Stabilizing role of multiplicative noise in nonconfining potentials

Ewan T. Phillips,[1,*] Benjamin Lindner,[2,3] and Holger Kantz[1]
[1] *Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Straße 38, 01187 Dresden, Germany*
[2] *Bernstein Center for Computational Neuroscience, Haus 2, Philippstraße 13, 10115 Berlin, Germany*
[3] *Department of Physics, Humboldt Universität zu Berlin, Newtonstraße 15, 12489 Berlin, Germany*

(Received 18 November 2024; accepted 17 March 2025; published 14 May 2025)

We provide a simple framework for the study of parametric (multiplicative) noise, making use of scale parameters. We focus on a large class of one-dimensional stochastic differential equations in which the deterministic drift pushes trajectories toward infinity. We show that increasing the multiplicative noise intensity surprisingly causes the mass of the stationary probability distribution to become increasingly concentrated around the points of minimal multiplicative noise strength. Under quite general conditions the trajectory exhibits intermittent burstlike jumps away from these minima. Our framework relies on first-term expansions, which become more accurate for larger noise intensities. In this work we show that the full width at half maximum in addition to the maximum is appropriate for quantifying the stationary probability distribution (instead of the mean and variance, which are often undefined). We define a corresponding kind of weak-sense stationarity. We end by applying these results to the problem of a double-well potential with multiplicative noise, where noise stabilizes unstable fixed points.

DOI: 10.1103/PhysRevResearch.7.023146

I. INTRODUCTION

Multiplicative noise (or parametric noise) is ubiquitous in real-world systems and (in contrast to additive noise) is known to play an important role in state transitions [1–8], including tipping points where hopping is induced from one attractor to another [9–12]. Such events have relevance for example in synchronization [13–16], human balancing tasks [17], financial crises [18,19], environmental tipping points [20–27], neuroplasticity [28,29], epidemics, and stochastic optimization [30].

A simple model where this effect is of relevance is that of a double-well potential with multiplicative noise. Such a model may describe Rayleigh-Bénard convection when the Rayleigh number is a noisy function of time [31,32] or for example serve as a simple model for the periodic transitions between the glacial and interglacial periods in Earth's history, where there is evidence that the noise is parametric [33].

Another important noise-induced phenomenon is that of on-off intermittency, where noise induces an aperiodic switching between static, so called laminar behavior and chaotic bursts [34–37]. This is due to the noise causing the bifurcation parameter to fluctuate around the bifurcation point [38]. Such behavior is often observed as a continuous route from regular behavior to chaotic motion [37,39,40]. Such intermittency has been observed for example in noisy laser systems, which may exhibit sporadic high-intensity pulses [41–43], Earth surface temperature at both weather and climate timescales [44–47], and synchronization of coupled systems of interacting dynamical units, where the bursting induces a phase slip [48,49].

Theoretical work on the topic of noise-induced transitions has generally taken two different approaches. In the first approach the effect of noise on a system is quantified according to the moments (mean, variance, etc.) of the probability distribution [32,50–53]. This approach, while useful in a small noise or additive noise limit, is however not able to describe noise-induced bifurcations, since mean zero noise has no affect on the mean in the Itô formulation.

The effect of noise on a system is instead to broaden and/or to skew the stationary probability distribution, generally leading to heavy tails and in some cases even leading to divergence of the moments. In the second class the sharp transition in the presence of noise is interpreted as the bifurcation of a maximum of the stationary probability distribution [1]. It remained unexplained, however, why only the maximum of the distribution should be observed in experiments [54].

In this paper we explore a class of Langevin stochastic differential equations (SDEs), which we show under quite general conditions exhibit on-off intermittency. We build on the second approach suggesting the use of the maximum as an alternative to the mean and the full width at half maximum (FWHM) as an alternative to the variance for systems with multiplicative noise. These two quantities collectively account for the body of the probability density function excluding the heavy tail. We find counterintuitively that by skewing the tail of the probability distribution the noise induces a kind of stabilization of the minima of the noise term.

[*] Contact author: tphillips@pks.mpg.de

It has been known for some time that unbiased parametric noise may stabilize an autonomous linear stochastic system of two or more dimensions in the sense of the moments, and that such noise is not sufficient to stabilize a one-dimensional SDE [55–57]. Here, however, we show that noise is sufficient to stabilize a one-dimensional SDE in the sense of the maximum and the FWHM.

