


Anomalous scaling and first-order dynamical phase transition in large deviations of the Ornstein-Uhlenbeck process

Naftali R. Smith *

*CNRS-Laboratoire de Physique Théorique de l'Ecole Normale Supérieure, 24 Rue Lhomond, F-75231 Paris Cedex, France
and Department of Solar Energy and Environmental Physics, Blaustein Institutes for Desert Research,
Ben-Gurion University of the Negev, Sede Boqer Campus, 8499000, Israel*



(Received 18 October 2021; accepted 6 January 2022; published 19 January 2022)

We study the full distribution of $A = \int_0^T x^n(t)dt$, $n = 1, 2, \dots$, where $x(t)$ is an Ornstein-Uhlenbeck process. We find that for $n > 2$ the long-time ($T \rightarrow \infty$) scaling form of the distribution is of the anomalous form $P(A; T) \sim e^{-T^\mu f_n(\Delta A/T^\nu)}$ where ΔA is the difference between A and its mean value, and the anomalous exponents are $\mu = 2/(2n - 2)$ and $\nu = n/(2n - 2)$. The rate function $f_n(y)$, which we calculate exactly, exhibits a first-order dynamical phase transition which separates between a homogeneous phase that describes the Gaussian distribution of typical fluctuations, and a “condensed” phase that describes the tails of the distribution. We also calculate the most likely realizations of $\mathcal{A}(t) = \int_0^t x^n(s)ds$ and the distribution of $x(t)$ at an intermediate time t conditioned on a given value of A . Extensions and implications to other continuous-time systems are discussed.

DOI: [10.1103/PhysRevE.105.014120](https://doi.org/10.1103/PhysRevE.105.014120)

I. INTRODUCTION

The study of fluctuations in stochastic systems is of central importance in nonequilibrium statistical mechanics and probability theory. Rare events, or large deviations, are of particular interest [1–13]. A standard way to characterize fluctuations in dynamical systems is by considering distributions of “dynamical observables” which are given by integrating the stochastic process over time,

$$A = \int_0^T u(X(t), \dot{X}(t))dt, \quad (1)$$

where $X(t)$ is the stochastic process and $u(X, \dot{X})$ is an arbitrary function. For ergodic systems, in the long-time limit $T \rightarrow \infty$ the time average A/T converges to its ensemble-average value, but the fluctuations from this value depend on the temporal correlations of the process, and can exhibit nonequilibrium features even if the system is in equilibrium. A general theory for the study of such fluctuations was developed, based on the Feynman-Kac formula [5–7, 13–15]. This theory is sometimes referred to as the Donsker-Varadhan (DV) theory. Under quite general conditions, capturing a broad class of physical systems, the theory predicts that at long times the probability density function (PDF) $P(A; T)$ of A obeys a large-deviation principle,

$$P(A; T) \sim e^{-TI(A/T)}, \quad T \rightarrow \infty, \quad (2)$$

i.e., the limit $-\lim_{T \rightarrow \infty} \ln P(aT^\nu; T)/T^\mu = I(a)$, with the standard exponents $\mu = \nu = 1$ exists with a “rate function” $I(a)$ that is nonnegative, convex, and vanishes when its argument $a = A/T$ equals its corresponding ensemble-average value. Thus, $I(a)$ quantifies large deviations of a which

become exponentially unlikely in T . The rate function is calculated from the dominant eigenvalue of the Feynman-Kac equation for the generating function of A . The calculation boils down to solving an auxiliary problem of finding the largest eigenvalue of a “tilted operator,” and then applying a Legendre-Fenchel transform to the result. In particular cases, the DV auxiliary problem can be cast as that of finding the ground-state energy of a quantum system consisting of a particle in a potential well [13].

However, the scaling (2) has recently been observed to break down in numerous instances [16–19], where “anomalous” scalings were found. Nickelsen and Touchette (NT) [16] considered an Ornstein-Uhlenbeck (OU) process,

$$\dot{x}(t) = -\gamma x(t) + \sigma \eta(t). \quad (3)$$

Here $x(t)$ is the position of the particle at time t , $\gamma > 0$ is the damping, $\eta(t)$ is white noise with $\langle \eta(t) \rangle = 0$ and $\langle \eta(t)\eta(t') \rangle = \delta(t - t')$ where angular brackets denote ensemble averaging, and $\sigma > 0$ is the noise intensity. They studied the distribution of the observable [20],

$$A = \int_0^T x^n(t)dt, \quad (4)$$

where $n = 1, 2, \dots$. For $n = 1$, A can represent the work that laser tweezers perform when pulling on a Brownian particle [21], or the power dissipated in a noisy circuit [22]. For $n = 2, 3$ and higher, A is related to studies of fluctuations in turbulence velocity fields and small-scale intermittency [23, 24].

NT probed the weak-noise limit $\sigma \rightarrow 0^+$ by employing the optimal fluctuation method (OFM) (sometimes also called the weak noise theory) [25–28]. They found that for $n > 2$, the scaling (2) breaks down and instead one observes an

*naftalismsith@gmail.com

anomalous scaling,

$$\mathcal{P}(A; T, \sigma, \gamma) \sim \exp \left[-c_n \frac{\gamma^{(n+2)/n} A^{2/n}}{\sigma^2} \right], \quad \sigma \rightarrow 0^+, \quad (5)$$

where c_n is an n -dependent constant that they calculated exactly. The physical mechanism behind this result is that the dominant contribution to $\mathcal{P}(A; T)$ comes from an instanton—an optimal (most likely) trajectory $x(t)$. The instanton is the minimizer of the dynamical action $\frac{1}{2} \int_0^T [\dot{x}(t) + x(t)]^2 dt$ constrained on a given value of A , and it was found explicitly in [18] [for completeness, we give the instanton in (C3)]. The instanton is (temporally) localized, that is, $x(t) \simeq 0$ at all times except for a short time window around $t = t_*$ for some $0 < t_* < T$. The weak-noise limit $\sigma \rightarrow 0^+$ corresponds to the far tail $|A| \rightarrow \infty$ of the distribution [18]. However, the full long-time distribution $\mathcal{P}(A; T)$ —including typical fluctuations in addition to the distribution tails—has been unknown, and its calculation is the main result of the current work.

We first point out that a similar anomaly can occur when considering distributions $P(S; N)$ of sums $S = \sum_{i=1}^N x_i$ of independent and identically distributed (i.i.d) random variables x_1, \dots, x_N , in the limit $N \gg 1$. Here, the usual scaling is $P(S; N) \sim e^{-NI(S/N)}$ in analogy with (2). However, if the distribution tails of each of the x_i 's decays slower than exponentially, the usual scaling breaks down [29], and in the far tail of $P(S; N)$ the “big-jump” principle holds: The dominant contribution to $P(S; N)$ comes from realizations where $S \simeq x_i$ for some $1 \leq i \leq N$ [30–34]. The big-jump principle has been observed in a wide range of systems including anomalous transport in quenched disorder [35–37] and Lévy walks [37–40]. A “condensation” transition that separates between the typical-fluctuation regime, where the central limit theorem applies, and the distribution tail(s) where the big-jump principle applies, is a general phenomenon that has been observed in many instances [41,42]. Examples include the zero-range process [43,44], the discrete nonlinear Schrödinger equation [45–49], economic and financial models [50–52], mass-transport models [53–61], and run-and-tumble active particles [62–64]. Above a critical point in the tail of $P(S; N)$, a condensate appears meaning that one of the x_i 's contributes a macroscopic fraction to S . In the far tail this fraction approaches unity, so the big-jump principle is recovered. The condensation transition was shown to be universal for i.i.d random variables whose distribution decays slower than exponentially [64,65].

