

# Reliable extraction of energy landscape properties from critical force distributions

Sudeep Adhikari<sup>\*</sup> and K. S. D. Beach<sup>†</sup>

Department of Physics and Astronomy, The University of Mississippi, University, Mississippi 38677, USA



(Received 10 January 2020; accepted 31 March 2020; published 3 June 2020)

The structural dynamics of a biopolymer is governed by a process of diffusion through its conformational energy landscape. In pulling experiments using optical tweezers, features of the energy landscape can be extracted from the probability distribution of the critical force at which the polymer unfolds. The analysis is often based on rate equations having Bell-Evans form, although it is understood that this modeling is inadequate and leads to unreliable landscape parameters in many common situations. Dudko *et al.* [Phys. Rev. Lett. **96**, 108101 (2006)] have emphasized this critique and proposed an alternative form that includes an additional shape parameter (and that reduces to Bell-Evans as a special case). Their fitting function, however, is pathological in the tail end of the pulling force distribution, which presents problems of its own. We propose a modified closed-form expression for the distribution of critical forces that correctly incorporates the next-order correction in pulling force and is everywhere well behaved. Our claim is that this new expression provides superior parameter extraction and is valid even up to intermediate pulling rates. We present results based on simulated data that confirm its utility.

DOI: [10.1103/PhysRevResearch.2.023276](https://doi.org/10.1103/PhysRevResearch.2.023276)

## I. INTRODUCTION

The contribution of explicitly quantum processes notwithstanding [1], classical energy landscape theory [2–5] provides a useful framework for describing the evolution of biopolymers between various folded and unfolded configurations through a process of thermally driven escape from local confining potentials [6]. Developing tools of analysis within this framework has become ever more pressing, given the profound developments in single-molecule biophysics [7–22]. One of the key practical problems is how to infer the energy landscape, or at least a projection of it onto an appropriate reaction coordinate, from experimentally measured quantities [23–32]. As is typical of inverse problems, recovery of the landscape from measured data is ill conditioned: it is highly sensitive to experimental uncertainties and to any assumptions that go into the forward model.

In pulling experiments using optical tweezers [33], the determination of landscape features has historically been carried out based on Bell-Evans phenomenological theory [34–38], which assumes that the rate constant  $k(F)$  scales up exponentially with applied force from its unperturbed, intrinsic value  $k_0$  according to the Arrhenius law,

$$k_{\text{BE}}(F) = k_0 e^{\beta F x^\ddagger}. \quad (1)$$

Here,  $\beta^{-1} = k_B T$  is the thermal energy scale set by the aqueous environment;  $x^\ddagger$  is the minimum-to-barrier distance of the

effective one-dimensional potential  $U_0(x)$ , a continuous (but not necessarily smooth) function of the end-to-end extension.

A common experimental situation involves the application of a pulling force  $F = KVt$  that grows linearly in time until the rupture force  $F_c$  is reached. Although other pulling protocols are sometimes employed [39–43], we focus on the case of constant pulling speed, and we ignore instrument-specific issues of compliance [44–46].

It is recognized that a description of pulling experiments based on the Bell-Evans formula for the force-induced rupture rate is in poor accord with results from numerical simulations [47]. The naive thermal-activation picture, represented by Bell-Evans theory, suffers from various inadequacies that are important to address. To begin, Eq. (1) is strictly applicable only in the limit of low pulling rate ( $KV \lesssim KV_{\text{min}} = k_0/\beta x^\ddagger$ ) and ultrahigh barrier ( $\Delta G^\ddagger \gg Fx^\ddagger, k_B T$ ). Even in the moderate pulling regime, it incorrectly predicts the rupture force distribution. It also ignores self-consistency effects in the sense that it does not account for the fact that the distance  $x^\ddagger$  and the energy barrier  $\Delta G^\ddagger$  are themselves force dependent and both diminish with increasing  $F$  as the energy landscape is tilted. Nor does it properly account for the shape of the barrier, which plays a vital role in establishing the escape rate and the nature of the escape trajectory for more modest, biologically relevant barrier heights.

Consequently, there are many situations in which the phenomenological theory incorrectly predicts the results of pulling experiments. It tends to overestimate the rate of rupture,  $k(F)$ , at a given force  $F$  and to underestimate the mean and most-probable rupture forces. Hence, when the Bell-Evans rate,  $k_{\text{BE}}(F)$ , is used as the basis for a fit to experimental data, the extracted parameters,  $\Delta G^\ddagger$ ,  $x^\ddagger$ , and  $k_0$ , may be incorrectly predicted. Our main concern here lies in the reliable extraction of these physical quantities.

<sup>\*</sup>sadhika6@go.olemiss.edu

<sup>†</sup>kbeach@olemiss.edu

Attempts have been made to improve on Bell-Evans theory by introducing additional fitting parameters [48,49], sometimes in an *ad hoc* way. Dudko and co-workers have tried to make the analysis more rigorous [50]. They calculated  $k(F)$  and the corresponding probability density of the rupture force  $p(F_c)$  within the framework of Kramers theory [51] for two specific free energy surfaces—the cusp surface and the linear cubic surface—and showed that these two examples can be subsumed into a single result (appearing as Eq. (3) in Ref. [50]),

$$k_D(F) = k_0 \left(1 - \frac{\nu F x^\ddagger}{\Delta G^\ddagger}\right)^{1/\nu-1} e^{\beta \Delta G^\ddagger [1 - (1 - \nu F x^\ddagger / \Delta G^\ddagger)^{1/\nu}]}, \quad (2)$$

with interpolation provided by a shape parameter  $\nu$ . This encompasses the Bell-Evans result, since  $k_D(F) \rightarrow k_{BE}(F)$  as  $\nu \rightarrow 1$ . It is clear, however, that for all  $\nu \neq 1$  Eq. (2) has a dangerous point of nonanalyticity. The vanishing of the rate  $k_D(F) \rightarrow 0$  as  $F \rightarrow \Delta G^\ddagger / x^\ddagger \nu$  (for shape parameters in the range  $0 < \nu < 1$ ) is manifestly unphysical; hence the Dudko expression is only appropriate for the pulling regime in which  $F \ll \Delta G^\ddagger / x^\ddagger \nu$ . In fact, the region of validity is more constrained still, since we should further require that the escape rate grow with pulling force. As it turns out, the function  $k_D(F)$  is monotonic increasing only for

$$F < \frac{\Delta G^\ddagger}{x^\ddagger \nu} \left[1 - \left(\frac{1 - \nu}{\beta \Delta G^\ddagger}\right)^\nu\right]. \quad (3)$$

We pursue a different approach that produces no non-analyticity and no obviously unphysical behavior. We compute  $\log k(F)/k_0$  order by order in the pulling force. Rather than truncate the expansion, we approximate the higher-order terms as a resummation by geometric series—similar in spirit to the random phase approximation or the infinite summation of ladder diagrams in many-body theory:

$$k(F) = k_0 \exp\left(\frac{\beta F x^\ddagger}{1 + F/2\kappa^\ddagger x^\ddagger}\right). \quad (4)$$

Here  $\kappa^\ddagger$  is the reduced curvature of the well and barrier. The route to Eq. (4) is nothing more than a mathematical trick, but it rather elegantly cures the ill behavior of a truncated expansion, and it fortuitously leads to a closed-form expression for the cumulative probability distribution.

