

Noise correction of large deviations with anomalous scaling

Daniel Nickelsen ^{1,*} and Hugo Touchette ^{2,†}

¹*African Institute for Mathematical Sciences (AIMS), Muizenberg 7950, South Africa*

²*Department of Mathematical Sciences, Stellenbosch University, Stellenbosch 7600, South Africa*



(Received 15 February 2022; accepted 15 May 2022; published 3 June 2022)

We present a path integral calculation of the probability distribution associated with the time-integrated moments of the Ornstein-Uhlenbeck process that includes the Gaussian prefactor in addition to the dominant path or instanton term obtained in the low-noise limit. The instanton term was obtained recently [D. Nickelsen and H. Touchette, *Phys. Rev. Lett.* **121**, 090602 (2018)] and shows that the large deviations of the time-integrated moments are anomalous in the sense that the logarithm of their distribution scales nonlinearly with the integration time. The Gaussian prefactor gives a correction to the low-noise approximation and leads us to define an instanton variance giving some insights as to how anomalous large deviations are created in time. The results are compared with simulations based on importance sampling, extending our previous results based on direct Monte Carlo simulations. We conclude by explaining why many of the standard analytical and numerical methods of large deviation theory fail in the case of anomalous large deviations.

DOI: [10.1103/PhysRevE.105.064102](https://doi.org/10.1103/PhysRevE.105.064102)

I. INTRODUCTION

We have shown recently [1] that time-integrated functions or observables of simple diffusions can have anomalous large deviations in the sense that their distribution decays exponentially with a scaling exponent that is nonlinear in the integration time. A simple model showing this behavior is the Ornstein-Uhlenbeck process (OUP) on \mathbb{R} defined by

$$dX_t = -\gamma X_t dt + \sigma dW_t, \quad (1)$$

where $\gamma > 0$ is the friction coefficient, $\sigma > 0$ is the noise amplitude, and W_t is a Brownian or Wiener motion representing the driving noise. Considering the integrated random variable

$$A_T = \frac{1}{T} \int_0^T X_t^\alpha dt, \quad (2)$$

which is an estimator of the α -moment of the stationary distribution of the OUP, we have shown that the probability density $p_T(a)$ of A_T scales for integers $\alpha > 2$ according to

$$p_T(a) \sim e^{-T^\xi \tilde{I}(a)/\sigma^2}, \quad (3)$$

with $\xi = 2/\alpha$ in the limit of large integration time ($T \rightarrow \infty$) and small noise amplitude ($\sigma \rightarrow 0$) [1]. This is to be contrasted with the scaling

$$p_T(a) \sim e^{-TI(a)/\sigma^2}, \quad (4)$$

which is usually expected to hold for random variables or observables that are integrated in time as in (2).

Other processes are known to have anomalous large deviations characterized by the scaling (3), including tracer

dynamics [2–4], the Kardar-Parisi-Zhang equation [5–7], branching processes [8–10], some non-Markovian processes [11–14], as well as random walk models arising in queueing theory [15–19]. The simplicity of the OUP makes it a useful model for understanding anomalous large deviations with analytical methods. For this process, we find normal large deviations that scale according to (4) for $\alpha = 1$ and $\alpha = 2$, but anomalous large deviations for $\alpha > 2$ with $\xi < 1$, so moments larger than 2 have fatter tails in time. This is confirmed by mathematical results that have been reported recently for a class of diffusions that includes the OUP as a special case [19].

In this paper, we extend these results by including the Gaussian prefactor in the instanton approximation of the path integral of $p_T(a)$, which underlies the low-noise approximation (3). This prefactor, which is expressed in terms of a functional determinant, not only gives a correction to the low-noise approximation, but can also be used in the path integral to define an instanton variance, which is useful for understanding how anomalous large deviations are created in time. To test these results, we present simulations based on importance sampling that extend the direct simulations previously reported [1].

The corrected $p_T(a)$ agrees remarkably well with the simulations and gives overall a good idea of the scaling of this density when considering only the long-time limit. The results on the instanton variance also support the conjecture that anomalous large deviations are created by a modified or effective process that is inherently time dependent [1]. By contrast, it is known that normal dynamical large deviations governed by the scaling (4) are created by a time-independent effective process, obtained by solving a spectral problem which happens to be ill defined for anomalous large deviations [20–22].

We explore these two fluctuation mechanisms in Sec. III by comparing the instanton and its variance for $\alpha = 1$ and $\alpha = 3$, after reviewing in Sec. II the theory behind the instanton

*danieln@aims.ac.za

†htouchette@sun.ac.za, htouchet@alum.mit.edu

approximation and its Gaussian correction. We then conclude in Sec. IV by explaining why many of the standard analytical and numerical techniques used in large deviation theory to study the long-time limit fail in the case of anomalous large deviations. The reasons are fundamental to the theory of large deviations and point to the need for new methods.

II. INSTANTON APPROXIMATION AND GAUSSIAN CORRECTION

Gaussian corrections to path integrals date back to the work of Gel'fand and Yaglom [23] in quantum mechanics and have been used for classical stochastic processes to derive corrections to instanton approximations of escape problems and transition pathways [24–29], yielding temperature-dependent corrections to the original Kramer's escape result, as well as for large deviations [30–33]. In this section, we present the standard approach to these corrections in which the path integral underlying $p_T(a)$ is discretized in the process space so as to perform a Gaussian integral around the instanton. In the continuum limit, the result of the Gaussian integral is expressed in terms of a functional determinant, calculated by solving a set of coupled linear differential equations with appropriate boundary conditions.

Since the determinant depends on the instanton, we start by recalling our results [1] about the low-noise approximation of $p_T(a)$ as well as the instanton underlying this approximation, and then present our results for the Gaussian correction based on the functional determinant. The detailed calculations leading to the determinant are presented in the appendices.

A. Instanton

The starting point of the low-noise or instanton approximation is the path representation of $p_T(a)$,

$$p_T(a) = \int \mathcal{D}[x] P[x] \delta(A_T[x] - a), \quad (5)$$

expressing this probability density as an integral over the path probability density $P[x]$ of all paths of the stochastic process leading to $A_T = a$. From the work of Onsager and Machlup [34], formalized in large deviation theory by Freidlin and Wentzell [35], we know that $P[x]$ can be expressed, up to a normalization constant, as

$$P[x] = e^{-S[x]/\sigma^2} \quad (6)$$

in terms of the action

$$S[x] = \int_0^T L(x, \dot{x}) dt, \quad (7)$$

where

$$L(x, \dot{x}) = \frac{1}{2}(\dot{x} + \gamma x)^2 \quad (8)$$

is the Lagrangian associated with the OUP. As a result, we can write

$$p_T(a) = \int \mathcal{D}[x] e^{-S[x]/\sigma^2} \delta(A_T - a). \quad (9)$$

This path integral is exponential with the noise amplitude σ , so it is natural to approximate it in the low-noise limit $\sigma \rightarrow 0$ using the path having the lowest action, similarly to

semiclassical approximations of quantum path integrals. The difference with the latter is that, apart from the fact that the path integral is real, we have to take into account the constraint $A_T = a$ using either a Lagrange parameter or by expressing the δ function in terms of its Laplace transform, which would add another integral in the path integral. The result of both procedures is the same: The optimal path or instanton having the lowest action, denoted by $\bar{x}(t)$, is found by minimizing the modified action,

$$S[x, \beta] = \beta Ta + \int_0^T \mathcal{L}(x, \dot{x}, \beta) dt, \quad (10)$$

which includes a Lagrange parameter β dual to the constraint $A_T = a$ in the Lagrange function,

$$\mathcal{L}(x, \dot{x}, \beta) = \frac{1}{2}(\dot{x} + \gamma x)^2 - \beta x^\alpha. \quad (11)$$

Equivalently, β can be seen as the parameter of the Laplace transform that represents the δ function in the path integral. In this case, the additional Laplace integral is further approximated by a specific value of β which is known to be equivalent to the Lagrange parameter $\beta(a)$ fixing the constraint $A_T = a$ (see [36, Appendix C.1]).

The minimization of the modified action proceeds at this point in the usual way using the Euler-Lagrange equation for the modified Lagrangian, which here takes the form

$$\ddot{x}(t) = \gamma^2 x(t) - \beta \alpha x(t)^{\alpha-1}. \quad (12)$$

The boundary conditions are

$$\dot{x}_0 - \gamma x_0 = 0, \quad \dot{x}_T + \gamma x_T = 0, \quad (13)$$

since we consider open terminal conditions in which x_0 and x_T are not *a priori* fixed.

