PCCP

Transition Path Times in Asymmetric Barriers

Journal Name

Transition Path Times in Asymmetric Barriers

Michele Caraglio,[a,b] Takahiro Sakaue,[c,d] and Enrico Carlon[a‡]

Biomolecular conformational transitions are usually modeled as barrier crossings in a free energy landscape. The transition paths connect two local free energy minima and transition path times (TPT) are the actual durations of the crossing events. The simplest model employed to analyze TPT and to fit empirical data is that of a stochastic particle crossing a parabolic barrier. Motivated by some disagreement between the value of the barrier height obtained from the TPT distributions as compared to the value obtained from kinetic and thermodynamic analyses, we investigate here TPT for barriers which deviate from the symmetric parabolic shape. We introduce a continuous set of potentials, that starting from a parabolic shape, can be made increasingly asymmetric by tuning a single parameter. The TPT distributions obtained in the asymmetric case are very well-fitted by distributions generated by parabolic barriers. The fits, however, provide values for the barrier heights and diffusion coefficients which deviate from the original input values. We show how these findings can be understood from the analysis of the eigenvalues spectrum of the Fokker-Planck equation and highlight connections with experimental results.

Introduction

Conformational transitions in complex molecular systems are typically described by the dynamics of a reaction coordinate in a free energy landscape. Two local free energy minima correspond to two distinct states of the system, say A and B. The transitions between the two occur through the crossing of a free energy barrier. In this process two different time scales can be identified. The first is the dwell time, which is the average time the system stays in A (B) before making a transition to B (A). A second time scale is the transition path time (TPT) which is the time needed to actually cross the barrier from A to B. The system spends most of the time fluctuating around the local minima A or B, while transitions between them are very rapid. The average dwell times can therefore be considerably longer that the average TPT, as indeed seen in experiments of protein and nucleic acids folding[1,2].

Considerable attention has been devoted in the past decade to study the properties of TPT, both in experiments[3–7] and in theories[8–16], see also Ref.[17] for a recent review. The analysis of TPT has attracted quite some interest in the past as it is expected to shed some light on the strong molecular rearrangements occur-

ing close to the top of the free energy barrier. A central model is that of a diffusive particle crossing a parabolic potential barrier $V(x) = -kx^2/2$, where x denotes a reaction coordinate and $k > 0$ (note that also paths crossings a linear symmetric cusped potential barrier have been recently considered[13,16]). Given a generic potential barrier $V(x)$ one defines transition paths as those originating at one side of the barrier in a position $x = x_1$ and end at the opposite side at $x = x_2$. Absorbing boundary conditions are imposed at $x = x_1$ and $x = x_2$ to eliminate the recrossing events at the begin and end points of the trajectories. In practice, however, free boundary conditions are used[8,14], as these are analytically simpler to implement and converge to the absorbing boundaries results in the high barrier limit ($\Delta U \gg k_B T$, with ΔU the barrier height, T the temperature and k_B the Boltzmann constant) where recrossings are extremely rare.

For parabolic barriers exact TPT distributions were derived in the free boundary conditions case[8], providing simple expressions which have been used to fit experimental data[5]. Calculations with parabolic barriers were also extended to the case of inertia[14] and in the presence of memory or active forces[12,18,19]. Absorbing boundary conditions were recently implemented in the parabolic model as well[15]. Although any smooth potential can be approximated by a parabola sufficiently close to its maximum, this approximation may hold only for a very small reaction coordinate range. It is therefore interesting to analyze some wider set of potentials departing from a perfect parabolic shape, which is the aim of the present paper. We focus here on a barrier which is composed of a parabolic part joined to a linear part. In the model the degree of asymmetry of the potential can be tuned starting from

a KU Leuven, Soft Matter and Biophysics Unit, Celestijnenlaan 200D, B-3001 Leuven, Belgium.

b Institut für Theoretische Physik, Universität Innsbruck, Technikerstraße 21A, A-6020 Innsbruck, Austria.

c Department of Physics and Mathematics, Aoyama Gakuin University, 5-10-1 Fuchinobe, Chuo-ku, Sagamihara, Kanagawa 252-5258, Japan.

d PRESTO, Japan Science and Technology Agency, 4-1-8 Honcho Kawaguchi, Saitama 332-0012, Japan.

‡ Corresponding author: enrico.carlon@kuleuven.be
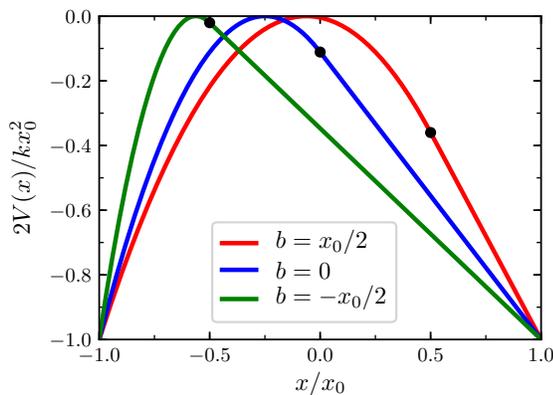
Fig. 1 Plot of the rescaled potential $2V(x)/kx_0^2$ vs. rescaled distance x/x_0 defined by Eq. (1) for $b = x_0/2$, $b = 0$ and $b = -x_0/2$. The symmetric parabolic potential corresponds to $b = x_0$ and, as the parameter b decreases, the potential becomes increasingly asymmetric. $V(x)$ is parabolic for $-x_0 \le x \le b$ and linear for $b \le x \le x_0$.

the fully symmetric parabolic shape to a strongly asymmetric linear barrier. TPT distributions are obtained by mapping the problem into one dimensional effective Schrödinger equation [20]. As both parabolic and linear potentials are solvable, one just needs to fix the matching conditions at the boundary between them, as done in quantum mechanical problems with discontinuous potentials. We show how TPT distributions from asymmetric potential can always be fitted by distributions obtained from parabolic potentials, showing that the shape of the distribution is quite universal, i.e., insensitive to the actual potential shape. The fit however leads to estimates of the barrier height and of the diffusion coefficient that deviate from the original input values. We discuss the origin of these effects and their relevance in view of contradicting estimates of barrier heights obtained in recent DNA folding experiments.

