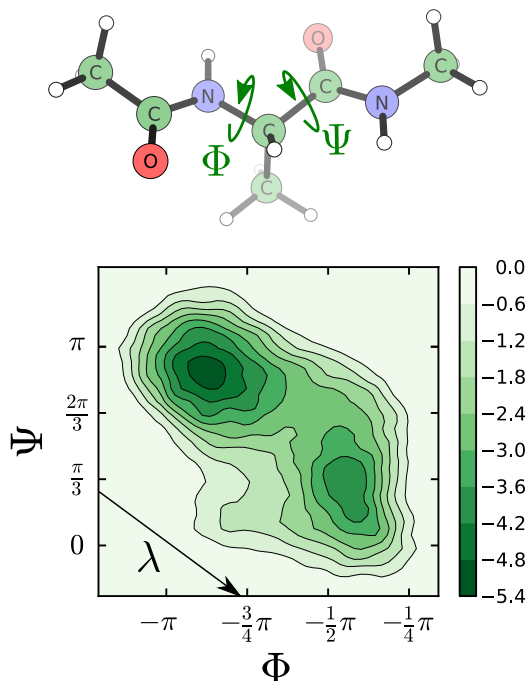
FIG. 2: Definition of the reaction coordinates (Φ, Ψ), and λ on the corresponding free energy surface (depicted in kcal/mol units) for the C5 $\rightleftharpoons$ C7$_{\text{eq}}$ conformational change in alanine-dipeptide.

FIG. 3: Deviation of the calculated rate constants of reaction C5 $\longrightarrow$ C7$_{\text{eq}}$ from the best estimate as a function of the number of generated trajectory segments for EPF (red, square) and DST (green, circles) methods. Inset: rate constants (in 10^{11} s^{-1} units) as a function of the number of generated trajectory segments for EPF (red, square) and DST (green, circles) methods. Dashed line: best estimate from direct MD.

actual position of the ds was assigned to 0.57 by the $p(\lambda_{ds}) = 0.5$ definition, which coincides with the FES maximum in this case. We have then determined the forward and backward reaction rates by our method. The SD-s were definded as $\lambda \in [0.40, 0.57]$ and $\lambda \in [0.57, 0.67]$ intervals for the forward and backward reactions, respectively. 1167 of the sampled configurations were in the forward SD (701 ps), whereas 833 were in the backward SD (362 ps). A 500 ns long unbiased MD simulation has been also carried out to obtain the reaction rates for comparison. The results are collected in Table I. We can see that the performance of the new method is very good. The discrepancies between the rate constants obtained with DST and those from long, unbiased trajectories are less than the standard errors of their estimation. It is important to note that the calculated rate constants are also in nice agreement with earlier calculations. In terms of transition times the earlier results for the backward reaction (C7$_{\text{eq}} \longrightarrow$ C5) are 0.1-3.0 ps (Ref. 45), 2.7 ps (Ref. 46) and 4.05 ps (Ref. 47) compare well with our value of $1/k_{-1}^a = 6.39$ ps. In sharp contrast, TST significantly overestimates the rates. While overestimation in general is expected from the fact that TST neglects the correlated recrossings, our data show that in this case their contributions are significant for the overall rate.
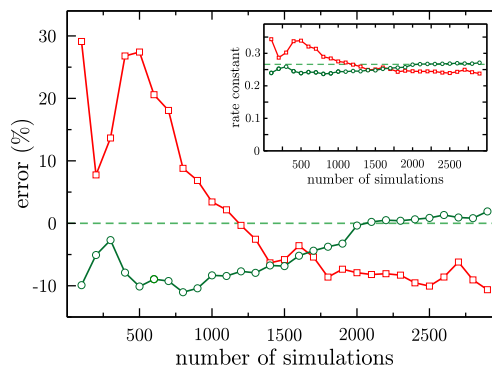
In Fig. 3 the efficiency of our method is compared with that of the EPF approach. We have generated a new NVT ensemble of configurations for the forward DS and for the ds (isocommittor region). By shooting off forward and backward directions with reversed momenta we have obtained the corresponding trajectory segments and calculated the rate constants by the two methods. Fig. 3 shows that both methods converge very quickly although in the present case DST behaves slightly better. Note however that EPF employs an order of magnitude less CPU time by following only a small fraction of the trajectories. Clearly, the efficiency of DST lies in its use to postprocess already generated trajectories.