These equations can be solved analytically in the $T \rightarrow \infty$ limit, as shown by Meerson [37], and lead to an explicit expression for the Lagrange parameter fixing the constraint $A_T = a$ [1],

$$\bar{\beta}(a) = \frac{1}{2} \gamma^{\frac{\alpha+2}{\alpha}} (aT)^{-\frac{\alpha-2}{\alpha}} \left(\frac{2\sqrt{\pi}}{\alpha-2} \frac{\Gamma(\frac{\alpha}{\alpha-2})}{\Gamma(\frac{3\alpha-2}{2\alpha-4})} \right)^{\frac{\alpha-2}{\alpha}}, \quad (14)$$

valid for $\alpha > 2$. With these two results, we then find an analytical expression for $S[\bar{x}]$, which scales with T according to $T^{2/\alpha}$, yielding the scaling (3) for $p_T(a)$ with

$$\bar{I}(a) = \lim_{T \rightarrow \infty} -\frac{\sigma^2}{T^\xi} \ln P[\bar{x}] = \lim_{T \rightarrow \infty} \frac{S[\bar{x}]}{T^\xi}, \quad (15)$$

where $\xi = 2/\alpha$. The exact expression of $\bar{I}(a)$ is too long to show (see Eq. (13) in [1]), but also scales like $|a|^\xi$.

We recall that the instanton physically represents the path most likely to be followed (measured or observed) if we condition the process X_t on the event $A_T = a$, that is, if we select only the paths of this process that realize this event. Figure 1 shows examples of instantons for the case $\alpha = 3$, obtained numerically for various values of a by solving the Euler-Lagrange equation with a relaxation method [38]. The time used, $T = 30$, is large enough for the action to converge [1].

The properties of the instanton were already discussed [1], so we only recall the main ones needed for the results to follow

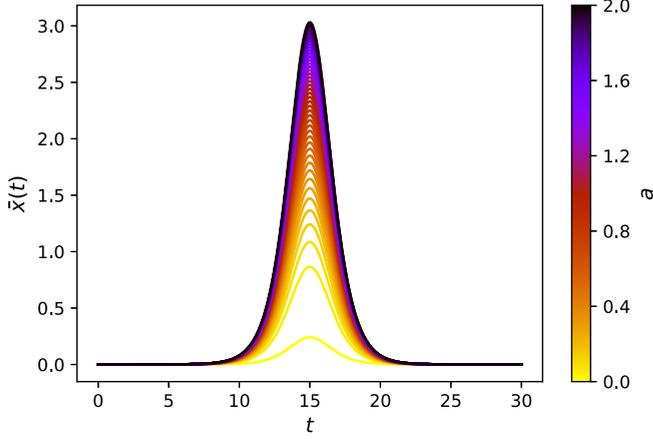


FIG. 1. Instantons obtained by numerically solving the Euler-Lagrange equation (12). Parameters: $\alpha = 3$, $\gamma = 1$, $\sigma = 0.5$, and $T = 30$.

when $\alpha > 2$:

(i) $\bar{x}(t)$ has a maximum in the middle $T/2$ of the simulated time interval and is symmetric with respect to this time, attaining a value very close to 0 for $t = 0$ and $t = T$. The symmetry follows from the fact that the OUP is a reversible diffusion.

(ii) The maximum of the instanton grows with a and T according to

$$\bar{x}_{\max}(a) = \left(\frac{\gamma^2}{2\bar{\beta}(a)\sigma^2} \right)^{\frac{1}{\alpha-2}} \propto (aT)^{1/\alpha}. \quad (16)$$

(iii) $\bar{x}(t)$ is well approximated by two symmetric exponentials: One growing with rate γ up to the maximum above, and another decaying back from this point with the same rate γ . This approximation does not capture the finite curvature of $\bar{x}(t)$ at its maximum, but does give the correct scaling of $\bar{I}(a)$ with a .

(iv) The instanton is localized over a time proportional to $1/\gamma$. This is consistent with the exponential approximation described above and explains why an integration time as short as $T = 30$ reaches the large deviation limit. For longer integration times, the instanton does not change much except for its height, and its tails close to 0 do not contribute significantly to the action for longer times. Much of the action, so to speak, happens in the localized region.

B. Gaussian correction

The instanton solution determines the scaling

$$p_T(a) \sim e^{-S[\bar{x}]/\sigma^2} \quad (17)$$

in the low-noise limit and, because of the time scaling of the action, the large deviation scaling shown in (3) with $\bar{I}(a)$ as given in (15). One way to correct this approximation is to expand the action to second order around the instanton and to carry out the resulting Gaussian path integral so as to obtain

$$p_T(a) \sim \frac{1}{\sqrt{D_0}} e^{-S[\bar{x}]/\sigma^2}, \quad (18)$$

where D_0 is a functional determinant corresponding to the continuous-time limit of the standard determinant that arises

in Gaussian integrals. We refer to the scaling above with D_0 as the Gaussian correction of the low-noise approximation (17), which does not mean, of course, that $p_T(a)$ is Gaussian.

For completeness, we present the full derivation of D_0 in Appendix A based on the discretization of the path integral. The end result is that D_0 is obtained from a set of four coupled linear differential equations [31]:

$$\ddot{A}(t) = 2\gamma\dot{A}(t) - \alpha(\alpha - 1)\bar{\beta}\bar{x}^{\alpha-2}A(t), \quad (19)$$

$$\dot{B}(t) = \gamma B(t) - \frac{\alpha}{T}\bar{x}^{\alpha-1}A(t), \quad (20)$$

$$\ddot{C}(t) = 2\gamma\dot{C}(t) - \alpha(\alpha - 1)\bar{\beta}\bar{x}^{\alpha-2}C(t) - \frac{2\alpha}{T}\bar{x}^{\alpha-1}B(t), \quad (21)$$

$$\dot{D}(t) = -\alpha(\alpha - 1)\bar{\beta}\bar{x}^{\alpha-2}C(t) - \frac{2\alpha}{T}\bar{x}^{\alpha-1}B(t), \quad (22)$$

with final values

$$1 = A(T),$$

$$0 = \dot{A}(T) = B(T) = C(T) = \dot{C}(T) = D(T). \quad (23)$$

The correction term D_0 corresponds to the value $D(0)$, obtained by integrating the equations above backwards in time from the terminal conditions in (23). The solution of these equations is the main result of this paper, which we study for specific parameter values in the next section.

C. Instanton variance

The expansion of the action up to second order around the instanton can be used to define a time-dependent function, denoted by $\bar{v}(t)$, which gives the local curvature of the path distribution $P[x]$ around the instanton and which is therefore interpreted as the variance of the path distribution along the instanton. The derivation of $\bar{v}(t)$ is outlined in Appendix B; the result is

$$\bar{v}(t) = \frac{\sigma^2 R(t)}{D_0} A(t)A(T-t), \quad (24)$$

where $A(t)$ is the solution of (19) and D_0 follows from (22). The factor $R(t)$ ensures that the constraint $A_T = a$ is met and is given by the integral

$$R(t) = \frac{\alpha}{T} \left[\int_0^t d\tau \bar{x}(\tau)^{\alpha-1} r_1(\tau) + \int_t^T d\tau \bar{x}(\tau)^{\alpha-1} r_2(\tau) \right], \quad (25)$$

where the auxiliary functions $r_1(\tau)$ and $r_2(\tau)$ both obey the same differential equation,

$$\ddot{r}(\tau) = [\gamma^2 - \alpha(\alpha - 1)\bar{\beta}\bar{x}(\tau)^{\alpha-2}] r(\tau) - \frac{\alpha}{T}\bar{x}^{\alpha-1}, \quad (26)$$

but differ in their boundary conditions,

$$\begin{aligned} 0 &= \gamma r_1(0) - \dot{r}_1(0), & 0 &= r_1(t), \\ 0 &= r_2(t), & 0 &= \gamma r_2(T) + \dot{r}_2(T), \end{aligned} \quad (27)$$

which make $r_1(\tau)$ and $r_2(\tau)$ dependent on t .

For large T , it can be shown that R becomes constant,

$$R = \frac{\alpha}{T} \int_0^T dt r(t)\bar{x}(t)^{\alpha-1}, \quad (28)$$

where $r(t)$ satisfies (26) and the boundary conditions

$$0 = \gamma r(0) - \dot{r}(0), \quad 0 = \gamma r(T) + \dot{r}(T). \quad (29)$$

We discuss this instanton variance and its meaning for specific parameters in the next section.

D. Importance sampling simulations

The Gaussian approximation of $p_T(a)$ shown in (18) needs to be compared with simulation results that should ideally cover a wide range of values of A_T . In our previous work [1], we simulated large numbers of trajectories of the process X_t and used them to directly estimate $p_T(a)$ and its corresponding rate function $I(a)$ by considering large simulation times [39]. This method is obviously limited in that, since $p_T(a)$ is exponentially small in both T and σ , an exponentially large sample is required to resolve the rate function over a wide range of values.