The model

We consider a particle diffusing in a potential $V(x)$ defined in the interval $-x_0 \le x \le x_0$ and parametrized as follows

$$V(x) = -\frac{k}{2}\left(\frac{x_0}{x_0 + \widetilde{x}}\right)^2 f(x) \qquad (1)$$

with

$$f(x) = \begin{cases} (x - \widetilde{x})^2 & -x_0 \le x \le b \\ (2x - b - \widetilde{x})(b - \widetilde{x}) & b \le x \le x_0 \end{cases} \qquad (2)$$

Here

$$\widetilde{x} = -\frac{(x_0 - b)^2}{4x_0} \le 0 \qquad (3)$$

is the location of the maximum of the potential $V(\widetilde{x}) = 0$. Figure 1 shows a plot of $V(x)$ for three different values of b. In the case $b = x_0$ the potential is parabolic. As b decreases, $V(x)$ becomes more asymmetric with an increasing negative curvature $\kappa_p = -kx_0^2/(x_0 + \widetilde{x})^2$ in the parabolic region and a decrease of slope $s_\ell = \kappa_p(b - \widetilde{x})$ in the linear region (see Fig. 1). The barrier height is independent of b and given by $\Delta U = \frac{1}{2}kx_0^2$. Note that, by con-

struction, $V(x)$ and $V'(x)$ are continuous in $x = b$.

The potential (1) is sufficiently simple so that one can infer analytically, as we shall see, the dominant rates of the process in the limiting cases of high and low asymmetry. The effect of asymmetry found for the parabolic-linear potential (1), is however quite general and holds also for other choices of potential barriers, as discussed in the Appendix. Note that in our description we choose the endpoints of a transition path, x_1 and x_2, to correspond to the same free energy level and we focus on the barrier region only, disregarding the potential minima at the two sides of the barrier. In typical experimental studies of TPT one tunes external parameters (as temperature or forces in optical tweezers), so that many back and forth transitions between the two local minima are observed [2]. The dwell times, the average times spent in each of the minima, are therefore set to be similar, hence the free energies of these states are very close to each other. In typical experimental settings, one can consider the free energies of the transition paths endpoints x_1 and x_2, to be close to each other as well. In the most generic case, the minima may have different free energies and also the endpoints of a transition path may be at different free energy levels. This introduces an additional source of variability, which is not discussed here.

The Fokker-Planck equation, describing the evolution of the distribution function $P(x,t)$, is given by:

$$\partial_t P(x,t) = D\partial_x \left[e^{-\beta V(x)} \partial_x \left(e^{\beta V(x)} P(x,t) \right) \right] \equiv L_{FP} P(x,t) \qquad (4)$$

where D is the diffusion coefficient $\beta = 1/k_B T$ and L_{FP} defines the Fokker-Planck operator. We impose absorbing boundary conditions, hence $P(\pm x_0, t) = 0$ at all times.

We use separation of variables and seek solutions of (4) of the form

$$P(x,t) = e^{-\lambda D t} e^{-\beta V(x)/2} q(x), \qquad (5)$$

The parameter λ determines the decay rate of a solution and from general properties of the Fokker-Planck equation one has $\lambda > 0$, see [20]. Substituting (5) in (4) we get

$$q''(x) - V_{eff}(x) q(x) = -\lambda q(x) \qquad (6)$$

where

$$V_{eff}(x) = -\frac{\beta}{2} V''(x) + \frac{\beta^2}{4} \left[V'(x) \right]^2 \qquad (7)$$

defines an effective potential [20]. We note that (6) is formally identical to the Schrödinger equation for a quantum particle in a potential $V_{eff}(x)$ (we have set the units $\hbar^2/2m = 1$, the potential thus defined has the dimension of a squared inverse length) and energy λ. The advantage of the transformation (5) is that (7) defines an eigenvalue problem with an hermitian operator (note that L_{FP} in (4) is not hermitian), hence eigenfunctions form a orthogonal complete set:

$$\int_{-x_0}^{+x_0} q_n(x) q_m(x)\, dx = \delta_{nm} \qquad (8)$$

with n and m "quantum" numbers. Plugging (1) into (7) we obtain

$$V_{\text{eff}}(x) = \begin{cases} \frac{\widetilde{k}}{2} + \frac{\widetilde{k}^2}{4}(x - \widetilde{x})^2 & -x_0 \leq x \leq b \\ \frac{\widetilde{k}^2}{4}(b - \widetilde{x})^2 & b \leq x \leq x_0 \end{cases} \quad (9)$$

where we have defined

$$\widetilde{k} = \beta k \left(\frac{x_0}{x_0 + \widetilde{x}}\right)^2 = \frac{16\beta k x_0^4}{(x_0 + b)^2(3x_0 - b)^2} \quad (10)$$

(using (3)). Note that $\widetilde{k} = -\beta \kappa_p$, hence the curvature of V_{eff} at $x \in [-x_0, b]$ is $\widetilde{k}^2/2 = (\beta \kappa_p)^2/2$, and the constant value in V_{eff} at $x \in [b, x_0]$ (bottom of (9)) is written as $(\beta s_\ell/2)^2$ using the curvature and the slope of the original potential, respectively. The curvature is proportional to $\widetilde{k}^2$ and hence positive, irrespectively on the sign of k. We also note that although $V(x)$ and $V'(x)$ are by construction continuous, $V_{\text{eff}}(x)$ has a discontinuous jump in $x = b$. The jump originates from the discontinuity of $V''(x)$ in $x = b$, which gives $\Delta V_{\text{eff}} = V_{\text{eff}}(b^+) - V_{\text{eff}}(b^-) = \widetilde{k}/2$. Although the potential in the Fokker-Planck problem is repulsive, the effective potential of the associated quantum mechanical problem is attractive and temperature dependent.

Figure 2 plots $V_{\text{eff}}(x)$ for three different values of b and fixed k, β and x_0. For weak asymmetry ($b^* \leq b \leq x_0$) the minimum of V_{eff} is in the parabolic part of $V(x)$. For $-x_0 < b \leq b^*$ the minimum appears in the constant part of $V_{\text{eff}}(x)$. We find

$$b^* = x_0 \frac{3 - \sqrt{\beta \Delta U}}{1 + \sqrt{\beta \Delta U}} \quad (11)$$

Note that when the asymmetry is increased, V_{eff} increases in the parabolic region, while it decreases in the linear region.

