


Experimental Measurement of Relative Path Probabilities and Stochastic Actions

Jannes Gladrow,^{1,2} Ulrich F. Keyser,² R. Adhikari,³ and Julian Kappler^{3,*}

¹Microsoft Research, Station Rd, Cambridge CB1 2FB, United Kingdom

²Cavendish Laboratory, University of Cambridge, JJ Thomson Ave,
Cambridge CB3 0HE, United Kingdom

³Department of Applied Mathematics and Theoretical Physics, Centre for Mathematical Sciences,
University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, United Kingdom

 (Received 4 June 2020; revised 11 May 2021; accepted 18 May 2021; published 27 July 2021)

For diffusive stochastic dynamics, the probability to observe any individual trajectory is vanishingly small, making it unclear how to experimentally validate theoretical results for ratios of path probabilities. We provide the missing link between theory and experiment by establishing a protocol to extract ratios of path probabilities from measured time series. For experiments on a single colloidal particle in a microchannel, we extract both ratios of path probabilities and the most probable path for a barrier crossing, and find excellent agreement with independently calculated predictions based on the Onsager-Machlup stochastic action. Our experimental results at room temperature are found to be inconsistent with the low-noise Freidlin-Wentzell stochastic action, and we discuss under which circumstances the latter action is expected to describe the most probable path. Furthermore, while the experimentally accessible ratio of path probabilities is uniquely determined, the formal path-integral action is known to depend on the time-discretization scheme used for deriving it; we reconcile these two seemingly contradictory facts by careful analysis of the time-slicing derivation of the path integral. Our experimental protocol enables us to probe probability distributions on path space and allows us to relate theoretical single-trajectory results to measurement.

DOI: [10.1103/PhysRevX.11.031022](https://doi.org/10.1103/PhysRevX.11.031022)

Subject Areas: Chemical Physics, Soft Matter,
Statistical Physics

I. INTRODUCTION

Stochastic effects are of fundamental relevance for statistical physics and beyond [1–9]. For example, diffusion processes are used to model colloidal particles [4,10,11], polymer dynamics [12–15], or active particles such as driven colloidal systems, cells, or bacteria [8,10]. Any stochastic dynamics is fully characterized by its path probabilities, which are also highly relevant in applications; examples are irreversibility in stochastic thermodynamics, which is expressed in terms of ratios of path probabilities [10,16], or transition pathways between metastable states [17,18], as relevant, e.g., for conformational transitions in biomolecules [19–21].

For diffusive dynamics, the probability to observe any given individual path is zero. Still, ratios of path probabilities can be quantified theoretically by stochastic actions [22–37]. The literature contains several proposals for stochastic actions, the prominent ones being associated with the names of Onsager and Machlup (OM) [22–25], as

well as Freidlin and Wentzell (FW) [38–40]. Since it is not straightforward to access probability-zero events in measurement, hitherto it was not clear how to directly validate the theoretical results for stochastic actions.

Here we overcome this difficulty by establishing an experimental protocol to determine the ratios of path probabilities from observed data without fitting a model to the stochastic dynamics. The key observable for our protocol is the sojourn probability [23,25,38], i.e., the probability that a stochastic trajectory remains within a tube of small-but-finite radius R around a reference path [see Fig. 3(a) for an illustration]. For a colloidal particle in a microchannel subject to a double-well potential, we directly measure the finite-radius sojourn probability for a pair of reference paths, and subsequently extrapolate the ratio of sojourn probabilities to the limit $R \rightarrow 0$. We demonstrate that this experimentally observed ratio of path probabilities is well described by the difference in OM Lagrangians along the two reference paths, thereby confirming classical theoretical results on the asymptotic sojourn probability [23–25] and transforming the OM action from a purely mathematical construct into a physical observable. We observe that our results for relative path probabilities at room temperature are markedly different from the predictions of the FW Lagrangian [38–40], which in the context of sojourn probabilities is associated with a

*jkappler@posteo.de

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

low-noise limit [38]. Considering the most probable path, or instanton, as the zero-radius limit of the most probable tube, we furthermore determine the instanton for experimental barrier-crossing paths in a double-well potential, which is again well described by the OM Lagrangian and different from the FW prediction. We discuss quantitatively for which system parameters the FW Lagrangian [38] is expected to describe the physical most probable path. Finally, we resolve the seeming contradiction that the asymptotic-tube Lagrangian is uniquely defined, whereas the formal stochastic action Lagrangian that appears in the path-integral formalism is not. Our results demonstrate that ratios of path probabilities can be inferred from experimental data without the need to fit a stochastic model.

II. THEORY AND EXPERIMENTAL RESULTS

A. Sojourn probability and stochastic action

We consider the approach to relative path probabilities via the sojourn probability $P_R^\varphi(t)$, i.e., the probability that a stochastic trajectory $X_t \equiv X(t)$ remains within a moving ball of radius R and with a center parametrized by a twice-differentiable reference path $\varphi_t \equiv \varphi(t)$ up to a time t [23–29,38]. The relative likelihood for two given reference paths $\varphi(t), \psi(t), t \in [t_i, t_f]$ is then quantified by a stochastic action S defined via [23–29,38]

$$\frac{e^{-S[\varphi]}}{e^{-S[\psi]}} \equiv \lim_{R \rightarrow 0} \frac{P_R^\varphi(t_f)}{P_R^\psi(t_f)}. \quad (1)$$

For Markovian dynamics, the action is the integral over a Lagrangian [25,29]

$$S[\varphi] = \int_{t_i}^{t_f} dt \mathcal{L}^\varphi(t), \quad (2)$$

and for overdamped Langevin dynamics with additive noise, Stratonovich was the first to show that the limiting ratio of sojourn probabilities is described by the OM Lagrangian [23]; this analytical result has since been confirmed by several subsequent derivations [24,25,27–29,41]. A notable exception is the derivation of Ventsel' and Freidlin [38], where the FW Lagrangian is obtained as the sojourn probability in a low-noise limit.

A different route to quantifying path probabilities via stochastic actions is given by the path-integral formalism, in which the propagator of a stochastic dynamics is formally represented as an integral over all paths connecting an initial and a final point [22,30,31,33–37]. In the standard time-slicing derivation of the path-integral formalism, a formal stochastic action is defined via the continuum limit of concatenated short-time propagators. The formal limit is not uniquely defined, but even for overdamped Langevin dynamics with constant diffusivity (additive noise), it depends on the discretization used for

the short-time propagators; this discretization-dependence of the short-time propagator leads to the FW and OM actions as two of the infinitely many equivalent representations of the path-integral action [31,35,36,42]. While this nonuniqueness seems in conflict with the definite finite-temperature limit found in Eq. (2), in Sec. III, we resolve this apparent contradiction by careful analysis of the time-slicing path-integral derivation.

The practical relevance of the asymptotic sojourn probability Eq. (1) is that it directly leads to a simple relation between the stochastic action Lagrangian and physical observables. To make this explicit, we substitute Eq. (2) into the logarithm of Eq. (1), differentiate the result with respect to t_f , and subsequently rename t_f as t . This calculation yields

$$\mathcal{L}^\varphi(t) - \mathcal{L}^\psi(t) = \lim_{R \rightarrow 0} (\alpha_R^\varphi(t) - \alpha_R^\psi(t)), \quad (3)$$

where the instantaneous exit rate at which stochastic trajectories first leave the ball of radius R around φ is given by $\alpha_R^\varphi(t) \equiv -\dot{P}_R^\varphi(t)/P_R^\varphi(t)$, where a dot denotes a derivative with respect to time t . For a finite radius R , each exit rate on the right-hand side of Eq. (3) can be measured directly experimentally via the ratio of recorded trajectories which remain within the threshold distance R to the respective reference path as a function of time. Thus, Eq. (3) yields a model-free experimental route to action Lagrangian differences via extrapolating exit-rate differences measured at finite radius to the limit $R \rightarrow 0$. In the following, we use this extrapolation strategy to measure stochastic Lagrangian differences, and hence to experimentally test theoretical predictions for ratios of path probabilities.

B. Experimental setup

In our experiments, we observe the motion of a colloidal particle confined to a microchannel. Our experimental setup illustrated in Fig. 1 consists of a holographic optical tweezer, which can autonomously capture colloidal particles and position them inside a microchannel filled with an aqueous salt solution. Because of the strong confinement created by the channel, the motion of the center point of the colloidal particle can be considered effectively one dimensional, i.e., along the channel axis. A spatial light modulator is used to form an optical landscape which gives rise to a potential-energy landscape that the colloid experiences inside the channel. We tune the modulator to create an approximate double-well potential shown in Fig. 1(c); on the scales we probe, this potential is time independent. The position of the colloidal particle is recorded at 1000 frames per second. In total, we record and analyze approximately 104 minutes of experimental measurements partitioned into short trajectories of variable length ranging from 10 to 60 seconds. The experimental setup is discussed in more detail in Refs. [43,44].

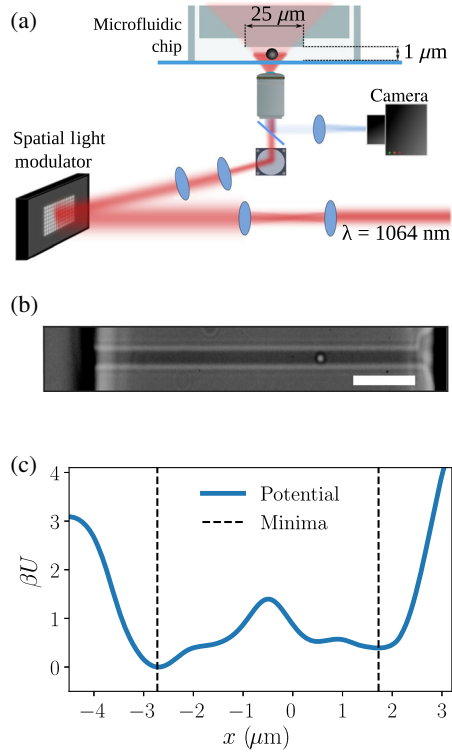


FIG. 1. (a) Experimental setup. In our experiments, following the general setup of Ref. [44], we observe the motion of a colloidal particle inside a microchannel. (b) Image of a colloidal particle in a microchannel. The motion of the colloidal particle can be considered effectively one dimensional. The horizontal scale bar in the lower right corner is $5 \mu\text{m}$ in length; the colloidal particle has a diameter of 500 nm . (a) and (b) are adapted from Ref. [44]. (c) Potential energy extracted from experimental time series. The blue solid line depicts the potential energy obtained from evaluating the first two Kramer-Moyal coefficients based on experimental data, and subsequent smoothing as explained in the Appendix A. The vertical dashed lines denote two local minima of the potential-energy landscape located at $x_0^{\min} \approx -2.7 \mu\text{m}$, $x_1^{\min} \approx 1.7 \mu\text{m}$.

C. Relative path likelihoods from experiment

We now compare the experimentally measured right-hand side of Eq. (3) to the corresponding difference in theoretical stochastic action Lagrangians. From recorded experimental time series of a colloidal particle in a microchannel, as illustrated in Fig. 1, we evaluate

$$\Delta\alpha_R(t) \equiv \alpha_R^\varphi(t) - \alpha_R^\psi(t) \quad (4)$$

for several finite values of R . Since the sojourn probabilities for the paths and radii we consider are so small that not a single recorded trajectory remains within the tube until the final time, we introduce a cloning algorithm to obtain the experimental exit rate based on concatenating short recorded trajectories; see Fig. 2 for an illustration and Appendix B for further details. For φ , we consider a path

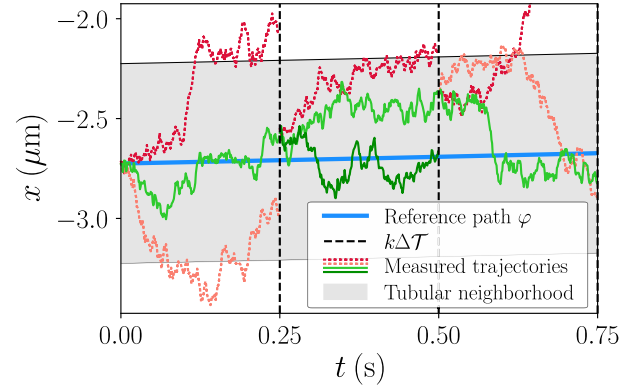


FIG. 2. Illustration of our algorithm for obtaining sojourn probabilities from measured time series. The solid blue line represents a reference path φ , around which a tube of radius $R = 0.5 \mu\text{m}$ is shown as the gray shaded area. We randomly select $M = 3$ measured trajectories which start in a small interval around $x_0^{\min} \approx -2.7 \mu\text{m}$, and follow them for a duration $\Delta\mathcal{T} = 0.25 \text{ s}$ (vertical dashed lines). Trajectories which leave the tube (red dotted lines) are discarded, and the final positions of those trajectories that stay in the tube (green solid lines) are collected. We then again randomly select $M = 3$ measured trajectories, which start in a small interval around any of the collected final positions, and repeat the process. The exit rate $\alpha_R^\varphi(t)$ which appears in Eq. (4) is the rate at which the red sample trajectories leave the tube for the first time. The small value $M = 3$ is chosen here for illustration; to calculate exit rates from experimental data, we use values of the order 10^4 , which are chosen dynamically; see Appendix B for details. Our algorithm assumes that the time series is Markovian; we verify in Appendix A that this holds approximately for our experimental data. To demonstrate that concatenating short measured trajectories does not artificially alter the dynamics, in Appendix B we vary the parameters of the algorithm, including $\Delta\mathcal{T}$ and M , and find that the inferred exit rate is independent of the particular choices. All trajectories shown here are actual experimental data, and the reference path φ is the same as in Fig. 3(a).

which moves from the left minimum of the experimental potential energy to the right minimum of the potential energy in $\Delta t_{\text{tot}} = 20 \text{ s}$, as illustrated in Fig. 3(a). For ψ we consider a constant path, which rests at the right minimum for the duration $\Delta t_{\text{tot}} = 20 \text{ s}$ shown as the upper horizontal dashed line in Fig. 3(a). This choice for ψ ensures that all time dependence in the measured exit-rate differences can be attributed to the exit rate pertaining to φ . In principle, arbitrary pairs of twice-differentiable paths of the same duration can be considered in Eq. (1); we illustrate this in Appendix B, where we show the results for several other path pairs. In Fig. 3(b), we show the exit-rate differences Eq. (4) for the paths φ, ψ for several finite values of R . We extrapolate to $R = 0$ as follows. Since the exit rate is invariant under a parity transformation around the instantaneous tube center $\varphi(t)$, for small radius the difference in exit rates scales as $\Delta\alpha_R(t) = \Delta\alpha^{(0)}(t) + R^2\Delta\alpha^{(2)}(t) + \mathcal{O}(R^4)$. For every time t , we therefore fit a quadratic function

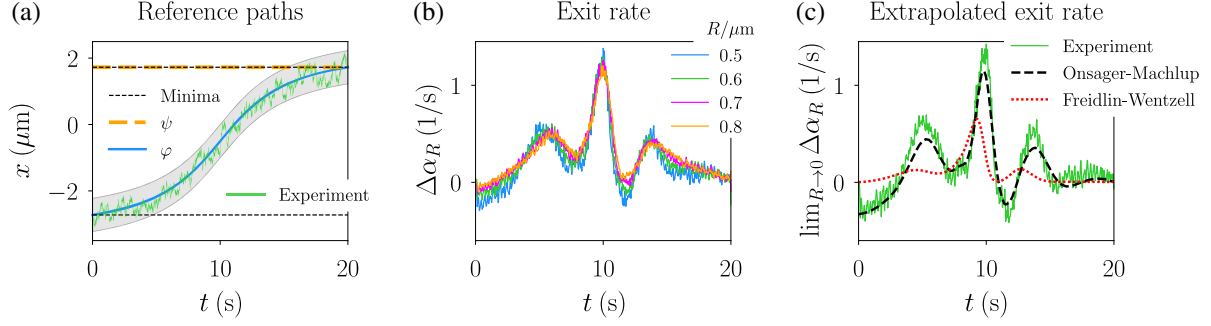


FIG. 3. (a) Reference paths used to extract relative path probabilities. The thin horizontal dashed lines denote the minima in the experimental potential-energy landscape; cf. Fig. 1. The upper dashed line additionally denotes the constant path ψ , as indicated by a thick orange dashed line. The blue solid line denotes a path φ which moves from the left potential-energy minimum to the right minimum in 20 seconds. The gray shaded region around the path φ indicates a tube of radius $R = 0.5 \mu\text{m}$; the green solid line depicts concatenated experimental time series obtained using the algorithm from Fig. 2. (b) Exit-rate differences for finite radius R . Colored solid lines denote the exit-rate difference Eq. (4) extracted directly from experimental time series for various values of the radius R , as indicated in the legend. The used reference paths φ, ψ are shown in (a); the shown exit rates are smoothed using a Hann window of width 0.1s. (c) Extrapolation of exit-rate differences to radius $R = 0$. The green solid line denotes the extrapolation to $R = 0$ of (the presmoothing versions of) finite-radius exit-rate differences as shown in (b). The shown extrapolated exit rate is smoothed using a Hann window of width 0.1s. The black dashed line and the red dotted line denote the difference in the OM and FW Lagrangians for the paths φ, ψ calculated using Eqs. (5) and (6) the diffusivity and force estimated from the experimental data; cf. Fig. 1.