Our attempts to benchmark Eq. (4) fall into two categories, prediction and parameter extraction, which correspond to the forward and inverse problems. In the forward direction, we determine the escape rates and the cumulative probability distribution of the critical force following the numerical method described in Sec. III. We compare the simulated behavior to the various analytical predictions. We find that our proposal outperforms the Bell-Evans and Dudko expressions, across many different choices of energy landscape and over a broad range of pulling rates. In the inverse direction, analytical forms for the cumulative probability distribution  $P(F_c)$  are fit to the simulated data to extract the optimal values of the intrinsic parameters  $k_0$ ,  $x^\ddagger$ , and  $\kappa^\ddagger$ .

The results we achieve are compelling. The values of the three parameters that we extract are in excellent agreement with the actual values that characterize the underlying energy landscape. Moreover, the agreement appears to hold over an

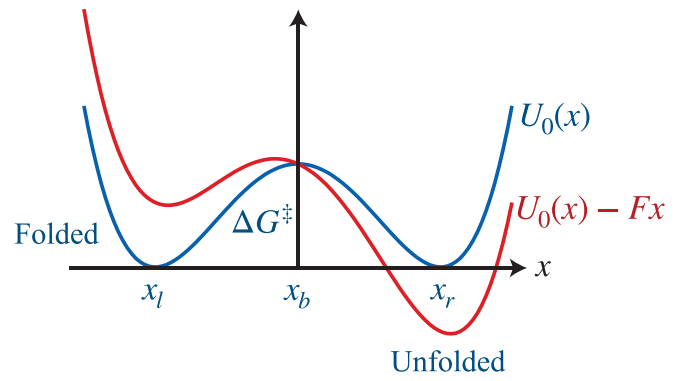


FIG. 1. Blue curve: The double-well potential has equilibrium positions at  $x_l$  and  $x_r$ , separated by a barrier at  $x_b$ . A particle escaping from left to right experiences a barrier of height  $\Delta G^\ddagger = U_0(x_b) - U_0(x_l)$ , peaked at a distance  $x^\ddagger = x_b - x_l$  from the bottom of the left well. Red curve: After application of a pulling force  $F$ , the energy landscape has tilted to favor the destination well on the right. Observe that the well positions have shifted and that the height of the barrier holding the particle in the left well has decreased.

unexpectedly large range of pulling rates, with  $KV/KV_{\min}$  spanning six or seven orders of magnitude.

In contrast, fits of simulation data to the Bell-Evans cumulative probability distribution, insofar as they are able to produce good values of  $k_0$  and  $x^\ddagger$  at all, only do so at the very slowest pulling rates. It is difficult to speak definitively of how well Dudko's expression performs, since in that context fits must be carried out in conjunction with a force cutoff somewhere below the point of nonanalyticity. This is an unwelcome complication. The cutoff itself introduces a significant element of uncertainty in the fit, since where best to put the cutoff cannot be determined if the landscape is not yet known.

## II. FORMAL DEVELOPMENT

Kramers theory tells us that the escape rate depends weakly (polynomially) on the curvature at the bottom of the well and the top of the barrier but strongly (exponentially) on the height of the apparent energy barrier in the direction of travel [51,52]. We consider a double well potential  $U_0(x)$ , with wells at positions  $x_l$  and  $x_r$  separated by a barrier at  $x_b$  ( $x_l < x_b < x_r$ ), as illustrated in Fig. 1. The well escape rate from left to right is given by  $k_0 \sim \exp(-\beta \Delta U_0)$ , where  $\Delta U_0 = U_0(x_b) - U_0(x_l)$ .

We allow for a pulling force  $F$  that tilts the potential landscape according to

$$U(x) = U_0(x) - Fx. \quad (5)$$

The corresponding rate equation becomes

$$k(F) \sim \exp[-\beta(U(x_b + \delta x_b) - U(x_l + \delta x_l))], \quad (6)$$

where  $\delta x_l$  and  $\delta x_r$  denote the shifts in the well positions as a result of the tilt. Taylor expansions of the extremal conditions  $U'(x_l + \delta x_l) = 0$  and  $U'(x_b + \delta x_b) = 0$  around  $x_l$  and  $x_b$  up to first order in  $\delta x_l$  and  $\delta x_b$  give  $\delta x_l = F/U_0''(x_l) = F/\kappa_l$  and  $\delta x_b = F/U_0''(x_b) = -F/\kappa_b$ . A further expansion of  $U(x_b + \delta x_b)$  and  $U(x_l + \delta x_l)$  around  $x_b$  and  $x_l$ , respectively, up to

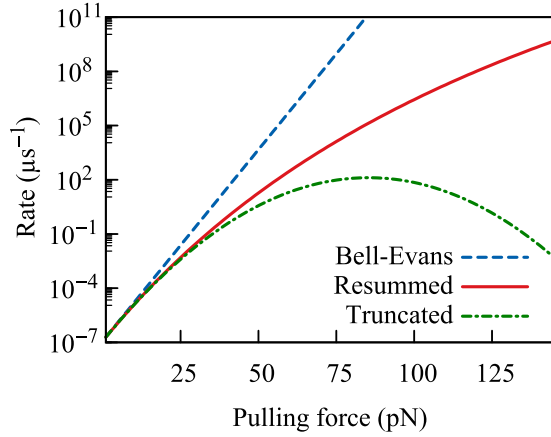


FIG. 2. The well escape rate  $k(F)$  is plotted against the applied pulling force  $F$ . The upper (dashed blue) curve corresponds to the Bell-Evans (BE) rate [Eq. (1)], the middle (solid red) to our proposed infinite-resummation expression [Eq. (4)], and the lower (dot-dashed green) to an expansion truncated at second order in the pulling force [Eq. (7)]. The BE result grows exponentially without bound (but shows as a straight line because of the log-linear scale). The truncated expression turns over and becomes unphysical around 80 pN. The resummed form strikes a middle course, growing monotonically but saturating at a large, finite value,  $k_0 \exp[2\beta\kappa^\ddagger(x^\ddagger)^2]$ .