Although V_{eff} has a discontinuous jump, the solution $q(x)$ of Eq. (6) and its derivative $q'(x)$ remain continuous in $x = b$. To see this we integrate (6) in an infinitesimal interval $[b - \varepsilon, b + \varepsilon]$:

$$q'(b + \varepsilon) - q'(b - \varepsilon) = \int_{b - \varepsilon}^{b + \varepsilon} [V_{\text{eff}}(x) - \lambda]\, q(x)\, dx \quad (12)$$

As $\varepsilon \to 0$ the integral in the right hand side vanishes, implying continuous derivative $q'(b^+) = q'(b^-)$ in $x = b$. In conclusion we solve the eigenvalue problem requiring that $q(x)$ and $q'(x)$ to be continuous in $x = b$.

The Eigenvalue Problem

We solve (6) separately in the two domains and then use the matching conditions of continuity of q and q' in $x = b$.

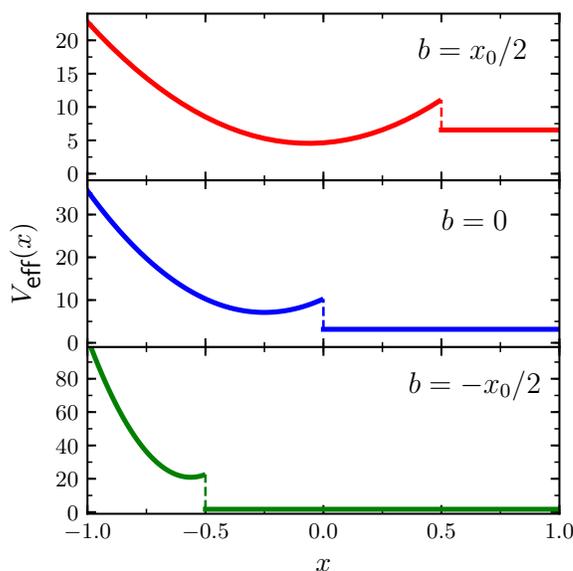**The solution in the linear region ($b \leq x \leq x_0$)**

In this region V_{eff} is constant, hence the solution vanishing at the boundary $q(x_0) = 0$ is

$$q(x) = B \sin[\kappa(x_0 - x)] \quad (13)$$

where B is a constant and

$$\kappa^2 = \lambda - \frac{\widetilde{k}^2}{4}(b - \widetilde{x})^2 \quad (14)$$

as obtained by plugging in (13) into (6).

Fig. 2 Plot of $V_{\text{eff}}(x)$ for $x_0 = 1$, $\beta k = 8$ and three different values of b ($b = 0.5, 0, -0.5$ from top to bottom). We have here $\beta \Delta U = 4$, hence from Eq. (11) we get $b^* = x_0/3$, corresponding to the value of b where the minimum of the parabolic part of V_{eff} is at the same level as the constant part of the potential.

The solution in the parabolic region ($-x_0 \leq x \leq b$)

As the parabolic potential is centered in $\widetilde{x}$ it is convenient to shift the x-coordinate and define

$$Y(x) \equiv q(x + \widetilde{x}) \quad (15)$$

Equation (6) then becomes

$$Y''(x) - \frac{\widetilde{k}^2 x^2}{4} Y(x) = -\left(\lambda - \frac{\widetilde{k}}{2}\right) Y(x) \quad (16)$$

The solutions of this equation are related to the Kummer confluent hypergeometric functions ${}_1F_1(a; b; z)$. The solutions even in x are

$$Y_+(x) = e^{-\widetilde{k}x^2/4} {}_1F_1\left(\frac{s_n + 1}{2}; \frac{1}{2}; \frac{\widetilde{k}x^2}{2}\right) \quad (17)$$

and the odd ones

$$Y_-(x) = \sqrt{\widetilde{k}}\, x\, e^{-\widetilde{k}x^2/4} {}_1F_1\left(\frac{s_n}{2} + 1; \frac{3}{2}; \frac{\widetilde{k}x^2}{2}\right) \quad (18)$$

We have defined here $s_n = -\lambda_n/\widetilde{k}$. The Kummer hypergeometric functions are defined by

$${}_1F_1(a; b; z) = \sum_{k=0}^{\infty} \frac{a^{(k)}}{b^{(k)}k!} z^k \quad (19)$$

where $a^{(0)} = 1$, $a^{(k)} = a(a+1)(a+2)\ldots(a+k-1)$. When the argument a is a negative integer, the sum in (19) becomes finite and reduces (apart some multiplicative factors) to Hermite polynomials.

The most general solution is obtained by a linear combination
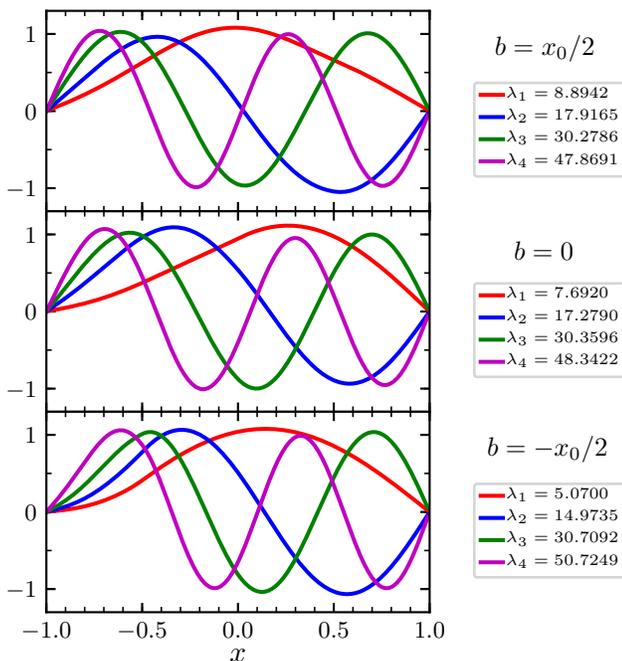
Fig. 3 Normalized eigenfunctions $q_n(x)$ for $n = 1, 2, 3, 4$ with $x_0 = 1$, $\beta k = 8$ (corresponding to $\beta \Delta U = 4$) and three different values of b. The asymmetry has a stronger influence on the ground state $q_1(x)$, while it has a milder effect on the higher excited states.