Barbaralane is a fluxional molecule featuring an intramolecular Cope rearrangement. During the reaction the molecule rearranges itself into chemically identical but configurationally different structure as shown in Fig. 4. At elevated temperature the reaction can be sampled very efficiently[48]. Curiously at this T only the Cope reaction takes place, no other process, such as fragmentation occurs within the simulation time. Hence it is an ideal system to test our method. The $\lambda \in [-0.20 \text{ bohr}, 0.00 \text{ bohr}]$ interval has been selected for the SD region. The FES and all averages were calculated by taking advantage of the inherent symmetry of the rearrangements. We generated trajectories from 200 phase points belonging to the SD. These simulations summed up to 16.3 ps overall simulation time and yielded a rate constant of $1.926 \cdot 10^{12}$ s^{-1}, ie. 0.52 ps transition time. In contrast, the unbiased MD run has been followed for 1 ns and gave a rate
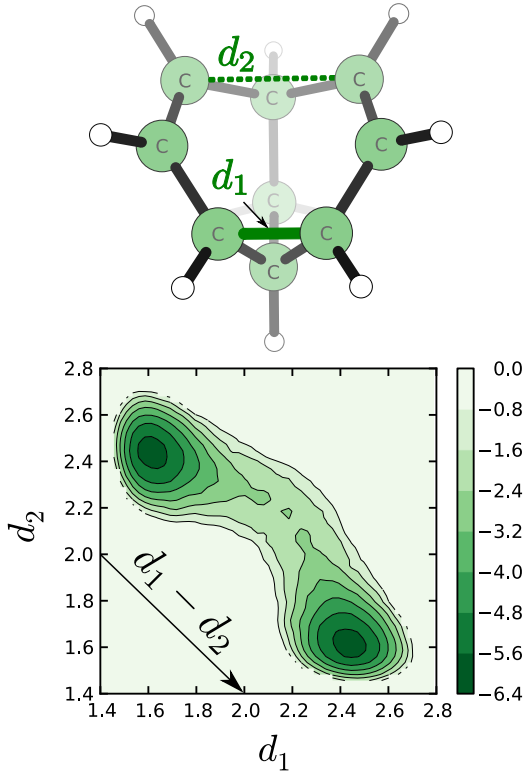
FIG. 4: Definition of the distances, the corresponding free energy surface and the 1D collective variable for the Cope rearrangement of barbaralene. Units are kcal/mol and bohr.

TABLE I: Comparison of the computed rate constants and equilibrium constants with those obtained from unbiased runs. Values are in 10^{11} s^{-1} units. Standard error is given in parenthesis.

	direct[c]	present method	TST[c]
k_1^a	0.266 (0.011)	0.257 (0.014)	1.474 (0.063)
k_{-1}^a	1.629 (0.033)	1.564 (0.068)	11.866 (0.464)
K^a	0.163 (0.006)	0.164 (0.011)	0.124 (0.007)
k^b	19.26 (0.51)	18.63 (1.65)	49.63 (6.25)

[a]alanine-dipeptide;
[b]barbaralene;
[c]direct: rate constants from unbiased MD runs, TST: from Eq. 1.

of barbaralane the K is unity by all methods owing to the degenerate nature of the its Cope rearrangement.

constant of $1.863 \cdot 10^{12}$ s^{-1} which is equivalent to a transition time of 0.54 ps. Similarly to the alanine-dipeptide calculations, the DST rate is in very good agreement with that obtained from direct calculations. TST overestimates the interconversion rate by a factor of three for barbaralane. Clearly, TST performs better for barbaralane than for alanine-dipeptide. This may be due to the fact that the reaction coordinates employed for barbaralane are of better quality and the corresponding transition region on the FES is narrower[49].

Additional information can be obtained from the simulations. The ratio of the forward and backward rate constants gives the equilibrium constant between reactant and product states (RS and PS):

$$K_{RS \rightleftharpoons PS} = \frac{k_1}{k_{-1}} \qquad (19)$$

The equilibrium constant obtained from DST for the C5 $\rightleftharpoons$ C7$_{eq}$ conformational change in alanine-dipeptide compares nicely with that obtained from direct calculations. The 25% underestimation of TST can be attributed to the limitations of the original TST equation (Eq. 1.). In case

VI. CONCLUSIONS

We have developed a method to calculate rate constant from equilibrium simulations in a relatively cheap manner. The method requires the free energy profile of the reaction, implying that a suitable reaction coordinate is already identified. The ideas behind the method are to define the *Saddle Domains* within the reactant and the product states and to recognize that the rate k_{SD} can be calculated very efficiently. The phenomenological rate constants can be easily recovered by reweighting k_{SD} with the statistical weight of the selected SD within the full reactant or product states. The method allows the accurate estimation of the rate of rare events when the mechanism is already known in terms of reaction coordinates. If necessary, the calculated rate constants can be used to derive activation free energies and related quantities. This can be useful to test the applicability of a kinetic model or to compare with previous results. Further studies are underway to assess the efficiency of DST and to compare it with other methods.