To improve the estimation, we can apply the idea of importance sampling by simulating a new process Y_t , different from X_t , chosen so as to make the event $A_T = a$ more likely and, ideally, to make it typical. Let $Q[x]$ denote the path distribution associated with this process. Then we can write

$$p_T(a) = \int \mathcal{D}[y] Q[y] W_T[y] \delta(A_T - a), \quad (30)$$

where

$$W_T[y] = \frac{P[y]}{Q[y]} = \exp\left[-\frac{1}{\sigma^2}(S[y] - S_Y[y])\right], \quad (31)$$

and $S_Y[y]$ is the action of the Y_t process. Thus, $p_T(a)$ can be estimated by simulating this process many times and by constructing a histogram of the samples of A_T obtained, including in the histogram the likelihood factor W_T computed as part of the simulation, in order to correct for the fact that we simulate Y_t rather than X_t [39–41]. Since Y_t is chosen so as to “hit” the event $A_T = a$ more often than X_t , it leads to a better estimation of $p_T(a)$ and, in turn, $I(a)$, sometimes with very few trajectories.

In practice, there are many processes that can be used to render $A_T = a$ typical. A natural one is obtained by guiding Y_t along the instanton using

$$dY_t = \dot{\bar{x}}(t)dt + \sigma dW_t, \quad (32)$$

with $Y_0 = \bar{x}(0)$, so that, in the limit $\sigma \rightarrow 0$, $Y(t) = \bar{x}(t)$. This change of process has been used before in various contexts [42–45], but was not found here to be accurate for sampling $p_T(a)$ as the noise drives trajectories far from the instanton over long times. To mitigate this effect, we guide a linear process with the same friction as X_t around the instanton using the stochastic differential equation,

$$dY_t = -\gamma(Y_t - \bar{x}(t))dt + \sigma dW_t. \quad (33)$$

This has the effect of producing trajectories that wander randomly around the instanton shown in Fig. 1. Other nonlinear friction terms were tested, but we found that the linear friction above, which defines another OUP that tracks the instanton, gives accurate results for the values of α considered, as it leads to a low variance for the likelihood factor W_T [41].

Note that importance sampling can be used to independently validate our theoretical results even if it uses the instanton because the estimator of $p_T(a)$ based on (30) is unbiased and consistent for any modified process Y_t . Thus any such process can be used, in principle, to estimate $p_T(a)$, including the original OUP which is not guided in any way, provided that the sample is large enough. What defines a good change of process is the variance of the resulting estimator, determined by the variance of W_T . For more details about the efficiency of importance sampling, we refer to [41].

III. RESULTS

We present in this section the results of the Gaussian correction and the instanton variance for the empirical moments of the OUP. We first consider $\alpha = 1$ to test our method for normal large deviations, and then $\alpha = 3$ to obtain results for anomalous large deviations. The value $\alpha = 3$ is representative of all integer values $\alpha \geq 3$ leading to anomalous large deviations [1].

A. $\alpha = 1$

For $\alpha = 1$, all the results can be obtained exactly. For the instantons, we obtain, from (12),

$$\bar{x}(t) = \frac{a\gamma T}{2\Omega_T^2} (2 - e^{-\gamma t} - e^{-\gamma(T-t)}), \quad (34)$$

with

$$\Omega_T^2 = \gamma T + e^{-\gamma T} - 1. \quad (35)$$

This predicts for $T \gg 1/\gamma$ that a fluctuation $A_T = a$ is created by a constant instanton evolving close to a for a time proportional to T . The exact action of this instanton is

$$S[\bar{x}] = \frac{a^2\gamma^3 T^2}{2\Omega_T^2} \quad (36)$$

and scales like $S[\bar{x}] \sim \gamma^2 a^2 T/2$ consistently with the fact that $\bar{x}(t) \approx a$ for $t \in [0, T]$. Consequently, $p_T(a)$ has the normal large deviation scaling (4) with

$$I(a) = \frac{\gamma^2 a^2}{2}. \quad (37)$$

To find the Gaussian correction to this result, we solve the coupled differential equations underlying the fluctuation determinant,

$$A(t) = 1, \quad (38)$$

$$B(t) = \frac{1 - e^{-\gamma(T-t)}}{\gamma T}, \quad (39)$$

$$C(t) = \frac{e^{-2\gamma(T-t)} - 4e^{-\gamma(T-t)} - 2\gamma(T-t) + 3}{2\gamma^3 T^2}, \quad (40)$$

$$D(t) = \frac{2\gamma(T-t) - 2 + 2e^{-\gamma(T-t)}}{T^2\gamma^2}. \quad (41)$$

Therefore,

$$D_0 = \frac{2\Omega_T^2}{\gamma^2 T^2}, \quad (42)$$

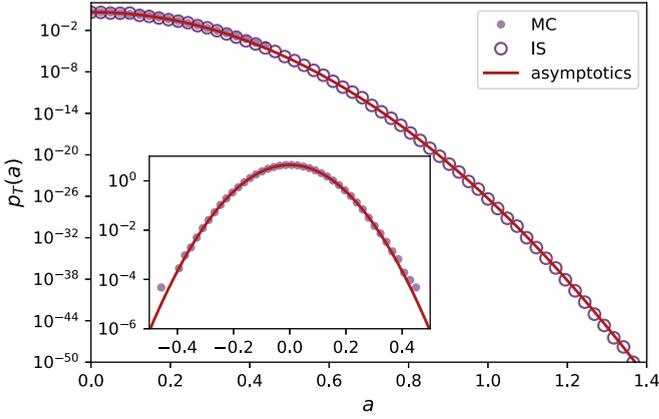


FIG. 2. Gaussian-corrected $p_T(a)$ (red line) compared with direct Monte Carlo (MC) simulations and importance sampling (IS) simulations. Parameters: $\gamma = 1$, $\sigma = 0.5$, and $T = 30$.

so that

$$p_T(a) \sim \sqrt{\frac{\gamma^2 T^2}{2\Omega_T^2}} \exp\left[-\frac{a^2 \gamma^3 T^2}{2\sigma^2 \Omega_T^2}\right], \quad (43)$$

which becomes, for large T ,

$$p_T(a) \sim \sqrt{\frac{\gamma T}{2}} \exp\left[-\frac{a^2 \gamma^2 T}{2\sigma^2}\right]. \quad (44)$$

The result on the right-hand side is actually the exact distribution of A_T if we normalize it properly, so the instanton calculation gives, in this case, the correct density for all noise amplitudes. This was already noted by Onsager and Machlup [34] and arises because the linear integral of a Gaussian process is also Gaussian.

We compare in Fig. 2 the theoretical result (44) with numerical results obtained from the direct Monte Carlo (MC) and importance sampling (IS) simulations of $p_T(a)$. The direct sampling, also shown in the inset, is naturally limited by the exponential concentration of $p_T(a)$ and the fact that this density is extremely small in the tails. Here we have used about 10^6 sample trajectories, leading to events seen in Fig. 2 to have a density of about 10^{-6} . The importance sampling overcomes this limitation by returning values that accurately match the theoretical distribution for values as low as 10^{-50} with a much smaller (and constant) sample size. The importance sampler in this case is chosen as

$$dY_t = -\gamma(Y_t - a)dt + \sigma dW_t \quad (45)$$

to sample A_T at the value a . This choice of dynamics, corresponding to an OUP recentered at a , is known to be optimal as it has a near-constant likelihood factor in the long-time limit [22], resulting in an estimator of $p_T(a)$ that has the least asymptotic variance [46].

This optimal property of Y_t is related to the instanton variance, which can also be calculated exactly. From (26), we

obtain the auxiliary functions

$$r_1(\tau) = \frac{2 - (2 - e^{-\gamma t})e^{-\gamma(t-\tau)} - e^{-\gamma\tau}}{2\gamma^2 T}, \quad (46)$$

$$r_2(\tau) = \frac{2 - (2 - e^{-\gamma(T-t)})e^{-\gamma(\tau-t)} - e^{-\gamma(T-\tau)}}{2\gamma^2 T}, \quad (47)$$

for the two boundary conditions in (27), giving

$$R(t) = \frac{2\gamma T - 6 + 4(e^{-\gamma t} + e^{-\gamma(T-t)}) - (e^{-2\gamma t} + e^{-2\gamma(T-t)})}{2\gamma^3 T^2}, \quad (48)$$

when inserted in (25). From the constant $A(t) = 1$, we thus find, with (24),

$$\bar{v}(t) = \frac{2\gamma T - 6 + 4(e^{-\gamma t} + e^{-\gamma(T-t)}) - (e^{-2\gamma t} + e^{-2\gamma(T-t)})}{4\gamma\Omega^2/\sigma^2}, \quad (49)$$

which reduces to

$$\bar{v}(t) = \frac{\sigma^2}{2\gamma} \quad (50)$$

in the limit $T \rightarrow \infty$. Hence, the path distribution $P[x]$ has a constant variance along the constant instanton $\bar{x}(t) = a$, which means physically that the fluctuation $A_T = a$ can be seen as being created by a linear process with stationary mean a and variance $\sigma^2/(2\gamma)$. These, as can be checked, are precisely the stationary mean and variance of the importance sampling process Y_t defined above, so that this process matches the local process determined from the path distribution around the instanton.