Fig. 4 Red Solid line: TPT distribution for $x_0 = 1$, $\beta \Delta U = 4$, $b = -x_0/2$ and $D = 2.5$ as obtained from Eq.(28). As we set the length to be dimensionless, also λ becomes dimensionless. As a consequence D has the dimension of an inverse time. Here we set both D and t to dimensionless units. Green bars: TPT histogram obtained from BD simulations. Blue Dashed line: TPT histogram fitted from a parabolic barrier with $x_0 = 1$. The fit reproduces very well the original data in the whole range, however with parameters ($\beta \Delta U = 1.8(1)$ and $D = 2.81(1)$) which deviate significantly from the original values. Inset: Comparison between the full distribution from Eq.(28) (red solid line) and truncations of the same expression to $n = 1, 2, 3$ terms.

of (17) and (18)

$$q(x) = Y_+(x - \widetilde{x}) + AY_-(x - \widetilde{x}) \tag{20}$$

(without loss of generality, except in the symmetric limit $b = x_0$, corresponding to $\widetilde{x} = 0$, one can set one of the two coefficients in (20) equal to 1). Absorbing boundary conditions impose a vanishing solution in $-x_0$ which yields

$$Y_+(x_0 + \widetilde{x}) = AY_-(x_0 + \widetilde{x}) \tag{21}$$

where we have used $Y_\pm(-x) = \pm Y_\pm(x)$.

Matching conditions
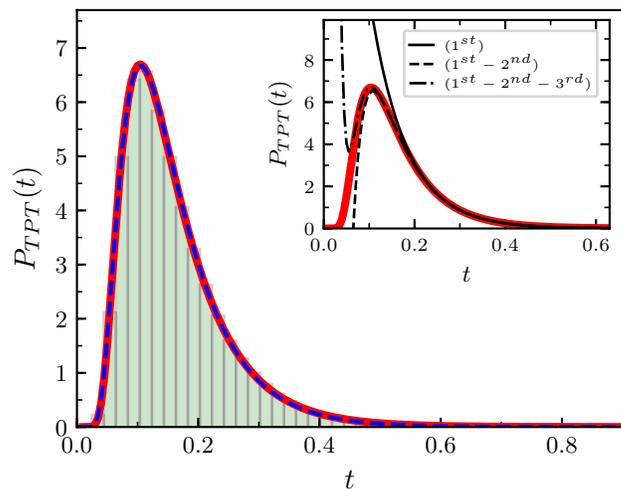
Continuity of $q(x)$ and $q'(x)$ in $x = b$ implies

$$Y_+(b - \widetilde{x}) + AY_-(b - \widetilde{x}) = B\sin[\kappa(x_0 - b)] \tag{22}$$

$$Y'_+(b - \widetilde{x}) + AY'_-(b - \widetilde{x}) = -\kappa B\cos[\kappa(x_0 - b)] \tag{23}$$

Combining (21), (22) and (23) we can eliminate A and B to obtain a single equation

$$\frac{Y_-(x_0 + \widetilde{x})Y_+(b - \widetilde{x}) + Y_+(x_0 + \widetilde{x})Y_-(b - \widetilde{x})}{Y_-(x_0 + \widetilde{x})Y'_+(b - \widetilde{x}) + Y_+(x_0 + \widetilde{x})Y'_-(b - \widetilde{x})} = -\frac{1}{\kappa}\tan[\kappa(x_0 - b)] \tag{24}$$

(24) needs to be solved numerically to calculate the spectrum of allowed values of λ, which is the only unknown in the equation. Once λ_n ($n = 1, 2 \ldots$) has been obtained one can use (21) to get A and either (22) or (23) to get B. The eigenfunctions are then normalized to fulfill (8).

Properties of the spectrum

Figure 3 shows a list of the lowest eigenvalues λ_n and the corresponding q_n for fixed k, x_0, β and for three different values of b. Note that the eigenfunctions have an increasing number of nodes as the eigenvalue increases, as well-known in one dimensional quantum mechanics. The nodes count provides a good check that no eigenvalues are missed in the numerical calculation of the spectrum. For sufficiently large n, quantum energies are high compared to the variations of V_{eff} in the interval $[-x_0, x_0]$ and the wave functions are close to that of a free particle in a box of width $2x_0$:

$$q_n(x) \approx \frac{1}{\sqrt{x_0}}\sin\left[\pi n\frac{x + x_0}{2x_0}\right] \tag{25}$$

with eigenvalues $\lambda_n \approx \frac{\pi^2 n^2}{4x_0^2}$. The low-lying part of the spectrum is crucially dependent on the shape of the potential, and this determines the long time behavior of the system.

TPT distributions

To obtain the TPT distribution for barrier crossing events from $-x_0$ to x_0 one solves the Fokker-Planck equation (4) using a δ source in $-x_0 + \varepsilon$. The probability current is given by

$$j_\varepsilon(x, t) = -D\left[e^{-\beta V(x)}\partial_x\left(e^{\beta V(x)}P_\varepsilon(x, t)\right)\right] \tag{26}$$

with $P_\varepsilon(x, t)$ the solution with the said initial condition. The TPT distribution is proportional to the current exiting from the oppo-

site boundary (x_0), hence

$$p_{\text{TPT}}(t) = \lim_{\varepsilon \to 0} \frac{j_\varepsilon(x_0, t)}{\int_0^\infty j_\varepsilon(x_0, t') \, dt'} \qquad (27)$$

This quantity can be written as follows (see e.g. Ref. [15] for a more detailed derivation):

$$p_{\text{TPT}}(t) = -\frac{1}{A} \sum_{n=1}^\infty q_n'(-x_0) q_n'(x_0) e^{-\lambda_n D t} \qquad (28)$$

where the constant

$$A = -\sum_{n=1}^\infty q_n'(-x_0) q_n'(x_0) \frac{1}{\lambda_n D} \qquad (29)$$

normalizes the distribution so that $\int_0^\infty p_{\text{TPT}}(t) \, dt = 1$. We note that the expression (28) is symmetric with respect to the direction of the paths, i.e. transitions $x_0 \to -x_0$ and $-x_0 \to x_0$ have the same distribution, reflecting a well-known time reversal symmetry[9] (holding also for asymmetric barriers). Equation (28) is a series with coefficients of alternating signs. This can be understood from the properties of eigenfunctions. The ground state has no nodes hence the derivatives at the two ends have opposite signs $q_1'(-x_0) q_1'(x_0) < 0$. The number of nodes of $q_n(x)$ is equal to $n-1$, therefore in general $q_k'(-x_0) q_k'(x_0) > 0$ for k even and $q_k'(-x_0) q_k'(x_0) < 0$ for k odd.