We observe a striking similarity between the big jump in discrete-time systems and the instanton of NT in the continuous-time system [16]. Both are localized events that dominate the contribution to the observable in question in the far distribution tail. It is therefore appealing to search for a condensation transition in the distribution $\mathcal{P}(A; T)$. Indeed, since the stationary distribution of the OU process is Gaussian, we can gain intuition by considering an analogous discrete-time problem of the distribution of $\sum_{i=1}^N x_i^n$ where x_1, \dots, x_N are i.i.d. Gaussian random variables. This analogous problem exhibits a condensation transition, as can be shown using the general results of [63–65]. As we now show, such a transition is indeed present in our system, too.

II. RESCALING AND SUMMARY OF MAIN RESULTS

We begin by rescaling $\gamma t \rightarrow t$, $x\sqrt{\gamma}/\sigma \rightarrow x$, leading to the rescaled Langevin equation $\dot{x} = -x + \eta$ and to the (exact) scaling form,

$$\mathcal{P}(A; T, \sigma, \gamma) = \frac{\gamma^{(n+2)/2}}{\sigma^n} P \left(\frac{\gamma^{(n+2)/2} A}{\sigma^n}; \gamma T \right), \quad (6)$$

of the distribution, where $P(A; T)$ is dimensionless and so are its (rescaled) arguments. The weak-noise limit $\sigma \rightarrow 0^+$ is mathematically equivalent to the limit $|A| \rightarrow \infty$ in $P(A; T)$ [18]. Therefore, the weak-noise results of [16] describe the far tail(s) $|A| \rightarrow \infty$ of the distribution.

Let us state the main result of this paper. We study, for $n > 2$, the distribution $P(A; T)$ in the long-time limit $T \gg 1$. For $A > \langle A \rangle$ for even n , and all A for odd n , we find that the distribution obeys a large-deviation principle with anomalous scaling exponents $\mu = 2/(2n - 2)$, and $\nu = n/(2n - 2)$:

$$P(A; T) \sim \exp \left[-T^{2/(2n-2)} f_n \left(\frac{\Delta A}{T^{n/(2n-2)}} \right) \right], \quad (7)$$

where $\Delta A = A - \langle A \rangle$,

$$f_n(y) = \min_{z \in [0, y]} F_n(y, z), \quad (8)$$

$$F_n(y, z) = c_n z^{2/n} + \beta_n (y - z)^2, \quad (9)$$

c_n was calculated in [16] [and for completeness, is given in Eq. (C2)] and we calculate β_n [see Eq. (11)] through a perturbative treatment of the DV auxiliary quantum problem; see Appendix B for the details. For instance, for $n = 3$, $c_3 = (9/10)^{1/3}$ [16] and we find $\beta_3 = 2/11$, whereas for $n = 4$, $c_4 = 2/\sqrt{3}$ [16] and we find $\langle A \rangle \simeq 3T/4$ and $\beta_4 = 2/21$. The functions $f_3(y)$ and $F_3(y, z)$ are plotted in Fig. 1. The result (7)–(9) is valid in the limit $T \rightarrow \infty$ with constant $\Delta A/T^{n/(2n-2)}$. Remarkably, the first derivative of the rate function $f_n(y)$ has a discontinuity that can be interpreted as a first-order dynamical phase transition. For example, for $n = 3$ this transition occurs at the critical values $y = \pm y_c$ where $y_c(n = 3) = \frac{11^{3/4}}{15^{1/4}} = 3.069\dots$, and for $n = 4$ the transition occurs at $y = y_c(n = 4) = \frac{3^{4/3} 7^{2/3}}{2^{5/3}} = 4.987\dots$. In the sub-critical regime $|y| < y_c$ the rate function is purely parabolic $f_n(y) = \beta_n y^2$, describing a Gaussian distribution of typical fluctuations. Interestingly, $f_n(y)$ is nonconvex.

As pointed out in [16], for even $n > 2$, DV theory is valid at $0 < A < \langle A \rangle$ [66], so the scaling (2) holds. We now turn to the derivation of the results (7)–(9). We treat only the anomalous case $n > 2$, as for $n \leq 2$ the scaling (2) holds at all A , and $I(a)$ is found from DV theory [16]. Our strategy in the derivation is first to use a perturbative DV approach in order to treat the regime of typical fluctuations, and then to use this result in conjunction with the result for the tail $|A| \rightarrow \infty$ which is known from [16]. Remarkably, this enables us to extract the entire intermediate regime $\Delta A \sim T^{n/(2n-2)}$. This is achieved by exploiting the separation of time scales between the duration of the instanton and the (much longer) duration T of the entire dynamics.