The structure of this text is as follows. We first explore a general class of SDEs. We then focus on the noise-stabilizing case where the trajectory turns out to exhibit on-off intermittency. We show how these systems can be analyzed in terms of scale parameters and how from this the FWHM can be derived. We then show why the mean is insufficient to arrive at conclusions about the systems with a finite number of trajectories. In the following section we explore the tails by considering first passage times (distribution of the bursts). We finish by generalizing the results and as an example of application exploring the effect of multiplicative noise in a double-well potential.

II. BASIC MODEL

The general one-dimensional Langevin SDE is in the Itô formulation [58]

$$\dot{x} = \widetilde{f}(x) + \sqrt{2D}g(x)\xi(t), \tag{1}$$

where $\widetilde{f}(x) = f(x) + \nu D g(x) g'(x)$. Here $\nu = 1$ refers to the Stratonovich interpretation and $\nu = 0$ to the Itô interpretation [59]. Here $\xi(t)$ represents Gaussian white noise with mean $\langle \xi(t) \rangle = 0$ and correlation $\langle \xi(t)\xi(t')\rangle = \delta(t - t')$. $\widetilde{f}(x)$ and $g(x)$ represent the drift and diffusion terms, respectively. The probability distribution of Eq. (1) evolves according to the Fokker-Planck equation given by

$$\frac{\partial}{\partial t} p(x,t) = -\frac{\partial}{\partial x}[\widetilde{f}(x)p(x,t)] + \frac{\partial^2}{\partial x^2}[Dg(x)^2 p(x,t)]. \tag{2}$$

We focus on the case where the drift $f(x)$ represents a nonconfining potential; i.e., the trajectories drift toward infinity in the absence of noise. We first consider the case with a natural boundary at $x = \infty$ and a kind of natural boundary at $x = 0$. Formally speaking the natural boundary would be at $x = -\infty$, but given a positive drift and a multiplicative noise term $x = 0$ will effectively act as a natural boundary. The stationary distribution is easily obtained from the condition $\partial_t p(x,t) = 0$ as

$$p_s(x) = \frac{N}{g^2(x)} \exp 2 \int^x \frac{\widetilde{f}(u)}{g^2(u)} du. \tag{3}$$

The extrema x_m of the stationary solution $p_s(x)$ of the Fokker-Planck equation (FPE) can be easily obtained [by setting the derivative to zero $p_s'(x) = 0$] as [1]

$$\widetilde{f}(x_m) - (2 - \nu) D g(x_m) g'(x_m) = 0. \tag{4}$$

Due to the symmetry of the noise around zero and its lack of temporal correlations we notice that we may replace $g(x)$ with $|g(x)|$, and thus assume $g(x) \geqslant 0$. For what follows we furthermore assume that $g(x)$ has only one minimum x_0 where $g(x_0) = 0$. From this equation we see immediately that as D increases the term $Dg(x)g'(x)$ becomes dominant, and thus that the x_m moves closer to the minimum x_0 of $g(x)$ as D increases. This suggests that for large D it may be sufficient to take the first term of a Taylor expansion around the solution x_0 of $g(x_0) = 0$.

We expand $g(x)$ as a series and take the first nonzero term of the expansion $g(x) \approx a_m x^m$ (where $m \in \mathbb{R}$ is the order of the lowest nonzero term of the series expansion). The details of the higher-order terms will often be unimportant, since we will be primarily interested in the case that the orbit spends most of the time in the vicinity of the minimum of $g(x)$. Considering for the moment that $g(x)$ has a minimum at $x_0 = 0$ (generalizations are discussed in the appendices) suggests studying the general class of first-order Langevin SDEs

$$\dot{x} = ax^n + \nu D x^{2m-1} + \sqrt{2D} x^m \xi(t), \tag{5}$$

where the drift $a \in \mathbb{R}_{>0}$ acts to create a nonconfining potential. $n, m \in \mathbb{R}$ although we will generally restrict ourselves to (real numbers) $n, m \leqslant 1$ in order to guarantee that the solution does not explode (see appendices). We assume here natural boundary conditions (this choice will also prove to be unimportant in the vicinity of $x = 0$). It can be obtained easily from Eq. (4) that the maximum of the stationary distribution changes with D according to