Figure 4 shows a plot of a TPT distribution calculated from Eq. (28) (red solid line) with $x_0 = 1$, $\beta \Delta U = 4$ and $D = 2.5$. Data from Brownian Dynamics (BD) simulations of barrier crossing events for the same set of parameters are shown as green bars. The overlap between the two sets confirms the validity of the numerical procedure to obtain the eigenvalues and eigenfunctions from (24). While the BD simulations need large statistics, the TPT distribution can be obtained more rapidly from (28), which is particularly advantageous in the high barrier limit. Moreover from (28) one can analyze separately the contribution of each eigenstate to the total sum. In the calculations the sums in (28) and (29) are truncated to $n = 100$, a threshold value which guarantees very good accuracy of the distribution. In Fig. 4 we also plotted as a blue dashed line the best fit to the data using a parabolic potential with $x_0 = 1$, where βk (or equivalently $\beta \Delta U$) and D are taken as fitting parameters. The fitted curve overlaps extremely well with the original data (showing that the TPT distribution for a strongly asymmetric barrier is for practical purposes indistinguishable from that generated by particles crossing a parabolic barrier), but yields $\beta \Delta U = 1.8(1)$ and $D = 2.81(1)$, which significantly deviate from the original parameters used in the asymmetric potential. In particular the barrier height is strongly underestimated, with respect to the original value $\beta \Delta U = 4$. In order to infer the contribution of the various terms in the sum (28) to the total TPT distribution we plotted in the inset of Fig. 4 again the full $p_{\text{TPT}}(t)$ from Eq. (28) (using the first 100 terms) and truncations to $n = 1$, 2 and 3 terms. The truncation is done only in the numerator of (28), while the normalization (29) is obtained from the full calculation up to $n = 100$. The analysis shows that the first three terms in (28) approximate well the whole decaying part of the distribution and even the region beyond, but close to
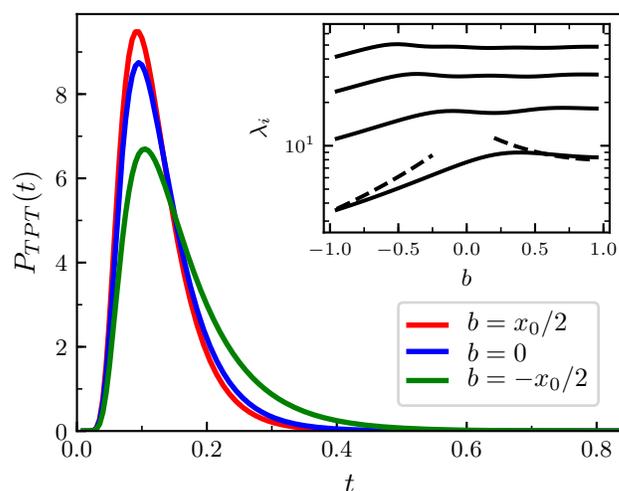
Fig. 5 Examples of TPT distributions $P_{TPT}(t)$ for $x_0 = 1$, $\beta \Delta U = 4$ ($\beta k = 8$) and $D = 2.5$. The average TPT increases with increasing skewness of the energy barrier (i.e. decreasing b). Typically, the distributions are obtained truncating the infinite series (28) to approx $n \leq 100$. Inset: behavior of λ_1, λ_2, λ_3 and λ_4 (respectively from bottom to top) as a function of b for $x_0 = 1$ and $\beta \Delta U = 4$. The dashed lines around $b \approx 1$ and $b \approx -1$ are respectively $\lambda_1 = \tilde{k}$ and Eq. (30) and approximate well the ground state energy around these two limits, see text.

the maximum.

Figure 5 shows the TPT distribution as obtained from Eq. (28) for the same parameters as in Fig. 4 and different values of b. In the strongest asymmetric case ($b = -x_0/2$) the decay is the slowest and the average TPT is the highest of the three cases analyzed. The inset of Fig. 5 plots the first four eigenvalues for $-0.95 \leq b/x_0 \leq 0.95$, showing that the ground state λ_1 is most strongly affected by variations of b, while the influence of b on higher eigenvalues is much weaker. For highly excited eigenstates λ_n converges to the free particle in a box spectrum $\lambda_n \approx (\pi n / 2 x_0)^2$, which is independent on $V_{\text{eff}}(x)$, and hence on b. The shape of the underlying potential hence influences few low lying eigenstates. We also note that λ_1 is a non-monotonic function of b, which is due to the two competing trends on $V_{\text{eff}}(x)$ visible in Fig. 2. As b decreases from the symmetric case $b = x_0$ the parameter $\tilde{k}$ defined in Eq. (10) increases. For $b \approx x_0$ we can approximate the ground state value to that of an harmonic oscillator in an infinite domain $\lambda_1 = \tilde{k}$. In the other limit of large asymmetry $b \approx -x_0$, the ground state wavefunction is localized in the constant part of V_{eff} in the interval $b \leq x \leq x_0$ (see (9)). In first approximation we can set the wavefunction to zero in $x = b$ and x_0, leading to the ground state of a free quantum particle localized in $[b, x_0]$, which is

$$\lambda_1 = \frac{\pi^2}{(x_0 - b)^2} + \frac{\tilde{k}^2}{4}(b - \tilde{x})^2 \qquad (30)$$

Both $\lambda_1 = \tilde{k}$ and Eq. (30) approximate well the ground state energy for $b \approx x_0$ and $b \approx -x_0$, as shown by the dashed lines in the inset of Fig. 5, and the transition between these two behaviors takes place around $b \simeq b^*$ (11), at which λ_1 is maximal.