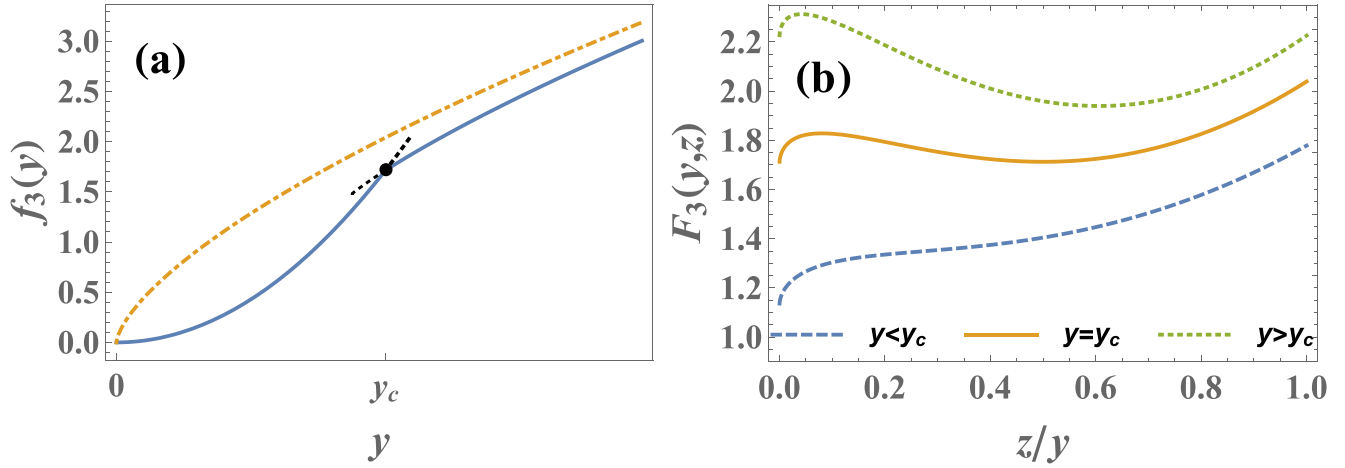


FIG. 1. (a) Solid line is the rate function $f_3(y)$ that describes the full distribution $P(A; T)$ for $n = 3$; see Eqs. (7)–(9). Dotted lines are the continuations of the two branches of $f(y)$ into the regimes in which they are not optimal. Dot-dashed line is the instanton’s action $c_3 y^{2/3}$ which is strictly larger than $f(y)$, but gives the correct leading asymptotic behavior in the far tail of the distribution $y \gg 1$. (b) $F_3(y, z)$ as a function of z/y for three different values of y : subcritical $y = 5/2$ (dashed), critical $y = y_c$ (solid), and supercritical $y = 7/2$ (dotted). In the subcritical regime, the minimum of $F_3(y, z)$ is at $z = 0$ whereas in the supercritical regime it is at $z = z_*(y) \neq 0$ given in Eq. (16).

III. TYPICAL FLUCTUATIONS

We begin by considering the regime $|\Delta A| \ll T^{n/(2n-2)}$ which includes typical fluctuations. The quantum potential whose ground-state energy one must calculate for a particle of unit mass (in units where $\hbar = 1$) when using the DV method is

$$V_k(x) = \frac{x^2}{2} - \frac{1}{2} - kx^n, \quad (10)$$

where k is the tilt parameter and is related to the DV rate function $I(a)$ through a Legendre-Fenchel transform. As pointed out in [16], the potential (10) has no ground state due to the x^n term (for $n > 2$), signaling that the DV method breaks down. However, we notice that as $|k| \ll 1$, the potential is effectively confining for $|x| \ll 1/|k|$, and its effective ground-state energy $-\lambda(k)$ can be found perturbatively in k . Using second-order perturbation theory, we find (see Appendix B) $\lambda(k) = \alpha_n k + k^2/4\beta_n + \dots$, where

$$\alpha_n = \langle 0|x^n|0\rangle, \quad \beta_n = \left(4 \sum_{m=1}^{\infty} \frac{|\langle m|x^n|0\rangle|^2}{E_m - E_0}\right)^{-1}. \quad (11)$$

Here the $|m\rangle$ ’s are the energy eigenstates of the unperturbed ($k = 0$) quantum oscillator with corresponding energies $E_m = m$. The Legendre-Fenchel transform then yields the quadratic approximation of the DV rate function around its minimum, $I(a) = \beta_n(a - \alpha_n)^2 + \dots$. This predicts Gaussian fluctuations,

$$P(A; T) \sim e^{-\beta_n(\Delta A)^2/T}, \quad (12)$$

around the mean value $\langle A \rangle \simeq \alpha_n T$ implying, in particular, that the variance of the distribution is $\text{Var}(A) \simeq T/2\beta_n$. Note that α_n is simply the n th moment of the Gaussian distribution $P_s(x) = e^{-x^2}/\sqrt{\pi}$, so $\alpha_n = 2^{n/2}\Gamma(\frac{n+1}{2})/\sqrt{\pi}$ in agreement with [16]. The prediction (12) shows excellent agreement with a computation of $P(A; T)$ using Monte Carlo simulations with $n = 3$ and $T = 100$; see Fig. 2(a). The simulations

were performed using an Itô discretization of the (rescaled) Langevin equation (3) with time steps of size 0.01. Finally, the Legendre-Fenchel transform also yields the connection $a = \alpha_n + k/2\beta_n + \dots$.

We now analyze the regime of validity of the perturbative DV result (12). The last term in the tilted potential (10) becomes of the same order as the first term at $x \sim 1/|k|^{1/(n-2)}$. One expects the DV calculation to be valid as long as the probability for the particle to reach this position is much smaller than $P(A; T)$ itself and therefore trajectories which reach such a position can be neglected when evaluating $P(A; T)$. Since the stationary distribution of the particle’s position (for the OU process) is $P_s(x) = e^{-x^2}/\sqrt{\pi}$, this validity condition yields $1/|k|^{2/(n-2)} \gg (\Delta A)^2/T$ which, using the connection $k \sim a - \alpha_n = \Delta A/T$, becomes $|\Delta A| \ll T^{n/(2n-2)}$. As we will show shortly, we found that the perturbative DV result (12) is actually valid in the entire subcritical regime $|\Delta A|/T^{n/(2n-2)} \leq y_c$. Finally, in this regime one easily checks that $|k| \ll 1$ (since $|\Delta A| \ll T$), justifying the use of perturbation theory above.

IV. FULL DISTRIBUTION

A. Mixed scenario

We now have $P(A; T)$ at hand in two distinct regimes: in the far tail(s) $|A| \rightarrow \infty$, the result $P(A; T) \sim e^{-c_n A^{2/n}}$ of [16], and for typical fluctuations, the perturbative DV result (12). As we show below, the transition between the two regimes occurs at $\Delta A \sim T^{n/(2n-2)}$ where the two results predict probabilities of the same order. In order to study the full distribution, it is therefore useful to take the scaling limit $T \rightarrow \infty$ with $\Delta A/T^{n/(2n-2)}$ constant. The two regimes correspond to two different physical mechanisms for creating the large deviation. The instanton is a localized burst where the dominant contribution to the integral (4) comes from a narrow temporal window. Defining