second order in  $F$ , yields a rate equation of the form

$$k(F) = k_0 \exp\left[\beta F x^\ddagger \left(1 - \frac{F}{2\kappa^\ddagger x^\ddagger}\right)\right]. \quad (7)$$

Here,  $x^\ddagger = x_b - x_l$ , and

$$\frac{1}{\kappa^\ddagger} = \frac{1}{U_0''(x_l)} - \frac{1}{U_0''(x_b)} = \frac{1}{\kappa_l} + \frac{1}{\kappa_b}. \quad (8)$$

Successive terms in the expansion of  $\log k(F)/k_0$  have alternating sign, which is important for proper convergence of the series. Indeed, there is no polynomial expression, arising as a truncation of the series at finite order, that does not either substantially over- or undershoot the true rate for large applied  $F$ . The negative-prefactor terms at even powers of  $F$  are particularly troublesome, because they lead to non-monotonicity. As a workaround, we make use of the idea of infinite resummation,  $1 - \epsilon + \epsilon^2 - \dots \approx 1/(1 + \epsilon)$ , which transforms Eq. (7) into Eq. (4), at least up to discrepancies at  $O(F^3)$ . The transformed expression is well behaved everywhere and displays no obviously unphysical behavior (see Fig. 2). Moreover, it leads to a closed-form expression for the cumulative probability distribution (with the correct normalization  $P(F_c) \rightarrow 1$  as  $F_c \rightarrow \infty$ ; Dudko's expression, in contrast, cannot be properly normalized).

In the usual adiabatic limit, the expression for the cumulative probability distribution of the rupture force is given by

$$P(F_c) = 1 - \exp\left[-\int_0^{F_c} \frac{dF}{\dot{F}} k(F)\right]. \quad (9)$$

Equations (1) and (9) together give the cumulative probability distribution of the rupture force as predicted by the Bell-Evans

phenomenological model,

$$P_{\text{BE}}(F_c) = 1 - \exp\left[\frac{k_0}{KV\beta x^\ddagger} (1 - e^{\beta F_c x^\ddagger})\right]. \quad (10)$$

If instead we put Eq. (4) into Eq. (9), we get a more complicated result, but one that is still simple enough to use for fitting (e.g., via the Marquardt-Levenberg method):

$$P(F_c) = 1 - \exp\left[-\frac{k_0}{KV} (F_1 + F_2 - 2x^\ddagger \kappa^\ddagger)\right]. \quad (11)$$

The quantities  $F_1$  and  $F_2$  have units of force and are explicit functions of the critical value  $F_c$ :

$$\begin{aligned} F_1 &= (F_c + 2x^\ddagger \kappa^\ddagger) \exp\left(\frac{2F_c x^\ddagger{}^2 \beta \kappa^\ddagger}{F_c + 2x^\ddagger \kappa^\ddagger}\right), \\ F_2 &= 4x^\ddagger{}^3 \beta \kappa^\ddagger{}^2 \exp(2x^\ddagger{}^2 \beta \kappa^\ddagger) \\ &\times \left[ \text{Ei}\left(-\frac{4x^\ddagger{}^3 \beta \kappa^\ddagger{}^2}{F_c + 2x^\ddagger \kappa^\ddagger}\right) - \text{Ei}(-2x^\ddagger{}^2 \beta \kappa^\ddagger) \right]. \end{aligned} \quad (12)$$

The exponential integral  $\text{Ei}(x) = -\int_{-x}^{\infty} dt t^{-1} e^{-t}$  is a standard special function that is available in most data analysis software.

The choice  $F = KVt$  is helpful here but not essential. Its main advantage is that the differential appearing in Eq. (9) simplifies to  $dF/\dot{F} = (KV)^{-1} dF$ , and hence the integration measure is trivial. The closed-form expression that we obtain in Eqs. (11) and (12) does depend on this choice. But any pulling schedule  $F(t)$  that is monotonic increasing (so that  $\dot{F}$  never vanishes or goes negative) and growing at most as a polynomial in  $t$  can be treated similarly.

We now comment on the connection to the prior work of Dudko and co-workers. The unperturbed potential  $U_0(x)$  can be expanded to quadratic order around the bottom of the well,  $U_{0,l}(x) = U_0(x_l) + (\kappa_l/2)(x - x_l)^2$ , and around the peak of the barrier,  $U_{0,b}(x) = U_0(x_b) - (\kappa_b/2)(x - x_b)^2$ . We identify the position  $x^* = (\kappa_b x_b + \kappa_l x_l)/(\kappa_l + \kappa_b)$  where the two approximations take a common slope and match the functions smoothly there. The resulting piecewise composite curve has a total rise of

$$\begin{aligned} U_{0,l}(x^*) - U_0(x_l) + U_{0,b}(x^*) - U_0(x_b) \\ = \frac{\kappa_l \kappa_b (x_b - x_l)^2}{2(\kappa_l + \kappa_b)} = \frac{1}{2} \kappa^\ddagger (x^\ddagger)^2, \end{aligned} \quad (13)$$

which differs from the true barrier height  $\Delta G^\ddagger = U_0(x_b) - U_0(x_l)$  by a factor that Dudko labels  $1/\nu$ . That is,

$$\frac{\Delta G^\ddagger}{\nu} = \frac{1}{2} \kappa^\ddagger (x^\ddagger)^2. \quad (14)$$

The equality  $\nu = 2/3$  holds for any degree-three polynomial. If the energy landscape is represented by a higher-degree polynomial, then the value of the shape parameter is idiosyncratic and should be viewed as drawn from a distribution with average  $\langle 1/\nu \rangle < 3/2$ . For smooth potentials (no cusps or discontinuities), typical values of the shape parameter  $\nu$  range between  $2/3$  and  $\approx 1.1$ . An advantage of working in terms of  $\nu$ , rather than the effective curvature  $\kappa^\ddagger$ , is that the former can be defined even if the derivatives  $U''(x_l)$  and  $U''(x_b)$  vanish