Table 1 summarizes the results of the analysis. TPT distri-

Table 1 Summary of the analysis of TPT distributions generated from Eq. (28) for $x_0 = 1$, $D = 2.5$ and different values of $\beta\Delta U$ and asymmetry parameter b. The second column gives the average TPT obtained from the first moment of the distribution. The last six columns report the parameters calculated from a fitting to a $p_{TPT}(t)$ generated by a symmetric parabolic potential. The three columns "ABC" refer to a fit with the absorbing boundary condition case, that is Eq. (28) with $b = 1$. As expected, the fit for $b = 1$ reproduce the input values for $\beta\Delta U$, and D. The last three columns are for a parabolic barrier with free boundary conditions given by Eq. 31. The quality of the fit is estimated by χ, as defined by Eq. (33). The overlapping distributions (solid red line and the fitted dashed blue line) of Fig. 4 have a $\chi = 10^{-3}$. Uncertainties on $\langle t_{TPT} \rangle$, and fitted $\beta\Delta U$, D are approximately ± 1 in the last digit.

$\beta\Delta U = 4, D = 2.5$		fit to ABC			fit to FBC		
b	$\langle t_{TPT} \rangle$	$\beta\Delta U$	D	χ	$\beta\Delta U$	D	χ
1.0	0.126	4.0	2.50	0.000	3.6	3.11	0.005
0.5	0.123	4.3	2.44	0.001	3.9	3.00	0.003
0	0.131	3.6	2.53	0.001	3.3	3.15	0.006
−0.5	0.160	1.8	2.81	0.001	2.0	3.66	0.015

$\beta\Delta U = 8, D = 2.5$		fit to ABC			fit to FBC		
b	$\langle t_{TPT} \rangle$	$\beta\Delta U$	D	χ	$\beta\Delta U$	D	χ
1.0	0.082	8.0	2.50	0.000	7.1	2.90	0.003
0.5	0.079	9.1	2.35	0.004	8.1	2.69	0.001
0	0.083	8.4	2.38	0.006	7.4	2.76	0.002
−0.5	0.104	4.5	2.82	0.001	4.0	3.35	0.010

butions were generated from Eq. (28) with parameters $x_0 = 1$, $D = 2.5$, $\beta\Delta U = 4$ and $\beta\Delta U = 8$ and four different values of b. These data were fitted to TPT distributions generated from (28) but with a parabolic potential, e.g. $b = x_0 = 1$. Another fit was performed against the distribution[8,14]

$$p_{TPT}(t) = -\frac{2}{\pi} \frac{G'(t) e^{-G^2(t)}}{1 - \text{Erf}[\sqrt{\beta\Delta U}]} \,, \tag{31}$$

where

$$G(t) = \sqrt{\beta\Delta U} \sqrt{\frac{e^{\Omega t} + 1}{e^{\Omega t} - 1}} \tag{32}$$

and where $\Omega = \beta k D$ is the characteristic rate of the process. Equation (31) is the TPT distribution for a particle crossing a parabolic potential, obtained from the free boundary condition case. This simpler analytical expression (compared to the infinite series (28)) is most commonly used to fit experimental data[5]. For very high barrier multiple crossing at boundaries are highly unlikely and the free boundary conditions distribution provides a very good approximation to the exact one. As check of the quality of the fits the following parameter was used

$$\chi = \int_0^\infty \left[p_{TPT}(t) - p_{TPT}^{(fit)}(t) \right]^2 p_{TPT}(t) \, dt \tag{33}$$

In all cases analyzed, and as shown in Fig. 4 the TPT distributions are well-fitted by distributions obtained from a parabolic potential, as demonstrated by the very low value of χ reported in Table 1. However, as discussed in the case of Fig. 4, the fitted values can significantly deviate from the original ones. Table 1 shows that the effect is most significant for estimates of the barrier height.

Discussion

Transition paths are events of very short duration (typically in the μs scale for protein and nucleic acid folding[1]), therefore their experimental study has been very challenging for long time. Recent advances in single molecule techniques, have allowed the determination of average TPT in biomolecular folding experiments[1,2] and in some cases also the full TPT distribution[5]. The comparison of experiments with theory has revealed that some quantities match quite well the simple barrier crossing model. In particular, estimates of the diffusion coefficient, as obtained from the analysis of average TPT, their distribution, shapes and velocities, turned out to be quite robust and consistent among each other[21]. However, a markable disagreement was found in the determination of the barrier height[5], where a fit of $p_{TPT}(t)$ (as given by Eq. (31)) gave a very small barrier height $\beta\Delta U \approx 0.5$, as compared to other thermodynamic measurements, from which $\beta\Delta U \approx 9$ was estimated. This large difference has triggered several theoretical studies to reconcile the measurements with theory discussing the entropy of the barrier[22], nonthermal noise[19], anomalous diffusion[18,19], memory effects in the dynamics[19,23] and boundary conditions[12,15]. In this paper we investigate the effect of the barrier shape on TPT distributions by considering potential barriers which deviate from the parabolic ones.

Our analysis showed that TPT distributions generated from asymmetric potentials are for all practical purposes indistinguishable from those obtained from a parabolic potential. Fitting the former against the latter returns two values as fitting parameters: the scaled barrier height $\beta\Delta U$ and the diffusion constant D. The results, summarized in Table 1, show that these parameters deviate from the input values. For instance, for the barrier (1) with $b = -0.5$, $\beta\Delta U = 4$, $D = 2.5$, the fit to a TPT distribution with a parabolic potential and absorbing boundaries, gives $\beta\Delta U = 1.8$ and $D = 2.8$. While the variation in D is of about 10%, there is more than a factor 2 of difference in the estimated barrier height. The effect is similar, but even stronger, for the potential considered in the Appendix: for an asymmetry parameter $a = -0.7$ $\beta\Delta U = 4$, $D = 1$ the fit with parabolic potential gives $\beta\Delta U = 0.3$ and $D = 1.2$ (see Table 2): again a huge discrepancy in the height estimate and a more limited variability in the diffusion constant. Our work suggests that, while $p_{TPT}(t)$ is somehow universal, to get a reliable estimate of $\beta\Delta U$ and D from a given empirical distribution one needs to know the shape of the underlying potential. This is particularly true when estimating $\beta\Delta U$ and can be understood as follows. At long times, $p_{TPT}(t)$ is determined by few low-lying eigenvalues of the associated Schrödinger equation (see inset of Fig. 4). These eigenvalues are most strongly affected by the shape of $V_{eff}(x)$. To understand this let us consider the average TPT for a parabolic potential and free boundaries