$$\Delta A(t) = \int_0^t [x^n(s) - \langle x^n(s) \rangle] ds \quad (13)$$

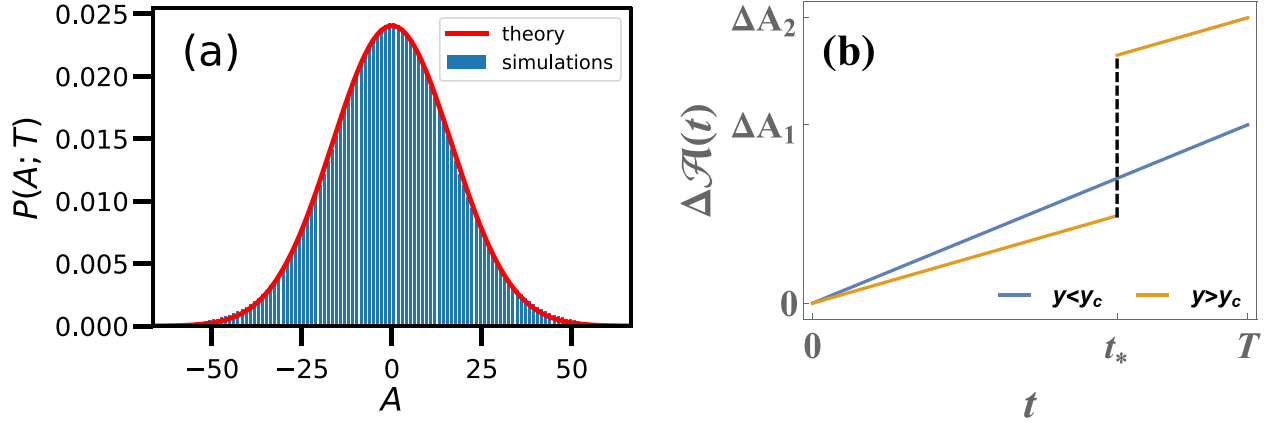


FIG. 2. (a) Bars represent a computation of $P(A; T)$ for $n = 3$ and $T = 100$ over 10^9 Monte Carlo simulations. Solid line is the Gaussian asymptotic (12) with the normalization factor accounted for. (b) An illustration of the optimal $\Delta\mathcal{A}(t)$ from (14) for two different values of ΔA : subcritical $0 < \Delta A_1 < y_c T^{3/4}$ for which $\Delta A_{\text{ins}} = 0$, and supercritical $\Delta A_2 > y_c T^{3/4}$ for which $\Delta A_{\text{ins}} \neq 0$. In the supercritical case, zooming in around the time t_* where the instanton occurs, the discontinuity in (14) is smoothed out over a time scale of order unity, in which the corresponding trajectory $x(t)$ is described by the instanton [16,18].

[so $\Delta A = \Delta\mathcal{A}(T)$], the OFM (weak-noise) prediction corresponds to trajectories for which $\Delta\mathcal{A}(t) = \Delta A \theta(t - t_*)$ where $t_* \in [0, T]$ is the time when the instanton occurs and $\theta(t)$ is the Heaviside function. To the contrary, the DV prediction corresponds to the opposite scenario, where the contribution to the integral (4) is homogeneous throughout the entire dynamics because in the DV formalism, the conditioned process [i.e., the process $x(t)$ conditioned on a given value of A] becomes stationary in the large- T limit. This leads to the linear behavior $\Delta\mathcal{A}(t) = \Delta A t / T$.

The optimal scenario, however, can combine the two scenarios described above, as we now show. One part ΔA_{ins} of the integral (4) could come from an instanton and the remainder, $\Delta A_{\text{DV}} = \Delta A - \Delta A_{\text{ins}}$, could come from the DV mechanism, corresponding to

$$\Delta\mathcal{A}(t) = \Delta A_{\text{ins}} \theta(t - t_*) + (\Delta A - \Delta A_{\text{ins}}) t / T; \quad (14)$$

see Fig. 2(b). Let us now calculate the probability of this mixed scenario. A key observation is that, since the instanton is temporally localized while the DV mechanism is homogeneous over the entire duration of the dynamics, the two mechanisms (instanton and DV) work independently from each other, so that the probability of this mixed scenario is

$$\sim \exp \left[-c_n \Delta A_{\text{ins}}^{2/n} - \frac{\beta_n (\Delta A - \Delta A_{\text{ins}})^2}{T} \right]. \quad (15)$$

Importantly, for $\Delta A_{\text{ins}} \sim \Delta A \sim T^{n/(2n-2)}$, the two terms in the exponent in (15) are of the same order. Integrating the probability (15) over ΔA_{ins} while using the saddle-point approximation (see Appendix A for details), we obtain our main result reported above in Eqs. (7)–(9), describing a large-deviation principle with anomalous scaling and a rate function $f_n(y)$ which exhibits a first-order phase transition. The optimal value of ΔA_{ins} in (14) is $\Delta A_{\text{ins}} = z \Delta A / y$ where z is the minimizer in Eq. (8). In the subcritical regime $|y| < y_c$, the minimizer is at $z = 0$, so $f_n(y) = \beta_n y^2$ is exactly parabolic, and the system is in a (temporally) homogeneous state. In the

supercritical regime $|y| > y_c$, the minimizer is at a nonzero value $z = z_*$, with the requirement $\partial F_n / \partial z = 0$ giving $f_n(y)$ in a parametric form,

$$y = \frac{c_n}{n\beta_n} z_*^{(2-n)/n} + z_*, \quad f_n(y) = c_n z_*^{2/n} + \frac{c_n^2}{n^2 \beta_n} z_*^{2(2-n)/n}, \quad (16)$$

and the system is in a “condensed” state. For $n = 3$, these equations can in fact be solved to find $f_n(y)$ explicitly but the result is very cumbersome so we do not give it here (a similar calculation was performed in [63]). At $y = y_c$, $f_n(y)$ is continuous but its first derivative is not; see Fig. 1 and Appendix A. The asymptotic behavior (obtained in Appendix A),

$$f_n(|y| \gg 1) \simeq c_n y^{2/n} - \frac{c_n^2}{n^2 \beta_n} y^{2(2-n)/n}, \quad (17)$$

describes the far tail(s) of the distribution, the leading-order term coinciding with the OFM prediction of [16].

We now show that higher-order corrections to the DV perturbative result do not affect the result (7)–(9). Higher-order perturbation theory in the DV formalism will produce correction terms of order $(\Delta A)^m / T^{m-1}$ with $m > 2$ which will be added to Eq. (7) in the exponent. However, when taking the limit $T \rightarrow \infty$ with $\Delta A / T^{n/(2n-2)}$ constant, these correction terms will have a vanishing contribution to f_n , i.e., the correction terms do not affect the limit $-\lim_{T \rightarrow \infty} \ln P(\langle A \rangle + y T^{n/(n-2)}; T) / T^{2/(n-2)} = f_n(y)$.

B. Optimality of the mixed scenario

Strictly speaking, the mixed scenario described above gives, at this point, only an upper bound for the rate function $f_n(y)$ since we haven’t yet shown that this scenario is optimal (i.e., that no other, likelier scenario exists). We now give a strong theoretical argument in favor of the optimality of the mixed scenario. Let us rewrite A in the form $A = A_1 + \dots +$

A_N where

$$A_i = \int_{(i-1)T/N}^{iT/N} x^n(t) dt, \quad (18)$$

and N is chosen such that $1 \ll N \ll T$. The A_i 's are correlated, but weakly so—the correlation between A_i and A_j decays with $|i - j|$ over a characteristic scale that is much smaller than N . It is natural to expect the general result of [65], obtained for i.i.d. random variables, to be valid for weakly correlated random variables, too. Indeed, using that the variance of each of the A_i 's is $\text{Var}(A_i) \simeq \frac{T}{2\beta_n N}$ and that the tail of the PDF's of their distributions $P_i(A_i)$ behaves as $P_i(A_i \rightarrow \infty) \sim e^{-c_n A_i^{2/n}}$, our Eqs. (7)–(9) are in perfect agreement with the calculation that uses the result of [65]; see Appendix D.