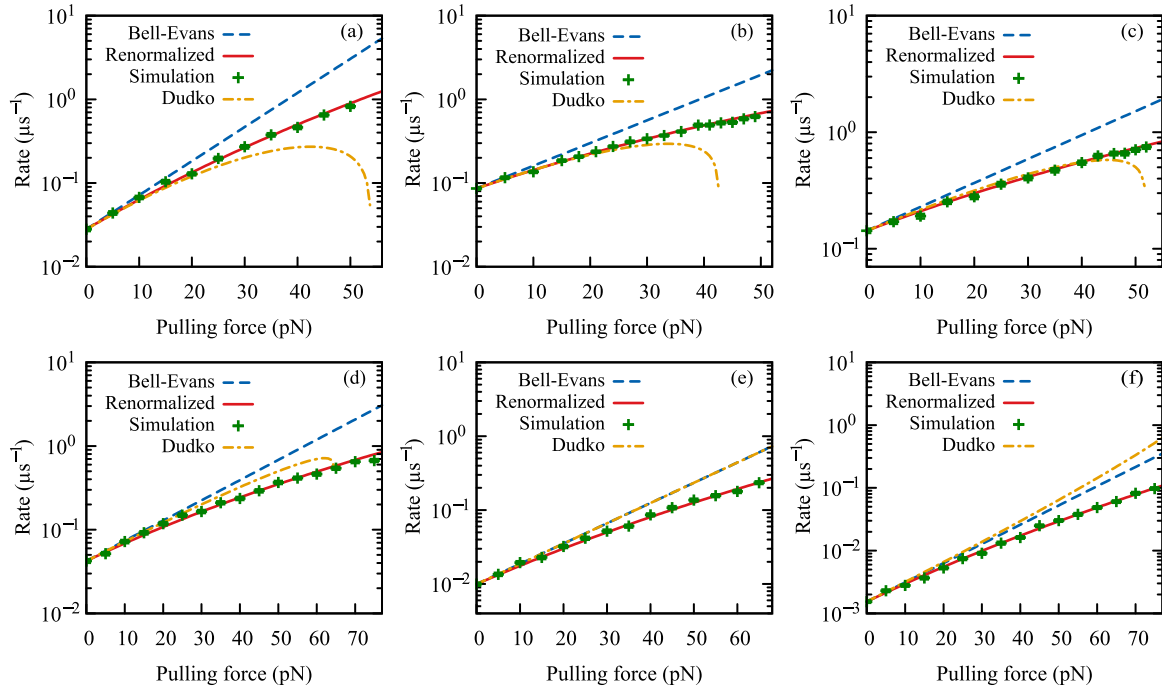


FIG. 3. Numerical measurements of the escape rate (green data points with error bars) are plotted versus the applied pulling force. Also shown for comparison are the predictions of the Bell-Evans approach [Eq. (1), dashed blue line], Dudko approach [Eq. (2), dot-dashed orange line], and our renormalized rate equation [Eq. (15) with a common value  $\alpha = 1/3$ , solid red line]. The simulations were carried out for potentials with various values of  $\nu = 2\Delta G^\ddagger/\kappa^\ddagger(x^\ddagger)^2$ , the shape parameter: (a)  $\nu = 0.66$ , (b)  $\nu = 0.75$ , (c)  $\nu = 0.82$ , (d)  $\nu = 0.9$ , (e)  $\nu = 1.0$ , and (f)  $\nu = 1.1$ . We note the remarkable agreement between simulation and the renormalized form. No fitting is involved.

(e.g., a quartic well or barrier) or are not well defined (e.g., a cusp barrier).

With Eq. (14) in mind, matching our resummed rate expression to that of Dudko order by order in the small-pulling-force, large-barrier-height limit suggests the form

$$k(F) = k_0 \exp \left[ \frac{\beta F x^\ddagger}{(1 + \alpha/\beta \Delta G^\ddagger)(1 + \nu F x^\ddagger/4\Delta G^\ddagger)} \right], \quad (15)$$

where  $\alpha > 0$  is a pure number with a weak dependence on the shape parameter. Equation (15) can be understood as a rewriting of Eq. (1), the Bell-Evans phenomenological rate, with an upward renormalization of the temperature,  $\beta \rightarrow \beta/(1 + \alpha/\beta \Delta G^\ddagger)$ , and a downward renormalization of the barrier distance  $x^\ddagger \rightarrow x^\ddagger/(1 + \nu F x^\ddagger/4\Delta G^\ddagger)$ . Unlike Eq. (2), Eq. (15) is well behaved everywhere.

In the case of an ultrahigh barrier, defined by the double limit  $\beta \Delta G^\ddagger \gg 1$  and  $\Delta G^\ddagger \gg F x^\ddagger$ , Eq. (15) reduces to Eq. (1). For more modest barriers or higher temperatures, one or both of the factors  $(1 + \alpha/\beta \Delta G^\ddagger)$  and  $(1 + \nu F x^\ddagger/4\Delta G^\ddagger)$  may differ appreciably from 1; this allows the rate expression to become aware of the details of the barrier's height and shape through the factor  $\Delta G^\ddagger/\nu$ .

The reliability of Eq. (15) was tested for six potential landscapes with different values of  $\nu$  using the simulation scheme described in the next section. In every test example (see Fig. 3), our renormalized equation closely tracked the empirical escape rate determined from simulations. It noticeably outperformed the Bell-Evans and Dudko escape rate equations.

### III. NUMERICAL SIMULATIONS

The reaction coordinate  $x$  was made to execute Langevin dynamics according to

$$m\ddot{x} = m\dot{v} = -\frac{\partial U}{\partial x} - \gamma v + \xi(t). \quad (16)$$

This was implemented using a modern reformulation [53] of the Verlet algorithm [54]. We mimicked the experimental situation by assuming stochastic motion of a molecule of effective mass  $m = 2$  pg in a biquadratic potential. The data that appear in Figs. 4–7 correspond to the choice  $U_0(x) = 4x^4 - 32x^2 + 64$  (with  $x$  measured in nm and  $U_0$  in pN nm). The molecule was assumed to be pulled from two ends along the reaction coordinate  $x$  by a laser potential with force constant  $K$  and pulling velocity  $V$ , i.e., with an instantaneous force  $F = K V t$  that increases linearly in time. For the given potential, the energy barrier was  $\Delta G^\ddagger = 64$  pN nm, the minimum-to-barrier distance  $x^\ddagger = 2$  nm, and the effective curvature  $\kappa^\ddagger = 42.7$  pN nm<sup>-1</sup>. The stochastic forces  $\xi(t)$  on the molecule were drawn randomly from a Gaussian distribution of width  $(2m\gamma k_B T \delta t)^{1/2}$  with  $k_B T = 4.1$  pN nm,  $\gamma = 7$  μs<sup>-1</sup>, and a discrete time step  $\delta t$  ranging from 10<sup>-2</sup> to 10<sup>-6</sup> μs.

Note that, for generality, small inertial effects were included in the numerics. The simulations were not run in the strongly overdamped, diffusion-only limit: parameter values were chosen to be physically plausible but also to produce a nonextreme limit (neither  $\gamma \ll \omega_b$  nor  $\gamma \gg \omega_b$ ) of the prefactor to the exponential in the Kramers rate [which will appear in Eq. (17)].