This result is expected. From recent works [20–22], it is known that the process X_t conditioned on realizing the fluctuation $A_T = a$ is equivalent in the long-time limit to another Markov process, called the effective or driven process, which happens here to be the process Y_t defined in (45) [21, Sec. 6.2]. The construction of the driven process is known when the large deviations of A_T are normal, in the sense of (4), and predicts in this case that the driven process is a homogeneous process. The low-noise limit of that process gives the instanton, which explains why we obtain here a constant instanton centered at a having a constant variance.

B. $\alpha = 3$

The instanton and fluctuation determinant cannot be found analytically for $\alpha = 3$ when the integration time T is finite, so we resort to obtaining them numerically. For the instanton, we solved the Euler-Lagrange equation (12) with a relaxation algorithm for different a , using a double exponential peaked at $t = T/2$ as the initial guess. Once we have the instantons for two contiguous a values, we extrapolate from these a new initial guess for the next a value. We repeat this procedure until we cover a desired range of a values. For the boundary solver, we use a minimal tolerance of 3×10^{-14} and a maximum of 10^5 mesh points. We also use $T = 30$ for the integration time, which appears to be enough to give results that are in the large deviation regime [1]. The solutions are shown again in Fig. 1, with the properties that we listed in the previous section, and were checked for a peak at $t = T/2$.

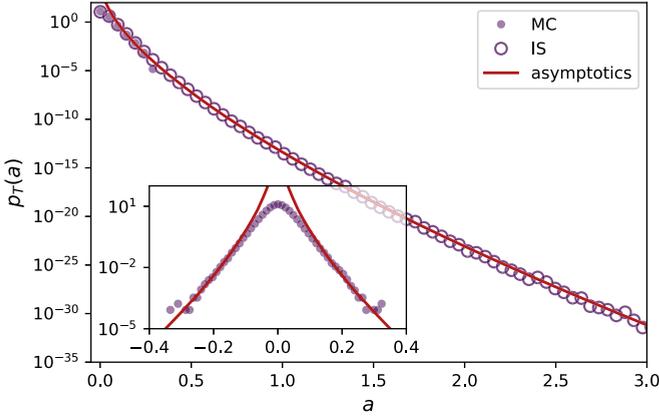


FIG. 3. Gaussian-corrected $p_T(a)$ (red line) for $\alpha = 3$ compared with MC and IS simulations. Parameters: $\gamma = 1$, $\sigma = 0.5$, and $T = 30$.

To obtain the fluctuation determinant, we numerically solve Eqs (19)–(22), feeding in the instantons as high resolution cubic interpolation functions and using a Runge-Kutta scheme with a minimum tolerance set to 3×10^{-14} to control numerical instabilities. We show in Fig. 3 the approximate $p_T(a)$ obtained from (18) with the resulting value for D_0 , as well as

$$S[\bar{x}] \sim T^{2/3} \bar{I}(a) \tag{51}$$

and

$$\bar{I}(a) = \left(\frac{9}{10}\right)^{\frac{1}{3}} \gamma^{\frac{5}{3}} a^{\frac{2}{3}}. \tag{52}$$

These results for the action and the rate function were found in our previous study [1]. We also show in Fig. 3 the results of the MC and IS simulations based on the modified OUP (33) tracking the instanton.

The simulation results show again that the Gaussian correction gives a good approximation of $p_T(a)$, except now near $a = 0$ where the low-noise approximation based on $\bar{I}(a)$ above predicts a peaked maximum at $a = 0$ which is actually smooth when σ is finite. This rounding effect is illustrated in Fig. 4,

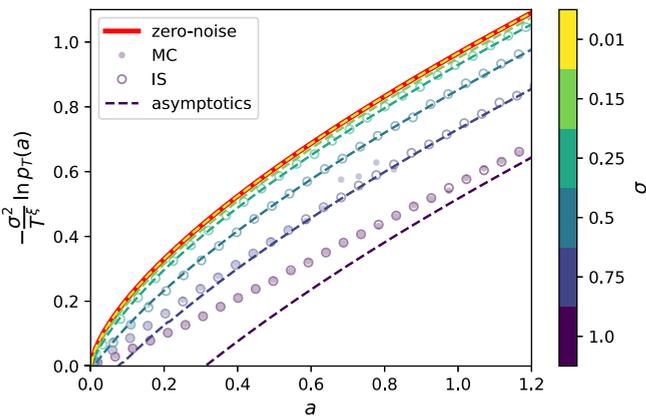


FIG. 4. Log-probability scaled with σ and T for different values of σ showing the convergence to the low-noise rate function $\bar{I}(a)$. Parameters: $\alpha = 3$, $\gamma = 1$, and $T = 30$.

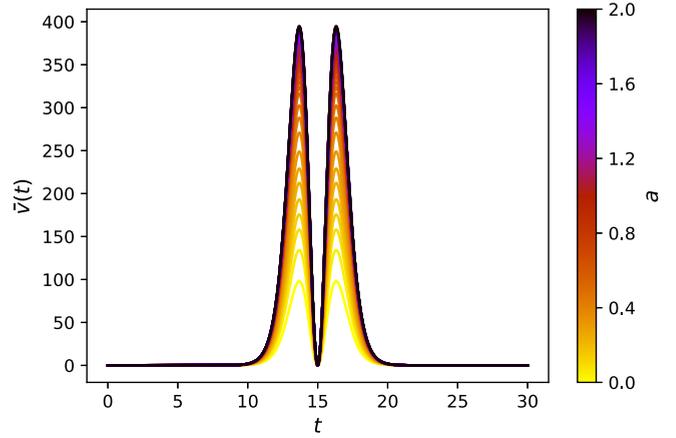


FIG. 5. Instanton variance $\bar{v}(t)$ for different values of a . Parameters: $\alpha = 3$, $\gamma = 1$, $\sigma = 0.5$, and $T = 30$.

which shows the same data on a different scale and different values of σ . The convergence to the low-noise rate function $\bar{I}(a)$, shown in red, and the emergence of a peak at $a = 0$ are clearly seen.

The fact that the instanton evolves in a time-dependent way for $\alpha = 3$, as seen in Fig. 1, makes this case very different from the case $\alpha = 1$ and is what gives rise to the scaling (3) describing anomalous large deviations. The instanton variance $\bar{v}(t)$ found from (24) is also time dependent, as shown in Fig. 5.

Obtaining $\bar{v}(t)$ is a challenging task since the values involved in Eqs. (26) and (25) are very small (of the order of 10^{-30} to 10^{-20}). For this reason, we took care to solve these equations numerically using different mesh points and interpolations for the instanton to see if the results were stable. We found that the maximum value of $\bar{v}(t)$, which sets the scale of the variance, cannot be relied on since it is sensitive to the order of approximation used for $\bar{x}(t)$ [47], but that the double-peak shape of $\bar{v}(t)$ seen in Fig. 5 is stable and so is quantitatively valid. Initially, the variance is low and starts to increase when the instanton itself starts increasing to its maximum. Unlike the instanton, however, the variance has a turning point before $t = T/2$, beyond which it decreases rapidly to a low value (close to 0 from numerical calculations) precisely at $t = T/2$. After this time, the same behavior is repeated, showing overall that the fluctuations of A_T are created by stochastic trajectories that follow the time-dependent instanton and fluctuate around that instanton, except at $t = T/2$, where they all converge and go through $\bar{x}(T/2) = \bar{x}_{\max}$ as a result of $\bar{v}(T/2) \approx 0$.

It is difficult to verify the instanton variance independently from simulations since it relies on rare trajectories underlying the large deviations of A_T whose variance differs from the variance of the trajectories simulated with importance sampling [48]. However, the behavior of $\bar{v}(t)$ shown in Fig. 5 agrees qualitatively with the fluctuation paths reported in our previous study [1, Fig. 3], which have reduced fluctuations at the instanton peak (shifted numerically at $t = T/2$ for comparison). The fact that $\bar{v}(t)$ is symmetric with respect to $t = T/2$ is also supported qualitatively from that figure and is consistent with the fact that the conditioning of the OUP

on the event $A_T = a$ is a reversible process since the OUP itself is reversible [21, Sec. 5.5]. This does not mean that all stochastic trajectories realizing the event $A_T = a$ have to be symmetric with respect to $t = T/2$. However, the time reversal of any such trajectory realizes the same value $A_T = a$ with the same probability, by virtue of the OUP being reversible, which means that the whole ensemble of trajectories realizing $A_T = a$ must define a reversible process.

This ensemble of trajectories realizing $A_T = a$ was studied extensively for normal large deviations [20–22]: It is defined mathematically as a conditioning of the path distribution $P[x]$ on the event $A_T = a$, and is known in the regime of normal large deviations to be a time-homogeneous and stationary Markov process, at least in the absence of dynamical phase transitions [49]. The case $\alpha = 1$ follows this result: The mean and variance of the instanton are time independent since the conditioning of $P[x]$ on $A_T = a$ is time independent in the long-time limit. For $\alpha = 3$, by contrast, we find a time-dependent instanton mean and variance, suggesting that the ensembles of paths realizing $A_T = a$ are described as a whole by a time-dependent process. Similar results are found in the context of simple random walks and jump processes arising in queueing theory [15–17].