C. Conditioned process

DV theory also predicts that the distribution $p(x|A)$ of $x = x(t)$ at some arbitrary intermediate time [67], conditioned on a given value of A , is given by the squared absolute value of the ground-state eigenfunction of the tilted potential (10) [13]. As stated above, for our system the tilted potential has no ground state, but using the same logic as we used above, we obtain the effective ground state by treating k perturbatively. This yields (see Appendix B for details)

$$p(x|A) \simeq P_s(x) - 2k \sum_{m=1}^{\infty} \frac{\langle m|x^n|0\rangle}{E_m - E_0} \langle m|x\rangle \langle 0|x\rangle + \text{instanton contribution}, \quad (19)$$

where using the connection between a and k found above, $k = 2\beta_n \Delta A_{\text{DV}}/T$. Due to the localization of the instanton, the last term in Eq. (19) is of order $1/T \ll 1$. For $n=3$ Eq. (19) reads $p(x|A) \simeq [1 - k(\frac{2}{3}x^3 + 2x)]e^{-x^2}/\sqrt{\pi} + \text{instanton contribution}$ (the calculation is performed explicitly in Appendix B).

V. SUMMARY AND DISCUSSION

We calculated the distribution $P(A; T)$ for $n > 2$ in the long-time limit $T \gg 1$ for $A > \langle A \rangle$ for even n , and all A for odd n ; see Eqs. (7)–(9). We showed that the two main generic tools in the study of large deviations, namely DV and the OFM, correspond to scenarios that describe the fluctuations of A in different regimes. We uncovered a remarkable first-order dynamical phase transition, corresponding to a jump in the first derivative of the rate function $f_n(y)$. In the subcritical regime the fluctuations of A are Gaussian, while in the supercritical regime the optimal scenario leading to a given value of A is a combination of the DV and instanton scenarios, described by (14), that dominates the contribution to $P(A; T)$. We also calculated the conditional distribution of $x(t)$ at an intermediate time, conditioned on a given value of A .

For even n , at $0 < A < \langle A \rangle$, $P(A; T)$ is described by Eq. (2) where $I(a)$ is found from DV theory [16]. At $A \simeq \langle A \rangle$ the two results (2) and (7) match smoothly due to the parabolic behaviors (with the same coefficient β_n) of the rate functions $f_n(y)$ and $I(a)$ around their minima.

It is worth noting that our anomalous exponents, $\mu = 2/(2n - 2)$ and $\nu = n/(2n - 2)$, are different to those found by NT [16]. In contrast to NT's result which is valid only in the very far tail, the rate function $f_n(y)$ given in the present work describes the entire distribution $P(A; T)$: from typical Gaussian fluctuations up to the far tail that is dominated by the instanton in the leading order.

It would be interesting to observe the regime of the transition in numerical simulations. As shown in Fig. 2(a), the theory shows good agreement with the simulations in the typical-fluctuations regime. The far-tail result showed good agreement with simulations in [16], which were performed with $T = 30$. However, capturing the transition regime is far more challenging: We found that $T = 100$ was not large enough to observe the convergence to (7) in the transition regime (not shown), so longer simulation times and special sampling methods [68] are needed in order to observe this regime.

It would be interesting to explore a possible universality of the condensation transition found here in a broader class of continuous-time systems, in analogy with that found in [63–65] for the distribution of sums of N i.i.d. random variables. Remarkably, our rate function $f_n(y)$ coincides, up to the constants c_n and β_n , with those found in [63–65]. This connection may be related to the argument given in Sec. IV B, formulating our problem in terms of the distribution of a sum of weakly correlated random variables. Our perturbative DV approach appears to be a general method for calculating the variance of typical, Gaussian fluctuations in a broad class of continuous-time systems even when the DV tilted operator has no ground state. In fact, the perturbative approach is well known in the context of quantum mechanics, as it describes long-lived bound states with a complex energy describing a decay [69,70]. Similarly, the OFM appears to be a general method for calculating the far tail(s) in such systems, and the two methods together give the full rate function.

In particular, one could search for a similar transition in a class of Gaussian, but not necessarily Markovian, processes (which includes the OU process as a particular case) studied in [18], whose fluctuations were also shown to exhibit anomalous scaling. For these processes the anomalous scaling $P(A; T) \sim e^{-T^\mu f(\Delta A/T^\nu)}$ was conjectured and it was shown that from the behavior in the far tail, one can deduce the relation $\nu = n\mu/2$ [18] which is indeed satisfied by our anomalous exponents in (7).

ACKNOWLEDGMENTS

I warmly acknowledge helpful discussions with Tal Agranov, Hugo Touchette, Satya N. Majumdar, Alex Kamenev, and Baruch Meerson, and I thank the latter for a critical reading of the manuscript. I acknowledge support from the Yad Hanadiv fund (Rothschild fellowship).

APPENDIX A: OBTAINING $F_n(y, z)$ AND $f_n(y)$ AND DERIVING SOME OF THEIR PROPERTIES

In order to obtain our results (7)–(9) from (15) (all in the main text), we first rewrite the latter equation in the form

$e^{-S(\Delta A_{\text{ins}})}$ where

$$\begin{aligned} S(\Delta A_{\text{ins}}) &= c_n \Delta A_{\text{ins}}^{2/n} + \beta_n \frac{(\Delta A - \Delta A_{\text{ins}})^2}{T} \\ &= T^{2/(2n-2)} F_n \left(\frac{\Delta A}{T^{n/(2n-2)}}, \frac{\Delta A_{\text{ins}}}{T^{n/(2n-2)}} \right). \end{aligned} \quad (\text{A1})$$

$P(A; T)$ is now obtained by integrating

$$\begin{aligned} P(A; T) &\sim \int_0^{\Delta A} e^{-S(\Delta A_{\text{ins}})} d\Delta A_{\text{ins}} \\ &\sim \int_0^y e^{-T^{2/(2n-2)} F_n(y, z)} dz, \end{aligned} \quad (\text{A2})$$

where we are considering here the large- T limit with $y = \Delta A/T^{n/(2n-2)}$ constant (because in this limit that the two terms in S are of the same order). In the long-time limit, the large parameter $T^{2/(2n-2)} \gg 1$ in the exponent ensures that the dominant contribution to the integral (A2) comes from saddle-point approximation, $P(A; T) \sim e^{-T^{2/(2n-2)} F_n(y, z_*)}$ where z_* is the minimizer of $F_n(y, z)$ over z , which is precisely Eqs. (7)–(9) in the main text.

We now analyze properties of the functions $F_n(y, z)$ and $f_n(y)$, and in particular we focus on the first-order dynamical phase transition in $f_n(y)$. For odd n , the distribution is clearly symmetric $P(A; T) = P(-A; T)$ so that $f_n(y) = f_n(-y)$, whereas for even n , the scaling form (7) in the main text is only valid at $y > 0$. Therefore, for simplicity, let us assume here that $y > 0$.