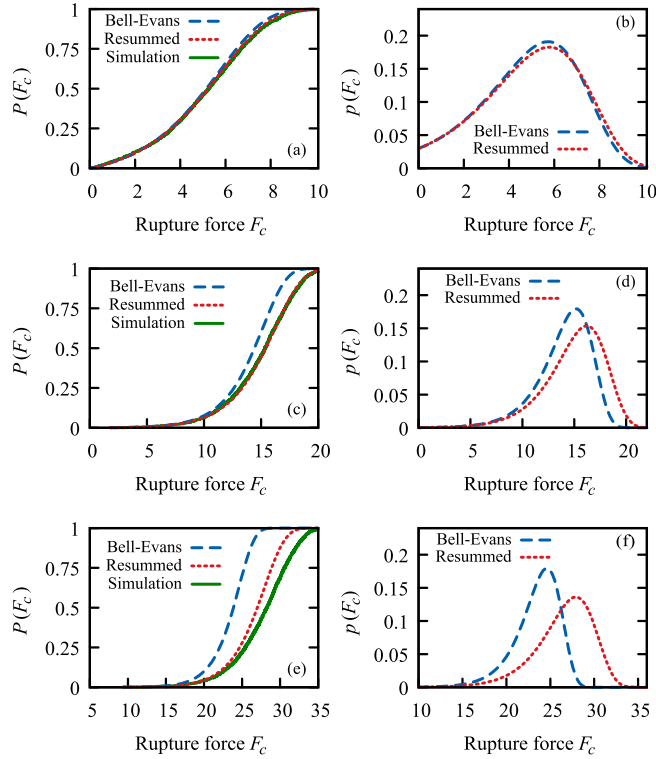


FIG. 4. Plots paired on the left and right show the cumulative probability distribution  $P(F_c)$  and the corresponding probability density  $p(F_c) = P'(F_c)$  of the rupture force. Each row shows results for successively faster pulling rates: (a, b)  $KV = 4 \times 10^{-6} \text{ pN } \mu\text{s}^{-1}$ , (c, d)  $KV = 4 \times 10^{-4} \text{ pN } \mu\text{s}^{-1}$ , and (e, f)  $KV = 4 \times 10^{-2} \text{ pN } \mu\text{s}^{-1}$ .

Pulling rates for the force  $F = KVt$  are measured with respect to  $KV_{\min} = k_0/\beta x^\ddagger$ , which is the minimum rate for effectual pulling. Below  $KV_{\min}$ , the probability density  $p(F_c)$  is peaked at  $F_c = 0$ ; the particle escapes the well on its own before the applied force has appreciably modified the energy landscape. On the other hand, for rates above  $KV/KV_{\min} \approx 10^6$ , the barrier vanishes too quickly, long before the particle has moved any significant distance. Accordingly, we worked in the regime of pulling rates between these two extremes.

The simulation was initialized in the left well by drawing starting values of velocity  $v$  and position  $x$  from the distributions  $e^{-\beta mv^2/2}$  and  $e^{-\beta U(x)}\Theta(x_b - x)$ , respectively, so that the each simulation began fully thermalized. The simulation flagged the value of pulling force at which  $x$  convincingly crossed the barrier or the barrier vanished; we took this to be the rupture or critical force  $F_c$ . For each value of the pulling rate  $KV$ , the simulation was carried out 2500 times, each run generating a unique value of the rupture force. The cumulative probability distribution  $P(F_c)$  was constructed in the standard way—by sorting the measured rupture forces in ascending order and then pairing them with a uniform grid of values running from zero to 1. The plot for  $P(F_c)$  so obtained was tested against Eq. (11) and against the Bell-Evans form, Eq. (10). The process was repeated for pulling rates ranging from  $KV = 10^{-7}$  to  $0.6 \text{ pN } \mu\text{s}^{-1}$  (roughly  $1 \lesssim KV/KV_{\min} \lesssim 10^7$ ) to determine how these expressions fare in the slow, intermediate, and fast pulling regimes.

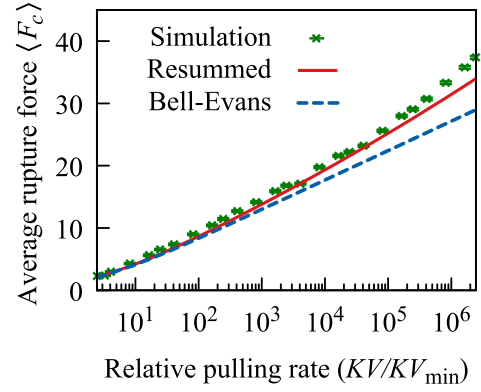


FIG. 5. The average rupture force  $\langle F_c \rangle = \int dF F P'(F)$ , determined by numerical integration with  $P(F)$  taken from Eqs. (10) (dashed blue line) and (11) (red line), is compared to the empirical values from simulation (green crosses).

In order to test parameter extraction, the original  $P(F_c)$  data set for each pulling rate was bootstrapped [55] 100 times to generate 100 new instantiations. These data were fitted with Eq. (11) to extract the intrinsic parameters of the potential landscape:  $k_0$ ,  $x^\ddagger$ , and  $\kappa^\ddagger$ . The spread in fit values was used to generate error estimates.

The data sets were also fitted to the Bell-Evans form given by Eq. (10) in order to extract the values of  $k_0$  and  $x^\ddagger$  ( $\kappa^\ddagger$  does not appear in the Bell-Evans expression). The bootstrap-average values of the extracted parameters were compared to their known values. The theoretical intrinsic rate  $k_0$  was computed according to the usual Kramers result,

$$k_0 = \frac{\omega_l}{2\pi} \frac{\sqrt{\gamma^2/4 + \omega_b^2} - \gamma/2}{\omega_b} \exp(-\beta \Delta G^\ddagger). \quad (17)$$

Our test potential corresponded to  $\omega_l = \sqrt{\kappa_l/m} = 8 \mu\text{s}^{-1}$  and  $\omega_b = \sqrt{\kappa_b/m} = 5.65 \mu\text{s}^{-1}$ . We verified that the theoretical value of  $k_0 = 1.192 \times 10^{-7} \mu\text{s}^{-1}$  was in agreement with numerical measurements of the escape rate for the nontilted energy landscape.

#### IV. RESULTS AND CONCLUSIONS

In the left column of Fig. 4, the rupture force distributions predicted by Eqs. (10) and (11) are compared to the results from simulation for three different pulling rates (corresponding to  $KV/KV_{\min} \approx 10^1, 10^3$ , and  $10^5$ ). For slow pulling (top row), the Bell-Evans theory and our resummed expression are well matched to each other and to the numerics. For intermediate pulling (middle row), the Bell-Evans result begins to deviate significantly, whereas our proposal continues to give accurate results (i.e., the solid green and dotted red lines coincide). Only at the highest pulling rates (bottom row) do we find significant deviation from the simulated rupture force distributions for both Eqs. (10) and (11); although, even there, our expression performs better and is in good agreement up to  $\sim 25 \text{ pN}$ .