VII. ACKNOWLEDGMENT

[1] Hänggi, P.; Talkner, P.; Borkovec, M. *Rev. Mod. Phys.* **1990**, *62*, 251-342.

[2] Bennett, C. H.; in *Algorithms for Chemical Computations*; R. Christofferson, Ed.; ACS Symposium Series No. 46 ACS, Washington, D.C., 1977, 63-97.;

[3] Chandler, D. *J. Chem. Phys.* **1978**, *68*, 2959-2970.

[4] Bolhuis, P.G. ; Dellago, C. in *Rev. Comp. Chem.* Lipkowitz, K. B. Ed. Wiley, New York, 2010, Vol. 27, pp. 111-202.

[5] Ruiz-Montero, M. J.; Frenkel, D.; Brey, J. J. *Mol. Phys.* **1977**, *90*, 925-941.

[6] Hummer, G. *J. Chem. Phys.* **2004**, *120*, 516-523.

[7] Dellago, C.; Bolhuis, P. G.; Csajka, F. S.; Chandler, D. *J. Chem. Phys.* **1998,** *108*, 1964-1977.

[8] van Erp, T. S.; Moroni, D.; Bolhuis, P. G. *J. Chem. Phys.*, **2003**, *118*, 7762-7774.

[9] Moroni, D.; Bolhuis, P. G.; van Erp, T. S. *J. Chem. Phys.* **2004**, *120*, 4055-4065.

[10] Allen, R. J.; Warren, P. B.; ten Wolde, P. R. *Phys. Rev. Lett.*, **2005**, *94*, 018104.

[11] E, W.; Vanden-Eijnden E. in *Multiscale Modelling and Simulation*, Attinger, S.; Koumoutsakos P.; Eds.; LNCSE 39; Springer: Berlin, 2004, pp. 35-68.

[12] Wigner, E. *J. Chem. Phys.* **1937**, *5*, 720-725.

[13] Keck, J. C. *Adv. Chem. Phys.* **1967**, *13*, 85-121.

[14] Pollak, E.; Grabert, H.; Hänggi, P. *J. Chem. Phys.* **1989**, *91*, 4073-4087.

[15] Fernandez-Ramos, A.; Ellingson, B. A.; Garrett, B. C.; Truhlar, D. G. in *Revs. Comp. Chem.* Lipkowitz, K. B.; Cundari, T. R.; Eds.; Vol. 23, 2007, pp. 125-232.

[16] Fernández-Ramos, A.; Miller, J. A.; Klippenstein, S. J.; Truhlar, D. G. *Chem. Rev.* **2006**, *106*, 4518-4584.

[17] Moroni, D. *Efficient sampling of rare event pathways: from simple models to nucleation.* Ph.D. Thesis, University of Amsterdam, 2005.

[18] Kozuch, S. *WIREs Comp. Mol. Sci.* **2012**, *2*, 795-815.

[19] Paranjothy, M.; Sun, R.; Zhuang, Y.; Hase, W. L. *WIREs Comp. Mol. Sci.* **2012**, *2*, 795-815.

[20] Nicolis, G.; Maes, D.; van Erp, T. *Advances in Chemical Physics*, **2012**, *151*, 27-60.

[21] see e.g. Boero, M.; Ikeda, T.; Ito, E.; Terakura, K. *J. Am. Chem. Soc.* **2006**, *128*, 16798-16807.; De Vivo, M.; Dal Peraro, M.; Klein, M. L.; *J. Am. Chem. Soc.* **2008**, *130*, 10955-10962.; Schreiner, E.; Nair, N. N.; Marx, D. *J. Am. Chem. Soc.* **2009**, *131*, 13668-13675.; Comas-Vives, A. Stirling, G. Ujaque, A. Lledós, *Chem. Eur. J.*, **2010**, *16*, 8738-8747.; Santarossa, G.; Vargas, A.; Iannuzzi, M.; Baiker, A. *Phys. Rev. B*, **2010**, *81*, 174205.; Kovács, G.; Stirling, A.; Lledós, A.; Ujaque, G. *Chem. Eur. J.*, **2012**, *18*, 5612-5619.; Dopieralski, P.; Ribas-Arino, J.; Anjukandi, P.; Krupicka, M.; Kiss, J.; Marx, D. *Nat. Chem.*, **2013**, *5*, 685-691.

[22] Du, R.; Pande, V. S.; Grosberg, A. Y.; Tanaka, T.; Shakhnovich, E. S. *J. Chem. Phys.* **1998**, *105*, 334-350.

[23] Bolhuis, P.; Chandler, D.; Dellago, C.; Geissler, P. *Annu. Rev. Phys. Chem.* **2002**, *53*, 291-318.