$f(t, R) = a(t) + R^2 b(t)$ to measured finite-radius exit rates, as the ones shown in Fig. 3(b), and extrapolate to vanishing radius as $\lim_{R \rightarrow 0} \Delta\alpha_R(t) \equiv a(t)$. Figure 3(c) compares the result with the theoretical difference in OM Lagrangians [22–25]

$$\mathcal{L}^\varphi(t) = \frac{\gamma}{4k_B T_0} \left[\dot{\varphi}(t) - \frac{1}{\gamma} F(\varphi(t)) \right]^2 + \frac{1}{2\gamma} (\partial_x F)(\varphi(t)), \quad (5)$$

where $\gamma = 1.75 \times 10^{-8}$ kg/s is the friction coefficient, $k_B T_0$ the thermal energy with k_B the Boltzmann constant and $T_0 = 294$ K the absolute temperature at which the experiment is carried out, and $F(x) = -(\partial_x U)(x)$ the force corresponding to the potential shown in Fig. 1; for details on the parametrization of $\gamma, U(x)$, see Appendix A. While in Fig. 3(c) there are some minor differences between theory and experimental extrapolation, the overall agreement is very good. This shows both that our protocol for extracting ratios of path probabilities from experiments yields meaningful results and confirms that relative path probabilities are indeed quantified by the OM Lagrangian [23]. On the other hand, the difference in FW Lagrangians [38] given by

$$\mathcal{L}_{\text{FW}}^\varphi(t) = \frac{\gamma}{4k_B T_0} \left[\dot{\varphi}(t) - \frac{1}{\gamma} F(\varphi(t)) \right]^2, \quad (6)$$

and also shown in Fig. 3(c), disagrees considerably with both the experimental data and the OM prediction. This illustrates very clearly that in the context of physically observable asymptotic sojourn probabilities at finite temperature, the two actions Eqs. (5) and (6) are not equivalent. They differ by a term proportional to $\partial_x F$, which in the nonlinear force profile we consider here, contributes

considerably to the asymptotic exit rates. This highlights that, while in the context of the path-integral formalism, the OM and FW actions are usually considered equivalent, this is not the case for asymptotic sojourn probabilities. We discuss this in more detail in Sec. III.

According to Eqs. (1)–(3), the ratio of path probabilities is given by the exponential of the negative integral over the data shown in Fig. 3(c). For the extrapolated experimental exit-rate difference, this yields a value of approximately 0.026 for the path-probability ratio. As expected from Fig. 3(c), this experimental ratio is closer to the OM value, which evaluates to 0.068, as compared to the FW value, which yields 0.132.

D. Most probable path from experiment

The most probable path φ^* , also called the instanton, connecting an initial point $\varphi^*(t_i) = x_i$ and a final point $\varphi^*(t_f) = x_f$, is given by

$$\varphi^* \equiv \lim_{R \rightarrow 0} \left[\operatorname{argmin}_{\varphi} \int_{t_i}^{t_f} dt \alpha_R^\varphi(t) \right], \quad (7)$$

where we minimize over all continuous paths with given end points $\varphi(t_i) = x_i, \varphi(t_f) = x_f$. This equation follows from maximizing the right-hand side of Eq. (1) with respect to φ for any fixed ψ , and states that the most probable path is the one where the integrated exit rate diverges slowest as $R \rightarrow 0$.

As in Fig. 3, we use $t_i = 0, t_f = 20$ s, and for x_i, x_f , we consider the two minima of the experimental potential energy; cf. Fig. 1. Using our experimental time series, we minimize the right-hand side of Eq. (7), but without the limit, to obtain the most probable tube for several

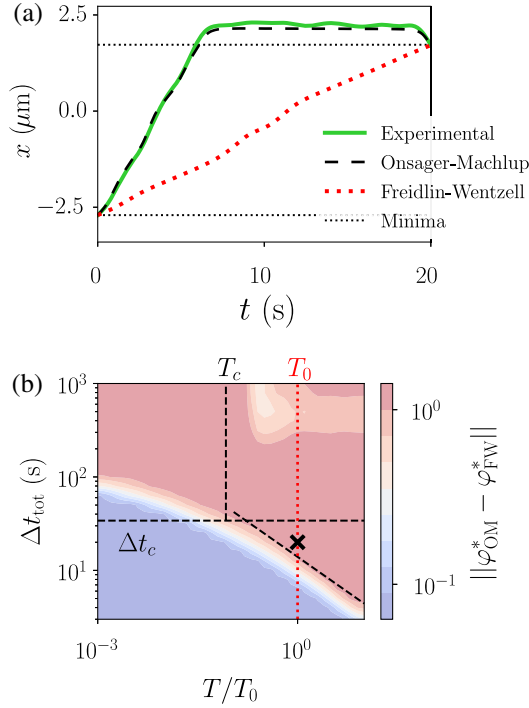


FIG. 4. (a) Most probable paths for barrier crossing. The green solid line denotes the most probable path extracted directly from experimental data; see Appendix C for details. The black dashed line is obtained by minimizing the integrated OM Lagrangian Eq. (5), and the red dotted line is obtained by minimizing the integrated FW Lagrangian Eq. (6). (b) Mean difference between the OM and FW instanton. The actions corresponding to the Lagrangians Eqs. (5) and (6) are minimized for various values of temperature T/T_0 and total transition time Δt_{tot} . The plot shows the dimensionless mean difference between the resulting instantons, as defined in Eq. (8); for technical details, see Appendix D. The horizontal and diagonal black dashed lines denote the crossover time $\Delta t_c \approx 34$ s defined in Eq. (9), and the right-hand side of Eq. (10). The black vertical dashed line denotes the crossover temperature $T/T_0 \approx 0.08$ defined in Eq. (11); the red vertical dotted line denotes the reference temperature $T_0 = 294$ K. The cross denotes the parameters $(T/T_0, \Delta t_{\text{tot}})$ used for (a).

finite values of R . We subsequently extrapolate the result to the limit $R \rightarrow 0$, to obtain the most probable path; for details, see Appendix C. In Fig. 4(a), we compare the resulting experimental instanton to the directly minimized OM action obtained by substituting Eq. (5) into Eq. (2). As the figure shows, the extrapolated most probable tube agrees very well with the OM instanton, demonstrating that the most probable path can be extracted from experimental data without fitting a model for the stochastic dynamics. The FW instanton obtained from minimizing the temporal integral over the Lagrangian Eq. (6) is also shown in Fig. 4(a) and disagrees significantly with the experimental data. This confirms experimentally that the OM Lagrangian yields the correct action to describe

the physically observable most probable paths at a finite temperature [24,36].

E. Range of validity of the FW Lagrangian

In the context of sojourn probabilities, the FW Lagrangian is derived for asymptotically low noise [38], which in the present case means asymptotically low temperature. Indeed, for fixed γ , F , the first term in Eq. (5) scales with the inverse temperature, whereas the second term is independent of the temperature. One might thus expect that for a sufficiently small temperature, the second term becomes irrelevant and the Lagrangian Eq. (5) reduces to Eq. (6). To quantitatively understand in which parameter regime the FW Lagrangian predicts an instanton which is in agreement with the OM instanton, we now compare the minima of the functionals obtained by substituting Eqs. (5) and (6) into Eq. (2) for a wide range of both the total duration Δt_{tot} and the temperature T . For all temperatures, we use the friction coefficient γ and the force profile F inferred from our experimental data at the temperature $T_0 = 294$ K. In reality, γ and F of the experimental system shown in Fig. 1(a) of course depend on the temperature [45]. Here we hold those parameters fixed because the focus of this section is not to understand the temperature dependence of our particular experimental system, but of the theoretical instantons predicted by Eqs. (5) and (6). Note that, while in the FW Lagrangian Eq. (6), the temperature appears only as an overall prefactor which does not affect the variational extremum, varying the temperature in the OM Lagrangian Eq. (5) amounts to changing the relative size of the two terms.

In Fig. 4(b), we show the numerically evaluated dimensionless average difference between FW and OM instanton

$$\|\varphi_{\text{FW}}^* - \varphi_{\text{OM}}^*\| \equiv \frac{1}{\Delta t_{\text{tot}} L} \int_{t_i}^{t_f} dt |\varphi_{\text{FW}}^*(t) - \varphi_{\text{OM}}^*(t)| \quad (8)$$

as a function of the total duration $\Delta t_{\text{tot}} = t_f - t_i$, and temperature T/T_0 , with the experimental temperature $T_0 = 294$ K indicated in the plot by a red vertical dotted line. For the typical length scale in Eq. (8), we use $L = 1 \mu\text{m}$. As can be seen, for short total duration Δt_{tot} , there is a parameter regime where the FW Lagrangian predicts the correct instanton. The FW Lagrangian has been derived as a low-temperature approximation to the sojourn probability [38], but for total times $\Delta t_{\text{tot}} \gtrsim 10^2$ s the FW and OM instantons disagree even at the lowest temperature considered here; we discuss the reason for this further below. In Appendix E, we show instantons from experimental and numerical data for several more points in the $(T/T_0, \Delta t_{\text{tot}})$ plane, all of which confirm that the OM instanton describes the measured instanton.

A quantitative estimate for the range of applicability of the FW action to determine the instanton is obtained by

investigating for which parameters the second term in Eq. (5) is negligible as compared to the first term. For this, we distinguish between the two limiting cases of short and long total duration. If the total duration Δt_{tot} is short, the precise meaning of which is quantified in the following, then the trajectory needs a large velocity to reach the given final position $x_f = x_i + \Delta x$. For this scenario, we estimate the typical velocity as $\dot{\varphi} \approx \Delta x / \Delta t_{\text{tot}}$, so that $\dot{\varphi}^2$ dominates the first term in Eq. (5) if the total time is much smaller than a crossover time

$$\Delta t_{\text{tot}} \ll \Delta t_c \equiv \frac{\gamma |\Delta x|}{\langle |F| \rangle}, \quad (9)$$

where the brackets $\langle \rangle$ denote a spatial average between x_i and x_f , so that $\langle |F| \rangle$ represents the typical magnitude of the force between x_i and x_f . Using the experimentally inferred values for γ , F , Eq. (9) yields $\Delta t_c \approx 34$ s, which is shown in Fig. 4(b) as a horizontal dashed line. For a fast transition, $\Delta t_{\text{tot}} \ll \Delta t_c$, the second term in the OM Lagrangian Eq. (5) is negligible as compared to the first term if

$$\Delta t_{\text{tot}} \ll \frac{1}{\sqrt{T/T_0}} \frac{|\Delta x|}{D_0 \sqrt{\beta_0} \langle |\partial_x F| \rangle}, \quad (10)$$

where we write $\beta_0^{-1} \equiv k_B T_0$ as well as $D_0 = k_B T_0 / \gamma$. For our example system, $|\Delta x| / (D_0 \sqrt{\beta_0} \langle |\partial_x F| \rangle) \approx 14$ s. The right-hand side of Eq. (10) is shown in Fig. 4(b) as a diagonal black dashed line, and indeed, in the lower left corner of the plot, where both Eqs. (9) and (10) are fulfilled, the FW action predicts the correct instanton. Surprisingly, even though the FW action is in the context of sojourn probabilities derived only for low noise, it predicts the correct instanton even at $T/T_0 = 1$ if the transition is fast enough; we illustrate this with an example in Appendix E. For long total duration $\Delta t_{\text{tot}} \gg \Delta t_c$, we observe in Fig. 4(b) that the FW instanton is different from the OM instanton even for the smallest temperatures considered. To understand this, we first discuss a simple scaling argument to determine at which temperature the second term in Eq. (5) might be negligible, and then rationalize why even in this regime we observe deviations between the FW and OM instanton. For $\Delta t_{\text{tot}} \gg \Delta t_c$, the first term in the OM Lagrangian Eq. (5) is expected to be of order $\langle |F| \rangle^2 / (4\gamma k_B T)$, so that the second term should be negligible if the temperature is much smaller than a crossover temperature T_c defined by

$$\frac{T}{T_0} \ll \frac{T_c}{T_0} \equiv \frac{1}{2} \frac{\langle \beta_0 F^2 \rangle}{\langle \beta_0 \partial_x F \rangle}. \quad (11)$$

For our system, $T_c/T_0 \approx 0.08$, which is shown as a vertical dashed line in Fig. 4(b); however, even in the parameter regime where both $\Delta t_{\text{tot}} \gg \Delta t_c$ and Eq. (11) are fulfilled, i.e., in the upper left corner of Fig. 4(b), the FW and OM

instantons are distinct. This is because for long enough Δt_{tot} , the OM instanton rests close to one of the minima of the potential-energy landscape, as observed in Fig. 4(a). That staying close to a potential minimum leads to a particularly small exit rate (and hence, large sojourn probability) is intuitively reasonable, as the positive curvature of the potential around the minimum hinders particle exit from the tube. Now if the trajectory rests close to a potential minimum, then both the velocity and the force are very small, and hence, the second term in Eq. (5) may not be negligible as compared to the first term even for rather low temperatures. Thus, even at low temperature the FW action is not expected to produce the OM instanton whenever the latter yields a path that temporarily rests close to a state where the deterministic force vanishes. In summary, Fig. 4(b) shows that while the FW instanton agrees with the OM instanton for short transitions, as quantified by Eqs. (9) and (10), a similar scaling argument for long transitions given by Eq. (11) fails. This is because even at the lowest temperature considered here, the second term in the OM action remains relevant for the instanton.

III. RELATION BETWEEN THE ASYMPTOTIC SOJOURN PROBABILITY AND THE PATH-INTEGRAL ACTION

In this section, we show that the asymptotic-tube action defined by Eq. (1) is consistent with the path-integral formalism, in the sense that both approaches lead to identical probability ratios for twice-differentiable paths.

By formally writing the sojourn probabilities on the right-hand side of Eq. (1) as path integrals, the equation reads

$$\frac{e^{-S[\varphi]}}{e^{-S[\psi]}} = \lim_{R \rightarrow 0} \frac{\int \mathcal{D}[X] \mathbb{1}_R^\varphi[X] e^{-S[X]}}{\int \mathcal{D}[X] \mathbb{1}_R^\psi[X] e^{-S[X]}}, \quad (12)$$

where S on the left-hand side denotes the OM action defined by Eqs. (2) and (5), $\mathcal{D}[X]$ stands for a fictitious uniform measure on the space of all continuous paths [28], $e^{-S[X]}$ is the corresponding path weight which defines the formal path-integral action \mathcal{S} , $\mathbb{1}_R^\varphi[X]$ denotes the indicator function on the set of all paths that remain within a tube of radius R around the reference path φ , and for each path integral, we consider the initial point of the respective reference path as an initial condition for the corresponding stochastic paths. To give operational meaning to the symbolic path-integral expressions on the right-hand side of Eq. (12), we interpret the expression as the continuum limit of the standard time-slicing procedure [22,30–37,42,46,47], as we discuss in more detail below.

The limiting ratio of sojourn probabilities Eq. (1) defines the action S on the left-hand side of Eq. (12); the path-integral action \mathcal{S} in the integrand on the right-hand side of the equation is defined as a formal limit of a time-slicing

procedure. From Eq. (12), one would assume that these two actions are equal: As R on the right-hand side of the equation tends to zero, only the immediate neighborhood of the reference path remains within the tube. One would thus expect that the right-hand side tends to $e^{-S[\phi]}/e^{-S[\psi]}$, which is equal to the left-hand side if $S = \mathcal{S}$, up to a possible path-independent additive contribution. However, the formal path-integral action on the right-hand side of Eq. (12) is not uniquely defined; in fact, infinitely many equivalent stochastic actions can be employed, including the OM and FW actions Eqs. (5) and (6) [31,35,36,42,46]. This indeterminacy is in sharp contrast to the limiting ratio of tube probabilities at a finite temperature, which is described by the OM action Eq. (5), as given on the left-hand side of Eq. (12) [23–25,27–29].

To show how these two seemingly inconsistent mathematical results are reconciled, we now consider the usual time-slicing derivation of the path-integral action [22,32,34–37,42,46]. The key insight in the following discussion is that the continuum limit of the discretized path-integral action depends on whether the action is evaluated on a typical realization of the Langevin equation or on a differentiable path. The ambiguity in the stochastic action discussed in the literature [31,35,36,42,46] emerges only in the former case; in the latter case, which is the appropriate scenario for comparing the time-slicing action to the limiting ratio Eq. (1), the OM action follows unambiguously, i.e., independent of the discretization scheme. The relation between the action of rough and smooth paths has been discussed from another perspective in previous works [30,32].

For a one-dimensional stochastic process X_t , we consider the Itô-Langevin equation

$$dX_t = a(X_t)dt + b dB_t, \quad (13)$$

where $a(x) = D_0\beta_0 F(x)$ is the drift, dB_t is the increment of the Wiener process, and the noise strength $b = \sqrt{2D_0}$ is independent of position. Since the noise is additive, the following derivation is independent of whether one interprets Eq. (13) in the Itô or Stratonovich sense [1,48].