It is instructive to look at the corresponding probability density of the rupture force,  $p(F_c) = P'(F_c)$ , obtained from Bell-Evans and our resummed form, as shown in the right

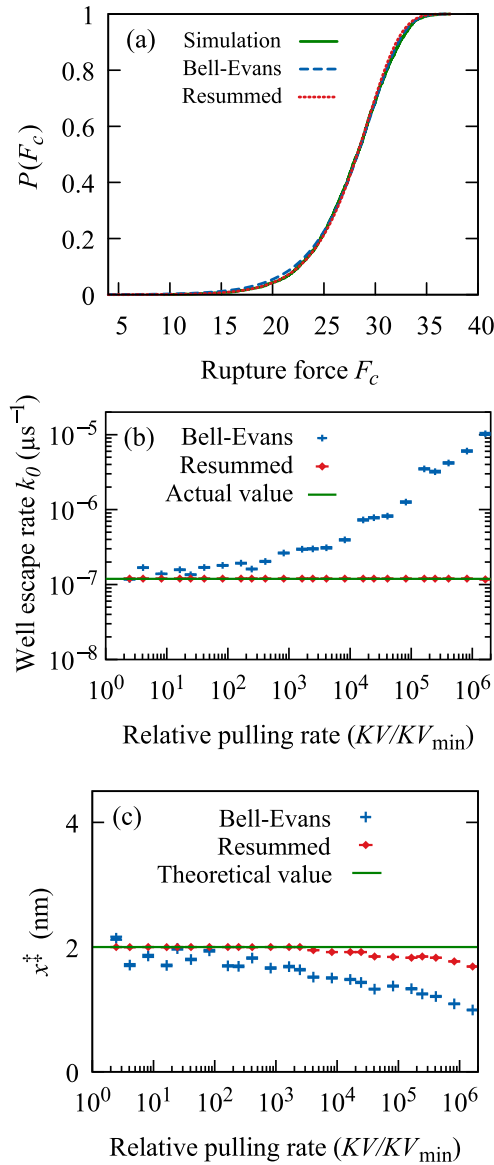


FIG. 6. (a) The cumulative probability distribution of the rupture force, computed from 2500 simulated pulling experiments at rate  $KV = 0.04 \text{ pN}/\mu s$  (green line), is plotted alongside the best fits for Eqs. (10) (dashed blue line) and (11) (dotted red line). The near indistinguishability of the curves illustrates the strong tendency toward overfitting. The Bell-Evans expression, though ill suited for describing the behavior at this high pulling rate, is able to mimic the numerical data—but at the cost of producing fitting parameters that have drifted far from their true values. This is in contrast to the poor agreement in Fig. 4(e), where there is no fitting and the known values of  $k_0$  and  $x^\ddagger$  are used. Estimates of (b) the intrinsic escape rate  $k_0$  and (c) the barrier distance  $x^\ddagger$ , as determined from fits of Eqs. (10) (blue crosses) and (11) (red diamonds) to simulation data over a range of pulling rates, are plotted alongside the actual value (green line).

column of Fig. 4. The Bell-Evans result systematically underestimates the pulling force required to traverse the barrier—and increasingly so for faster pulling. One observes that both its peak (typical rupture force) and its overall weight (mean rupture force) are positioned too far to the left (toward low force values). The same information is contained in

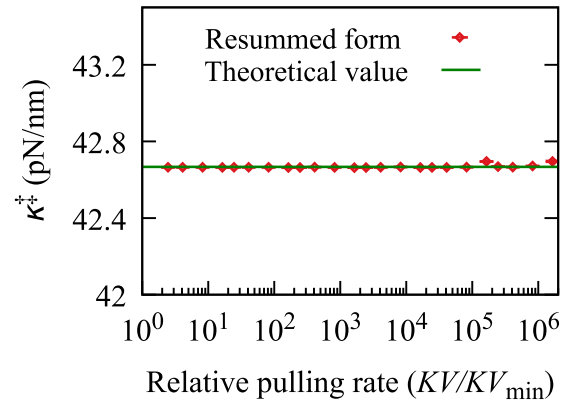


FIG. 7. The plotted points (red diamonds) are estimates of the reduced curvature  $\kappa^\ddagger$ , as determined from fits of Eq. (11) to simulation data. They compare favorably to the actual value (green line) over a range of pulling rates spanning many decades.

the average critical force  $\langle F_c \rangle$ , which we obtained from the cumulative probability distributions, Eqs. (10) and (11), by numerical integration. Figure 5 shows a plot of  $\langle F_c \rangle$  as a function of the relative pulling rate. One can readily identify an intermediate regime ( $10^2 \lesssim KV/KV_{min} \lesssim 10^5$ ) in which the curve computed from the resummed rate tracks the true, numerically determined values of the average rupture force. In that same regime, the Bell-Evans curve deviates significantly.

The second part of the numerical analysis focused on the inverse problem. Here, the simulated data were fitted using Eq. (11), and the intrinsic parameters of the energy landscape, viz.,  $k_0$ ,  $x^\ddagger$ , and  $\kappa^\ddagger$ , were determined by minimizing discrepancies between theory and data in the least-squares sense. The process was repeated for Eq. (10), but only with  $k_0$  and  $x^\ddagger$  (since  $\kappa^\ddagger$  does not appear in the fitting function). We found unambiguously that the parameter extraction is much more reliable using our resummed form. Indeed, use of the Bell-Evans theory was often quite misleading, because it would produce an apparently good fit that corresponded to incorrect values of the landscape parameters.

The top panel of Fig. 6, which shows a fast-pulling example with  $KV = 4 \times 10^{-2} \text{ pN}/\mu s$ , emphasizes this point. The cumulative probability distribution appears to be equally well fit by Eqs. (10) and (11). The middle and bottom panels reveal this to be illusory. In the Bell-Evans analysis, the value of  $k_0$  is systematically overestimated and  $x^\ddagger$  underestimated, and both ever more so as the pulling rate is ramped up. On the other hand, the analysis based on our resummed form yields values consistent with the correct landscape parameters. Moreover, even at low pulling rates, where Bell-Evans performs not too badly, our proposal is more reliable and produces less scatter in the parameter values.