[24] Peters. B. *J. Chem. Phys.* **2006**, *125*, 241101.

[25] Anderson, J. B. *Adv. Chem. Phys.* **1995**, *91*, 381-431.

[26] van Erp, T. S. *Solvent effects on chemistry with alcohols.* PhD thesis, Universiteit van Amsterdam, **2003**.

[27] Vanden-Eijnden E.; Tal, F. A. *J. Chem. Phys.* **2005**, *123*, 184103.

[28] van Erp, T. S.; Bolhuis, P. G. *J. Comp. Phys.* **2005**, *205*, 157-181.

[29] Generalization of the method for several λ-s is straightforward. For discussion about how to find accurate reaction coordinates see e.g. Ref. 24, Ma. A.; Dinner, A. R. *J. Phys. Chem. B* **2005**, *109*, 6769-6779.; Peters, B.; Beckham, G. T.; Trout, B. L. *J. Chem. Phys.* **2007**, *125*, 034109.; Peters. B. *Chem. Phys. Lett.* **2012**, *554-*, 248-253. Peters, B.; Bolhuis, P. G.; Mullen, R. G.; Shea, J.-E. *J. Chem. Phys.* **2013**, *138*, 054106.

[30] The actual value of the mean lifetime (or mean residence time) depends on the definition of a reactive event.

[31] Ref. 6 also discusses the use of lifetime for rate constant calculation in TPS context.

[32] In principle the separation of the two adjacent SD regions can be at arbitrary value of λ, without affecting the validity of the equations. However to make compare with experiments $\lambda = ds$ is the optimal choice.

[33] Vanden-Eijnden, E. *Lect. Notes Phys.* **2006**, *703*, 439–478.

[34] Note that shifting the indicator line is analogous with varying the time parameter in the original Bennett-Chandler formulation[3,4]. In particular, the insensitivity of k_{SD} to a small variation of the position of the indicator lines corresponds to the plateau of $k(t)$ calculated from the time correlation function $C(t)$.

[35] The maximum of FES along a λ does not necessarily coincide with the ds. See e.g. Refs. 4,11.

[36] CP2K version 2.2.372; CP2K Developers Group, 2010. CP2K is freely available from http://www.cp2k.org/ (accessed date Jan 13. 2014).

[37] Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz Jr, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.*, **1995**, *117*, 1995, 5179-5196.

[38] Laio, A; Parrinello, M. *Proc. Natl. Acad. Sci. U.S.A.*, **2002**, *20*, 12562-12566.; Iannuzzi, M.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.*, **2003**, *90*, 238302.; Laio, A.; Parrinello, M. *Lect. Not. Phys.*, **2006**, *707*, 315-347. Laio, A.; Gervasio, F. L. *Rep. Prog. Phys.*, **2008**, *71*, 126601.

[39] Henkelman G.; Jonsson, H. *J. Chem. Phys.* **2000**, *113*, 9978-9985.
Henkelman, G.; Uberuaga, B. P.; Jonsson, H. *J. Chem. Phys.*, **2000**, *113*, 9901-9904.

[40] Becke, A. *Phys. Rev. A* **1988**, *38*, 3098-3100; Lee, C.; Yang, W.; Parr, R. *Phys. Rev. B* **1988**, *37*,

785-789.

[41] Goedecker, S.; Teter, M.; Hutter, J. *Phys. Rev. B* **1996**, *54*, 1703-1710.

[42] Martyna, G. J.; Tuckerman, M. E. *J. Chem. Phys.*, **1999**, *110*, 2810-2821.

[43] Hub, J. S.; de Groot, B. L.; van der Spoel, F. *J. Chem. Theory Comput.*, **2010**, *6*, 3713-3720.

[44] Stryer, L. *Biochemistry*, 4th edition, Freeman W. H.; & Company: New York, 1995, 420-421.

[45] Vedell, P.; Wu, Z. *Int. J. Num. Anal. Mod.*, **2008**, *10*, 920-942.

[46] Chun, H. M.; Padilla, C. E.; Chin, D. N.; Watanabe, M.; Karlov, V. I.; Alper, H. E.; Soosaar, K.; Blair, K. B.; Becker, O. M.; Caves, L. S. D.; Nagle, R.; Haney, D. N.; Farmer, B. L. *J. Comp. Chem.*, **2000**, *21*, 159-184.

[47] Velez-Vega, C.; Borrero, E. E.; Escobedo, F. A. *J. Chem. Phys.*, **2009**, *130*, 225101.

[48] Rozgonyi, T.; Bartók-Pártay, A.; Stirling, A. *J. Phys. Chem. A*, **2010**, *114*, 1207-1211.

[49] We thank the anonymous referee for raising this issue.

VIII. TOC GRAPHICS