We consider a given continuous path Y_t , $t \in [0, t_f]$, and aim to quantify its probability according to the Itô-Langevin Eq. (13). We begin by discretizing the time interval $[0, t_f]$ into N equally large slices of duration $\Delta t = t_f/N$, and denote the position of the path at time $t_i \equiv i\Delta t$ by $Y_i \equiv Y_{t_i}$. Using the Markov property, we rewrite the joint probability density that a realization of the Langevin Eq. (13), after starting at Y_0 at time t_0 , is at the points Y_i at times t_i , as

$$P(Y_N, t_N; Y_{N-1}, t_{N-1}; \dots; Y_1, t_1 | Y_0, t_0) \quad (14)$$

$$= \prod_{i=0}^{N-1} P(Y_{i+1}, t_{i+1} | Y_i, t_i) \quad (15)$$

$$\equiv \mathcal{N}_N \exp \left[- \sum_{i=0}^{N-1} \Delta \mathcal{S}_i \right], \quad (16)$$

where we define the normalization constant $\mathcal{N}_N \equiv (2\pi b^2 \Delta t)^{-N/2}$. We now derive an expression for the discretized action $\sum_i \Delta \mathcal{S}_i \equiv \sum_i \Delta \mathcal{S}(Y_{i+1}, Y_i, \Delta t)$, which is defined by Eq. (16), by calculating the short-time propagator of the Langevin dynamics. Equation (16) is the usual starting point for the time-slicing derivation of the path integral [22,32,34–37,42,46].

At this point, we do not require that the given continuous path Y_t be a realization of the Langevin equation. We assume only that the increments $\Delta Y_i \equiv Y_{i+1} - Y_i$ fulfill $\Delta Y_i = \mathcal{O}(\Delta t^{1/2})$ for small Δt . This condition holds if Y_t is a typical realization of the Langevin equation, but it is also true if Y_t is a continuously differentiable path, for which in fact the stronger condition $\Delta Y_i = \mathcal{O}(\Delta t)$ holds.

For the Itô-Langevin Eq. (13), the increment $\Delta X \equiv X_{\Delta t} - X_0$ for a short time interval Δt , and with initial condition X_0 , follows via an Itô-Taylor expansion as [48],

$$\Delta X = \Delta t^{1/2} b \Delta W + \Delta t a(X_0) + \Delta t^{3/2} a'(X_0) b \Delta Z + \mathcal{O}(\Delta t^2), \quad (17)$$

where a' denotes the spatial derivative of the drift a , and where $(\Delta W, \Delta Z)$ are distributed according to a two-dimensional Gaussian distribution [48]

$$P(\Delta W, \Delta Z) = \frac{1}{2\pi \sqrt{\det \Sigma}} e^{-\frac{1}{2}(\Delta W, \Delta Z) \Sigma^{-1} (\Delta W, \Delta Z)^T}, \quad (18)$$

with

$$\Sigma = \begin{pmatrix} 1 & 1/2 \\ 1/2 & 1/3 \end{pmatrix}. \quad (19)$$

By $\mathcal{O}(\Delta t^2)$, in Eq. (17) we subsume both random and deterministic terms that scale at least as Δt^2 for small Δt .

The probability density for observing the i th increment ΔY_i of the given continuous path Y_t in a realization of the Langevin dynamics is now obtained as [3]

$$\begin{aligned} P(Y_{i+1}, t_{i+1} | Y_i, t_i) \\ \equiv \int d\Delta W \int d\Delta Z [\delta(\Delta Y_i - \Delta X(\Delta W, \Delta Z)) P(\Delta W, \Delta Z)], \end{aligned} \quad (20)$$

where the increment as a function of the noise $\Delta X(\Delta W, \Delta Z)$ is given by Eq. (17), with the initial condition $X_0 \equiv Y_i$, and δ is the Dirac-delta distribution. The delta distribution can be used to directly evaluate the integral over ΔW ; the remaining integral over ΔZ in Eq. (20) is a Gaussian integral, which evaluates to

$$P(Y_{i+1}, t_{i+1} | Y_i, t_i) = \frac{1}{\sqrt{2\pi b^2 \Delta t}} e^{-\Delta \mathcal{S}_i}, \quad (21)$$

where

$$\Delta \mathcal{S}_i \equiv \Delta \mathcal{S}(Y_{i+1}, Y_i, \Delta t) \quad (22)$$

$$\begin{aligned} &\equiv \frac{\Delta t}{2b^2} \left(\frac{\Delta Y_i}{\Delta t} - a(Y_i) \right)^2 - \frac{a'(Y_i)}{2b^2} (\Delta Y_i^2 - b^2 \Delta t) \\ &\quad + \mathcal{O}(\Delta t^{3/2}). \end{aligned} \quad (23)$$

From the appearance of $a'(Y_i)$ in Eq. (23), it is apparent why we consider Eq. (17) beyond linear order in the time increment: The term of order $\Delta t^{3/2}$ in the discretized Langevin equation in fact contributes a term of order Δt in the short time propagator.

In Eq. (23), the drift a and its derivative are evaluated at the initial point of each time slice. Equivalently, we can use any other point in the interval $[Y_i, Y_{i+1}]$, and we parametrize the choice by a parameter $\xi \in [0, 1]$ as $\bar{Y}_i^\xi \equiv Y_i + \xi(Y_{i+1} - Y_i)$. Taylor expanding a, a' around \bar{Y}_i^ξ , substituting the result into Eq. (22), and rearranging, we obtain

$$\begin{aligned} \Delta \mathcal{S}_i &\equiv \frac{\Delta t}{2b^2} \left(\frac{\Delta Y_i}{\Delta t} - a(\bar{Y}_i^\xi) \right)^2 \\ &\quad - \frac{a'(\bar{Y}_i^\xi)}{2b^2} [(1-2\xi)\Delta Y_i^2 - b^2 \Delta t] + \mathcal{O}(\Delta t^{3/2}), \end{aligned} \quad (24)$$

where we use that $\Delta Y_i = \mathcal{O}(\Delta t^{1/2})$. This equation is, to order Δt , equivalent to Eq. (22), which is recovered from Eq. (24) by Taylor expanding a, a' around Y_i again. Therefore, despite the explicit appearance of ξ in Eq. (24), the expression is to order Δt independent of this parameter [31,35,36,42].

Equations (16) and (24) describe an N -dimensional probability density, which can be evaluated on any continuous path that obeys $\Delta Y_i = \mathcal{O}(\Delta t^{1/2})$. Since the sum over the $\Delta \mathcal{S}_i$ in Eq. (16) has N terms, the sum over the terms of order $\Delta t^{3/2} \sim N^{-3/2}$ from Eq. (24) is of order $\mathcal{O}(N^{-1/2})$. Thus, while the parameter ξ appears explicitly in Eq. (24), the discretized action in the probability density Eq. (14) depends on ξ only to subleading order $\mathcal{O}(N^{-1/2})$, which becomes irrelevant in the limit $N \rightarrow \infty$. Equations (16) and (24) thus constitute a one-parameter family parametrized by $\xi \in [0, 1]$ of asymptotically equivalent expressions for the N -point probability density evaluated on a given continuous path that obeys $\Delta Y_i = \mathcal{O}(\Delta t^{1/2})$ [31,35,36,42].

We first consider the formal continuum limit $N \rightarrow \infty$ of the exponent in Eq. (16) for the case where Y_t is a realization of the Langevin Eq. (13), which we denote by $Y_t \equiv X_t$. For a realization of the Langevin equation, the formal continuum limit of the action defined by Eqs. (16) and (24) follows as

$$\mathcal{S}[X] \equiv \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} \Delta \mathcal{S}(X_{i+1}, X_i, \Delta t) \quad (25)$$

$$= \int_0^{t_f} dt \left[\frac{1}{2b^2} (\dot{X}_t - a(X_t))^2 + \xi a'(X_t) \right]; \quad (26)$$

see Appendix F for a derivation of the second term in Eq. (26) from Eq. (24). For special choices of ξ , Eq. (26) yields the Freidlin-Wentzell ($\xi = 0$), Onsager-Machlup ($\xi = 1/2$), or Hänggi-Klimontovich ($\xi = 1$) action. That the formal continuum limit Eq. (26) depends on the parameter ξ is precisely the well-studied fact that the formal path-integral action is not uniquely defined [31,35,36,42]. In practice, this indeterminacy is not an issue: As it stands, the formal expression Eq. (26) is not mathematically well defined. More explicitly, for a realization of the Langevin equation, it holds that

$$\frac{1}{2b^2} \int dt \dot{X}_t^2 \equiv \frac{1}{2b^2} \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} \Delta t \frac{\Delta X_i^2}{\Delta t^2} \quad (27)$$

$$= \frac{1}{2} \lim_{N \rightarrow \infty} \left[\sum_{i=0}^{N-1} \Delta W_i^2 + \mathcal{O}(N^{-1/2}) \right], \quad (28)$$

where we use $\Delta X_i^2 / \Delta t = b^2 \Delta W_i^2 + \mathcal{O}(\Delta t^{3/2}) = b^2 \Delta W_i^2 + \mathcal{O}(N^{-3/2})$, which follows from Eq. (17). Because $\langle \Delta W_i^2 \rangle = 1$ and all the noise increments ΔW_i are independent, the expected value of the limit Eq. (28) is infinite. Consequently, Eq. (26) should be interpreted in its discretized form Eq. (24), which is to leading order independent of ξ [31,35,36,42]. Furthermore, as we detail in Appendix F, if the diverging square term Eq. (27) is subtracted from the discretized action before taking the limit $N \rightarrow \infty$, a well-defined expression is obtained, which is given by the Girsanov formula [2,49,50] and independent of the value of ξ used in Eq. (24).

Second, we consider the continuum limit $N \rightarrow \infty$ of the exponent in Eq. (16) for the case where Y_t is a continuously differentiable path with a square-integrable derivative, which we denote by $Y_t \equiv \varphi_t$. While any such path occurs with probability zero as the realization of the Langevin equation, it is of course possible to evaluate the N -point probability density Eq. (14) on a given set of positions parametrized by φ_t . For a continuously differentiable path, we have $\Delta \varphi_i = \mathcal{O}(\Delta t)$, so that $\Delta Y_i^2 \equiv \Delta \varphi_i^2 = \mathcal{O}(\Delta t^2)$, and hence, the continuum limit of Eqs. (16) and (24) follows as

$$\mathcal{S}[\varphi] \equiv \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} \Delta \mathcal{S}(\varphi_{i+1}, \varphi_i, \Delta t) \quad (29)$$

$$= \int_0^{t_f} dt \left[\frac{1}{2b^2} (\dot{\varphi}_t - a(\varphi_t))^2 + \frac{1}{2} a'(\varphi_t) \right]. \quad (30)$$

In contrast to Eq. (25), this limit is independent of the choice of ξ and always given by the OM action Eqs. (2) and (5) [30], which is seen by substituting $a(x) = F(x)/\gamma$, $b = \sqrt{2k_B T_0/\gamma}$ into Eq. (30). Furthermore, because the path φ_t is continuously differentiable with square-integrable derivative, the integral over $\dot{\varphi}_t^2$ is finite, and hence, the functional Eq. (30) is well defined.

The two continuum limits Eqs. (26) and (30) show how it is mathematically consistent that the formal path-integral action is not unique [31,35,36,42], whereas the limiting ratio Eq. (1) is [23–25,27–29]. More explicitly, the ambiguity in the continuum limit of the discretized action Eqs. (16) and (24) is observed only if the discretized action is evaluated on a typical realization of the Langevin Eq. (13), in which case, the formal expression Eq. (26) is obtained. However, in Eq. (1), we consider two twice-differentiable reference paths, for which the limiting ratio of N -point probability densities follows from Eq. (30) unambiguously as

$$\lim_{N \rightarrow \infty} \frac{P(\varphi_N, t_N; \varphi_{N-1}, t_{N-1}; \dots; \varphi_1, t_1 | \varphi_0, t_0)}{P(\psi_N, t_N; \psi_{N-1}, t_{N-1}; \dots; \psi_1, t_1 | \psi_0, t_0)} = \frac{e^{-S[\varphi]}}{e^{-S[\psi]}}, \quad (31)$$

which is identical to the limiting ratio of tube probabilities Eq. (1) [23–25,27–29].

IV. DISCUSSION

A. Summary of results

In this work, we establish a protocol to determine the ratios of path probabilities from a measured time series, without fitting a stochastic model to the data. Applying this protocol to the time series of a colloidal particle in a microchannel, we find that the Onsager-Machlup action Lagrangian [22–25] describes both ratios of path probabilities and the most probable path extracted from our experimental data, validating classical theoretical results [23–25,27–29]. The Freidlin-Wentzell action [38–40] disagrees with our finite-temperature experimental results, and we quantify for which parameters it is expected to predict the correct most probable path. By careful analysis of the continuum limit of the usual time-slicing approach to the path-integral formalism, we resolve the apparent inconsistency that the formal path-integral action is not unique, whereas both theoretical and experimental results point to a definitive probability ratio for differentiable paths.

Our results constitute a model-free experimental measurement of relative likelihoods of stochastic trajectories and demonstrate that from a physical point of view there is no ambiguity as to which stochastic action describes observable relative path probabilities, as defined in Eq. (1).

B. Probability vs probability density

The key idea in our approach is to measure finite-radius sojourn probabilities and to extrapolate the results to the vanishing-radius limit using Eq. (1). Our measurements hence focus on *probabilities*, as opposed to *probability densities*, which is because only nonzero probabilities, such as the sojourn probability at finite radius, can be directly measured in an experiment.

For overdamped Langevin dynamics with additive noise, the sojourn probability for small-but-finite radius can be calculated as [23,27,41]

$$P_R^\varphi(t_f) = P_R^{\text{W}}(t_f)[e^{-S[\varphi]} + \mathcal{O}(R)], \quad (32)$$

where $S[\varphi]$ is the OM action defined by Eqs. (2) and (5), and where $P_R^{\text{W}}(t_f)$ is the probability that a rescaled Wiener process starting at the origin and with the same increment variance as the random force in the Langevin dynamics remains within a tube of radius R around the origin until the final time t_f . Substituting the asymptotic sojourn probability Eq. (32) for both φ, ψ in Eq. (1), and noting that $P_R^{\text{W}}(t_f)$ is for additive noise independent of the reference path [23,27,41], the left-hand side of Eq. (1) is obtained. In that sense, our experiments directly probe the measure induced on the space of all continuous paths by the observed stochastic dynamics. For overdamped Langevin dynamics, Eq. (1) is straightforwardly generalized to higher dimensions, where a tube consists of a ball of radius R around the reference path [23,27,41].

Following Ref. [24], Eq. (32) can be compared to the probability that a one-dimensional real random variable X is within a small interval $I_R^x \equiv [x - R, x + R]$, $P(X \in I_R^x)$. If the probability measure associated with X has a density $\rho(x)$ with respect to the usual Lebesgue measure λ , then the probability can be written as

$$P(X \in I_R^x) = \lambda(I_R^x)[\rho(x) + \mathcal{O}(R)], \quad (33)$$

where, due to the translation invariance of the Lebesgue measure, $\lambda(I_R^x) = |[x - R, x + R]| = 2R$ is independent of x .

Comparing Eqs. (32) and (33), one might be inclined to think of $e^{-S[\varphi]}$ as a (unnormalized) probability density with respect to the Wiener measure. It is, however, apparent that this cannot be true [24], because the OM functional is not even well defined for a typical realization of the Langevin equation; cf. Eq. (28). While the measure induced on the space of all continuous paths by overdamped additive-noise Langevin dynamics does have a density with respect to the Wiener measure, this density is not described by the OM functional, but rather by the Girsanov formula [2,49,50]; cf. Appendix F.

C. Extension to other types of stochastic dynamics

In general, for any reaction coordinate X_t , independent of the precise nature of its stochastic dynamics, the probability to remain within an asymptotically small tube is the experimentally relevant characterization of its path probabilities. After all, it is precisely the realizations of X_t that one observes, and the asymptotic sojourn probability very concretely characterizes how these realizations behave. This means that the sojourn probability is also a physically relevant observable for other variants of Langevin dynamics, which are typically obtained from projecting high-dimensional dynamics onto a low-dimensional reaction coordinate [51,52]. If there is timescale separation between fast orthogonal degrees of freedom and the reaction coordinate, then such a projection leads to a Langevin equation with multiplicative noise [53], i.e., configuration-dependent diffusivities. If there is no timescale separation, then the reaction coordinate is described by non-Markovian Langevin dynamics. Physical examples where memory effects are relevant include conformational transitions in small molecules [54–56], colloidal particles in solution on very short timescales [57–59], or the motion of cells [60,61].

Whether a limiting process as described in Eq. (1) leads to a finite result for an arbitrary reaction coordinate X_t depends on the details of its stochastic dynamics. For example, for overdamped Langevin dynamics with multiplicative noise, Dürr and Bach have shown that the limit in general does not exist [24]. This is conceptually similar to the fact that, also on finite-dimensional spaces, not every physically relevant probability measure has a probability *density* with respect to the Lebesgue measure. However, even if the limiting ratio of sojourn probabilities is not finite, which means that one asymptotically small tube is infinitely more likely than the other, the sojourn probability for small-but-finite radius is still an experimentally accessible quantity that describes the behavior of a reaction coordinate [62].

Theoretically calculating the finite-radius sojourn probability for a given stochastic dynamics, reference path, and small-but-finite tube radius is a conceptually straightforward mathematical task: This is an absorbing-boundary problem with moving boundaries for the reaction coordinate X_t [41,62]. Whether absorbing boundaries should also be introduced for orthogonal degrees of freedom, such as the velocity in the case of inertial Langevin dynamics, or more generally, memory degrees of freedom, will in general depend on the physical question under investigation.