The local minimum of $F_n(y, z)$ at the nonzero value $z = z_*$ only exists at $y > y_\ell$ where

$$\left. \frac{\partial^2 F_n(y_\ell, z)}{\partial z^2} \right|_{z=z_\ell \equiv z_*(y_\ell)} = 0, \quad (\text{A3})$$

leading to

$$y_\ell = \frac{2n-2}{n-2} \left[\frac{n^2 \beta_n}{(n-2)c_n} \right]^{n/(2-2n)}. \quad (\text{A4})$$

This local minimum becomes the *global* minimum at $y > y_c$ which we find from the continuity of $f_n(y)$ at $y = y_c$,

$$y_c = \frac{c_3}{3\beta_3} z_c^{-1/3} + z_c, \quad \beta_3 y_c^2 = c_3 z_c^{2/3} + \frac{c_3^2}{9\beta_3} z_c^{-2/3}. \quad (\text{A5})$$

We find

$$z_c = \left[\frac{(n-2)c_n}{n\beta_n} \right]^{n/(2n-2)}, \quad (\text{A6})$$

$$y_c = \frac{n-1}{n-2} \left[\frac{(n-2)c_n}{n\beta_n} \right]^{n/(2n-2)}. \quad (\text{A7})$$

For $n \in \{3, 4\}$ this gives the critical values

$$y_c(n=3) = 2 \left(\frac{c_3}{3\beta_3} \right)^{3/4} = \frac{11^{3/4}}{15^{1/4}} = 3.069 \dots, \quad (\text{A8})$$

$$y_c(n=4) = \frac{3}{2^{5/3}} \left(\frac{c_4}{\beta_4} \right)^{2/3} = \frac{3^{4/3} 7^{2/3}}{2^{5/3}} = 4.987 \dots \quad (\text{A9})$$

At $y = y_c$, $f_n(y)$ is continuous but its first derivative jumps. Indeed, $f'_n(y_c^-) = 2\beta_n y_c$ while, in the supercritical regime one

finds from the chain rule,

$$\frac{df_n}{dy} = \frac{df_n/dz_*}{dy/dz_*} = \frac{2c_n}{n} z_*^{(2-n)/n}, \quad (\text{A10})$$

so $f'_n(y_c^+) = (2c_n/n) z_c^{(2-n)/n}$. Using Eq. (A7) one finds $f'_n(y_c^-) = (n-1)f'_n(y_c^+)$ so the two one-sided derivatives indeed differ. Finally, at $y \gg 1$, Eq. (16) yields $z_* \simeq y - \frac{c_n}{n\beta_n} y^{(2-n)/n}$ leading to the asymptotic behavior,

$$f_n(y \gg 1) \simeq c_n y^{2/n} - \frac{c_n^2}{n^2 \beta_n} y^{2(2-n)/n}, \quad (\text{A11})$$

given also in Eq. (17) of the main text.

APPENDIX B: PERTURBATION THEORY ON THE DV QUANTUM POTENTIAL

The Donsker-Varadhan formalism [14] reduces the problem of finding the rate function $I(a)$ from Eq. (2) of the main text, to that of finding the largest eigenvalue $\lambda(k)$ of the “tilted” operator $\frac{1}{2} \partial_x^2 - V_k(x)$. Equivalently, $-\lambda(k)$ is the ground-state energy for a quantum particle of unit mass in the “tilted” potential $V_k(x)$ (in units where $\hbar = 1$). According to the Gärtner-Ellis theorem [15], the DV rate function $I(a)$ is recovered from $\lambda(k)$ by applying a Legendre-Fenchel transform,

$$I(a) = \sup_{k \in \mathbb{R}} [ka - \lambda(k)]. \quad (\text{B1})$$

However, for the Ornstein-Uhlenbeck process and A defined as in Eq. (4) in the main text, the potential is given by

$$V_k(x) = \frac{x^2}{2} - \frac{1}{2} - kx^n. \quad (\text{B2})$$

As pointed out in [16], for $n > 2$, this potential clearly has no ground state for $k \neq 0$, signaling that DV breaks down here. Still, following the argument given in the main text, we can calculate the effective ground-state energy perturbatively in the parameter $|k| \ll 1$, and we expect the result to correctly describe typical fluctuations of A .

At $k = 0$ the ground-state energy vanishes and the corresponding wave function is $\langle 0|x\rangle = \pi^{-1/4} e^{-x^2/2}$. Second-order perturbation theory yields

$$\lambda(k) = k \langle 0|x^n|0\rangle + k^2 \sum_{m=1}^{\infty} \frac{|\langle m|x^n|0\rangle|^2}{E_m - E_0} + \dots, \quad (\text{B3})$$

where $E_m = m$ are the unperturbed energy levels. The matrix elements $\langle m|x^n|0\rangle$ are straightforward to calculate by expressing $x = (a + a^\dagger)/\sqrt{2}$ using creation and annihilation operators, and one finds that the sum in (B3) includes only a finite number of nonzero terms. For $n = 3$ one finds, using $(a + a^\dagger)^3|0\rangle = 3|1\rangle + \sqrt{6}|3\rangle$, that $\lambda(k) = 11k^2/8 + \dots$, and the Legendre-Fenchel transform (B1) then gives the connection $a = \lambda'(k) = 11k/4 + \dots$ leading to $I(a) = 2a^2/11 + \dots$, so $\beta_3 = 2/11$. Similarly, for $n = 4$, $\lambda(k) = 3k/4 + 21k^2/8 + \dots$, $a(k) = 3/4 + 21k/4 + \dots$ and $I(a) = 2(a - 3/4)^2/21 + \dots$, so $\beta_4 = 2/21$. Here the minimum of $I(a)$ is at $a = 3/4$ which corresponds to the mean value $\langle A \rangle \simeq 3T/4$, and it equals the value predicted by the stationary distribution

$\langle A \rangle = T \int_{-\infty}^{\infty} x^4 P_s(x) dx$. For general n one has

$$I(a) = \beta_n (a - \langle 0|x^n|0 \rangle)^2, \quad \beta_n = \left(4 \sum_{m=1}^{\infty} \frac{|\langle m|x^n|0 \rangle|^2}{E_m - E_0} \right)^{-1}, \tag{B4}$$

which is essentially equivalent to Eq. (11) of the main text. Finally, the ground state, including its first-order perturbative correction, is given by

$$|\psi\rangle = |0\rangle - k \sum_{m=1}^{\infty} \frac{\langle m|x^n|0 \rangle}{E_m - E_0} |m\rangle + \dots \tag{B5}$$

The conditional distribution $p(x|A)$ defined in the main text is then given by $\langle x|\psi\rangle^2$, which, using $P_s(x) = |\langle 0|x\rangle|^2$ and keeping the leading-order term in k , yields Eq. (19) of the main text without the last term on the right-hand side (which is not accounted for by DV theory). For $n = 3$, using $(a + a^\dagger)^3|0\rangle = 3|1\rangle + \sqrt{6}|3\rangle$ and plugging the wave functions of the (unperturbed) harmonic oscillator,

$$\langle 0|x\rangle = \pi^{-1/4} e^{-x^2/2}, \tag{B6}$$

$$\langle 1|x\rangle = \pi^{-1/4} \sqrt{2} x e^{-x^2/2}, \tag{B7}$$

$$\langle 3|x\rangle = \pi^{-1/4} \frac{8x^3 - 12x}{\sqrt{48}} e^{-x^2/2}, \tag{B8}$$

into Eq. (19) of the main text, we obtain the result reported just below it.