$$\langle t_{TP} \rangle = \tau_{diff} \frac{\log(\beta\Delta U)}{\beta\Delta U} \tag{34}$$

where $\tau_{diff} = 4x_0^2/D$ is the characteristic diffusion time for a free particle to cross the interval $[-x_0, x_0]$. This relation, valid for $\beta\Delta U \gg 1$ and originally derived by A. Szabo[1], can also be obtained as first moment of the distribution (31), see e.g. Ref. [14].

Assuming that the average TPT is fully determined by the lowest eigenvalue so that $\langle t_{\text{TP}} \rangle \sim \lambda_1^{-1}$, Eq. (34) then implies that for a parabolic potential $\lambda_1 \sim \beta \Delta U$. As we have shown here, the shape of the potential landscape $V(x)$ and as a consequence of $V_{\text{eff}}(x)$, has a rather strong influence on λ_1 (see Fig. 3 and inset of Fig. 5), it is therefore not surprising that fits of empirical TPT may lead to rather strong deviations of $\beta \Delta U$. In this work we found that the typical effect of asymmetry is to an understimation of barrier height (see also the case discussed in the Appendix), although the potential (1) shows a slight increase in λ_1 for weak asymmetry, which can be understood (see dashed lines in the inset of Fig. (5)).

The diffusion constant is instead much less influenced by the shape of the potential. Although the average TPT depends on D, see Eq. (34), its value in fitting empirical data is mostly determined by the short time behavior of the TPT distribution[8,14], which is in turn dependent on the spectrum of eigenvalues λ_n for large n. But these eigenvalues, as discussed before, are only very weakly affected by the shape of $V(x)$, and essentially fixed by the vanishing wavefunction at the absorbing boundaries, thus converge to the free particle in a box spectrum.

In conclusion our work shows that a reliable estimate of the barrier height from the analysis of $p_{\text{TPT}}(t)$ can only be obtained if one has some knowledge of the shape of the potential. Note that this is less problematic when estimating $\beta \Delta U$ from reaction rates. Kramers' formula for barrier crossing rates[24] shows that the shape of the potential only weakly affect the total rate compared to the exponential dependence on $\beta \Delta U$. Recent experiments report a $\beta \Delta U = 9$ from Kramers theory and $\beta \Delta U \approx 0.5$ from a fit of TPT distribution. This discrepancy is larger than those reported in Table 1 for the barrier (1), but note that the case discussed in the Appendix gives a much larger drop in $\beta \Delta U$. It is likely that other effects also play a role in giving rise to this large difference, see [12,18,22], but certainly barrier asymmetry should be taken into account in coming studies.

Conflicts of interest
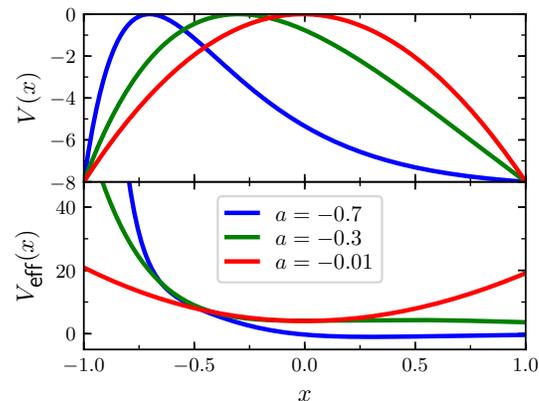
There are no conflicts to declare.

Acknowledgements

Appendix: Morse potential barrier

We consider here a Morse-like potential barrier defined by

$$V(x) = -\frac{\Delta U}{\Gamma^2} \left[e^{-\lambda(x-a)} - 1 \right]^2 \qquad (35)$$

Fig. 6 Top: Plot of Morse barriers (35) for $\Delta U = 8$, $\beta = 0.5$, $x_0 = 1$ and three values of a. Bottom: Corresponding V_{eff} obtained from the potentials of the top panel.

with $-x_0 \leq x \leq x_0$. The potential vanishes at its maximum $V(a) = 0$ and we impose that $V(\pm x_0) = -\Delta U$, which give

$$e^{-\lambda(\pm x_0 - a)} = 1 \pm \Gamma. \qquad (36)$$

These two equations constrain the parameters λ, a and Γ, so that only one of the three can be freely tuned. For convenience we choose here a, the value of x at the barrier top, as a tunable parameter and vary it in the interval $-x_0 < a \leq 0$. The limiting case $a, \Gamma, \lambda \to 0$, so that $\Gamma/\lambda \to 1$ corresponds to a symmetric parabolic potential centered in $x = 0$. The asymmetry increases as a approaches the lower bound $-x_0$, as shown in the top panel of Fig. 6. The corresponding effective potential V_{eff}, as obtained from Eq. (7), is shown in the bottom panel of Fig. 6. As (35) is analytic, $V_{\text{eff}}(x)$ is a smooth function of x, see Fig. 6(bottom). As in the case of linear-parabolic barrier, the increase in asymmetry leads to an increase of V_{eff} in the region close to the potential maximum $x = a$, due to an increase of the barrier curvature. In the region $a < x < x_0$, if the asymmetry is sufficiently high, the potential has an inflection point where the curvature is zero $V''(x_{\text{infl.}}) = 0$. The presence of a positive curvature region $V'' > 0$, is a feature of the Morse-like barrier, which leads to a strong decrease of V_{eff}, compared to the barrier (1).