$$x_m = \left(\frac{(2-\nu) m D}{a} \right)^{\frac{1}{n-2m+1}}. \tag{6}$$

We now obtain by straightforward calculation of the second derivative that the calculated extremum of Eq. (5) is a maximum $[p_s''(x_{\max}) < 0]$ in both the Itô and Stratonovich interpretations if and only if the condition $n - 2m + 1 < 0$ holds. We note that these equations also hold exactly even if there is additional additive noise in Eq. (5). The stationary distribution of Eq. (5) is given by

$$p_s(x; \beta) = N x^{\nu - 2m} \exp \frac{a}{D} \frac{x^{n-2m+1}}{n - 2m + 1}. \tag{7}$$

This distribution is not always normalizable. The normalization constant is calculated in the Itô interpretation as

$$N^{-1} = \begin{cases} \text{const.}, & \text{if } n - 2m < -1, \\ \infty, & \text{if } n - 2m \geqslant -1. \end{cases} \tag{8}$$

This result also holds in the Stratonovich interpretation provided $m > 1$ (see appendices for details). We now restrict ourselves to the study of Eq. (5) in the Itô interpretation $\nu = 0$, although we remark that qualitatively both interpretations lead to the same basic results. We have so far seen that for $n - 2m + 1 < 0$ the stationary distribution has a single maximum and is normalizable. We now state furthermore in this regime that this distribution has a scale parameter β, i.e.,

$$p_s(x; \beta) = \frac{p_s(x/\beta; 1)}{\beta}, \tag{9}$$

with the scaling parameter given by

$$\beta = (D/a)^{\frac{1}{n-2m+1}}. \tag{10}$$

The proof (outlined in detail in the appendices) follows easily from the stationary distribution. Using scale parameter properties we can obtain the effect of the parameter on the

FWHM, the height, and the cumulative probability distribution. We start by noting that since the scale parameter serves only to change the scale of the distribution we must in general have a relationship of proportionality

$$x_m \propto \beta, \tag{11}$$

which is indeed what we observe for both D and a comparing Eqs. (6) and (10). We note that since m for example is not a scale parameter we do not expect the same proportionality relationship to hold for varying m.

The FWHM is the value $(x_H - x_{-H})$ at which the stationary probability density has half its value compared to at the maximum x_m, i.e., $p(x_{\pm H}) = \frac{1}{2}p(x_m)$, where x_H is on the left side and x_H is on the right side of x_m, respectively. Although x_H and x_{-H} may in general be different distances from the x_m, the way these differences scale proportional to β is the same. Analogously the locations of the FWHM $x_{\pm H}$ compared to x_m are clearly

$$p_s(x_{\pm H};\beta) \propto p_s(x_{\pm H}/\beta;1) = \tfrac{1}{2}p_s(x_m/\beta;1) \tag{12}$$

and thus

$$\begin{aligned}\text{FWHM} &= (x_H - x_{-H})_1 \\ &= (\beta x_H - \beta x_{-H})_\beta \propto \beta,\end{aligned} \tag{13}$$

where $(\cdot)_\beta$ represents consideration of the distribution $p_s(\cdot;\beta)$. The cumulative distribution $F(x) = \int^x p(x')\,dx'$ is affected by the scale parameter according to

$$F(x;\beta) = F(x/\beta;1). \tag{14}$$

A. Cases

In the following we consider $n, m \geqslant 0$. We define $\gamma = 1/(n - 2m + 1)$.

The regime $n - 2m < -1$ $(\gamma < 0)$ is shown in Figs. 1(a) and 1(b). This distribution does exist and is normalizable; however it has diverging moments and a diverging tail (see appendices). This regime is characterized by a well defined finite maximum and FWHM, as well as on-off intermittency. In this regime increasing D surprisingly causes the stationary probability density to become more concentrated around the noise-induced maximum both in the sense of the maximum $\propto D^{-|\gamma|}$ and the FWHM $\propto D^{-|\gamma|}$ of Eqs. (6) and (13). It could be said that the noise has an attracting effect, attracting the trajectory to the point x_0 where $g(x_0) = 0$. In fact (as shown in the appendices) as $D \to \infty$ the stationary distribution surprisingly converges to a delta function. This regime will be the focus of the remaining sections of this paper. We first briefly discuss the two other regimes.