APPENDIX C: ADDITIONAL DETAILS REGARDING THE INSTANTON

In [16] the instanton’s action,

$$\frac{1}{2} \int_0^T [\dot{x}(t) + x(t)]^2 dt = c_n (\Delta A_{\text{ins}})^{2/n}, \tag{C1}$$

was calculated. The expression for c_n given there can in fact be simplified a little and written as

$$c_n = \frac{nB_n}{8}, \quad B_n = 2 \left[\frac{2\sqrt{\pi} \Gamma(\frac{n}{n-2})}{(n-2)\Gamma(\frac{3n-2}{2(n-2)})} \right]^{(n-2)/n}, \tag{C2}$$

where $\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt$ is the gamma function. In [18] the instanton itself was also found:

$$x(t) = (2B_n)^{1/(2-n)} (\Delta A_{\text{ins}})^{1/n} \times \left[\text{sech} \left(\frac{(n-2)(t - t_*)}{2} \right) \right]^{2/(n-2)}. \tag{C3}$$

As described in the main text, one indeed finds that the instanton is (temporally) localized, that is, $x(t) \simeq 0$ except for

a short time window (whose duration, of order unity, is much shorter than T) around the time t_* . Note that Eqs. (C1) and (C3) were originally given (in [16] and [18], respectively) with ΔA_{ins} replaced by A . The reason for the difference is that they worked in the low noise limit $\sigma \rightarrow 0^+$ in the physical variables, which, as shown in the present work, is equivalent to the far distribution tail(s) $|A| \rightarrow \infty$. In their limit $|A| \gg \langle A \rangle$ and, as argued in the main text, the instanton dominates the contribution to ΔA , so $A \simeq \Delta A \simeq \Delta A_{\text{ins}}$. In this work, however, here we work in the scaling limit $\Delta A \sim T^{n/(2n-2)}$ and thus it is important to leave ΔA_{ins} as it is in Eqs. (C1) and (C3).

APPENDIX D: COMPARISON WITH BROSSET ET AL.

In Ref. [65], distributions of sums $S = X_1 + \dots + X_N$ of i.i.d. random variables were studied. Assuming that each of the X_i ’s has zero mean and variance V and that the distribution tails are stretched exponentials,

$$\ln \mathbb{P}(X \geq x) \sim -qx^{1-\epsilon}. \tag{D1}$$

They proved that for all $C > 0$,

$$\begin{aligned} \lim_{N \rightarrow \infty} \frac{N}{C^2 N^{2/(1+\epsilon)}} \ln \mathbb{P}(S \geq CN^{1/(1+\epsilon)}) \\ = - \inf_{0 \leq t \leq 1} \left\{ \frac{q(1-t)^{1-\epsilon}}{C^{1+\epsilon}} + \frac{t^2}{2V} \right\}. \end{aligned} \tag{D2}$$

In order to compare with our results, it is convenient to rewrite this in the form,

$$\begin{aligned} \mathbb{P} \left(\frac{S}{N^{1/(1+\epsilon)}} \geq C \right) \\ \sim \exp \left[- \frac{C^2 N^{2/(1+\epsilon)}}{N} \inf_{0 \leq t \leq 1} \left\{ \frac{q(1-t)^{1-\epsilon}}{C^{1+\epsilon}} + \frac{t^2}{2V} \right\} \right], \end{aligned} \tag{D3}$$

which is valid in the limit $N \gg 1$.

Using the argument in Sec. IV B, we now show that our Eqs. (7)–(9) are in perfect agreement with [65]. For convenience, let us first consider odd n , for which the means of the A_i ’s vanish. Replacing $S \rightarrow A$, $X_i \rightarrow A_i$ [defined in Eq. (18)], $V \rightarrow \frac{T}{2\beta_n N}$, $1 - \epsilon \rightarrow 2/n$, and $q \rightarrow c_n$, and replacing the cumulative distribution functions by the PDFs (because, in the large-deviation regime they are equal up to a subleading prefactor), Eq. (D3) becomes

$$P(A) \sim \exp \left[- \inf_{0 \leq t \leq 1} \left\{ c_n (1-t)^{1-\epsilon} A^{1-\epsilon} + \frac{\beta_n t^2 A^2}{T} \right\} \right], \tag{D4}$$

which indeed coincides with our Eqs. (7)–(9). For even n , the same argument with the replacement $S = \Delta A$, $X_i = A_i - \langle A_i \rangle$ works.

[1] S. S. Varadhan, *Large Deviations and Applications*, CBMS-NSF Regional Conference Series in Applied Mathematics, No. 46 (SIAM, Philadelphia, 1984).
 [2] Y. Oono, *Prog. Theor. Phys. Suppl.* **99**, 165 (1989).

[3] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, 2nd ed. (Springer, New York, 1998).
 [4] F. den Hollander, *Large Deviations*, Fields Institute Monographs, Vol. 14 (AMS, Providence, 2000).