Based on these results, it is natural to conjecture that a necessary condition for observing anomalous large deviations is that the ensemble of trajectories or process realizing $A_T = a$ is time dependent. In other words, if the large deviations of A_T are anomalous, then the conditioned process realizing those large deviations is explicitly time dependent. This is suggested not only by our instanton results, which provide partial information about the mean and variance of that process, but also by the fact that many techniques used for obtaining large deviations do not work for anomalous large deviations, either because they assume or predict that that process is time independent in the long-time limit. We discuss this point in more detail in the next section and suggest ideas for dealing with anomalous large deviations.

IV. CONCLUDING REMARKS

In principle, the Gaussian correction of the path integral is not expected to describe $p_T(a)$ for arbitrary noise amplitudes since it is a further approximation of the path integral in terms of σ . However, the good agreement that we find between this correction and the simulation results shows that it recovers much of $p_T(a)$, especially in the tails (see Fig. 4), giving us some information about the anomalous large deviations of A_T in the limit where $T \rightarrow \infty$ with σ finite. In this regime, $p_T(a)$ is expected to scale according to

$$p_T(a) \sim e^{-T^\xi J(a)} \quad (53)$$

as $T \rightarrow \infty$, so the limit function

$$J(a) = \lim_{T \rightarrow \infty} -\frac{1}{T^\xi} \ln p_T(a) \quad (54)$$

should exist. Mathematical estimates of this function have been reported recently for a class of diffusions that include the OUP [19], although it is not clear whether they involve the low-noise limit. With this extra limit, the rate function that

one obtains is

$$\bar{I}(a) = \lim_{\sigma \rightarrow 0} \lim_{T \rightarrow \infty} -\frac{\sigma^2}{T^\xi} \ln p_T(a) = \lim_{\sigma \rightarrow 0} \sigma^2 J(a), \quad (55)$$

which follows from the instanton approximation, as well as logarithmic corrections of this function coming from the Gaussian prefactor [50].

The reason for considering the low-noise limit, as we have argued before [1], is that many techniques that are standard in large deviation theory do not work in the case of anomalous large deviations. In particular, we cannot obtain $J(a)$ by applying the contraction principle to the level 2 or level 2.5 rate functions (see [51] for details), as these all are defined in the normal scaling regime and thus predict normal large deviations for A_T when the contraction has a nontrivial solution. We also cannot obtain $J(a)$ as the Legendre transform of the scaled cumulant generating function (SCGF), defined as

$$\lambda(k) = \lim_{T \rightarrow \infty} \frac{1}{T} \ln E[e^{T k A_T}], \quad (56)$$

since the exponential in the expectation above has the wrong scaling in T and, therefore, does not capture the anomalous scaling (53). In fact, the SCGF, as defined above, diverges for all $k \neq 0$. This follows because the SCGF is related, in the case of normal large deviations, to the ground-state energy of a quantumlike potential [52], which is not confining and has no lower bound when $\alpha > 2$ and $k \neq 0$ [1].

To circumvent this problem, one can attempt to regularize the related quantum problem, e.g., by considering a limited range of the potential, as done recently by Smith [53]. This approach is able to recover the central, Gaussian part of $p_T(a)$, but seems insufficient to obtain $J(a)$ since it is based on approximating the SCGF, which is again formally divergent, and predicts a normal rather than anomalous scaling of large deviations because of the effective confinement introduced.

Another idea is to redefine the SCGF by the limit,

$$\lambda_\xi(k) = \lim_{T \rightarrow \infty} \frac{1}{T^\xi} \ln E[e^{T^\xi k A_T}], \quad (57)$$

to match the limit (53) capturing the anomalous scaling of $p_T(a)$. This modified SCGF is covered by large deviation theory (see [36, Appendix D]), although little is known about its properties, especially its connection with the Feynman-Kac equation and long-time solutions of this equation. In our case, $\lambda_\xi(k)$ also diverges when $\alpha > 2$ because the right tail of $J(a)$, which asymptotically matches that of $\bar{I}(a)$, is nonconvex, but there might be other processes and observables for which the modified SCGF is finite, provided that their large deviations are anomalous and have a convex rate function.

These considerations affect not only analytical methods for obtaining rate functions, but also numerical methods. For instance, the divergence of $\lambda(k)$ should be seen in runs of the cloning algorithm since this algorithm gradually estimates the limit (56). In this case, one could attempt to modify the algorithm to compute the modified SCGF in (57), but the precise form of this modification is yet to be investigated.

Similarly, it is not clear how importance sampling methods should be modified to account for anomalous large deviations. From our results, it seems that the appropriate way to use this method is to use a change of process that is inherently

time dependent, but finding processes that are efficient for sampling anomalous large deviations is also an open problem. The change of process used here, which has the effect of recentering the OUP, gives good results, as we have seen, but it is not expected to be optimal in the sense that it minimizes the asymptotic variance [40]. One way to solve this problem is to include time-dependent controls in the optimal control formalism developed for normal large deviations [22]. This leads to time-dependent Hamilton-Jacobi-Bellman equations that could be solved numerically, if not analytically.

ACKNOWLEDGMENTS

The work of D.N. is supported by the Oppenheimer Memorial Trust (postdoctoral fellowship).

APPENDIX A: CALCULATION OF FLUCTUATION DETERMINANT

We explain in this section how the fluctuation determinant D_0 is obtained in the continuous-time limit. The starting point is Laplace's method applied to the finite-dimensional integral,

$$F = \int_{\mathbb{R}^N} dz_1 \dots dz_N e^{-\frac{1}{\sigma^2} f(z_1, \dots, z_N)}. \quad (\text{A1})$$

We assume that $f: \mathbb{R}^N \rightarrow \mathbb{R}$ is such that the integral exists and has a unique minimum at the point $\bar{z} = (\bar{z}_1, \dots, \bar{z}_N)$ satisfying $\nabla f(\bar{z}_1, \dots, \bar{z}_N) = 0$. Expanding f to second order around \bar{z} , we obtain, after carrying out the Gaussian integral,

$$F \sim \frac{(2\pi\sigma^2)^{\frac{N}{2}}}{\sqrt{\det H}} e^{-\frac{1}{\sigma^2} f(\bar{z}_1, \dots, \bar{z}_N)}, \quad (\text{A2})$$

in the limit $\sigma \rightarrow 0$, where the fluctuation determinant enters as the determinant of the Hessian,

$$H_{kl} = \left. \frac{\partial^2 f(z_1, \dots, z_N)}{\partial z_k \partial z_l} \right|_{z=\bar{z}}. \quad (\text{A3})$$

This is the Gaussian-corrected form of the Laplace approximation. Additional corrections can be obtained by considering more terms in the Taylor expansion of f around \bar{z} beyond the second-order term [54].

In our problem, we apply Laplace's method to the path integral representation of $p_T(a)$, given in (9), replacing the

Dirac δ function by its Laplace transform [30,31],

$$p_T(a) = \int \frac{dx_0}{Z} \int dx_T \int \frac{dq}{2\pi\sigma^2} \int_{(0,x_0)}^{(T,x_T)} \mathcal{D}[x] e^{-S[x,\beta]/\sigma^2}, \quad (\text{A4})$$

where $\beta = iq/T$ and

$$S[x, \beta] = \gamma x_0^2 + \beta Ta + \int_0^T \mathcal{L}(x, \dot{x}, \beta) dt \quad (\text{A5})$$

is the modified action. Note that compared with (10), we now integrate explicitly over the final state x_T and the initial state x_0 with the stationary density

$$p(x_0) = \frac{e^{-\gamma x_0^2/\sigma^2}}{Z}, \quad Z = \sqrt{\frac{\pi\sigma^2}{\gamma}}. \quad (\text{A6})$$

These added terms do not influence the approximation significantly, so we do not include them in the text.

Discretizing the path integral into N time slices or steps, $t_j = j\epsilon$, $T = N\epsilon$, $x_j = x(t_j)$, we obtain

$$p_N(a) = \int \frac{dq}{2\pi\sigma^2} \int \frac{dx_0}{Z} \prod_{j=1}^N \int \frac{dx_j}{\sqrt{2\pi\sigma^2\epsilon}} e^{-\frac{1}{\sigma^2} \mathcal{S}_N(x_0, \dots, x_N, \beta)}, \quad (\text{A7})$$

with the discretized action

$$\mathcal{S}_N = \gamma x_0^2 + \epsilon \sum_{j=0}^{N-1} \left[\frac{1}{2} \left(\frac{x_{j+1} - x_j}{\epsilon} + \gamma x_j \right)^2 - \beta x_j^\alpha \right] + \beta a. \quad (\text{A8})$$

Applying (A2) to the discretized path integral, we then obtain

$$p_N(a) \sim \frac{e^{-\mathcal{S}_N(\bar{x}_0, \dots, \bar{x}_N, \bar{\beta})/\sigma^2}}{Z \sqrt{\epsilon^N \det H}}, \quad (\text{A9})$$

where H is the $(N+2) \times (N+2)$ Hessian with elements

$$H_{kl} = \left. \frac{\partial^2 \mathcal{S}_N(x_0, \dots, x_N, \beta)}{\partial x_k \partial x_l} \right|_{x_j=\bar{x}_j, \beta=\bar{\beta}}, \quad (\text{A10})$$

using $x_{N+1} = q$. Note that the exponent can also be expressed in terms of the bare action S , since the instanton with $\bar{\beta}$ enforces the constraint $A_T = a$, so that $\mathcal{S}(\bar{x}_0, \dots, \bar{x}_N, \bar{\beta}) = S(\bar{x}_0, \dots, \bar{x}_N)$.