From an experimental point of view, measuring the sojourn probability for a finite-radius tube is also a well-defined problem. As we do in this work, one needs to count which fraction of the recorded trajectories (all of which start inside the tube) leave the tubular neighborhood around the reference path for the first time at each recorded time.

One key aspect of our algorithm is how we overcome the exponential temporal decay of the number of sample trajectories that have never left the tube: We periodically increase the number of sample trajectories at each multiple of ΔT by drawing new measured trajectories with initial conditions according to the instantaneous distribution of the current trajectories inside the tube. As it assumes only Markovianity, this algorithm is directly applicable also for overdamped Langevin dynamics with multiplicative noise. If the dynamics of the reaction coordinate is non-Markovian, our algorithm needs to be amended. A possible extension would be to not only use the instantaneous distribution of trajectories at multiples of ΔT to draw new sample trajectories. Instead, at iteration k , a multiple-time distribution could be considered, which includes the positions of trajectories at several times between $k\Delta T - \tau$ and $k\Delta T$, where τ denotes the longest memory timescale of the system. A newly drawn trajectory would then need to approximately share the same recent history with a current trajectory so that the two might be considered at approximately the same non-Markovian state.

The dynamics of the system we consider in the present work is time homogeneous, meaning that both the external force and the strength of the random force are independent of time. While theoretically calculating the finite-radius sojourn probability for a time-dependent force protocol is a conceptually clear task [41,62], extending our data analysis to time-dependent forces means that more data need to be collected. More explicitly, in concatenating measured trajectories, as illustrated in Fig. 2, one needs to make sure that the external and random forces consistently follow the time-dependent protocol. In practice, this can be achieved by measuring a large number of independent experimental trajectories, each with the given protocol, and considering only the value of each trajectory at the current state of the protocol in the cloning algorithm. In contrast, in the algorithm in Appendix B, the time t that appears in the sojourn probability does not need to agree with the physical time that has passed since the recording of a trajectory started.

Both the extrapolated exit rate and the theoretical model used for predictions in Fig. 3(c) are based on measurements pertaining to comparable length- and timescales, and on which the motion of the colloidal particle is well described by an overdamped Langevin model. This observational scale set by the experimental apparatus explains why even in the limit of asymptotically small tube radius, inertial effects do not play a role in our results. If one were able to experimentally observe the dynamics in the colloidal-particle system shown in Fig. 1(b) with arbitrary precision, one would inevitably observe inertial effects or noise correlations on some small scale [57,63]. However, any measurement can of course only directly probe the physics down to scales that are resolvable by the measurement

apparatus. This means that, very generally, a limiting ratio for the sojourn probability obtained directly from measured data should always be compared to a theory that describes the dynamics on the scale used for measuring the finite-radius exit rate. That any measurement fixes a modeling scale is of course not a deficiency of our particular approach, but a fundamental property of the physical sciences.

D. Entropy production and further applications

Equations (3) and (7) relate path properties to experiment by formulating them as vanishing-radius limits of tube properties. This is a general strategy to relate questions on individual paths to measurement: If a given single-path statement can be reformulated in terms of a limit of tubes, the resulting finite-radius expression is accessible in measurement. The corresponding finite-radius observables can then be either investigated directly or used to infer a vanishing-radius limit. This method to experimentally access single-path properties will be particularly important for the field of stochastic thermodynamics, which extensively employs the concept of individual trajectories [4,10]. For example, entropy production is in this context typically quantified by the probability ratio of forward and backward paths [10], which can be accessed experimentally by considering the ratios of sojourn probabilities for forward and backward finite-radius tubes [64].

In this study, we probe the probability distribution on the space of all trajectories at a finite temperature. However, our measurement algorithm is directly applicable also in the low-noise limit, as considered in the Freidlin-Wentzell theory [38,39], or in the study of switching transitions, where the full probability distribution in path space induced by Langevin dynamics may be very concentrated around a single path [17,18,65–67]. One concrete example for possible future research on switching transitions is conformational changes of small molecules [54,56]. For such molecules, measuring sojourn probabilities of dihedral angle reaction coordinates will allow us to quantitatively study transitions between metastable states, without the need to fit a stochastic model to the data.

In Sec. III, we show that the time-slicing approach to path integrals [31,35,36,42], is consistent with the mathematical results on the asymptotic ratio of sojourn probabilities Eq. (1) [23–25,27–29,41], in the sense that both unambiguously lead to the Onsager-Machlup action. From a theoretical perspective, it will be interesting to explore the relation between these two approaches to the stochastic action beyond the single-path limit of vanishing tube radius. More explicitly, as formulated in Eq. (12), one can think of the finite-radius sojourn probability $P_R^o(t_f)$ for a tube of radius R around a reference path φ as a path integral over an indicator function [68,69]. It will be interesting to see whether the theoretical results like the finite-radius sojourn probability [41] can also be calculated

directly as path-integral averages within the time-slicing path-integral formalism.

To summarize again, with Eq. (3) we provide an intuitive and experimentally useful relation between observable exit rates and the stochastic action Lagrangian. We use this relation to show experimentally that the Onsager-Machlup Lagrangian characterizes physical ratios of path probabilities. More generally, our work demonstrates that the asymptotic sojourn probability provides a direct and experimentally accessible characterization of path properties in stochastic dynamics. We believe that both the theoretical and experimental study of this observable will be valuable for relating theoretical single-trajectory results to measurement in stochastic dynamics.

ACKNOWLEDGMENTS

We thank Prof. Mike Cates, Dr. Yongjoo Baek, Dr. Jules Guioth, Dr. Rob Jack, Dr. Patrick Pietzonka, and MAST Lukas Kikuchi for stimulating discussions. In particular, we thank Dr. Yongjoo Baek for pointing out that the results of Refs. [35,42] can be formulated elegantly via Eq. (24). Work was funded in part by the European Research Council under the EU's Horizon 2020 Program, Grant No. 740269, and by an Early Career Grant to R. A. from the Isaac Newton Trust. J. G. and U. F. K. were supported by the European Union's Horizon 2020 research and innovation program under European Training Network (ETN) Grant No. 674979-NANOTRANS. U. F. K. acknowledges funding from an European Research Council Consolidator Grant (DesignerPores 647144).

APPENDIX A: PARAMETRIZING THE OVERDAMPED LANGEVIN EQUATION

We consider the overdamped one-dimensional Itô-Langevin equation

$$dX_t = D\beta F(X_t)dt + \sqrt{2D}dB_t \quad (\text{A1})$$

with D the diffusivity, $\beta^{-1} = k_B T$ the thermal energy with k_B the Boltzmann constant and T the absolute temperature, $F(x) = -\nabla U$ an external force with a potential U , and dB_t the increment of the Wiener process. Equation (A1) is identical to Eq. (13) with $a(x) = D\beta F(x)$, $b = \sqrt{2D}$, $T = T_0$. As we explain in the following, we parametrize Eq. (A1) by locally calculating the first two Kramers-Moyal coefficients based on the experimental time series. While this parametrization allows for a position-dependent diffusivity $D(x)$, we see below that for our experimental system the diffusivity is well approximated by a spatially constant diffusivity. The approximate position-independence of the diffusivity in particular implies that, while we use the Itô interpretation for Eq. (A1), the choice of stochastic integral does not lead to any ambiguity in our results because for constant diffusivity D , the Itô and

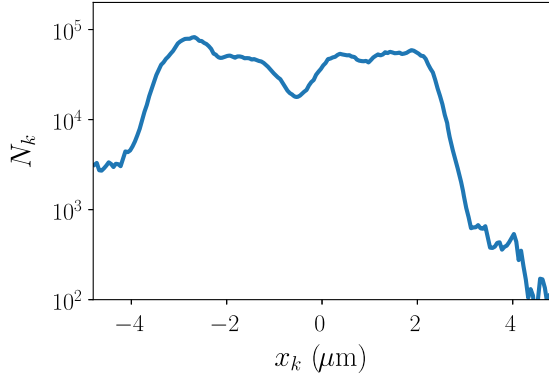


FIG. 5. Number of experimental data points per discretization bin. The solid line denotes the number of experimental data points per bin, as defined in Eq. (A5). The bin center x_k is the center of the bin B_k defined in Eq. (A3).

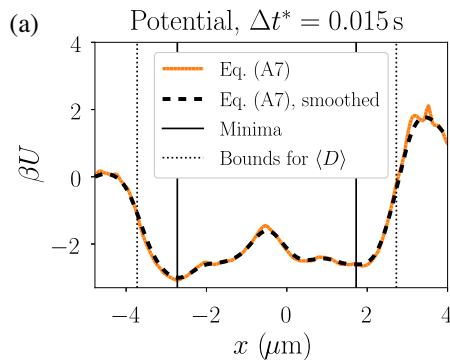
Stratonovich interpretations of Eq. (A1) are equivalent [1,48].

From 104 minutes of experimental measurements, we obtain $N = 230$ uncorrelated discrete time series

$$X_i(t_j) \equiv X_i(j\Delta t) \equiv X_{ij}, \quad (\text{A2})$$

where $i \in I = \{1, \dots, N\}$ labels the time series, and the maximal time index $j \in \{0, \dots, J_i\}$ for each time series depends on i , meaning the recorded time series are of variable length. All time series have identical time step $\Delta t = 0.001$ s; the lengths of the time series range from 10 to 60 s. We divide space into bins of width $\Delta x = 0.05$ μm , with the k th bin

$$B_k = [\hat{x}_L + k\Delta x, \hat{x}_L + (k+1)\Delta x), \quad (\text{A3})$$



where for the left boundary we use $\hat{x}_L = -4.8$ μm , and $k \in \{0, \dots, K\}$ with $K = 192$, so that $\hat{x}_R \equiv \hat{x}_L + K\Delta x = 4.8$ μm . The center of the k th bin denoted by x_k is located at $x_k \equiv \hat{x}_L + (k + 1/2)\Delta x$. The positions \hat{x}_L, \hat{x}_R are still well within the experimental microchannel, meaning that a colloid starting at \hat{x}_L, \hat{x}_R is very unlikely to leave the channel within 1 second. For every bin B_k , we create a list of all the experimentally recorded tuples (i, j) such that $X_{ij} \in B_k$; i.e., we for every k construct the set

$$B_k = \{(i, j) | X_{ij} \in B_k\}. \quad (\text{A4})$$

We denote the total number of data points in bin B_k by

$$N_k \equiv |\{X_{ij} \in B_k\}| \equiv |B_k|, \quad (\text{A5})$$

and show a plot of N_k as a function of the bin center x_k in Fig. 5.

To parametrize the overdamped Langevin Eq. (A1), we locally estimate both the diffusivity and the force via discretized Kramers-Moyal coefficients [1]. At the bin centered on x_k , we obtain

$$D(x_k) = \frac{1}{2\Delta t^*} [\langle \Delta X^2(\Delta t^*) \rangle_k - \langle \Delta X(\Delta t^*) \rangle_k^2], \quad (\text{A6})$$

$$\beta F(x_k) = \frac{\langle \Delta X(\Delta t^*) \rangle_k}{D(x_k)\Delta t^*}, \quad (\text{A7})$$

where the symbol $\langle \bullet \rangle_k$ denotes the average over all N_k experimental time series which start in the bin B_k . In the evaluation of Eqs. (A6) and (A7), we furthermore use the lag time $\Delta t^* = 15\Delta t = 0.015$ s; a discussion of the dependence of our results on lag time is given further below. From the force, we calculate a potential as

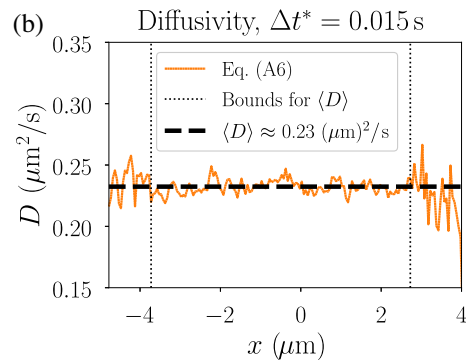


FIG. 6. Local potential and diffusivity extracted from experimental time series. (a) The orange solid line depicts the potential energy as obtained from Eqs. (A7) and (A8) for $\Delta t^* = 0.015$ s. The black dashed solid line is a smoothed version of the orange line obtained via a Hann-window average using 20 data points at each point x_k . The vertical solid lines denote local minima $x_0^{\min} \approx -2.725$ μm , $x_1^{\min} \approx 1.725$ μm of the smoothed potential energy. The vertical dashed lines indicate the bounds of the interval $[x_0^{\min} - 1$ $\mu\text{m}, x_1^{\min} + 1$ $\mu\text{m}]$ over which the average diffusivity $\langle D \rangle$ is calculated in (b). (b) The orange line shows the local diffusivity as obtained from Eq. (A6). The horizontal dashed line depicts the average over the diffusivity inside the interval $[x_0^{\min} - 1$ $\mu\text{m}, x_1^{\min} + 1$ $\mu\text{m}]$, as indicated by the two vertical lines; cf. (a).

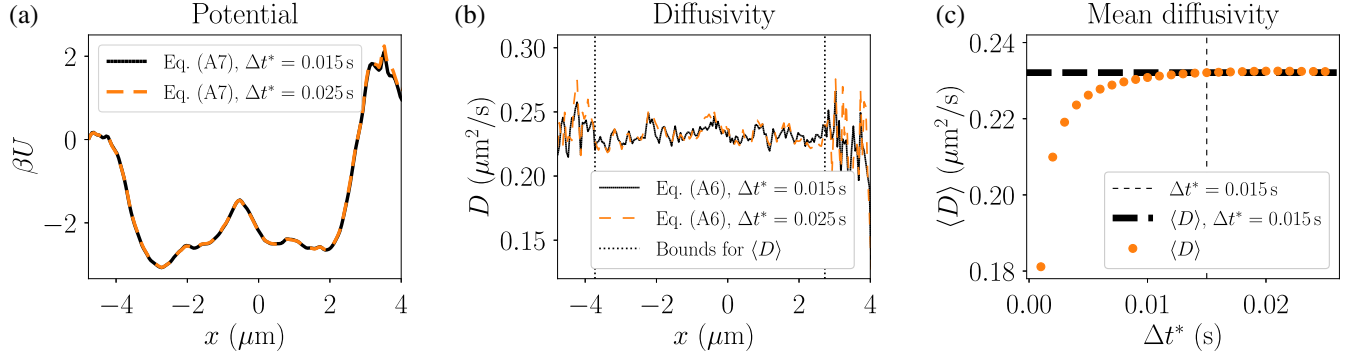


FIG. 7. Lag-time dependence of potential and diffusivity extracted from experiments. (a) The potential energy as obtained from Eqs. (A7) and (A8) is shown for $\Delta t^* = 0.015$ s as a black solid line and for $\Delta t^* = 0.025$ s as an orange dashed line. (b) The diffusivity as obtained from Eq. (A6) is shown for $\Delta t^* = 0.015$ s as a black solid line and for $\Delta t^* = 0.025$ s as an orange dashed line. The vertical dashed lines depict the boundary of the interval $[x_0^{\min} - 1 \mu\text{m}, x_1^{\min} + 1 \mu\text{m}]$, where $x_0^{\min} \approx -2.725 \mu\text{m}$, $x_1^{\min} \approx 1.725 \mu\text{m}$ are two local minima of the potential energy; cf. Fig. 6. (c) Mean diffusivity $\langle D \rangle$ averaged over $[x_0^{\min} - 1 \mu\text{m}, x_1^{\min} + 1 \mu\text{m}]$ as a function of the lag time Δt^* used in Eq. (A6). The horizontal thick dashed line denotes the value $\langle D \rangle$ for $\Delta t^* = 0.015$ s, as indicated by the vertical dashed line.

$$\beta U(x_k) = - \int_{\hat{x}_L}^{x_k} dx' \beta F(x'), \quad (\text{A8})$$

where we use the trapezoidal rule to perform the integral on the right-hand side numerically; the result of this integration is furthermore smoothed using a Hann window that at each x_k incorporates the 20 closest data points. The smoothed potential is then interpolated using polynomial splines of degree 3; this polynomial interpolation is used in evaluations of the stochastic action to calculate the force F and its derivative $\partial_x F$.

The diffusivity and potential-energy profiles obtained from Eqs. (A6) and (A7) are shown in Fig. 6. The potential energy in Fig. 6(a) shows two local minima at $x_0^{\min} \approx -2.725 \mu\text{m}$, $x_1^{\min} \approx 1.725 \mu\text{m}$ separated by a barrier at $x \approx -0.5 \mu\text{m}$. Note that in the main text, a constant is added to the potential, such that the potential vanishes at x_0^{\min} . From Fig. 6(b), we conclude that the diffusivity is almost independent of position within the interval $[x_0^{\min} - 1 \mu\text{m}, x_1^{\min} + 1 \mu\text{m}]$, with an average value

$$\langle D \rangle \approx 0.232 \frac{(\mu\text{m})^2}{\text{s}}. \quad (\text{A9})$$

The dependence of the inferred potential and diffusivity Eqs. (A6) and (A7) on the lag time Δt^* is shown in Fig. 7. Figures 7(a) and 7(b) show that both the potential and the diffusivity for the two lag times $\Delta t^* = 0.015, 0.025$ s agree with each other. Figure 7(c) shows the average diffusivity $\langle D \rangle$ as a function of the lag time Δt^* . For short lag times $\Delta t^* \lesssim 0.01$ s, the mean diffusivity slightly depends on the lag time (note the scaling on the y axis), which we attribute to inaccuracies of the centroid algorithm which we use to estimate colloidal positions. For lag times $\Delta t^* \gtrsim 0.01$ s, the

mean diffusivity is independent of the lag time, which justifies our choice $\Delta t^* = 0.015$ s.