In the case of $n - 2m > -1$ (or $\gamma > 0$) it is clear from Eq. (8) that the distribution is nonstationary. It is visible from Fig. 1(e) that the noise plays a more diffusive role causing x to fluctuate around the deterministic trajectory.

The boundary between the two discussed cases is at $n - 2m = -1$ $(\gamma = 0)$. It is shown in the appendices that in the Itô as well as in the Stratonovich interpretation the noise induces a drift term $g(x)g'(x)$, which may counteract the drift term $f(x)$. The equality holds if and only if the deterministic drift term and the stochastic drift term have the same dependency on x, i.e., $f(x) \propto g(x)g'(x)$. This means that no maxima/minima can exist as can be seen from Eq. (4). There is no maximum and no scale parameter. The distribution is in that sense "scale" free. It has been shown by Kaulakys *et al.* [60,61] that such SDEs (with $n - 2m = -1$) correspond to $1/f$ noise. This noise has the stationary distribution $p_s(x) \propto 1/x^\lambda$ and spectral density $S(f) \propto 1/f^\alpha$ where $\alpha = (4m - 5 - a/D)/(2m - 2)$. A number of papers have pointed out the arising of power-law noise in SDEs containing white noise [62–64]. Power-law spectra are also well known to arise in analogous systems of discrete maps, which by contrast exhibit chaotic intermittency [64–66].

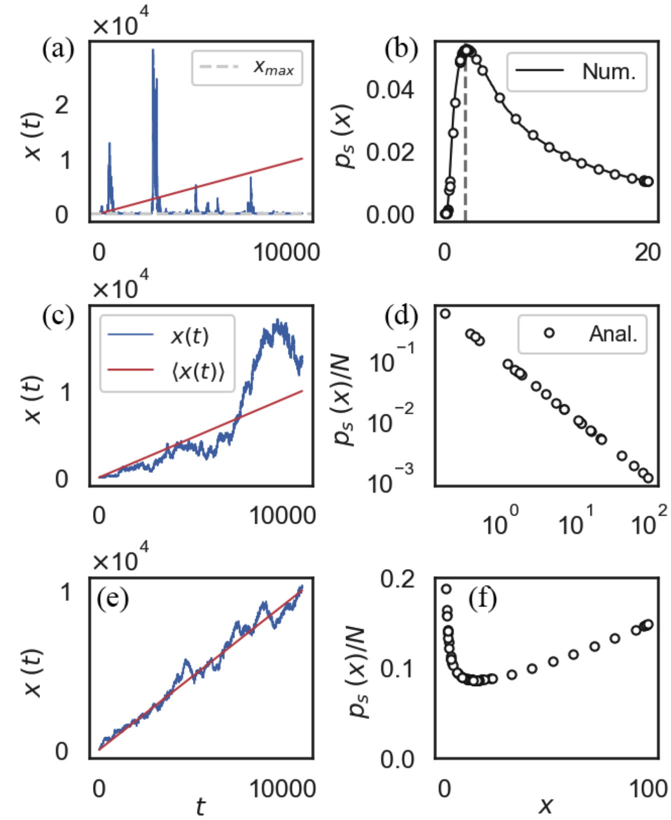

FIG. 1. Time series and (unnormalized) stationary probability distribution for Eq. (5), where $n = 0$, $a = 1$, and $m = 0.8$, $D = 0.4$ [(a), (b)], $m = 0.5$, $D = 1$ [(c), (d)], $m = 0.35$, $D = 5$ [(e), (f)]. For (d) and (f) $p_s(x)/N$ plotted (where N is the normalization constant), since $p_s(x) = 0$.

B. Stationarity of $n - 2m < -1$

Is the regime $n - 2m < -1$ of Eq. (5) stationary? A (strict-sense) stationary process is one whose density function is invariant to time shifts. Typically this is taken to mean that the moments are all finite and are time independent. A wide-sense stationary process is one where first and second order properties are finite and independent of time. Generally this is understood to mean that the first two moments $\langle x(t)\rangle$ and $\langle x(t+\tau)x(t)\rangle$ are independent of time t.