We remark that fits of the simulation data to Eq. (11) yield astonishingly good values of  $\kappa^\ddagger$ , the effective curvature (see Fig. 7). In almost every case, regardless of pulling rate, the predicted value of  $\kappa^\ddagger$  coincides with the true value. This suggests to us that our inclusion of higher-order corrections in the rate equation plays an important role in improving the overall quality of the parameter extraction.

To conclude, our work highlights the known inadequacies of the Bell-Evans phenomenological well escape rate. It also

suggests that the celebrated equation due to Dudko and co-workers is not an adequate fix. We propose a new expression, Eq. (4), that improves on the Bell-Evans expression by including beyond-Arrhenius contributions from the shape of the energy potential. Equation (4) clearly outperforms the Bell-Evans and Dudko expressions in terms of predicting the well escape rate (as is evident from Fig. 3). Crucially, it avoids the unphysical behavior that plagues Dudko's rate equation at large pulling force.

Of particular utility is that Eq. (4) integrates to give a manageable, closed-form expression for the cumulative probability distribution. The resulting Eq. (11) is

straightforward to implement as a fitting function and can be incorporated into existing workflows with little additional effort. Rigorous numerical tests (illustrated in Figs. 6 and 7) confirm that fits to Eq. (11) can be used to reliably extract the parameters that characterize the underlying energy landscape.

## ACKNOWLEDGMENTS

One of the authors (S.A.) thanks Thomas Perkins (JILA, NIST, and the University of Colorado Boulder) for helpful discussions.

- 
- [1] M. Stöhr and A. Tkatchenko, Quantum mechanics of proteins in explicit water: The role of plasmon-like solute-solvent interactions, *Sci. Adv.* **5**, eaax0024 (2019).
  - [2] J. D. Bryngelson and P. G. Wolynes, Spin glasses and the statistical mechanics of protein folding, *Proc. Natl. Acad. Sci. USA* **84**, 7524 (1987).
  - [3] J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, Theory of protein folding: The energy landscape perspective, *Annu. Rev. Phys. Chem.* **48**, 545 (1997).
  - [4] O. V. Galzitskaya and A. V. Finkelstein, A theoretical search for folding/unfolding nuclei in three-dimensional protein structures, *Proc. Natl. Acad. Sci. USA* **96**, 11299 (1999).
  - [5] J. N. Onuchic, H. Nymeyer, A. E. García, J. Chahine, and N. D. Socci, The energy landscape of protein folding: Insights into folding mechanisms and scenarios, *Adv. Protein Chem.* **53**, 87 (2000).
  - [6] P. Talkner and D. Ryter, Lifetime of a metastable state at low noise, *Phys. Lett. A* **88**, 162 (1982).
  - [7] R. Merkel, P. Nassoy, A. Leung, K. Ritchie, and E. Evans, Energy landscapes of receptor-ligand bonds explored with dynamic force spectroscopy, *Nature (London)* **397**, 50 (1999).
  - [8] J. Liphardt, B. Onoa, S. B. Smith, I. Tinoco, and C. Bustamante, Reversible unfolding of single RNA molecules by mechanical force, *Science* **292**, 733 (2001).
  - [9] P. C. Li and D. E. Makarov, Simulation of the mechanical unfolding of ubiquitin: Probing different unfolding reaction coordinates by changing the pulling geometry, *J. Chem. Phys.* **121**, 4826 (2004).
  - [10] S. Kirmizialtin, L. Huang, and D. E. Makarov, Topography of the free-energy landscape probed via mechanical unfolding of proteins, *J. Chem. Phys.* **122**, 234915 (2005).
  - [11] P. Hinterderfer, and Y. F. Dufrene, Detection and localization of single molecular recognition events using atomic force microscopy, *Nat. Methods* **3**, 347 (2006).
  - [12] Y. Gilbert, M. Deghorain, L. Wang, B. Xu, P. D. Pollheimer, H. J. Gruber, J. Errington, B. Hallet, X. Haulot, C. Verbelen, P. Hols, and Y. F. Dufrene, Single-molecule force spectroscopy and imaging of the vancomycin/D-Ala-D-Ala interaction, *Nano Lett.* **7**, 796 (2007).
  - [13] W. J. Greenleaf, M. T. Woodside, and S. M. Block, High-resolution, single-molecule measurements of biomolecular motion, *Annu. Rev. Biophys. Biomol. Struct.* **36**, 171 (2007).
  - [14] K. Neupane, H. Yu, D. A. N. Foster, F. Wang, and M. T. Woodside, Single-molecule force spectroscopy of the adenine riboswitch relates folding to regulatory mechanism, *Nucleic Acids Res.* **39**, 7677 (2011).
  - [15] D. E. Souza, Pulling on single molecules, *Nat. Methods* **9**, 873 (2012).
  - [16] M. S. Bull, R. M. A. Sullan, H. Li, and T. T. Perkins, Improved single molecule force spectroscopy using micromachined cantilevers, *ACS Nano* **8**, 4984 (2014).
  - [17] D. T. Edwards, J. K. Faulk, A. W. Sanders, M. S. Bull, R. Walder, M.-A. LeBlanc, M. C. Sousa, and T. T. Perkins, Optimizing 1- $\mu$ s-resolution single-molecule force spectroscopy on a commercial atomic force microscope, *Nano Lett.* **15**, 7091 (2015).
  - [18] W. J. Van Patten, R. Walder, A. Adhikari, S. R. Okoniewski, R. Ravichandran, C. E. Tinberg, D. Baker, and T. T. Perkins, Improved free-energy landscape quantification illustrated with a computationally designed protein-ligand interaction, *ChemPhysChem* **19**, 19 (2017).
  - [19] S. R. Okoniewski, L. Uyetake, and T. T. Perkins, Force-activated DNA substrates for probing individual proteins interacting with single-stranded DNA, *Nucleic Acids Res.* **45**, 10775 (2017).
  - [20] H. Yu, M. G. W. Siewny, D. T. Edwards, A. W. Sanders, and T. T. Perkins, Hidden dynamics in the unfolding of individual bacteriorhodopsin proteins, *Science* **355**, 945 (2017).
  - [21] R. Walder, W. J. Van Patten, A. Adhikari, and T. T. Perkins, Going vertical to improve the accuracy of atomic force microscopy based single-molecule force spectroscopy, *ACS Nano* **12**, 198 (2018).
  - [22] R. Walder, W. J. Van Patten, D. B. Ritchie, R. K. Montange, T. W. Miller, M. T. Woodside, and T. T. Perkins, High-precision single-molecule characterization of the folding of an HIV RNA hairpin by atomic force microscopy, *Nano Lett.* **18**, 6318 (2018).
  - [23] C. Jarzynski, Nonequilibrium Equality of Free Energy Differences, *Phys. Rev. Lett.* **78**, 2690 (1997).
  - [24] G. Hummer and A. Szabo, Free energy reconstruction from nonequilibrium single-molecule pulling experiments, *Proc. Natl. Acad. Sci. USA* **98**, 3658 (2001).
  - [25] N. C. Harris, Y. Song, and C. H. Kiang, Experimental Free Energy Reconstruction from Single Molecule Force Spectroscopy Using Jarzynski's Inequality, *Phys. Rev. Lett.* **99**, 068101 (2007).
  - [26] A. N. Gupta, A. Vincent, K. Neupane, H. Yu, F. Wang, and M. T. Woodside, Experimental validation of