Using the Einstein relation, the friction coefficient γ follows from Eq. (A9) as

$$\gamma = \frac{k_B T}{\langle D \rangle} \approx 1.75 \times 10^{-8} \frac{\text{kg}}{\text{s}}, \quad (\text{A10})$$

where k_B is the Boltzmann constant and $T = 294$ K is the experimental temperature.

Figures 6 and 7 demonstrate that, on the millisecond timescale, the dynamics of the colloidal particle along the channel axis is approximately Markovian and well described by an overdamped Langevin equation with additive noise. This fact in particular implies that both hydrodynamic interactions with the channel walls, as well as temporal noise correlations, are irrelevant on the scales we consider. We now briefly discuss that this is consistent with other experimental studies involving colloidal particles.

For colloidal particles in bulk water, deviations from white-noise behavior of the thermal force in fluids have been reported and characterized by Franosch *et al.* [57]. While the hydrodynamic-memory timescale τ_f is of the order of $1 \mu\text{s}$ [57,63], hydrodynamic effects can be observed on much larger timescales because hydrodynamic noise correlations decay with a power-law tail. More explicitly, in Ref. [57] it was observed that colored deviations from white noise start to become relevant in bulk water on timescales slightly below 1 ms for beads of size 2–3 μm . For smaller colloids with diameter 500 nm, as we use in our present work, we expect the onset of colored noise, and hence, of hydrodynamic effects, to be on the order of 0.1 ms. We expect this timescale to decrease even more in strong confinement (as compared to the corresponding bulk value), so that hydrodynamic-memory

effects are irrelevant on the millisecond timescale probed in our measurements. In this context, we furthermore note that Ref. [63] also shows that the inertial timescale τ_p of our colloidal particle is below $1 \mu\text{s}$, which explains why inertial effects can be neglected in the overdamped Langevin model we use for our recorded data on the millisecond timescale.

The diffusivity of a colloidal particle in a confining microchannel is well characterized experimentally [70]. In particular, the diffusivity is approximately position independent in the interior of the channel, i.e., sufficiently far away from the channel ends. The region of the channel shown in Fig. 1(c), which we use for our measurements, is well within the channel, so that as can be seen by the almost position-independent diffusivity in Fig. 6(b), the boundary effects described in Ref. [70] are irrelevant here.

APPENDIX B: EXTRACTING SOJOURN PROBABILITIES FROM EXPERIMENTAL TIME SERIES

1. Algorithm

We now explain how we extract sojourn probabilities and exit rates from experimental time series. We assume as given several uncorrelated time series, a reference path $\varphi(t)$, and a radius R . In essence, the algorithm we use concatenates randomly sampled short recorded trajectories.

We assume that the dynamics is time homogeneous and that the time series are Markovian. The former assumption holds in our experimental data because the external force is time independent. We discuss the validity of the latter assumption in Appendix A. We furthermore assume that the time series are indexed as described in the beginning of Appendix A.

At the initial time t_i , we choose an initial probability density inside the tube. In the discretization of space described in Appendix A, this probability density is represented by a normalized histogram that is only nonzero in the approximately $2R/\Delta x$ bins which intersect with the tube at time t_i , which is given by the interval $[\varphi(t_i) - R, \varphi(t_i) + R]$. To estimate the sojourn probability for a short time interval $\Delta\mathcal{T}$, we proceed as follows.

- (i) From the histogram representing the initial condition, we draw M sample bins (with replacement) $\{B_{k_1}, B_{k_2}, \dots, B_{k_M}\}$; for the definition of a bin, see Eq. (A3). Each sample bin represents an initial condition for a sample trajectory starting inside the tube.
- (ii) For each sample bin B_{k_i} , we draw one of the N_{k_i} measured data points inside this bin (with replacement, and using a uniform distribution on the set of all measured data points inside the bin), where N_i is defined in Eq. (A5). If the bin B_{k_i} only partly intersects the tube interior, and the drawn data point lies outside the tube, a new data point is drawn. The drawn data point belongs to a recorded time series, and we assume that this time series extends at least until a duration $\Delta\mathcal{T}$ beyond the

recording time of the drawn datapoint (this requirement can always be ensured by reducing the maximal index J_i corresponding to the trajectory X_i , and removing trajectories X_i that are shorter than $\Delta\mathcal{T}$).

- (iii) We follow each of the M randomly drawn time series from step 2 for the duration $\Delta\mathcal{T}$ and discard each trajectory as soon as it first leaves the tube. The number of trajectories left in the tube at each time step, which we denote by M_j , yields an estimate for the sojourn probability via $P_R^\varphi(j\Delta t) \equiv P_j \equiv M_j/M$, subject to the given initial condition, and for a duration $\Delta\mathcal{T}$.
- (iv) By creating a histogram from the final positions of those trajectories that stay inside the tube until time $\Delta\mathcal{T}$, a new initial distribution is obtained, and the algorithm can be repeated from step (i) for another time interval $\Delta\mathcal{T}$.

Figure 2 illustrates the algorithm for an initial distribution $P(x) = \delta(x - \varphi(t_i))$, $\Delta\mathcal{T} = 0.25$ s, $M = 3$ (to obtain a reliable estimate for the sojourn probability, of course much larger values for M need to be used). For the analysis of the experimental data, we use $\Delta x = 0.05 \mu\text{m}$, $\Delta\mathcal{T} = 0.25$ s; at the end of this section, we show that the results of this algorithm are independent of our particular choices for Δx and $\Delta\mathcal{T}$.

From the discrete time series P_j for the sojourn probability, the exit rate α_R^φ is obtained by discretizing

$$\alpha_R^\varphi(t) = -\frac{\dot{P}_R^\varphi(t)}{P_R^\varphi(t)}. \quad (\text{B1})$$

For the first iteration of steps (i)–(iii) of the algorithm outline above, we choose $M = 10^5$ and as an initial condition, a smeared-out delta peak at the tube center consisting of a uniform distribution on the three bins closest to the tube center. For each subsequent iteration of steps (i)–(iii), we estimate the number of trajectories M based on the recent trend of the exit rate. More explicitly, assuming we are at the k th repetition of steps (i)–(iii) (where $k > 1$), we fit a linear function

$$\alpha_{\text{fit}}(t) = a(t - (t_i + k\Delta\mathcal{T})) + b \quad (\text{B2})$$

to the exit rate in the time interval $[t_i + k\Delta\mathcal{T} - \Delta t_{\text{fit}}, t_i + k\Delta\mathcal{T}]$, where $\Delta t_{\text{fit}} = \min\{0.4 \text{ s}, \Delta\mathcal{T}\}$. Using the fitted Eq. (B2), we estimate the number M such that the expected number of trajectories inside the tube at the final time of the k th iteration step is approximately N_{final} , which yields

$$N_{\text{final}} = M \exp \left[- \int_{t_i + k\Delta\mathcal{T}}^{t_i + (k+1)\Delta\mathcal{T}} ds \alpha_{\text{fit}}(s) \right], \quad (\text{B3})$$

$$\Leftrightarrow M = N_{\text{final}} \exp \left[a \frac{\Delta\mathcal{T}^2}{2} + b \Delta\mathcal{T} \right]. \quad (\text{B4})$$

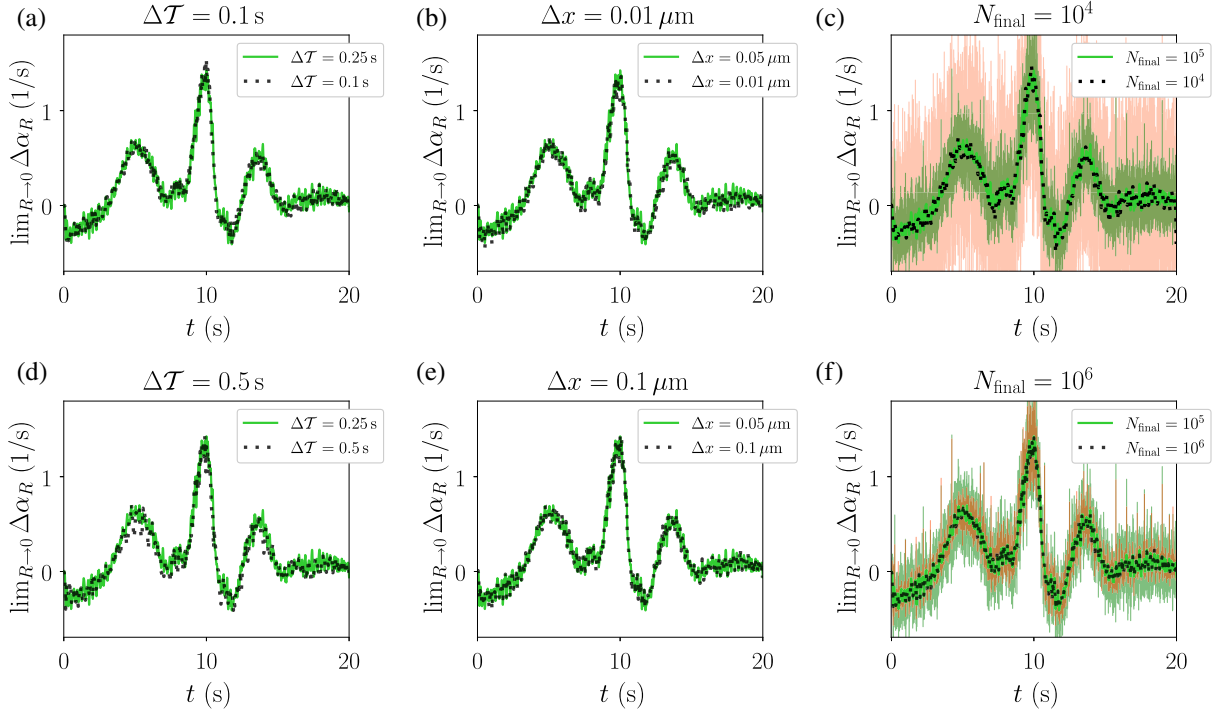


FIG. 8. Invariance of extrapolated exit-rate differences under variation of analysis parameters. All subplots feature a replot of the extrapolated exit rate difference from Fig. 3(c) of the main text (green solid line), which is obtained using the algorithm from Appendix B with $\Delta T = 0.25$ s, $\Delta x = 0.05$ μm , $N_{\text{final}} = 10^5$, and the pair of reference paths shown in Fig. 3(a) of the main text. (a),(d) compare the extrapolated exit rate for $\Delta T = 0.25$ s (green solid line) to results obtained using (a) $\Delta T = 0.1$ s and (d) $\Delta T = 0.5$ s (black dotted line), with all other parameters for both analyses identical. (b),(e) show both the extrapolated exit rate for $\Delta x = 0.05$ μm (green solid line) and the corresponding results extracted using (b) $\Delta x = 0.01$ μm and (e) $\Delta x = 0.1$ μm (black dotted line), with all other parameters identical. (c),(f) compare extrapolated exit-rate differences obtained for (c) $N_{\text{final}} = 10^4$ and (f) $N_{\text{final}} = 10^6$ to the results from Fig. 3(c) of the main text, which use $N_{\text{final}} = 10^5$. For the latter data, the running average is plotted as a thick green solid line, and the corresponding unsmoothed time series is shown as a thin green solid line (and looks like a shaded area because of the short-timescale fluctuations in the time series). The running average corresponding to $N_{\text{final}} = 10^4$, 10^6 is shown as a black dotted line, and the corresponding full unsmoothed time series is plotted as a thin red line. Except for the unsmoothed data in (c),(f), all shown data are smoothed using a running Hann-window average with window width 0.1s.

Unless noted otherwise, we use $N_{\text{final}} = 10^5$ for all exit rates shown in this paper; we demonstrate further below that our results are independent of the precise value used for N_{final} (as long as it is sufficiently large). For the minimization leading to the most probable path, we also use smaller values for N_{final} , as described in detail in Appendix C.

2. Relative path likelihood for a pair of paths

To infer the ratio of path probabilities for a pair of paths φ, ψ , we use the algorithm described just above to measure the exit rate for finite tube radius $R = 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8$ μm . Subsequently, we extrapolate the corresponding finite-radius exit-rate difference

$$\Delta\alpha_R(t) \equiv \alpha_R^\varphi(t) - \alpha_R^\psi(t) \quad (\text{B5})$$

to the limit $R \rightarrow 0$, as described in the main text. For Fig. 3 in the main text, the path φ is parametrized as

$$\varphi(t) = \frac{x_f - x_i}{2 \arctan(\kappa/2)} \arctan \left[\frac{\kappa}{t_f} (t - t_f/2) \right] + \frac{x_f + x_i}{2}, \quad (\text{B6})$$

where $t = [0, t_f]$ with $t_f = 20$ s, and where $x_i \equiv x_0^{\text{min}} \approx -2.725$ μm , $x_f \equiv x_1^{\text{min}} \approx 1.725$ μm are two minima of the potential energy. The path Eq. (B6) describes a barrier crossing starting at time $t_i = 0$ at the left minimum and arriving at the right minimum at time $t_f = 20$ s, with the parameter κ controlling the maximal path velocity during barrier crossing; for the results shown in Fig. 3, we use $\kappa = 5$. For ψ , we consider a path that rests at the right minimum $\psi(t) \equiv x_1^{\text{min}}$; see Fig. 3(a) for an illustration. As we demonstrate further below, pairs in which both paths are time dependent can also be considered; the advantage of considering one constant path is that then all time dependence in the exit rate can be attributed to the nonconstant path. While in principle, arbitrary paths can be considered,

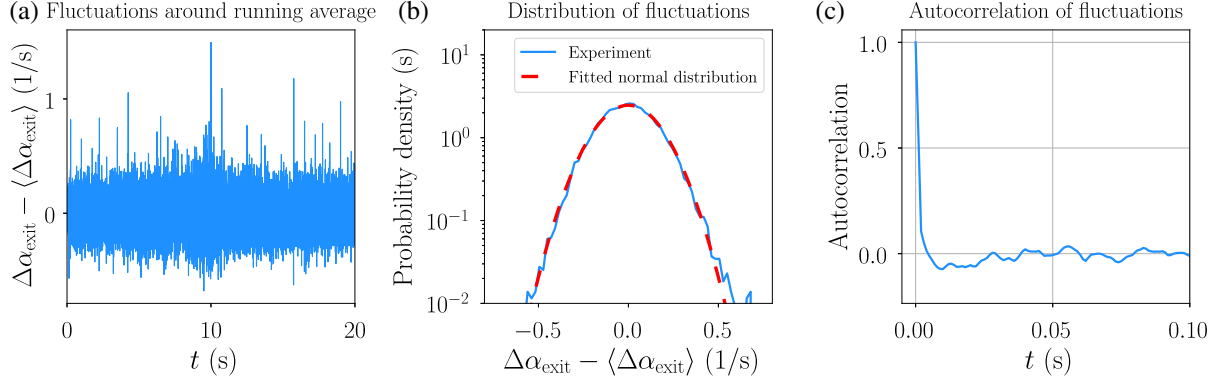


FIG. 9. Analysis of the noise eliminated by our smoothing procedure for extrapolated exit-rate differences. All data shown in this plot pertain to the smoothed extrapolated exit-rate difference considered in Fig. 3(c) of the main text, which is obtained using the algorithm described in Appendix B with parameters $\Delta x = 0.05 \mu\text{m}$, $\Delta\mathcal{T} = 0.25 \text{ s}$, $N_{\text{final}} = 10^5$. (a) shows the deviation of the extrapolated exit-rate difference from its running average; the latter is obtained from the former via a moving average with Hann window of width 0.1s. (b) shows the distribution of the time series from (a), together with a normalized fit of a Gaussian with zero mean; the fitting interval is $[-0.5/\text{s}, 0.5/\text{s}]$, and the resulting variance is $\sigma = 0.16/\text{s}^{-1}$. (c) shows the normalized autocorrelation of the time series shown in (a) as a function of time.

any path φ should of course move so slowly that given the experimental time resolution Δt of the data, the exit rate from a tube of radius R can be reliably inferred, i.e., $\dot{\varphi}\Delta t \ll R$.

3. Invariance of algorithm under variation of $\Delta\mathcal{T}$

For the results shown in the main text, we use $\Delta\mathcal{T} = 0.25 \text{ s}$. To demonstrate that exit rates obtained using the algorithm described above are independent of this particular choice of the parameter $\Delta\mathcal{T}$, we now consider the exit-rate difference of the pair of paths used in Fig. 3 of the main text for two other values of the parameter $\Delta\mathcal{T}$. In Figs. 8(a) and 8(d), we compare extrapolated exit-rate differences obtained for $\Delta\mathcal{T} = 0.1, 0.5 \text{ s}$ to results obtained using $\Delta\mathcal{T} = 0.25 \text{ s}$. All curves show excellent agreement, so that we conclude that our results are independent of $\Delta\mathcal{T}$.