In the regime $n - 2m < -1$ of Eq. (5) all of the moments $\langle x^l\rangle$ with $l = 1, 2, \ldots$ diverge to infinity. This is clear in the Itô formalism since the evolution of the mean is the same as that of the deterministic system, which clearly diverges for

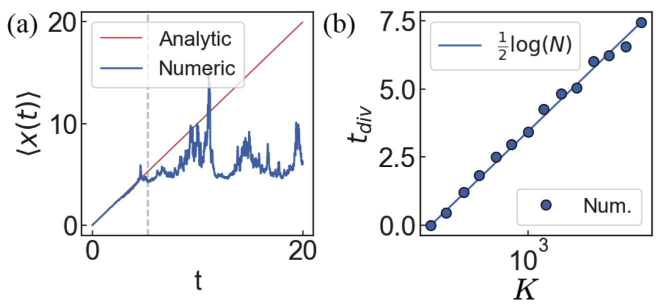

FIG. 2. (a) Time plot of Eq. (5) averaged over $= 10^6$ trajectories, $a = 1, n = 0, m = 1, D = 2$, and $dt = 10^{-4}$. Red line shows theoretical mean value. (b) Divergence point of theoretical and measured mean [gray line in (a)].

$\dot{x} = ax^n$ with $a > 0$ and $n \geqslant 0$ (and from this follows also the divergence of the higher moments). Thus, in this sense (for $a > 0$) the distribution is not stationary, even in a wide sense. The distribution does however exist in this regime, as evidenced by Eq. (7). For this reason we argue the moments are inappropriate to characterize the system of Eq. (5). If instead the first and second order properties are taken to be the maximum and FWHM then this distribution is in fact stationary in a wide sense. We will refer to this as M-wide-sense stationarity, defined by the existence of a stable maximum and a FWHM of this density. The stationary probability density of Eq. (5) when $n - 2m < -1$ is M-wide-sense stationary, but not wide-sense stationary in the sense of moments. Let us now explore this type of stationarity in some more detail.

The trajectories of the SDE of Eq. (5) when $n - 2m < -1$ exhibit bursts (see Fig. 1). If we observe a single trajectory, such as that of Fig. 1(a), we see that regardless of initial conditions the particle will quickly return to the vicinity of the maximum. Due to this, regardless of the initial condition, the correct maximum and FWHM can be established even for relatively short trajectories. This does not change over time. The mean on the other hand cannot in general be established for long trajectories (see Fig. 2).

Due to the heavy tail of the probability distribution the larger bursts dominate the smaller bursts in size by orders of magnitude [Fig. 2(a)]. Since the probability of such a burst having happened increases over time, this means that for a finite ensemble of trajectories the probability of the bursting distorting the mean increases over time. This means that larger sample sizes are increasingly necessary to determine the mean accurately as time evolves. The fact that the trajectory keeps returning to the vicinity of the maximum is not in contradiction to the fact that the mean continually increases. This apparent contradiction can be explained by the fact that as time evolves, the probability of having had a large burst increases.

To see why it is most appropriate to consider the distribution in terms of its maximum and FWHM instead of its moments, we run a large ensemble of trajectories (up to $K = 10^7$) of Eq. (5) with initial conditions at $x_0 = 0$ [Fig. 2(b)]. We see clearly that the average of a small ensemble may represent the theoretical mean faithfully only up to a certain time $t_{\rm div}$, after which the two values diverge. This amount of time $t_{\rm div}$ can be clearly observed to increase logarithmically with system size K. This means that it is unrealistic to describe any more than a fairly short trajectory $t_{\rm div} \approx 10$ in terms of the mean using only standard computers.

III. LAMINAR DISTRIBUTION

We now look at the statistical properties of the bursting for the on-off intermittency case $n - 2m < -1$. The particle has a laminar length, defined as the amount of time (or length l) between two bursts. A burst is defined as quick event where the trajectory exceeds a certain threshold x_c. To calculate this distribution we begin with the time-dependent Fokker-Planck equation of Eq. (5) and then have

$$\frac{\partial p(x,t)}{\partial t} = -\frac{\partial}{\partial x}[ax^n p(x,t)] + D\frac{\partial^2}{\partial x^2}[x^{2m} p(x,t)]. \quad (15)$$