We now consider the continuous-time limit by writing the Hessian as

$$H = \begin{pmatrix} a_0/\epsilon & -b_0/\epsilon & 0 & \dots & 0 & i u_0 \\ -b_0/\epsilon & a_1/\epsilon & -b_1/\epsilon & \dots & 0 & i u_1 \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & \dots & 0 & -b_{N-1}/\epsilon & a_N/\epsilon & i u_N \\ i u_0 & i u_1 & \dots & i u_{N-1} & i u_N & 0 \end{pmatrix}, \quad (\text{A11})$$

with

$$a_N = 1, \quad (\text{A12})$$

$$a_j = 2 - 2\epsilon\gamma + \epsilon^2\gamma^2 - \epsilon^2\alpha(\alpha-1)\bar{\beta}\bar{x}_j^{\alpha-2}, \quad (\text{A13})$$

$$a_0 = 1 + \epsilon^2\gamma^2 - \epsilon^2\alpha(\alpha-1)\bar{\beta}\bar{x}_0^{\alpha-2}, \quad (\text{A14})$$

$$b_j = 1 - \epsilon\gamma, \tag{A15}$$

$$u_j = \epsilon\alpha\bar{x}_j^{\alpha-1}/T, \tag{A16}$$

$$u_N = 0, \tag{A17}$$

recalling that $\beta = \frac{iq}{T}$. The matrix elements are rescaled with ϵ such that the determinant

$$D_0 = - \lim_{\substack{\epsilon \rightarrow 0 \\ N \rightarrow \infty}} \begin{vmatrix} a_0 & -b_0 & 0 & \cdots & 0 & u_0 \\ -b_0 & a_1 & -b_1 & \cdots & 0 & u_1 \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & -b_{N-1} & a_N & u_N \\ u_0 & u_1 & \cdots & u_{N-1} & u_N & 0 \end{vmatrix} \tag{A18}$$

exists in the continuous limit, so that (A9) becomes

$$p_T(a) \sim \frac{e^{-S[\bar{x}]/\sigma^2}}{Z\sqrt{D_0}}. \tag{A19}$$

To arrive at this result, which differs from (18) by the added Z term, we have retrieved the Hessian in (A10) from the matrix in (A18) by multiplying the first $N + 1$ rows with $1/\epsilon$, the last column with ϵ , and the last row and column with i .

To perform the continuous limit for D_0 , we define the minor

$$C_j = \begin{vmatrix} a_j & -b_j & 0 & \cdots & 0 & u_j \\ -b_j & a_{j+1} & -b_{j+1} & \cdots & 0 & u_{j+1} \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & -b_{N-1} & a_N & u_N \\ u_j & u_{j+1} & u_{j+2} & u_{N-1} & u_N & 0 \end{vmatrix}, \tag{A20}$$

which results from dropping the first j rows and columns, and, similarly, the two auxiliary minors

$$B_j = \begin{vmatrix} -b_j & 0 & 0 & \cdots & 0 & u_j \\ a_{j+1} & -b_{j+1} & 0 & \cdots & 0 & u_{j+1} \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & -b_{N-2} & a_{N-1} & -b_{N-1} & u_{N-1} \\ 0 & \cdots & 0 & -b_{N-1} & a_N & u_N \end{vmatrix} \tag{A21}$$

and

$$A_j = \begin{vmatrix} a_j & -b_j & 0 & \cdots & 0 \\ -b_j & a_{j+1} & -b_{j+1} & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & -b_{N-2} & a_{N-1} & -b_{N-1} \\ 0 & \cdots & 0 & -b_{N-1} & a_N \end{vmatrix}. \tag{A22}$$

We expand the determinant C_j in the following way:

$$C_j = a_j C_{j+1} + b_j \begin{vmatrix} -b_j & -b_{j+1} & 0 & \cdots & 0 & u_{j+1} \\ 0 & a_{j+2} & -b_{j+2} & \cdots & 0 & u_{j+2} \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & -b_{N-1} & a_N & u_N \\ u_j & u_{j+2} & u_{j+3} & \cdots & u_N & 0 \end{vmatrix} \tag{A23}$$

$$-(-1)^{N-j} u_j \begin{vmatrix} -b_j & a_{j+1} & -b_{j+1} & 0 & \cdots & 0 \\ 0 & -b_{j+1} & a_{j+2} & -b_{j+2} & \cdots & 0 \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & -b_{N-1} & a_N \\ u_j & u_{j+1} & \cdots & u_{N-2} & u_{N-1} & u_N \end{vmatrix} \tag{A24}$$

$$= a_j C_{j+1} + b_j [-b_j C_{j+2} + (-1)^{N-j} u_j B_{j+1}] \tag{A23}$$

$$-(-1)^{N-j} u_j [-b_j B_{j+1}^T + (-1)^{N-j} u_j A_{j+1}] \tag{A24}$$

to arrive at the recursion formula,

$$C_j = a_j C_{j+1} - b_j^2 C_{j+2} + 2(-1)^{N-j} u_j b_j B_{j+1} - u_j^2 A_{j+1}. \tag{A25}$$

Similarly, we find for the two auxiliary minors the recursion formulas,

$$B_j = -b_j B_{j+1} + (-1)^{N-j} u_j A_{j+1} \quad (\text{A26})$$

and

$$A_j = a_j \cdot A_{j+1} - b_j^2 \cdot A_{j+2}. \quad (\text{A27})$$

Together with the final conditions

$$C_N = u_N^2 = 0, \quad (\text{A28})$$

$$C_{N+1} = 0, \quad (\text{A29})$$

$$B_N = u_N = 0, \quad (\text{A30})$$

$$A_{N-1} = a_{N-1} a_N - b_{N-1}^2, \quad (\text{A31})$$

$$A_N = a_N, \quad (\text{A32})$$

we can iterate backwards to obtain the full fluctuation determinant C_0 in the discretized approximation.

The fluctuation determinant D_0 is obtained by turning the recursion formulas (A25)–(A27) into differential equations. Special care must be taken to ensure convergence.

As a first step, we eliminate the alternating factor $(-1)^{N-j}$ by the replacement $B_j \mapsto (-1)^{N-j} B_j$, obtaining

$$C_j = a_j C_{j+1} - b_j^2 C_{j+2} - 2u_j b_j B_{j+1} - u_j^2 A_{j+1}, \quad (\text{A33})$$

$$B_j = b_j B_{j+1} + u_j A_{j+1}. \quad (\text{A34})$$

Plugging in the coefficients a_j , b_j , and u_j and inspecting the order in ϵ , it turns out that ϵC_j is of the order of $O(1)$. Multiplying (A25) by ϵ and rearranging, we get

$$\begin{aligned} & \frac{(\epsilon C_{j+2}) - 2(\epsilon C_{j+1}) + (\epsilon C_j)}{\epsilon^2} \\ &= 2\gamma \frac{(\epsilon C_{j+2}) - (\epsilon C_{j+1})}{\epsilon} - \alpha(\alpha - 1) \bar{\beta} \bar{x}_j^{\alpha-2} (\epsilon C_{j+1}) \\ & \quad - \frac{2\alpha}{T} \bar{x}_j^{\alpha-1} B_{j+1} + O(\epsilon). \end{aligned} \quad (\text{A35})$$

Similarly, we can rearrange (A26) and (A27) to obtain

$$\frac{B_{j+1} - B_j}{\epsilon} = \gamma B_{j+1} - \frac{\alpha}{T} \bar{x}_j^{\alpha-1} A_{j+1} \quad (\text{A36})$$

and

$$\begin{aligned} & \frac{(A_{j+2}) - 2(A_{j+1}) + (A_j)}{\epsilon^2} \\ &= 2\gamma \frac{A_{j+2} - A_{j+1}}{\epsilon} - \alpha(\alpha - 1) \bar{\beta} \bar{x}_j^{\alpha-2} A_{j+1} + O(\epsilon). \end{aligned} \quad (\text{A37})$$

Taking the continuous limit, we then recover the differential equations (19)–(21), noted earlier, with the final conditions (A28)–(A32).