4. Invariance of results under variation of bin width Δx

In the algorithm described in this appendix, particle positions are binned repeatedly, and new samples of trajectories starting from those bins are drawn. We now demonstrate that our results are independent of the particular bin width used for all results in the main text, $\Delta x = 0.05 \mu\text{m}$. For this, we consider the pair of paths from Fig. 3 of the main text, and infer ratios of path probabilities from the experimental data using the bin widths $\Delta x = 0.01 \mu\text{m}$ and $\Delta x = 0.1 \mu\text{m}$, with all other parameters identical to the $\Delta x = 0.05 \mu\text{m}$ scenario. The resulting extrapolated exit-rate differences are shown in Figs. 8(b) and 8(e), where we observe that the result is indeed independent of the bin width. Note that the bin width $\Delta x = 0.01 \mu\text{m}$ is in fact of the order of the measuring error for the particle position (the experimental accuracy is about $0.05 \mu\text{m}$); according to Fig. 8(b), this additional error does

not influence the inferred extrapolated exit rate. On the other hand, for $\Delta x = 0.1 \mu\text{m}$ the bin width is 10% of the smallest tube diameter considered, $2R = 1 \mu\text{m}$; according to Fig. 8(e), this binning is still accurate enough to infer the extrapolated exit rate.

5. Invariance of results under variation of N_{final}

For the parameter N_{final} , which via Eq. (B4) determines the number of trajectorial samples we consider in each iteration step to measure the exit rate, we use $N_{\text{final}} = 10^5$ in this work unless noted otherwise. To show that our results do not depend on this particular choice, in Figs. 8(c) and 8(f), we show extrapolated exit-rate differences based on Figs. 8(c) $N_{\text{final}} = 10^4$ and 8(f) $N_{\text{final}} = 10^6$, with all other parameters identical to the $N_{\text{final}} = 10^5$ scenario. As expected, we observe that for smaller N_{final} , the fluctuations around the running average of the extrapolated exit-rate difference are increased. The running averages themselves agree very well for all values of N_{final} considered, so that we conclude that our results are independent of this parameter.

6. Analysis of the noise eliminated by temporal averaging (smoothing)

As can be seen in Figs. 8(c) and 8(f), the experimental extrapolated exit-rate difference fluctuates significantly around its running average. We now show that these fluctuations are approximately normally distributed and correlated only on very short timescales, which indicates that they are approximately described by Gaussian white noise. This observation justifies that we average over these fast fluctuations, which originate from the fact that only a finite amount of experimental data are available. For our

analysis, we consider the data used for Fig. 3(c) of the main text, which is based on analysis parameters $\Delta x = 0.05 \mu\text{m}$, $\Delta T = 0.25 \text{ s}$, and $N_{\text{final}} = 10^5$, and a moving Hann-window average with width 0.1s. For the discussion in this paragraph, we denote the unsmoothed extrapolated exit-rate difference by $\Delta\alpha_{\text{exit}}$, and its running average by $\langle\Delta\alpha_{\text{exit}}\rangle$ [this average corresponds to the green line shown in Fig. 3(c) in the main text]. In Fig. 9(a), we show the fluctuations of $\Delta\alpha_{\text{exit}}$ around its running average as a function of time. Surprisingly, the fluctuations are almost independent of time and only slightly larger at around $t = 10 \text{ s}$, when the extrapolated exit-rate difference is maximal; cf. Fig. 3(c). We note that there are discrete peaks, which occur at multiples of ΔT . These originate from the instantaneous binning of the trajectories at these times performed by our algorithm. This binning slightly perturbs the instantaneous probability distribution inside the tube; the following relaxation of this perturbation, which happens on a very short timescale, leads to the observed peaks in the exit rate. In Fig. 9(b), we plot the distribution of the fluctuations from Fig. 9(a). The distribution is approximately Gaussian, as shown by the included fit of a Gaussian distribution to the region $[-0.5/\text{s}, 0.5/\text{s}]$. This interval contains 99.4% of the fluctuations, which shows that the peaks observed at multiples of ΔT in Fig. 9(a) are statistically insignificant. Figure 9(c) displays the normalized autocorrelation function calculated from the time series that is shown in Fig. 9(a). We observe that the autocorrelation decays on a timescale comparable to the time step of our data, and is completely uncorrelated for most of the duration of our averaging-window width 0.1s. The short correlation time observed in the figure furthermore corroborates our assumption that the experimental time series is approximately Markovian; cf. Appendix A. In conclusion, Fig. 9 shows that the noise we eliminate by the running average is approximately described by a stationary stochastic process with Gaussian steady-state distribution and quickly decaying autocorrelation. Our smoothing procedure thus basically eliminates Gaussian white noise, and thus, does not introduce any spurious effects or bias into the recorded data.

7. Extrapolated and finite-radius exit rates from Langevin simulations

In Figs. 8(b) and 8(e), we show that the results of our algorithm are independent of the particular choice of the binning parameter Δx . As discussed before, the smallest feasible value for Δx in our experimental analysis is determined by the accuracy of our measurement apparatus. To demonstrate that even if there is no spatial averaging at all in our algorithm, i.e., in the limit $\Delta x \rightarrow 0$, the extrapolated exit rate is described by the OM Lagrangian, we use a variation of our cloning algorithm to measure tubular exit rates from Langevin dynamics. For this variation, we apply the algorithm to numerical simulations, but with two

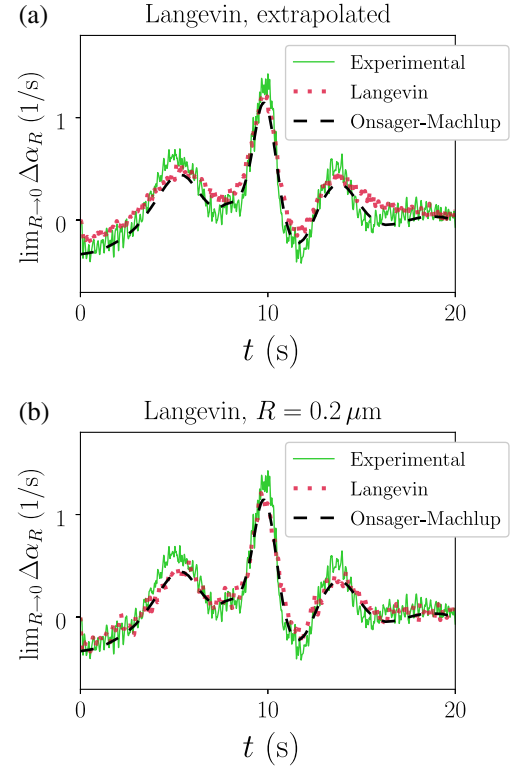


FIG. 10. Extrapolated and finite-radius exit-rate differences from Langevin simulations. This plot compares the extrapolated exit-rate difference obtained directly from the experimental data to results from Langevin simulations. (a) shows a numerical extrapolated exit rate obtained using a variation of the algorithm described in Appendix B, where sample trajectories are generated using Langevin simulations and no spatial binning is performed at multiples of $\Delta T = 0.25 \text{ s}$; see Appendix B for details. (b) shows the exit rate measured from Langevin simulations at finite radius $R = 0.2 \mu\text{m}$. All data are smoothed using a Hann-window average; for the numerical data in (b), we use an averaging window of width 0.2s, and for all other shown data, the averaging window has width 0.1s.

differences as compared to the analysis of experimental data. First, we do not aggregate an ensemble of trajectories beforehand, but generate each sample trajectory on the fly via an independent Langevin simulation in the potential-energy landscape shown in Fig. 3(c) of the main text using the constant diffusivity given by Eq. (A9), and using an Euler-Maruyama integration scheme with a time step $\Delta t_{\text{num}} = 10^{-4} \text{ s}$. Second, instead of binning positions after each iteration time $\Delta T = 0.25 \text{ s}$, we directly sample the initial conditions for the next iteration from the final positions of those stochastic trajectories that have remained inside the tube; this direct sampling of initial conditions corresponds to a binning with vanishingly small bin width, $\Delta x \rightarrow 0$. Using this algorithm, we measure the exit rate for the same values for R and extrapolate to $R \rightarrow 0$ as for the experimental data. The resulting extrapolated exit rate is compared to the experimental data in Fig. 10(a); we observe

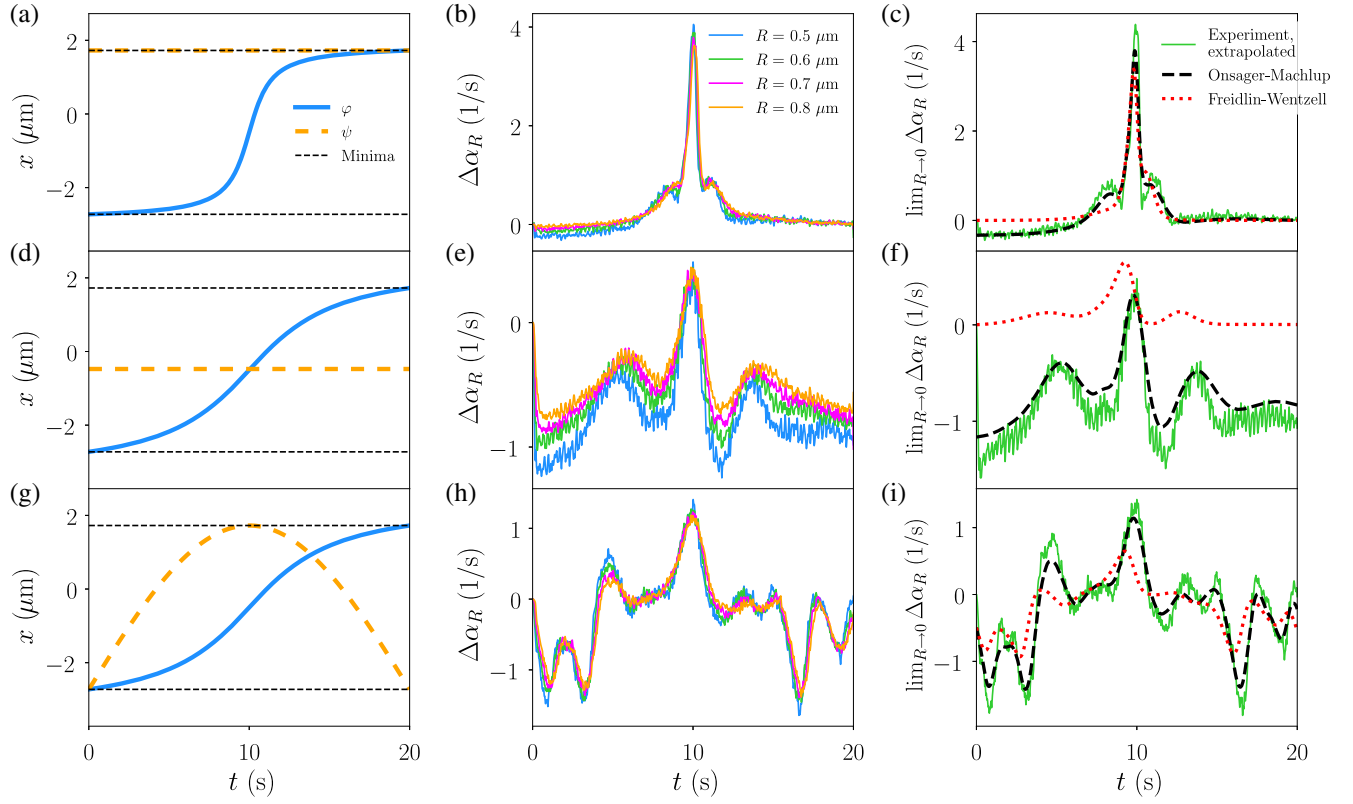


FIG. 11. (Extrapolated) exit-rate differences for various paths. Each subplot in the first column shows a pair of paths φ (blue solid line), ψ (orange dashed line), as well as two local minima of the potential landscape from Fig. 1(c). Each subplot in the second column depicts the measured exit-rate differences Eq. (B5) (colored solid lines), as extracted directly from experimental time series for the paths from the first column of each row, and for various values of the radius R , as indicated in the legend. The third column shows the result of extrapolating the measured finite-radius exit-rate differences for the paths from each row to the limit $R \rightarrow 0$, as described in the main text. Each subplot furthermore features the difference of FW and OM Lagrangians evaluated along the corresponding path pairs. Each exit rate shown in this figure is smoothed using a Hann window of width $\Delta T_{\text{smooth}} = 0.1$ s.

good agreement between the two, showing that our results are robust even in the limit $\Delta x \rightarrow 0$ (assuming that the stochastic process is described by an overdamped Langevin equation with additive noise). As mentioned before, the limiting factor for resolving the exit rate for small radius is both the temporal and spatial resolution; for our numerical simulations, the temporal resolution is given by the integration time step Δt_{num} , and the spatial resolution is basically the numerical integration error (which for the Euler-Maruyama algorithm and additive noise scales to leading order as Δt_{num}). Since the time step in the simulations is a factor of 10 smaller than the resolution of the experimental data, using simulations we can measure the exit rate also at smaller radius. Figure 10(b) compares the exit-rate difference measured in Langevin simulations at the finite radius $R = 0.2 \mu\text{m}$ to both the extrapolated experimental data and the OM Lagrangian; we observe that the finite-radius numerical result agrees well with the theoretical OM Lagrangian, indicating that for the system parameters and paths considered here, for the radius $R = 0.2 \mu\text{m}$ the limit $R \rightarrow 0$ is almost realized. This agreement

can be seen as direct numerical validation of the analytically calculated limit first obtained by Stratonovich [23].

8. Relative path likelihoods for several pairs of paths

To demonstrate that relative path likelihoods inferred using our algorithm agree with the OM Lagrangian Eq. (5) for arbitrary pairs of reference paths φ, ψ , we now consider three more pairs of reference paths, which are illustrated in the first column of Fig. 11. In each line, the second column shows the respective finite-radius exit-rate difference Eq. (B5) obtained from experimental data. The third column compares the resulting vanishing-radius extrapolations of the measured exit-rate differences to the Lagrangians Eqs. (5) and (6). As can be seen, for all pairs of paths the experimental result agrees reasonably well with the OM Lagrangian and shows clear disagreement with the FW Lagrangian.

Note that also the OM Lagrangian sometimes does not agree perfectly with the experimental data, which is most clearly observed in Fig. 11(f). We believe that this is because even the smallest radii of our finite-radius tubes are

not always small enough to perfectly capture the vanishing-radius limit. Consequently, the agreement between the experimental and theoretical results can presumably be improved upon by including measurements at smaller tube radius to the extrapolation. Measuring the exit rate for smaller tube radius, however, requires a higher temporal and spatial resolution in the recorded time series, which is ultimately limited by the experimental measurement apparatus.

APPENDIX C: CALCULATING THE MOST PROBABLE PATH FROM EXPERIMENTAL DATA

To extract the most probable path from experimental data, we evaluate

$$\varphi_R^* \equiv \operatorname{argmin}_{\varphi} \int_{t_i}^{t_f} dt \alpha_R^\varphi(t) \quad (\text{C1})$$

for the finite values $R = 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8 \mu\text{m}$, and then extrapolate to $R = 0$.

For each R , the minimization in Eq. (C1) is over all continuous paths with given end points $\varphi(t_i) = x_i$, $\varphi(t_f) = x_f$, so that the minimization is over an infinite-dimensional space of functions. To approximate this infinite-dimensional function space by a finite-dimensional space of dimension N , we parametrize φ as

$$\varphi(t) = x_i + \frac{t - t_i}{t_f - t_i} (x_f - x_i) + \sum_{n=1}^N \frac{a_n}{n^2} \sin\left(\pi n \frac{t - t_i}{t_f - t_i}\right). \quad (\text{C2})$$

Note that for any given set of coefficients $(a_1, \dots, a_N) \in \mathbb{R}^N$, Eq. (C2) fulfills the boundary conditions $\varphi(t_i) = x_i$, $\varphi(t_f) = x_f$. Employing this approximate parametrization, the minimization in Eq. (C1) is, for given R , over \mathbb{R}^N . Using our experimental data to evaluate the exit rate, we minimize the right-hand side of Eq. (C1), for $N = 20$ and $R = 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8 \mu\text{m}$ using a standard minimization algorithm [71]. For each evaluation of the sojourn probability, we employ the algorithm detailed in Appendix B. Since the algorithm presented there is based on stochastic sampling of recorded stochastic trajectories, the sojourn probability obtained from a single evaluation of our algorithm is also stochastic. Using a larger value for N_{final} decreases the variance of the inferred exit rate, but increases the computational time necessary to evaluate the exit rate for a given reference path.

For computational efficiency, we proceed in several steps to minimize Eq. (C1) for each given R . First, we perform four independent minimizations using $N_{\text{final}} = 2000$. For two of these minimizations, we use as the initial condition for the modes a_n , $n = 1, \dots, 20$, independent samples from a uniform distribution in $(-1, 1)$. For the other two, we use

as the initial condition the minimum of the analytical FW and OM action (each obtained as the lowest of ten independent minimizations of the respective action using the potential energy and friction coefficient inferred from the measured data). Each minimization has a starting variance $\sigma_0 = 0.5$ for the minimization algorithm, and we truncate the minimization after at most 2000 iterations of the algorithm (during each iteration, the sojourn probability is evaluated 12 times); typically, the minimization converges before that. After these four minimizations, the sojourn probability for each of the four minima is evaluated again using $N_{\text{final}} = 10^5$, and the path with the largest sojourn probability is chosen as φ_R^* .