- [5] S. N. Majumdar and A. J. Bray, *Phys. Rev. E* **65**, 051112 (2002).
- [6] S. N. Majumdar, *Current Science* **89**, 2076 (2005).
- [7] H. Touchette, *Phys. Rep.* **478**, 1 (2009).
- [8] B. Derrida, *J. Stat. Mech.* (2011) P01030.
- [9] B. Meerson and S. Redner, *Phys. Rev. Lett.* **114**, 198101 (2015).
- [10] P. Vivo, *Eur. J. Phys.* **36**, 055037 (2015).
- [11] M. Assaf and B. Meerson, *J. Phys. A: Math. Theor.* **50**, 263001 (2017).
- [12] S. N. Majumdar and G. Schehr, *ICTS Newsletter 2017*, Vol. 3, Issue 2 (ICTS, Bangalore, 2017).
- [13] H. Touchette, *Physica A* **504**, 5 (2018).
- [14] M. D. Donsker and S. R. S. Varadhan, *Commun. Pure Appl. Math.* **28**, 1 (1975); **28**, 279 (1975); **29**, 389 (1976); **36**, 183 (1983).
- [15] J. Gärtner, *Theory Probab. Appl.* **22**, 24 (1977); R. S. Ellis, *Ann. Probab.* **12**, 1 (1984).
- [16] D. Nickelsen and H. Touchette, *Phys. Rev. Lett.* **121**, 090602 (2018).
- [17] R. L. Jack and R. J. Harris, *Phys. Rev. E* **102**, 012154 (2020).
- [18] B. Meerson, *Phys. Rev. E* **100**, 042135 (2019).
- [19] Ž. Krajnik, E. Ilievski, and T. Prosen, [arXiv:2109.13088](https://arxiv.org/abs/2109.13088).
- [20] To complete the definition of the problem, one should specify an initial condition $x(t = 0) = x_0$. However, in the long-time limit that we are considering, the initial condition becomes unimportant.
- [21] R. van Zon and E. G. D. Cohen, *Phys. Rev. E* **67**, 046102 (2003).
- [22] R. van Zon, S. Ciliberto, and E. G. D. Cohen, *Phys. Rev. Lett.* **92**, 130601 (2004).
- [23] T. Matsumoto and M. Takaoka, *Phys. Rev. E* **87**, 013008 (2013).
- [24] D. Nickelsen, *J. Stat. Mech.* (2017) 073209.
- [25] L. Onsager and S. Machlup, *Phys. Rev.* **91**, 1505 (1953).
- [26] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer, New York, 1984).
- [27] M. I. Dykman and M. A. Krivoglaz, in *Soviet Physics Reviews*, edited by I. M. Khalatnikov (Harwood Academic, New York, 1984), Vol. 5, pp. 265–441.
- [28] R. Graham, in *Noise in Nonlinear Dynamical Systems*, edited by F. Moss and P. V. E. McClintock, Vol. 1 (Cambridge University Press, Cambridge, 1989), p. 225.
- [29] A. V. Nagaev, *Theory Probab. Appl.* **14**, 51 (1969); **14**, 193 (1969).
- [30] V. P. Chistyakov, *Theory Probab. Appl.* **9**, 640 (1964).
- [31] S. Foss, D. Korshunov, and S. Zachary, *An Introduction to Heavy Tailed and Subexponential Distributions* (Springer, Berlin/Heidelberg, 2013).
- [32] D. Denisov, A. B. Dieker, and V. Shneer, *Ann. Probab.* **36**, 1946 (2008).
- [33] J. Geluk and Q. Tang, *J. Theor. Probab.* **22**, 871 (2009).
- [34] E. Bertin and M. Clusel, *J. Phys. A: Math. Theor.* **39**, 7607 (2006).
- [35] R. Burioni, L. Caniparoli, and A. Vezzani, *Phys. Rev. E* **81**, 060101(R) (2010).
- [36] R. Burioni, E. Ubaldi, and A. Vezzani, *Phys. Rev. E* **89**, 022135 (2014).
- [37] A. Vezzani, E. Barkai, and R. Burioni, *Phys. Rev. E* **100**, 012108 (2019).
- [38] A. Vezzani, E. Barkai, and R. Burioni, *Sci. Rep.* **10**, 2732 (2020).
- [39] W. Wang, A. Vezzani, R. Burioni, and E. Barkai, *Phys. Rev. Research* **1**, 033172 (2019).
- [40] R. Burioni, G. Gradenigo, A. Sarracino, A. Vezzani, and A. Vulpiani, *J. Stat. Mech.* (2013) P09022.
- [41] M. R. Evans and T. Hanney, *J. Phys. A: Math. Gen.* **38**, R195 (2005).
- [42] S. N. Majumdar, in *Exact Methods in Low-dimensional Statistical Physics and Quantum Computing: Lecture Notes of the Les Houches Summer School* Vol. 89 (Oxford University Press, Oxford, 2010), p. 407.
- [43] S. Großkinsky, G. M. Schütz, and H. Spohn, *J. Stat. Phys.* **113**, 389 (2003).
- [44] O. Hirschberg, D. Mukamel, and G. M. Schütz, *J. Stat. Mech.* (2015) P11023.
- [45] K. Ø. Rasmussen, T. Cretegnny, P. G. Kevrekidis, and N. Grønbech-Jensen, *Phys. Rev. Lett.* **84**, 3740 (2000).
- [46] S. Iubini, A. Politi, and P. Politi, *J. Stat. Mech.* (2017) 073201.
- [47] G. Gradenigo, S. Iubini, R. Livi, and S. N. Majumdar, *J. Stat. Mech.* (2021) 023201.
- [48] G. Gradenigo, S. Iubini, R. Livi, and S. N. Majumdar, *Eur. Phys. J. E* **44**, 29 (2021).
- [49] G. Gotti, S. Iubini, and P. Politi, *Phys. Rev. E* **103**, 052133 (2021).
- [50] J-P. Bouchaud and M. Mézard, *Physica A* **282**, 536 (2000).
- [51] Z. Burda, D. Johnston, J. Jurkiewicz, M. Kamiński, M. A. Nowak, G. Papp, and I. Zahed, *Phys. Rev. E* **65**, 026102 (2002).
- [52] M. Filiasi, E. Zarinelli, E. Vesselli, and M. Marsili, [arXiv:1309.7795](https://arxiv.org/abs/1309.7795).
- [53] S. N. Majumdar, M. R. Evans, and R. K. P. Zia, *Phys. Rev. Lett.* **94**, 180601 (2005).
- [54] M. R. Evans, S. N. Majumdar, and R. K. P. Zia, *J. Stat. Phys.* **123**, 357 (2006).
- [55] M. R. Evans, T. Hanney, and S. N. Majumdar, *Phys. Rev. Lett.* **97**, 010602 (2006).
- [56] M. R. Evans and S. N. Majumdar, *J. Stat. Mech.* (2008) 05004.
- [57] M. R. Evans, S. N. Majumdar, I. Pagonabarraga, and E. Trizac, *J. Chem. Phys.* **132**, 014102 (2010).
- [58] J. Szavits-Nossan, M. R. Evans, and S. N. Majumdar, *Phys. Rev. Lett.* **112**, 020602 (2014).
- [59] J. Szavits-Nossan, M. R. Evans, and S. N. Majumdar, *J. Phys. A: Math. Theor.* **47**, 455004 (2014).
- [60] J. Szavits-Nossan, M. R. Evans, and S. N. Majumdar, *J. Phys. A: Math. Theor.* **50**, 024005 (2016).
- [61] G. Gradenigo and E. Bertin, *Entropy* **19**, 517 (2017).
- [62] G. Gradenigo and S. N. Majumdar, *J. Stat. Mech.* (2019) 053206.
- [63] F. Mori, G. Gradenigo, and S. N. Majumdar, *J. Stat. Mech.* (2021) 103208.
- [64] F. Mori, P. Le Doussal, S. N. Majumdar, and G. Schehr, *Phys. Rev. E* **103**, 062134 (2021).
- [65] F. Brosset, T. Klein, A. Lagnoux, and P. Petit, [arXiv:2007.8164](https://arxiv.org/abs/2007.8164).
- [66] $\langle A \rangle$ is easy to calculate from the stationary distribution, $\langle A \rangle \simeq T \int_{-\infty}^{\infty} x^n P_s(x) dx$, using the fact that the system is ergodic and therefore time-averaged quantities converge, for large times, to their ensemble-averaged values.

- [67] According to DV theory, $p(x|A)$ is independent of t except for narrow temporal boundary layers near the beginning and end of the dynamics [13].
- [68] A. K. Hartmann, *Phys. Rev. E* **65**, 056102 (2002).
- [69] A. I. Vainshtein, “*Decaying systems and divergence of the series of perturbation theory*,” December 1964 Report (Budker Institute of Nuclear Physics, Novosibirsk, 1964)
- [70] A. I. Baz, Ya. B. Zeldovich, and A. M. Perelomov, *Dispersion, Reactions and Disintegrations in the Nonrelativistic Quantum Mechanics* (Israel Program for Scientific Translations, Jerusalem, 1971).