Since the first (and last) row of H deviates from the other rows, as seen in (A11), we have

$$\lim_{\epsilon \rightarrow 0} \epsilon C_0 = 0, \quad (\text{A38})$$

and a final step is necessary to derive D_0 . Inspecting the last recursion step from (A25), we find

$$\begin{aligned} D_0 &= \lim_{\epsilon \rightarrow 0} (-a_0 C_1 + b_0^2 C_2 + 2u_0 b_0 B_1 - u_0^2 A_1) \\ &= \lim_{\epsilon \rightarrow 0} \left[\frac{(\epsilon C_2) - (\epsilon C_1)}{\epsilon} - 2\gamma \epsilon C_2 + O(\epsilon) \right] \end{aligned} \quad (\text{A39})$$

$$= \dot{C}(0) - 2\gamma C(0). \quad (\text{A40})$$

The solution for $C(t)$ can be plugged into the equation above to obtain D_0 . From a numerical perspective, however, it is better to use $\dot{D}(t) = \dot{C}(t) - 2\gamma \dot{C}(t)$ as in (22) to avoid cancellation of small numbers involving $\dot{C}(0)$.

APPENDIX B: CALCULATION OF INSTANTON VARIANCE EQUATIONS

The basis of the Gaussian correction is the Taylor expansion of the action around the instanton, which defines a multivariate Gaussian distribution in discrete time,

$$\begin{aligned} p_N(a) &\sim e^{-\frac{1}{2\sigma^2} S_N(\bar{x})} \int \prod_{i=0}^{N+1} dx_i \\ &\times \exp \left[-\frac{1}{2\sigma^2} \sum_{kl} (x_k - \bar{x}_k) H_{kl} (x_l - \bar{x}_l) \right]. \end{aligned} \quad (\text{B1})$$

Here, we have dropped the normalization constants and use $x_{N+1} = \beta$. From this expression, we see that the instanton given by the components \bar{x}_j in time is the mean vector of the multivariate Gaussian, while the Hessian gives the inverse covariance matrix,

$$\Sigma = \sigma^2 H^{-1}, \quad (\text{B2})$$

describing the Gaussian fluctuations about the instanton. As a result, it is natural to define the variance of the instanton as

$$\bar{v}(t_k) = \Sigma_{kk}. \quad (\text{B3})$$

To find the diagonal elements of H^{-1} , we use Cramer's rule,

$$(H^{-1})_{kl} = \frac{\det H_{(kl)}}{\det H}, \quad (\text{B4})$$

where $H_{(kl)}$ denotes the matrix that results from dropping the k th row and l th column in H . Expressing this element, as defined in (A11), in terms of the matrix

$$C = \begin{pmatrix} a_0 & -b_0 & 0 & \cdots & 0 & u_0 \\ -b_0 & a_1 & -b_1 & \cdots & 0 & u_1 \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & -b_{N-1} & a_N & u_N \\ u_0 & u_1 & \cdots & u_{N-1} & u_N & 0 \end{pmatrix}, \quad (\text{B5})$$

which underlies (A20), we can rewrite (B4) as

$$(H^{-1})_{kl} = \frac{\epsilon \det C_{(kl)}}{\det C}. \quad (\text{B6})$$

To further simplify $\det C_{(kl)}$, we focus on the diagonal elements ($k = l$) and write, in block form,

$$\det C_{(kk)} = \det \left(\begin{array}{c|c|c} A_{0,k-1} & 0 & U_{0,k-1} \\ \hline 0 & A_{k+1,N} & U_{k+1,N} \\ \hline U_{0,k-1} & U_{k+1,N} & 0 \end{array} \right), \quad (\text{B7})$$

with

$$A_{0,k-1} = \begin{pmatrix} a_0 & -b_0 & 0 & \cdots & 0 \\ -b_0 & a_1 & -b_1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & -b_{k-3} & a_{k-2} & -b_{k-2} \\ 0 & \cdots & 0 & -b_{k-2} & a_{k-1} \end{pmatrix}, \quad (\text{B8})$$

$$A_{k+1,N} = \begin{pmatrix} a_{k+1} & -b_{k+1} & 0 & \cdots & 0 \\ -b_{k+1} & a_{k+2} & -b_{k+2} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & -b_{N-2} & a_{N-1} & -b_{N-1} \\ 0 & \cdots & 0 & -b_{N-1} & a_N \end{pmatrix}, \quad (\text{B9})$$

$$U_{0,k-1} = (u_0, \dots, u_{k-1}), \quad (\text{B10})$$

$$U_{k+1,N} = (u_{k+1}, \dots, u_N). \quad (\text{B11})$$

Making use of the Schur complement, we can write

$$\det C_{(kk)} = -\det A_{0,k-1} \cdot \det A_{k+1,N} \in (U_{0,k-1}^T A_{0,k-1}^{-1} U_{0,k-1} + U_{k+1,N}^T A_{k+1,N}^{-1} U_{k+1,N}). \quad (\text{B12})$$

Recognizing $\det A_{k+1,N}$ as A_{k+1} in (A22), and noting that in the continuous limit the initial conditions for the forward determinant starting at $t = 0$ are the same as in (A32) for $t = T$, we find that

$$\det A_{0,k-1} \det A_{k+1,N} \rightarrow A(T-t) \cdot A(t) \quad (\text{B13})$$

for $\epsilon \rightarrow 0$.

For the quadratic form $\in U_{0,k-1}^T A_{0,k-1}^{-1} U_{0,k-1}$, we define an auxiliary vector,

$$r_1 = \epsilon A_{0,k-1}^{-1} U_{0,k-1}, \quad (\text{B14})$$

and similarly define

$$r_2 = \epsilon A_{k+1,N}^{-1} U_{k+1,N} \quad (\text{B15})$$

for the second quadratic form. Knowing r_1 and r_2 , we obtain the value of the quadratic form via the dot product

$$R_k = U_{0,k-1}^T \cdot r_1 + U_{k+1,N}^T \cdot r_2. \quad (\text{B16})$$

To get r_1 , we multiply (B14) with $\frac{1}{\epsilon} A_{0,k-1}$ from the left and obtain a linear set of equations,

$$A_{0,k-1} r = \epsilon U_{0,k-1}, \quad (\text{B17})$$

where we temporarily dropped the index of r_1 . Plugging in the coefficients (A12)–(A17) and rearranging, we find

$$\frac{r_{j+1} - 2r_j + r_{j-1}}{\epsilon^2} = [\gamma^2 - \alpha(\alpha - 1) \bar{\beta} \bar{x}_j^{\alpha-2}] r_j - \frac{\alpha}{T} \bar{x}_j^{\alpha-1}, \quad (\text{B18})$$

up to terms of the order of ϵ , which becomes the differential equation

$$\ddot{r}(\tau) = [\gamma^2 - \alpha(\alpha - 1) \bar{\beta} \bar{x}(\tau)^{\alpha-2}] r(\tau) - \frac{\alpha}{T} \bar{x}(\tau)^{\alpha-1} \quad (\text{B19})$$

in the limit $\epsilon \rightarrow 0$. This differential equation is completed by the two boundary conditions

$$\gamma r_1(0) - \dot{r}_1(0) = 0, \quad r_1(T) = 0, \quad (\text{B20})$$

resulting from evaluating the first and last equations of (B17).

The same steps apply to r_2 in (B15), leading to the same differential equation as in (B19), but with the boundary conditions

$$r_2(t) = 0, \quad \gamma r_2(T) + \dot{r}_2(T) = 0. \quad (\text{B21})$$

To obtain $R(t)$, we then insert the solutions $r_1(\tau)$ and $r_2(\tau)$ into the integrals that result from taking the continuous limit of (B16),

$$R(t) = \frac{\alpha}{T} \left[\int_0^t d\tau \bar{x}(\tau)^{\alpha-1} r_1(\tau) + \int_t^T d\tau \bar{x}(\tau)^{\alpha-1} r_2(\tau) \right]. \quad (\text{B22})$$

For large T , the two solutions $r_1(\tau)$ and $r_2(T - \tau)$ approach each other (as boundary values approach zero) and (B22) becomes

$$R(t) = \frac{\alpha}{T} \int_0^T d\tau \bar{x}(\tau)^{\alpha-1} r(\tau). \quad (\text{B23})$$

As a result, the determinants in (B12) become equal in the continuous limit to $-R(t)A(t)A(T-t)$. Together with D_0 from (A40) as the continuous limit of the determinant $\det C$, we finally arrive at

$$\bar{v}(t) = \frac{\sigma^2}{D_0} R(t)A(t)A(T-t). \quad (\text{B24})$$

-
- [1] D. Nickelsen and H. Touchette, Anomalous Scaling of Dynamical Large Deviations, *Phys. Rev. Lett.* **121**, 090602 (2018).
- [2] P. L. Krapivsky, K. Mallick, and T. Sadhu, Large Deviations in Single-File Diffusion, *Phys. Rev. Lett.* **113**, 078101 (2014).
- [3] T. Sadhu and B. Derrida, Large deviation function of a tracer position in single file diffusion, *J. Stat. Mech.* (2015) P09008.
- [4] T. Imamura, K. Mallick, and T. Sasamoto, Large Deviations of a Tracer in the Symmetric Exclusion Process, *Phys. Rev. Lett.* **118**, 160601 (2017).
- [5] P. Le Doussal, S. N. Majumdar, and G. Schehr, Large deviations for the height in 1D Kardar-Parisi-Zhang growth at late times, *Europhys. Lett.* **113**, 60004 (2016).
- [6] P. Sasorov, B. Meerson, and S. Prolhac, Large deviations of surface height in the 1 + 1-dimensional Kardar-Parisi-Zhang equation: Exact long-time results for $\lambda H < 0$, *J. Stat. Mech.* (2017) 063203.
- [7] I. Corwin, P. Ghosal, A. Krajenbrink, P. Le Doussal, and L.-C. Tsai, Coulomb-Gas Electrostatics Controls Large Fluctuations of the Kardar-Parisi-Zhang Equation, *Phys. Rev. Lett.* **121**, 060201 (2018).