As the coordinate of the system stays for a long time in the region $x < x_c$, one can suppose that the probability density may find the form of a metastable distribution decaying for a long period of time. The relaxation process of the probability density to this metastable state is supposed to be very fast in comparison with the time of the metastable distribution decay; therefore one can neglect the transient $0 \leqslant t \leqslant t_{tr}$ [67–69]. Under the assumptions above the probability density may be written as

$$p(x,t) = A(t)q(x) \quad (16)$$

for small x. $q(x)$ can be solved in much the same way as the stationary distribution. The decrease of $A(t)$ should be determined by the probability distribution taken in the boundary point x_c, i.e., $dA(t)/dt \propto -p(x_c, t)$. This assumption, which is also equivalent to neglecting the time correlation of the orbit, may be rewritten as

$$\frac{dA(t)}{dt} = -kA(t)\frac{1}{Dx_c^{2m}} \exp \frac{a}{D}\frac{x_c^{n-2m+1}}{n-2m+1}, \quad (17)$$

where k is a proportionality coefficient. In the absence of a further constraint it may be that both k and $A(t)$ are functions of D. Evidently the decrease of $A(t)$ is described by the exponential law,

$$A(t) = A(0) \exp -k\eta t, \quad (18)$$

where

$$\eta = \frac{k}{Dx_c^{2m}} \exp \frac{a}{D}\frac{x_c^{n-2m+1}}{n-2m+1}. \quad (19)$$

This is equivalent to the exponential law for the laminar phase distribution, since this distribution is defined as

$$l(t) = -\int_0^\infty \frac{\partial p(x,t)}{\partial t}\, dx, \quad (20)$$

and thus defining $\widetilde{T}^{-1} = \eta k$ we have

$$l(t) = \widetilde{T}^{-1} \exp(-t/\widetilde{T}). \quad (21)$$

This exponential decay of the laminar lengths is shown in Fig. 3.

Mean first passage time

For the purposes of further theoretical treatment we will explore the mean first passage time (MFPT). The first passage

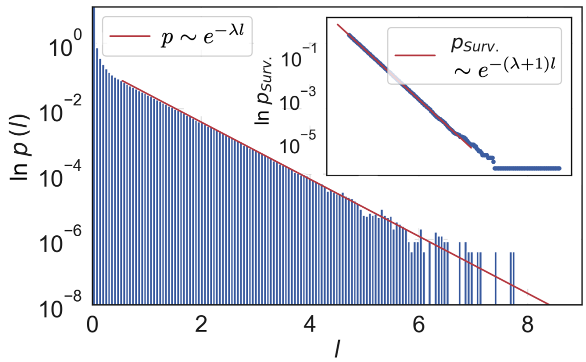

FIG. 3. Stationary probability distribution of the laminar length l. l obtained from a single long-time series by measuring the amount of time that elapses between trajectory passing below threshold x_c (toward $x = 0$) and passing once again above threshold x_c (toward $x \to \infty$). Plotted logarithmically and fitted with exponential fit $p(l) \propto e^{-\lambda l}$. Inset plot is survival probability density over l.

time is defined as the time taken to travel from x_0 to the threshold x_c. We choose $x_0 = 0.1$. Given that bursts quickly return back to the vicinity of $g(x) = 0$ we expect the distribution of laminar lengths to be equivalent to the distribution of first passage times. The mean of this distribution can be calculated using numerical methods (see appendices) considering reflecting boundaries at $x = 0$ as [58]

$$T(x_0) = \frac{1}{D}\int_{x_0}^{x_c} \frac{dy}{\psi(y)} \int_0^y \frac{\psi(z)}{z^{2m}}\,dz, \tag{22}$$

where we have defined $\psi(x) := \exp \frac{a}{D}\frac{x^{n-2m+1}}{n-2m+1}$. A plot of the mean first passage time for different m is shown in Figs. 4(a)–4(c). We see that a power law becomes an increasingly good approximation of the MFPT as the threshold boundary x_c is increased [compare Figs. 4(a) and 4(c)]. We see that (for $n > 0$) the frequency of the bursts T^{-1} (equivalent to l) decreases with D. Increasing the multiplicative noise intensity thus stabilizes the system also in the sense of longer laminar lengths (in addition to in the sense of the maximum and FWHM). We