Having obtained the most probable tube for several finite values of R , we subsequently extrapolate the corresponding modes $a_n(R)$, to $R = 0$ by fitting a function $f_n(R) = A_n + R^2 B_n$ to the finite-radius minimization results, and defining the corresponding expansion coefficients of the most probable path φ^* as

$$a_n^* \equiv \lim_{R \rightarrow 0} f_n(R) = A_n. \quad (\text{C3})$$

To minimize the OM and FW actions obtained by integrating Eqs. (5) and (6) along a path, we also use the parametrization Eq. (C2), with $N = 40$; the resulting instantons are shown in Fig. 4(a). Since the OM instanton agrees very well with the experimental extrapolation, for which we use $N = 20$, we conclude that $N = 20$ modes are indeed sufficient to characterize the most probable path for the transition considered.

APPENDIX D: PROTOCOL FOR THE COMPARISON OF FW AND OM INSTANTONS

In Fig. 4(b), we show a contour plot of Eq. (8) as a function of T/T_0 and $\Delta t_{\text{tot}} = t_f - t_i$. To obtain the figure, the actions corresponding to the OM and FW Lagrangians defined in Eqs. (5) and (6) are minimized using the experimental friction coefficient and force from Appendix A, using a path parametrized via Eq. (C2) with $N = 40$ modes. Finding the most probable path for each action is then a nonlinear minimization problem in \mathbb{R}^N . To carry out this minimization problem numerically, we employ a covariance-matrix-adaptation evolution-strategy algorithm [71] using the following protocol. To ensure we find the global minimum for each parameter combination $(T/T_0, \Delta t_{\text{tot}})$, we minimize each action in total 30 times. For every odd numbered of these 30 minimizations, we employ as the initial condition for the minimizer a random initial condition where all the a_n , $n = 1, \dots, 40$ are independent samples from a uniform distribution in $(-1, 1)$; for every even numbered of the 30 minimizations, we use the most probable OM or FW path from all the previous minimizations as the initial condition for the respective other action. In all cases, the initial variance for the

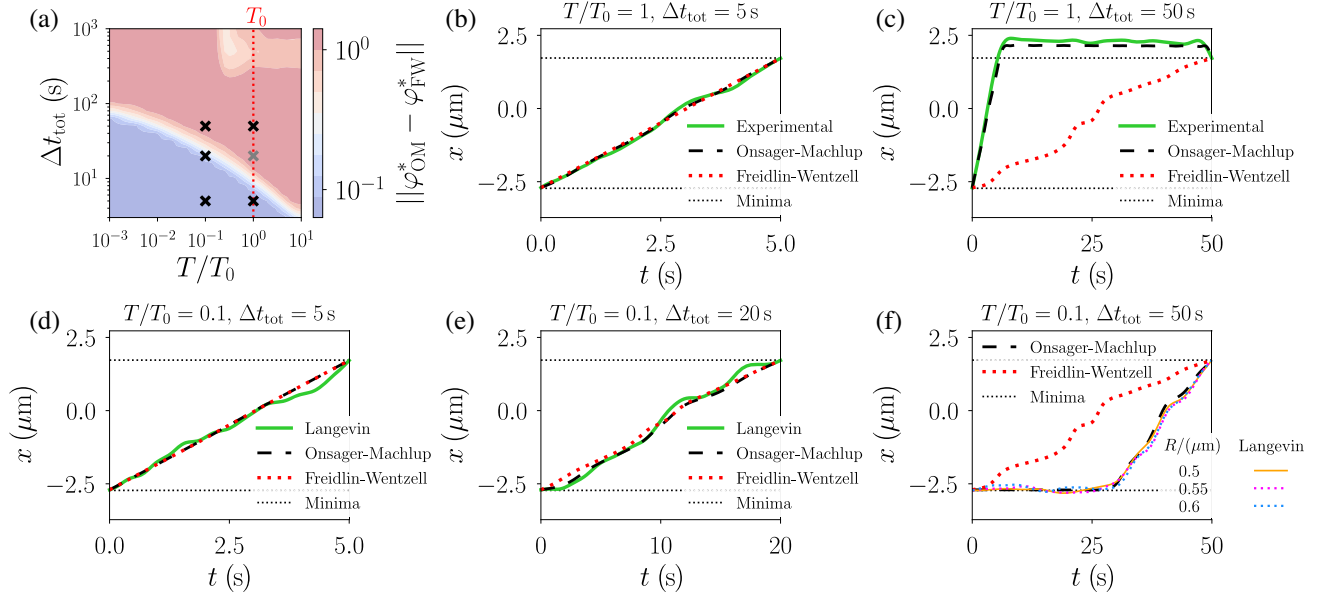


FIG. 12. Most probable paths for various temperatures T/T_0 and total durations Δt_{tot} . (a) is a replot of Fig. 4(b) of the main text and shows the mean difference between the OM and FW instanton as a function of T/T_0 and Δt_{tot} . While the gray cross denotes the parameters considered in Fig. 4(a) of the main text, the black crosses depict the parameters considered in (b)–(f) of the present figure. (b)–(f) show the most probable paths extracted from the data using Eq. (C3) (green solid line), together with the results of minimizing the integrated OM Lagrangian Eq. (5) (black dashed line), and minimizing the integrated FW Lagrangian Eq. (6) (red dotted line). In each plot, the horizontal dashed lines denote the local minima $x_0^{\min} \approx -2.725 \mu\text{m}$, $x_1^{\min} \approx 1.725 \mu\text{m}$ of the inferred potential, which are the start and end points of the paths. (b),(c) show the most probable paths extracted from the experimental data measured at $T_0 = 294$ K. (d)–(f) are obtained by analysis of an ensemble of numerical trajectories, which is generated using Langevin simulations based on the friction coefficient and force profile measured at $T_0 = 294$ K, but with the random force scaled down to represent the temperature $T = T_0/10 = 29.4$ K; see Appendix E for details.

minimization algorithm is chosen as $\sigma_0 = 0.5$. For each value of $(T/T_0, \Delta t_{tot})$, the respective action (FW or OM) is finally evaluated on all the results from the 30 minimizations, and the path with the smallest action is used in Eq. (8). For Fig. 4(b), the resulting 2D array of data are subsequently smoothed using a Gaussian filter.

APPENDIX E: INFERRED INSTANTONS FOR VARIOUS TEMPERATURES AND TOTAL DURATIONS

In Fig. 4(b) of the main text, we discuss in which parameter regime FW and OM Lagrangians predict the same instanton. In Fig. 4(a) of the main text, we compare the functional minima for parameters $T/T_0 = 1$ and $\Delta t_{tot} = 20$ s to the instanton extracted from experimental data; i.e., we consider one particular point of Fig. 4(b) of the main text (which is denoted by a black cross in the figure). We observe that, as expected from Figs. 3 and 4(a) of the main text, the experimental result agrees with the OM instanton and disagrees with the FW instanton. In this Appendix, we consider several more points in the $(T/T_0, \Delta t_{tot})$ plane, namely, $(T/T_0, \Delta t_{tot}/\text{s}) = (1, 5), (1, 50), (0.1, 5), (0.1, 20), (0.1, 50)$ as indicated in Fig. 12(a). For each of these parameters, we extract the

instanton from time series data and compare to the theoretical predictions from the OM and FW Lagrangians.

For $T/T_0 = 1$, we apply the algorithm from Appendix C to extract the instantons for $\Delta t_{tot} = 5$ s and $\Delta t_{tot} = 50$ s from the experimental data. For $T/T_0 = 1/10$, no experimental data are available. Indeed, as we discuss in the main text, if our experimental system is cooled down to $T = T_0/10 = 29.4$ K, the physics of the system would be radically different from what we observe at room temperature, and presumably not described by the friction coefficient and force profile obtained at $T_0 = 294$ K. To obtain instantons based on trajectorial data also at $T/T_0 = 1/10$, we generate an ensemble of Langevin trajectories, which we then analyze. To generate the data, we use our inferred friction coefficient and force profile (inferred at temperature T_0), a diffusion coefficient that is rescaled to $T = T_0/10$, and run 10^4 independent Langevin simulations of duration 2s each, with a time step 10^{-4} s. For each simulation, the initial condition is an independent sample of the uniform distribution in $[x_0^{\min} - 0.5 \mu\text{m}, x_1^{\min} + 0.5 \mu\text{m}] \approx [-3.225 \mu\text{m}, 2.225 \mu\text{m}]$. We treat this ensemble of numerically generated trajectories similar to the recorded experimental data and apply the algorithm from Appendix C to find instantons for $\Delta t_{tot} = 5, 20, 50$ s. Because for

lower temperature the exit rate during barrier crossing is quite large, we now consider a smaller refill time $\Delta\mathcal{T} = 0.025$ s; reducing $\Delta\mathcal{T}$ here is unproblematic because the data are by construction perfectly Markovian. For $\Delta t_{\text{tot}} = 5, 20$ s, we proceed exactly as with the experimental data, and extrapolate inferred instantons for finite tube radius $R = 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8$ μm to vanishing radius. However, for $\Delta t_{\text{tot}} = 50$ s, we observe that starting for $R \gtrsim 0.65$ μm , the finite-radius most probable tube starts to deviate a lot from the results for $R = 0.5, 0.55, 0.6$ μm . This indicates that for the larger radii, we are too far away from the asymptotic vanishing-radius behavior of the exit rate to properly infer the limit $R \rightarrow 0$. Therefore, instead of extrapolating to vanishing radius, for this case we consider the finite-radius most probable tubes for $R = 0.5, 0.55, 0.6$ μm .

In Figs. 12(b)–12(f), we compare the instantons obtained from experimental data [Figs. 12(b) and 12(c)] and Langevin data [Figs. 12(d)–12(f)] with the corresponding functional minima of the OM and FW action. In all subplots, the experimental or Langevin instanton shows good agreement with the OM instanton and agrees only with the FW instanton in the regime where the OM and FW instanton agree with each other; cf. Fig. 12(a).

APPENDIX F: CONTINUUM LIMIT OF THE ACTION FOR A REALIZATION OF THE LANGEVIN EQUATION

Here we discuss in more detail the continuum limit of the discretized action defined via Eqs. (26) and (24) for the case that we evaluate the action on a realization of the Langevin Eq. (13), which we write as $Y_t \equiv X_t$. The results derived in this Appendix are well known in the literature [2,42,46,49,50] and are included here for completeness.

Substituting Eq. (24) into the sum in Eq. (25), and expanding the square, the limit we are interested in can be written as

$$\begin{aligned} S[X] &\equiv \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} \Delta S_i & (F1) \\ &= \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} \frac{1}{b^2} \left[\frac{1}{2} \frac{\Delta X_i^2}{\Delta t} - a(\bar{X}_i^\xi) \Delta X_i + \frac{1}{2} a(\bar{X}_i^\xi)^2 \Delta t \right] \\ &\quad - \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} \frac{a'(\bar{X}_i^\xi)}{2b^2} [(1 - 2\xi) \Delta X_i^2 - b^2 \Delta t]. & (F2) \end{aligned}$$

As discussed in Eq. (27), the expectation value for the sum over the term $\Delta X_i^2 / \Delta t$ in Eq. (F2) diverges in the continuum limit $N \rightarrow \infty$. However, according to Eq. (28) we have

$$\frac{1}{2b^2} \sum_{i=0}^{N-1} \frac{\Delta X_i^2}{\Delta t} = \frac{1}{2b^2} \sum_{i=0}^{N-1} \frac{\Delta \mathcal{B}_i^2}{\Delta t} + \mathcal{O}(N^{-1/2}), \quad (F3)$$

where $\Delta \mathcal{B}_i \equiv b \Delta B_i \equiv b \sqrt{\Delta t} \Delta W_i$ is the increment of the Wiener process rescaled to the increment variance at which it enters as a random force in the Langevin equation, $\mathcal{B}_t \equiv b B_t$. The diverging term in Eq. (F2) can therefore be removed by subtracting the discretized action $\mathcal{S}^{(0)}[\mathcal{B}]$ of \mathcal{B} ,

$$\sum_{i=0}^{N-1} \Delta \mathcal{S}_i^{(0)} \equiv \frac{1}{2b^2} \sum_{i=0}^{N-1} \Delta t \left(\frac{\Delta \mathcal{B}_i}{\Delta t} \right)^2, \quad (F4)$$

before taking the continuum limit $N \rightarrow \infty$. Instead of Eq. (F2), we therefore consider the limit

$$S[X] - \mathcal{S}^{(0)}[\mathcal{B}] \quad (F5)$$

$$\begin{aligned} &\equiv \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} \left[\Delta S_i(X_{i+1}, X_i, \Delta t) - \frac{1}{2b^2} \frac{\Delta \mathcal{B}_i^2}{\Delta t} \right] & (F6) \\ &= - \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} \frac{1}{b^2} a(\bar{X}_i^\xi) \Delta X_i + \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} \frac{1}{2b^2} a(\bar{X}_i^\xi)^2 \Delta t \\ &\quad - \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} \frac{a'(\bar{X}_i^\xi)}{2b^2} [(1 - 2\xi) \Delta X_i^2 - b^2 \Delta t]. & (F7) \end{aligned}$$

This action difference amounts to considering the limiting ratio

$$\lim_{N \rightarrow \infty} \frac{P(X_N, t_N; \dots; X_1, t_1 | X_0, t_0)}{P^{\mathbb{W}}(\mathcal{B}_N, t_N; \dots; \mathcal{B}_1, t_1 | \mathcal{B}_0, t_0)} \equiv e^{-(S[X] - \mathcal{S}^{(0)}[\mathcal{B}])}, \quad (F8)$$

where the N -point probability density in the numerator of Eq. (F8) is with respect to the Langevin dynamics, whereas the N -point probability density in the denominator is with respect to the rescaled Wiener process \mathcal{B} , as indicated by the superscript \mathbb{W} . For the two diverging sums in Eq. (F6) to cancel, the Langevin realization in the numerator of Eq. (F8) must correspond to the same random force realization as used in the denominator; because the expression depends only on the increments of the Wiener process, the initial value of the Wiener process \mathcal{B}_0 is in fact irrelevant.

The limiting ratio Eq. (F8) quantifies the probability ratio of a realization X_t of the Langevin Eq. (13) and the corresponding realization of the Brownian motion \mathcal{B}_t . The limit Eq. (F7) therefore relates the probability distributions induced on the space of continuous paths by the Langevin Eq. (13) to the probability distribution induced on the same space by the (rescaled) Wiener process. This relation is described by the Girsanov formula, and indeed we see further below that the limit Eq. (F8) is precisely the Radon-Nikodym derivative from the Girsanov theorem [2,49,50].

Even more, the time-slicing approach employed here was used by Cameron and Martin to derive an early variant of the Girsanov theorem [49].

To calculate $\mathcal{S}[X] - \mathcal{S}^{(0)}[\mathcal{B}]$, and to see explicitly that the result is in fact independent of ξ , we now consider the limit for each of the sums in Eq. (F7) separately. Conceptually, it is clear without calculation that the final result of the following calculation must not depend on ξ , as the sum over the terms Eq. (24) is asymptotically independent of ξ .

The first sum in Eq. (F7) is simply the definition of the stochastic integral [1] using the convention of evaluating the integrand at an intermediate point given by ξ , which we write as

$$\lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} a(\bar{X}_i^\xi) \Delta X_i \equiv \int_0^{t_f} \mathcal{E}(\xi) dX_t a(X_t), \quad (\text{F9})$$

where the symbol $\mathcal{E}(\xi)$ indicates the convention for the stochastic integral. For example, for $\xi = 0$ we obtain the Itô stochastic integral, whereas for $\xi = 1/2$, we obtain the Stratonovich stochastic integral [1]. For any value of ξ , the stochastic integral along a solution of the Langevin equation can be rewritten in terms of an Itô integral [1,46]. This change in stochastic integral is achieved by Taylor expanding $a(\bar{X}_i^\xi)$ around X_i ,

$$a(\bar{X}_i^\xi = X_i + \xi \Delta X_i) = a(X_i) + \xi \Delta X_i a'(X_i) + \mathcal{O}(\Delta t), \quad (\text{F10})$$

where we use that from Eq. (17) it follows that $\Delta X_i = \mathcal{O}(\Delta t^{1/2})$. Substituting Eq. (F10) into the left-hand side of Eq. (F9), we obtain

$$\begin{aligned} & \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} a(\bar{X}_i^\xi) \Delta X_i \\ &= \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} a(X_i) \Delta X_i + \xi \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} a'(X_i) \Delta X_i^2 \\ & \quad + \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} \mathcal{O}(\Delta t) \Delta X_i. \end{aligned} \quad (\text{F11})$$

The first sum in Eq. (F11) is the Itô stochastic integral, which we denote by

$$\lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} a(X_i) \Delta X_i \equiv \int_0^{t_f} dX_t a(X_t). \quad (\text{F12})$$

Substituting $\Delta X_i = b \Delta t^{1/2} \Delta W_i + \mathcal{O}(\Delta t)$, which follows from Eq. (17), into the second sum in Eq. (F11), we obtain

$$\lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} a'(X_i) \Delta X_i^2 = b^2 \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} a'(X_i) \Delta t \Delta W_i^2 \quad (\text{F13})$$

$$= b^2 \int_0^{t_f} dt a'(X_t). \quad (\text{F14})$$

The last equality follows via the same argument as employed in proving the Itô formula; see, e.g., Chap. 4 of Ref. [2]. As stated in Ref. [2], we note that because $\Delta B_i \equiv \sqrt{\Delta t} \Delta W_i$ is the increment of the Wiener process, the limit Eq. (F14) is often expressed by the formula $dB_i^2 = dt$. Because $\mathcal{O}(\Delta t) \Delta X_i = \mathcal{O}(\Delta t^{3/2}) = \mathcal{O}(N^{-3/2})$, the last sum in Eq. (F11) scales as $\mathcal{O}(N^{-1/2})$ and hence vanishes in the limit $N \rightarrow \infty$.