- [8] J. T. Cox and D. Griffeath, Occupation times for critical branching Brownian motions, *Ann. Probab.* **13**, 1108 (1985).
- [9] O. Luidor and W. Perkins, Large deviations for the empirical distribution in the branching random walk, *Electron. J. Probab.* **20**, 18 (2015).
- [10] B. Derrida and Z. Shi, Slower deviations of the branching Brownian motion and of branching random walks, *J. Phys. A: Math. Theor.* **50**, 344001 (2017).
- [11] A. Dembo, Y. Peres, and O. Zeitouni, Tail estimates for one-dimensional random walk in random environment, *Commun. Math. Phys.* **181**, 667 (1996).
- [12] N. Gantert and O. Zeitouni, Quenched sub-exponential tail estimates for one-dimensional random walk in random environment, *Commun. Math. Phys.* **194**, 177 (1998).
- [13] O. Zeitouni, Random walks in random environments, *J. Phys. A: Math. Gen.* **39**, R433(R) (2006).
- [14] R. J. Harris and H. Touchette, Current fluctuations in stochastic systems with long-range memory, *J. Phys. A: Math. Theor.* **42**, 342001 (2009).
- [15] K. R. Duffy and S. P. Meyn, Most likely paths to error when estimating the mean of a reflected random walk, *Perform. Eval.* **67**, 1290 (2010).
- [16] J. Blanchet, P. Glynn, and S. Meyn, Large deviations for the empirical mean of an $m/m/1$ queue, *Queue. Syst.* **73**, 425 (2013).
- [17] K. R. Duffy and S. P. Meyn, Large deviation asymptotics for busy periods, *Stoch. Syst.* **4**, 300 (2014).
- [18] M. Bazhba, J. Blanchet, C.-H. Rhee, and B. Zwart, Sample path large deviations for unbounded additive functionals of the reflected random walk, [arXiv:2003.14381](https://arxiv.org/abs/2003.14381).
- [19] M. Bazhba, J. Blanchet, R. J. A. Laeven, and B. Zwart, Large deviations asymptotics for unbounded additive functionals of diffusion processes, [arXiv:2202.10799](https://arxiv.org/abs/2202.10799).
- [20] R. Chetrite and H. Touchette, Nonequilibrium Microcanonical and Canonical Ensembles and Their Equivalence, *Phys. Rev. Lett.* **111**, 120601 (2013).
- [21] R. Chetrite and H. Touchette, Nonequilibrium Markov processes conditioned on large deviations, *Ann. Henri Poincaré* **16**, 2005 (2015).
- [22] R. Chetrite and H. Touchette, Variational and optimal control representations of conditioned and driven processes, *J. Stat. Mech.* (2015) P12001.
- [23] I. M. Gel'fand and A. M. Yaglom, Integration in functional spaces and its applications in quantum physics, *J. Math. Phys.* **1**, 48 (1960).
- [24] N. Berglund, Kramer's law: Validity, derivations, and generalisations, *Markov Proc. Relat. Fields* **19**, 459 (2013).
- [25] Y. Lu, A. Stuart, and H. Weber, Gaussian approximations for transition paths in Brownian dynamics, *SIAM J. Math. Anal.* **49**, 3005 (2017).
- [26] G. Corazza and M. Fadel, Normalized Gaussian path integrals, *Phys. Rev. E* **102**, 022135 (2020).
- [27] G. Ferré and T. Grafke, Approximate optimal controls via instanton expansion for low temperature free energy computation, *Multiscale Model. Simul.* **19**, 1310 (2021).
- [28] T. Schorlepp, T. Grafke, and R. Grauer, Gel'fand–Yaglom type equations for calculating fluctuations around instantons in stochastic systems, *J. Phys. A: Math. Theor.* **54**, 235003 (2021).
- [29] T. Grafke, T. Schäfer, and E. Vanden Eijnden, Sharp asymptotics estimates for expectations, probabilities, and mean-first passage times in stochastic systems with small noise, [arXiv:2103.04837](https://arxiv.org/abs/2103.04837).
- [30] A. Engel, Asymptotics of work distributions in nonequilibrium systems, *Phys. Rev. E* **80**, 021120 (2009).
- [31] D. Nickelsen and A. Engel, Asymptotics of work distributions: The pre-exponential factor, *Eur. Phys. J. B* **82**, 207 (2011).
- [32] P. Pietzonka, E. Zimmermann, and U. Seifert, Fine-structured large deviations and the fluctuation theorem: Molecular motors and beyond, *Europhys. Lett.* **107**, 20002 (2014).
- [33] V. R. Fatalov, On the Laplace method for Gaussian measures in a Banach space, *Theor. Probab. Appl.* **58**, 216 (2014).
- [34] L. Onsager and S. Machlup, Fluctuations and irreversible processes, *Phys. Rev.* **91**, 1505 (1953).
- [35] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Grundlehren der Mathematischen Wissenschaften (Springer, New York, 1984), Vol. 260.
- [36] H. Touchette, The large deviation approach to statistical mechanics, *Phys. Rep.* **478**, 1 (2009).
- [37] B. Meerson, Anomalous scaling of dynamical large deviations of stationary Gaussian processes, *Phys. Rev. E* **100**, 042135 (2019).
- [38] We use the SCIPY function `integrate.solve_bvp` to solve the Euler-Lagrange equation with boundary conditions on x_0 and x_T , and β as an additional variable fixing the constraint.
- [39] H. Touchette, A basic introduction to large deviations: Theory, applications, simulations, in *Modern Computational Science 11: Lecture Notes from the 3rd International Oldenburg Summer School*, edited by R. Leidl and A. K. Hartmann (BIS-Verlag der Carl von Ossietzky Universität Oldenburg, Oldenburg, 2011).
- [40] J. A. Bucklew, *Introduction to Rare Event Simulation* (Springer, New York, 2004).
- [41] S. Asmussen and P. W. Glynn, *Stochastic Simulation: Algorithms and Analysis*, Stochastic Modelling and Applied Probability (Springer, New York, 2007).
- [42] M. Cottrell, J.-C. Fort, and G. Malgouyres, Large deviations and rare events in the study of stochastic algorithms, *IEEE Trans. Automat. Contr.* **28**, 907 (1983).
- [43] P. Dupuis and H. Kushner, Stochastic systems with small noise, analysis and simulation; a phase locked loop example, *SIAM J. Appl. Math.* **47**, 643 (1987).
- [44] D. M. Zuckerman and T. B. Woolf, Dynamic reaction paths and rates through importance-sampled stochastic dynamics, *J. Chem. Phys.* **111**, 9475 (1999).
- [45] E. Vanden-Eijnden and J. Weare, Rare event simulation of small noise diffusions, *Comm. Pure Appl. Math.* **65**, 1770 (2012).
- [46] A. Guyader and H. Touchette, Efficient large deviation estimation based on importance sampling, *J. Stat. Phys.* **181**, 551 (2020).
- [47] For some low approximations of $\bar{x}(t)$, for example, we obtain a maximum variance of the order of 10^4 , whereas for the most accurate instanton that we have with 10^5 points we obtain a maximum variance of 400, which is the value reported in Fig. 5.
- [48] Estimators based on importance sampling are unbiased, but generally have different variance.
- [49] P. T. Nyawo and H. Touchette, Dynamical phase transition in drifted Brownian motion, *Phys. Rev. E* **98**, 052103 (2018).
- [50] In the case $\alpha = 1$, we find $D_0 \sim 1/T$, so the correction is in $\ln T$ in the exponent. For $\alpha > 2$, simulation results suggest a different scaling, namely, $D_0 \sim T^{-2/\alpha} a^{2-2/\alpha}$, which is still logarithmic in the exponential.

- [51] J. Hoppenau, D. Nickelsen, and A. Engel, Level 2 and level 2.5 large deviation functionals for systems with and without detailed balance, *New J. Phys.* **18**, 083010 (2016).
- [52] H. Touchette, Introduction to dynamical large deviations of Markov processes, *Physica A* **504**, 5 (2018).
- [53] N. R. Smith, Anomalous scaling and first-order dynamical phase transition in large deviations of the Ornstein-Uhlenbeck process, *Phys. Rev. E* **105**, 014120 (2022).
- [54] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers* (McGraw-Hill, New York, 1978).