To analyze the potential (35) we discretized the Schrödinger equation (6) and diagonalized the corresponding "Hamiltonian" matrix to find low lying eigenvalues and eigenvectors. Having found for the parabolic-linear potential that only the first few eigenstates significantly contribute to the TPT distribution, we limited the calculation to ~ 20 low lying states of the spectrum. Figure 7 shows the TPT distributions obtained from Eq. (28) for three values of the asymmetry parameter a. The inset shows the first four eigenvalues calculated numerically as a function of the asymmetry parameter a. Differently from the case discussed in Fig. 4, we do not have analytical arguments to predict the behavior of the ground state here. The numerics shows that λ_1 is monotonic in a, and drops strongly as a is decreased.

Table 2 shows the results of fitting the generated TPT distributions with those from parabolic potentials. As found in the main text, a sufficiently strong asymmetry leads to a severe underesti-
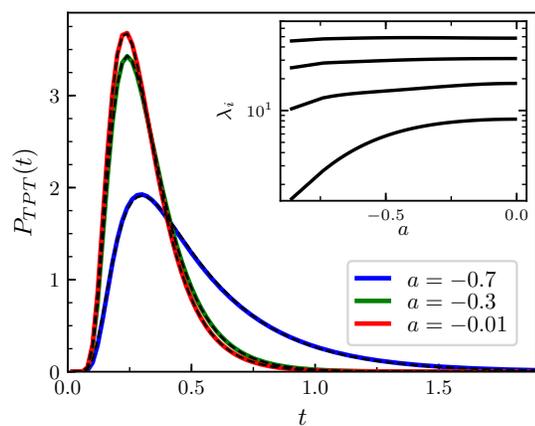
Fig. 7 Solid lines: Plot of TPT distributions for the three cases of Fig. 6 and $D = 1$ as obtained from numerical discretization of the Schrödinger equation (6). Dashed lines: Fits with distributions for a parabolic potential and absorbing boundaries. Inset: Plot of the first four eigenvalues as function of the parameter a.

Table 2 Summary of the analysis of TPT distributions generated with a Morse Potential for $x_0 = 1$, $D = 1.0$ and different values of the asymmetry parameter a. The second column gives the average TPT obtained from the first moment of the distribution. The last six columns report the parameters calculated from a fitting to a $p_{\text{TPT}}(t)$ generated by symmetric parabolic potential. The three columns "ABC" refer to a fit with the absorbing boundary condition case, which is Eq. (28) with $b = 1$. The quality of the fit is estimated by χ, as defined by Eq. (33). Uncertainties on $\langle t_{\text{TPT}} \rangle$, and fitted $\beta \Delta U$, D are approximately ± 1 in the last digit.

$\beta \Delta U = 4, D = 1.0$		fit to ABC			fit to FBC		
a	$\langle t_{TPT} \rangle$	$\beta \Delta U$	D	χ	$\beta \Delta U$	D	χ
-0.01	0.314	4.0	1.00	0.000	3.6	1.23	0.001
-0.3	0.327	3.5	1.02	0.000	3.2	1.28	0.001
-0.7	0.528	0.3	1.18	0.001	2.4	1.59	0.002

mate of the barrier height. For the Morse-like potential the effect is even stronger that (1), due to the positive curvature of (35). The variability in the fitted D as a function of the asymmetry is of about 20%.

Notes and references

1 H. S. Chung, K. McHale, J. M. Louis and W. A. Eaton, *Science*, 2012, **335**, 981.

2 K. Neupane, D. B. Ritchie, H. Yu, D. A. N. Foster, F. Wang and M. T. Woodside, *Phys. Rev. Lett.*, 2012, **109**, 068102.

3 H. S. Chung, J. M. Louis and W. A. Eaton, *Proc. Natl. Acad. Sci. USA*, 2009, **106**, 11837–11844.

4 K. Truex, H. S. Chung, J. M. Louis and W. A. Eaton, *Phys. Rev. Lett.*, 2015, **115**, 018101.

5 K. Neupane, D. A. Foster, D. R. Dee, H. Yu, F. Wang and M. T. Woodside, *Science*, 2016, **352**, 239–242.

6 N. Q. Hoffer, K. Neupane, A. G. Pyo and M. T. Woodside, *Proc. Natl. Acad. Sci. USA*, 2019, **116**, 8125–8130.

7 J. Gladrow, M. Ribezzi-Crivellari, F. Ritort and U. Keyser, *Nature Comm.*, 2019, **10**, 55.

8 B. W. Zhang, D. Jasnow and D. M. Zuckerman, *J. Chem. Phys.*, 2007, **126**, 074504.

9 S. Chaudhury and D. E. Makarov, *J. Chem. Phys.*, 2010, **133**, 034118.

10 H. Orland, *J. Chem. Phys.*, 2011, **134**, 174114.

11 W. K. Kim and R. R. Netz, *J. Chem. Phys.*, 2015, **143**, 224108.

12 E. Pollak, *Phys. Chem. Chem. Phys.*, 2016, **18**, 28872.

13 A. M. Berezhkovskii, L. Dagdug and S. M. Bezrukov, *J. Phys. Chem. B*, 2017, **121**, 5455.

14 M. Laleman, E. Carlon and H. Orland, *J. Chem. Phys*, 2017, **147**, 214103.

15 M. Caraglio, S. Put, E. Carlon and C. Vanderzande, *Phys. Chem. Chem. Phys.*, 2018, **20**, 25676.

16 A. M. Berezhkovskii, L. Dagdug and S. M. Bezrukov, *J. Phys. Chem. B*, 2019, **123**, 3786–3796.

17 N. Q. Hoffer and M. T. Woodside, *Curr. Opin. Chem. Biol.*, 2019, **53**, 68–74.

18 R. Satija, A. Das and D. E. Makarov, *J. Chem. Phys.*, 2017, **147**, 152707.

19 E. Carlon, H. Orland, T. Sakaue and C. Vanderzande, *J. Phys. Chem. B*, 2018, **122**, 11186.

20 H. Risken, *The Fokker-Planck Equation*, Springer, 1996.

21 K. Neupane, N. Q. Hoffer and M. T. Woodside, *Phys. Rev. Lett.*, 2018, **121**, 018102.

22 D. E. Makarov, *J. Chem. Phys.*, 2017, **146**, 071101.

23 E. Medina, R. Satija and D. E. Makarov, *J. Phys. Chem. B*, 2018, **122**, 11400–11413.

24 P. Hänggi, P. Talkner and M. Borkovec, *Rev. Mod. Phys.*, 1990, **62**, 251.