- free-energy-landscape reconstruction from non-equilibrium single-molecule force spectroscopy measurements, *Nat. Phys.* **7**, 631 (2011).
- [27] Q. Zhang, J. Brujic, and E. V. Eijnden, Reconstructing free energy profiles from nonequilibrium relaxation trajectories, *J. Stat. Phys.* **144**, 344 (2011).
- [28] M. C. Engel, D. B. Ritchie, D. A. N. Foster, K. S. D. Beach, and M. T. Woodside, Reconstructing Folding Energy Landscape Profiles from Nonequilibrium Pulling Curves with an Inverse Weierstrass Integral Transform, *Phys. Rev. Lett.* **113**, 238104 (2014).
- [29] A. P. Manuel, J. Lambert, and M. T. Woodside, Reconstructing folding energy landscapes from splitting probability analysis of single-molecule trajectories, *Proc. Natl. Acad. Sci. USA* **112**, 7183 (2015).
- [30] P. R. Heenan, H. Yu, M. G. W. Siewny, and T. T. Perkins, Improved free-energy landscape reconstruction of bacteriorhodopsin highlights local variations in unfolding energy, *J. Chem. Phys.* **148**, 123313 (2018).
- [31] N. J. L. Alamilla, M. W. Jack, and K. J. Challis, Analysing single-molecule trajectories to reconstruct free-energy landscapes of cyclic motor proteins, *J. Theor. Biol.* **462**, 321 (2019).
- [32] N. J. L. Alamilla, M. W. Jack, and K. J. Challis, Reconstructing free-energy landscapes for cyclic molecular motors using full multidimensional or partial one-dimensional dynamic information, *Phys. Rev. E* **100**, 012404 (2019).
- [33] R. I. Litvinov, H. Shuman, J. S. Bennett, and J. W. Weisel, Binding strength and activation state of single fibrinogen-integrin pairs on living cells, *Proc. Natl. Acad. Sci. USA* **99**, 7426 (2002).
- [34] G. I. Bell, Models for the specific adhesion of cells to cells, *Science* **200**, 618 (1978).
- [35] E. Evans, D. Berk, and A. Leung, Detachment of agglutinin-bonded red blood cells. I. Forces to rupture molecular-point attachments, *Biophys. J.* **59**, 838 (1991).
- [36] E. Evans and K. Ritchie, Dynamic strength of molecular adhesion bonds, *Biophys. J.* **72**, 1541 (1997).
- [37] M. Rief, M. Gautel, F. Oesterhelt, J. M. Fernandez, and H. E. Gaub, Reversible unfolding of individual titin immunoglobulin domains by AFM, *Science* **276**, 1109 (1997).
- [38] M. Rief, J. M. Fernandez, and H. E. Gaub, Elastically Coupled Two-Level Systems as a Model for Biopolymer Extensibility, *Phys. Rev. Lett.* **81**, 4764 (1998).
- [39] M. T. Woodside, W. M. Behnke-Parks, K. Larizadeh, K. Travers, D. Herschlag, and S. M. Block, Nanomechanical measurements of the sequence-dependent folding landscapes of single nucleic acid hairpins, *Proc. Natl. Acad. Sci. USA* **103**, 6190 (2006).
- [40] M. T. Woodside and S. M. Block, Reconstructing folding energy landscapes by single-molecule force spectroscopy, *Annu. Rev. Biophys.* **43**, 19 (2014).
- [41] V. Barsegov and D. Thirumalai, Probing Protein-Protein Interactions by Dynamic Force Correlation Spectroscopy, *Phys. Rev. Lett.* **95**, 168302 (2005).
- [42] A. Maitra and G. Arya, Model Accounting for the Effects of Pulling-Device Stiffness in the Analyses of Single-Molecule Force Measurements, *Phys. Rev. Lett.* **104**, 108301 (2010).
- [43] D. B. Ritchie and M. T. Woodside, Probing the structural dynamics of proteins and nucleic acids with optical tweezers, *Curr. Opin. Struct. Biol.* **34**, 43 (2015).
- [44] M. T. Woodside, J. Lambert, and K. S. D. Beach, Determining intrachain diffusion coefficients for biopolymer dynamics from single-molecule force spectroscopy measurements, *Biophys. J.* **107**, 1647 (2014).
- [45] D. E. Makarov, Communication: Does force spectroscopy of biomolecules probe their intrinsic dynamic properties? *J. Chem. Phys.* **141**, 241103 (2014).
- [46] G. M. Nam and D. E. Makarov, Extracting intrinsic dynamic parameters of biomolecular folding from single-molecule force spectroscopy experiments, *Protein Sci.* **25**, 123 (2016).
- [47] G. Hummer and A. Szabo, Kinetics from nonequilibrium single-molecule pulling experiments, *Biophys. J.* **85**, 5 (2003).
- [48] A. Garg, Escape-field distribution for escape from a metastable potential well subject to a steadily increasing bias field, *Phys. Rev. B* **51**, 15592 (1995).
- [49] R. W. Friddle, Unified Model of Dynamic Forced Barrier Crossing in Single Molecules, *Phys. Rev. Lett.* **100**, 138302 (2008).
- [50] O. K. Dudko, G. Hummer, and A. Szabo, Intrinsic Rates and Activation Free Energies from Single-Molecule Pulling Experiments, *Phys. Rev. Lett.* **96**, 108101 (2006).
- [51] H. A. Kramers, Brownian motion in a field of force and the diffusion model of chemical reactions, *Physica* **7**, 284 (1940).
- [52] P. Hänggi, P. Talkner, and M. Borkovec, Reaction-rate theory: Fifty years after Kramers, *Rev. Mod. Phys.* **62**, 251 (1990).
- [53] N. Grønbech-Jensen and O. Farago, A simple and effective Verlet-type algorithm for simulating Langevin dynamics, *Mol. Phys.* **111**, 983 (2013).
- [54] L. Verlet, Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules, *Phys. Rev.* **159**, 98 (1967).
- [55] B. Efron and R. J. Tibshirani, *An Introduction to the Bootstrap*, 1st ed. (Chapman & Hall, New York, 1993).