Combining Eqs. (F9), (F11), (F12), and (F14), we obtain

$$\lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} a(\bar{X}_i^\xi) \Delta X_i \equiv \int_0^{t_f} \mathcal{E}(\xi) dX_t a(X_t) \quad (\text{F15})$$

$$= \int_0^{t_f} dX_t a(X_t) + b^2 \xi \int_0^{t_f} dt a'(X_t), \quad (\text{F16})$$

where the last equality is the standard formula that relates different definitions of the stochastic integral along the solution of the Langevin equation to the Itô convention [1,46].

The second sum in Eq. (F7) yields

$$\lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} \frac{1}{2b^2} a(\bar{X}_i^\xi)^2 \Delta t = \frac{1}{2b^2} \int_0^{t_f} dt a(X_t)^2. \quad (\text{F17})$$

That the result is independent of ξ follows by Taylor expanding $a(\bar{X}_i^\xi)$ around X_i ; cf. Eq. (F10). Similar to the third sum in Eq. (F11), the sum on the left-hand side of Eq. (F17) then depends on ξ only via terms that vanish in the continuum limit $N \rightarrow \infty$.

The continuum limit of the third sum in Eq. (F7) is given by

$$\begin{aligned} & \lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} \frac{a'(\bar{X}_i^\xi)}{2b^2} [(2\xi - 1) \Delta X_i^2 + b^2 \Delta t] \\ &= \xi \int_0^{t_f} dt a'(X_t). \end{aligned} \quad (\text{F18})$$

To obtain this limit, we first note that

$$\lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} a'(\bar{X}_i^\xi) \Delta t = \int_0^{t_f} dt a'(X_t), \quad (\text{F19})$$

which follows from the same arguments as used in Eq. (F17). The other limiting sum in Eq. (F18) is given by

$$\lim_{N \rightarrow \infty} \sum_{i=0}^{N-1} a'(\bar{X}_i^\xi) \Delta X_i^2 = b^2 \int_0^{t_f} dt a'(X_t), \quad (\text{F20})$$

which follows from the same argument as used in Eq. (F14) [2]. Upon substituting the two limits Eqs. (F19) and (F20) into the left-hand side of Eq. (F18), the right-hand side of the equation follows.

Finally substituting Eqs. (F9), (F17), and (F18) into Eq. (F7), we obtain [46]

$$\mathcal{S}[X] - S^{(0)}[\mathcal{B}] = -\frac{1}{b^2} \int_0^{t_f} \xi dX_t a(X_t) + \frac{1}{2b^2} \int_0^{t_f} dt a(X_t)^2 + \xi \int_0^{t_f} dt a'(X_t). \quad (\text{F21})$$

While the symbol ξ appears explicitly in Eq. (F21), the expression is in fact independent of ξ . The independence of $\mathcal{S}[X_t] - S_0[\mathcal{B}_t]$ on ξ becomes apparent by transforming the stochastic integral in the equation to the Itô convention via Eq. (F16), which leads to

$$\mathcal{S}[X] - S^{(0)}[\mathcal{B}] = -\frac{1}{b^2} \int_0^{t_f} dX_t a(X_t) + \frac{1}{2b^2} \int_0^{t_f} dt a(X_t)^2. \quad (\text{F22})$$

From this manifestly ξ -independent expression, it is also apparent that the limit Eq. (F8) is the Radon-Nikodym derivative of the Langevin trajectory X with respect to its corresponding noise realization \mathcal{B} , as described by the Girsanov formula [2,49,50].

-
- [1] C. W. Gardiner, *Stochastic Methods: A Handbook for the Natural and Social Sciences*, 4th ed., Springer Series in Synergetics (Springer, Berlin, 2009).
- [2] B. K. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, 6th ed. (Springer, Berlin, 2007).
- [3] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry*, 3rd ed. (Elsevier, Amsterdam, 2007).
- [4] K. Sekimoto, *Stochastic Energetics*, Lecture Notes in Physics No. 799 (Springer, Heidelberg, 2010).
- [5] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed., Stochastic Modelling and Applied Probability Vol. 38 (Springer, Berlin, 2010).
- [6] R. Chetrite and H. Touchette, *Nonequilibrium Markov Processes Conditioned on Large Deviations*, *Ann. Inst. Henri Poincaré* **16**, 2005 (2015).
- [7] B. C. Nolting and K. C. Abbott, *Balls, Cups, and Quasi-Potentials: Quantifying Stability in Stochastic Systems*, *Ecology* **97**, 850 (2016).
- [8] D. B. Brückner, A. Fink, C. Schreiber, P. J. F. Röttgermann, J. O. Rädler, and C. P. Broedersz, *Stochastic Nonlinear Dynamics of Confined Cell Migration in Two-State Systems*, *Nat. Phys.* **15**, 595 (2019).
- [9] *Large Deviations and Asymptotic Methods in Finance*, Springer Proceedings in Mathematics & Statistics Vol. 110, edited by P. K. Friz, J. Gatheral, A. Gulisashvili, A. Jacquier, and J. Teichmann (Springer International Publishing, Cham, 2015).
- [10] U. Seifert, *Stochastic Thermodynamics, Fluctuation Theorems, and Molecular Machines*, *Rep. Prog. Phys.* **75**, 126001 (2012).
- [11] S. Bera, S. Paul, R. Singh, D. Ghosh, A. Kundu, A. Banerjee, and R. Adhikari, *Fast Bayesian Inference of Optical Trap Stiffness and Particle Diffusion*, *Sci. Rep.* **7**, 41638 (2017).
- [12] G. Wilemski and M. Fixman, *Diffusion-Controlled Intra-chain Reactions of Polymers. I Theory*, *J. Chem. Phys.* **60**, 866 (1974).
- [13] A. Szabo, K. Schulten, and Z. Schulten, *First Passage Time Approach to Diffusion Controlled Reactions*, *J. Chem. Phys.* **72**, 4350 (1980).
- [14] M. Doi and S. F. Edwards, *The Theory of Polymer Dynamics*, International Series of Monographs on Physics Vol. 73 (Clarendon Press, Oxford, 2007).
- [15] H. Vandebroek and C. Vanderzande, *On the Generalized Langevin Equation for a Rouse Bead in a Nonequilibrium Bath*, *J. Stat. Phys.* **167**, 14 (2017).
- [16] U. Seifert, *Entropy Production along a Stochastic Trajectory and an Integral Fluctuation Theorem*, *Phys. Rev. Lett.* **95**, 040602 (2005).
- [17] D. G. Luchinsky, P. V. E. McClintock, and M. I. Dykman, *Analogue Studies of Nonlinear Systems*, *Rep. Prog. Phys.* **61**, 889 (1998).
- [18] H. B. Chan, M. I. Dykman, and C. Stambaugh, *Paths of Fluctuation Induced Switching*, *Phys. Rev. Lett.* **100**, 130602 (2008).
- [19] E. Weinan, W. Ren, and E. Vanden-Eijnden, *String Method for the Study of Rare Events*, *Phys. Rev. B* **66**, 052301 (2002).
- [20] W. Ren, E. Vanden-Eijnden, P. Maragakis, and E. Weinan, *Transition Pathways in Complex Systems: Application of the Finite-Temperature String Method to the Alanine Dipeptide*, *J. Chem. Phys.* **123**, 134109 (2005).
- [21] E. Weinan, W. Ren, and E. Vanden-Eijnden, *Transition Pathways in Complex Systems: Reaction Coordinates, Isocommittor Surfaces, and Transition Tubes*, *Chem. Phys. Lett.* **413**, 242 (2005).
- [22] L. Onsager and S. Machlup, *Fluctuations and Irreversible Processes*, *Phys. Rev.* **91**, 1505 (1953).
- [23] R. L. Stratonovich, *On the Probability Functional of Diffusion Processes*, *Selected Trans. Math. Stat. Prob.* **10**, 273 (1971).
- [24] D. Dürr and A. Bach, *The Onsager-Machlup Function as Lagrangian for the Most Probable Path of a Diffusion Process*, *Commun. Math. Phys.* **60**, 153 (1978).
- [25] T. Fujita and S.-I. Kotani, *The Onsager-Machlup Function for Diffusion Processes*, *Journal of mathematics of Kyoto University* **22**, 115 (1982).
- [26] W. Horsthemke and A. Bach, *Onsager-Machlup Function for One Dimensional Nonlinear Diffusion Processes*, *Z. Phys. B* **22**, 189 (1975).
- [27] H. Ito, *Probabilistic Construction of Lagrangean of Diffusion Process and Its Application*, *Prog. Theor. Phys.* **59**, 725 (1978).

- [28] Y. Takahashi and S. Watanabe, in *Stochastic Integrals*, edited by D. Williams (Springer, Berlin, 1981), Vol. 851, pp. 433–463.
- [29] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, 2nd ed., North-Holland Mathematical Library Vol. 24 (North-Holland, Amsterdam, 1989).
- [30] R. Graham, *Path Integral Formulation of General Diffusion Processes*, *Z. Phys. B* **26**, 281 (1977).
- [31] F. Langouche, D. Roekaerts, and E. Tirapegui, *Functional Integral Methods for Stochastic Fields*, *Physica (Amsterdam)* **95A**, 252 (1979).
- [32] H. Dekker, *On the Functional Integral Representation of General Continuous Markov Processes*, *Physica (Amsterdam)* **94A**, 339 (1978).
- [33] H. Dekker, *On the Path Integral for Diffusion in Curved Spaces*, *Physica (Amsterdam)* **103A**, 586 (1980).
- [34] M. F. Weber and E. Frey, *Master Equations and the Theory of Stochastic Path Integrals*, *Rep. Prog. Phys.* **80**, 046601 (2017).
- [35] C. Wissel, *Manifolds of Equivalent Path Integral Solutions of the Fokker-Planck Equation*, *Z. Phys. B* **35**, 185 (1979).
- [36] A. B. Adib, *Stochastic Actions for Diffusive Dynamics: Reweighting, Sampling, and Minimization*, *J. Phys. Chem. B* **112**, 5910 (2008).
- [37] L. F. Cugliandolo, V. Lecomte, and F. van Wijland, *Building a Path-Integral Calculus: A Covariant Discretization Approach*, *J. Phys. A* **52**, 50LT01 (2019).
- [38] A. D. Ventsel' and M. I. Freidlin, *On Small Random Perturbations of Dynamical Systems*, *Russ. Math. Surv.* **25**, 1 (1970).
- [39] H. Touchette, *The Large Deviation Approach to Statistical Mechanics*, *Phys. Rep.* **478**, 1 (2009).
- [40] T. Grafke and Eric Vanden-Eijnden, *Numerical Computation of Rare Events via Large Deviation Theory*, *Chaos* **29**, 063118 (2019).
- [41] J. Kappler and R. Adhikari, *Stochastic Action for Tubes: Connecting Path Probabilities to Measurement*, *Phys. Rev. Research* **2**, 023407 (2020).
- [42] H. Haken, *Generalized Onsager-Machlup Function and Classes of Path Integral Solutions of the Fokker-Planck Equation and the Master Equation*, *Z. Phys. B* **24**, 321 (1976).
- [43] J. Gladrow, M. Ribezzi-Crivellari, F. Ritort, and U. F. Keyser, *Experimental Evidence of Symmetry Breaking of Transition-path Times*, *Nat. Commun.* **10**, 55 (2019).
- [44] M. Chupeau, J. Gladrow, A. Chepelianskii, U. F. Keyser, and E. Trizac, *Optimizing Brownian Escape Rates by Potential Shaping*, *Proc. Natl. Acad. Sci. U.S.A.* **117**, 1383 (2020).
- [45] E. J. G. Peterman, F. Gittes, and C. F. Schmidt, *Laser-Induced Heating in Optical Traps*, *Biophys. J.* **84**, 1308 (2003).
- [46] K. L. C. Hunt and J. Ross, *Path Integral Solutions of Stochastic Equations for Nonlinear Irreversible Processes: The Uniqueness of the Thermodynamic Lagrangian*, *J. Chem. Phys.* **75**, 976 (1981).
- [47] L. F. Cugliandolo and V. Lecomte, *Rules of Calculus in the Path Integral Representation of White Noise Langevin Equations: The Onsager-Machlup Approach*, *J. Phys. A* **50**, 345001 (2017).
- [48] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations* (Springer, Berlin, 1992).
- [49] R. H. Cameron and W. T. Martin, *Transformations of Wiener Integrals under Translations*, *Ann. Math.* **45**, 386 (1944).
- [50] I. V. Girsanov, *On Transforming a Certain Class of Stochastic Processes by Absolutely Continuous Substitution of Measures*, *Theory Probab. Appl.* **5**, 285 (1960).
- [51] R. Zwanzig, *Memory Effects in Irreversible Thermodynamics*, *Phys. Rev.* **124**, 983 (1961).
- [52] H. Mori, *Transport, Collective Motion, and Brownian Motion*, *Prog. Theor. Phys.* **33**, 423 (1965).
- [53] A. Berezhkovskii and A. Szabo, *Time Scale Separation Leads to Position-Dependent Diffusion along a Slow Coordinate*, *J. Chem. Phys.* **135**, 074108 (2011).
- [54] D. M. Zuckerman and T. B. Woolf, *Transition Events in Butane Simulations: Similarities across Models*, *J. Chem. Phys.* **116**, 2586 (2002).
- [55] W. Min, G. Luo, B. J. Cherayil, S. C. Kou, and X. S. Xie, *Observation of a Power-Law Memory Kernel for Fluctuations within a Single Protein Molecule*, *Phys. Rev. Lett.* **94**, 198302 (2005).
- [56] J. O. Daldrop, J. Kappler, F. N. Brüning, and R. R. Netz, *Butane Dihedral Angle Dynamics in Water Is Dominated by Internal Friction*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 5169 (2018).
- [57] T. Franosch, M. Grimm, M. Belushkin, F. M. Mor, G. Foffi, L. Forró, and S. Jeney, *Resonances Arising from Hydrodynamic Memory in Brownian Motion*, *Nature (London)* **478**, 85 (2011).
- [58] J. O. Daldrop, B. G. Kowalik, and R. R. Netz, *External Potential Modifies Friction of Molecular Solutes in Water*, *Phys. Rev. X* **7**, 041065 (2017).
- [59] B. Kowalik, J. O. Daldrop, J. Kappler, J. C. F. Schulz, A. Schlaich, and R. R. Netz, *Memory-Kernel Extraction for Different Molecular Solutes in Solvents of Varying Viscosity in Confinement*, *Phys. Rev. E* **100**, 012126 (2019).
- [60] D. Selmeczi, S. Mosler, P. H. Hagedorn, N. B. Larsen, and H. Flyvbjerg, *Cell Motility as Persistent Random Motion: Theories from Experiments*, *Biophys. J.* **89**, 912 (2005).
- [61] B. G. Mitterwallner, C. Schreiber, J. O. Daldrop, J. O. Rädler, and R. R. Netz, *Non-Markovian Data-Driven Modeling of Single-Cell Motility*, *Phys. Rev. E* **101**, 032408 (2020).
- [62] J. Kappler and R. Adhikari, *Sojourn Probabilities in Tubes and Pathwise Irreversibility for Itô Processes*, [arXiv:2009.04250](https://arxiv.org/abs/2009.04250).
- [63] B. Lukić, S. Jeney, Ž. Sviben, A. J. Kulik, E.-L. Florin, and L. Forró, *Motion of a Colloidal Particle in an Optical Trap*, *Phys. Rev. E* **76**, 011112 (2007).
- [64] J. Kappler and R. Adhikari, *Irreversibility and Entropy Production along Paths as a Difference of Tubular Exit Rates*, [arXiv:2007.11639](https://arxiv.org/abs/2007.11639).
- [65] M. I. Dykman, P. V. E. McClintock, V. N. Smelyanski, N. D. Stein, and N. G. Stocks, *Optimal Paths and the Prehistory Problem for Large Fluctuations in Noise-Driven Systems*, *Phys. Rev. Lett.* **68**, 2718 (1992).

- [66] D. Nickelsen and H. Touchette, *Anomalous Scaling of Dynamical Large Deviations*, *Phys. Rev. Lett.* **121**, 090602 (2018).
- [67] C. Herbert and F. Bouchet, *Predictability of Escape for a Stochastic Saddle-Node Bifurcation: When Rare Events Are Typical*, *Phys. Rev. E* **96**, 030201(R) (2017).
- [68] S. N. Majumdar, *Persistence in Nonequilibrium Systems*, *Curr. Sci.* **77**, 370 (1999), <https://www.jstor.org/stable/24102955>.
- [69] A. J. Bray, S. N. Majumdar, and G. Schehr, *Persistence and First-Passage Properties in Nonequilibrium Systems*, *Adv. Phys.* **62**, 225 (2013).
- [70] S. L. Dettmer, S. Pagliara, K. Misiunas, and U. F. Keyser, *Anisotropic Diffusion of Spherical Particles in Closely Confining Microchannels*, *Phys. Rev. E* **89**, 062305 (2014).
- [71] N. Hansen, Y. Akimoto, and P. Baudis, CMA-ES/PYCMA: r2.7.0, 2019, [10.5281/zenodo.2559634](https://zenodo.org/record/2559634).