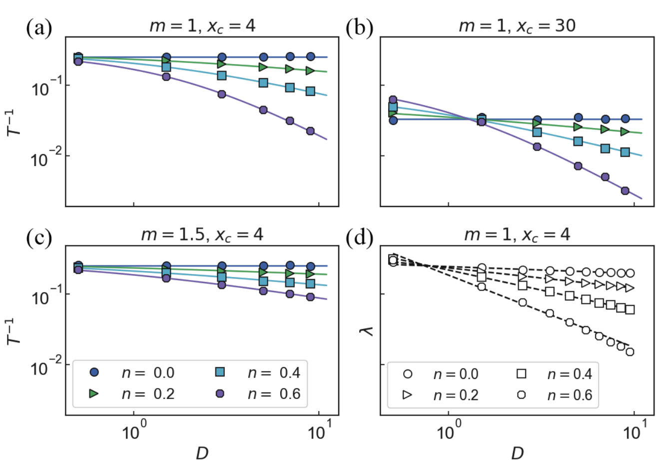

FIG. 4. (a)–(c) Inverse MFPT. Lines are analytical results obtained from Eq. (22) obtained with Simpson's rule; circles are numerics. (d) Circles are λ obtained from exponential fit shown in Fig. 3. Lines are power-law fit.

see by comparing Figs. 4(a) and 4(c) that multiplicative noise D has less of an effect for larger m. This is presumably due to the trajectory being drawn closer to the minimum $g(x) = 0$ for larger m where the noise term $\sqrt{2D}x^m$ is smaller.

In Fig. 4(d) numerical results obtained from the decay constant λ of the exponential distribution are shown. These are obtained by measuring a survival probability density determined in stochastic simulations (see Fig. 3). In calculating this density we do not consider laminar lengths under a given threshold $l_0 < 0.28$, which are very sensitive to the choice of initial conditions and as such do not obey an exponential distribution. In Fig. 4(d) it is clearly visible that λ decays with D_m according to a power law $\lambda \propto D_m^{-\alpha}$. The agreement between the inverse MFPT T^{-1} and λ as x_c becomes large is to be expected, since the higher the upper boundary the less sensitive is T^{-1} to the initial conditions.

IV. GENERALIZATIONS

A. Generalized drift and Stratonovich interpretation

The first generalization is the equation

$$\dot{x} = a_1 x^n + a_2 x^{2m-1} + \sqrt{2D}x^m \xi(t). \tag{23}$$

The maximum of this equation can be easily obtained from the derivative of the stationary probability density as

$$x_m = \left(\frac{(2-\nu)mD - a_2}{a_1}\right)^{\frac{1}{n-2m+1}}. \tag{24}$$

We see that (assuming $a_1 > 0$) the maximum is only well defined if $a_2 < (2-\nu)mD$ as well as $n - 2m + 1 < 0$. Equation (23) can use the following substitution,

$$y(x) = \alpha^{-1}x^{\alpha} \quad \text{where} \quad \alpha = 1 - (a_2/D), \tag{25}$$

and Itô's lemma [68] can be brought into the original form

$$\dot{y} = a_1(\alpha y)^{\frac{n+\alpha-1}{\alpha}} + \sqrt{2D}(\alpha y)^{\frac{m+\alpha-1}{\alpha}}\xi(t). \tag{26}$$

Since $y(x)$ is proportional to x^{α} we can see that from M-stability of y must also follow M-stability of x. Considering the exponents of Eq. (26) instead of those of Eq. (5) we see that the α terms cancel out and that we have M-stability for $n - 2m + 1 \leqslant 0$.

We note that if we instead interpret Eq. (5) using the Stratonovich interpretation then we may rewrite this in the Itô formulation by considering the Stratonovich drift term

$$\begin{aligned}\dot{x} &= a_1 x^n + a_2 x^{2m-1} + \sqrt{2D}x^m \circ \xi(t)\\ &= a_1 x^n + (a_2 + mD)x^{2m-1} + \sqrt{2D}x^m \xi(t),\end{aligned} \tag{27}$$

$\sqrt{2D}x^m\xi(t)$, where $\circ$ indicates that $x^m\xi(t)$ is interpreted in Stratonovich sense. We see from this and the previous discussion about substitution and use of scale parameters that the FWHM is well defined and stable for this equation as long as $a_2 < mD$ and $n - 2m + 1 < 0$. In other words if Eq. (23) is M-stable in the Itô interpretation then it is also M-stable in the Stratonovich interpretation as long as the condition $a_2 < (2-\nu)mD$ holds. We have seen that FWHM $\propto (1/a_1)^{\frac{1}{n-2m+1}}$.

It is interesting to ask more generally whether D behaves as a scale parameter for the more general Eq. (23) (where a_2

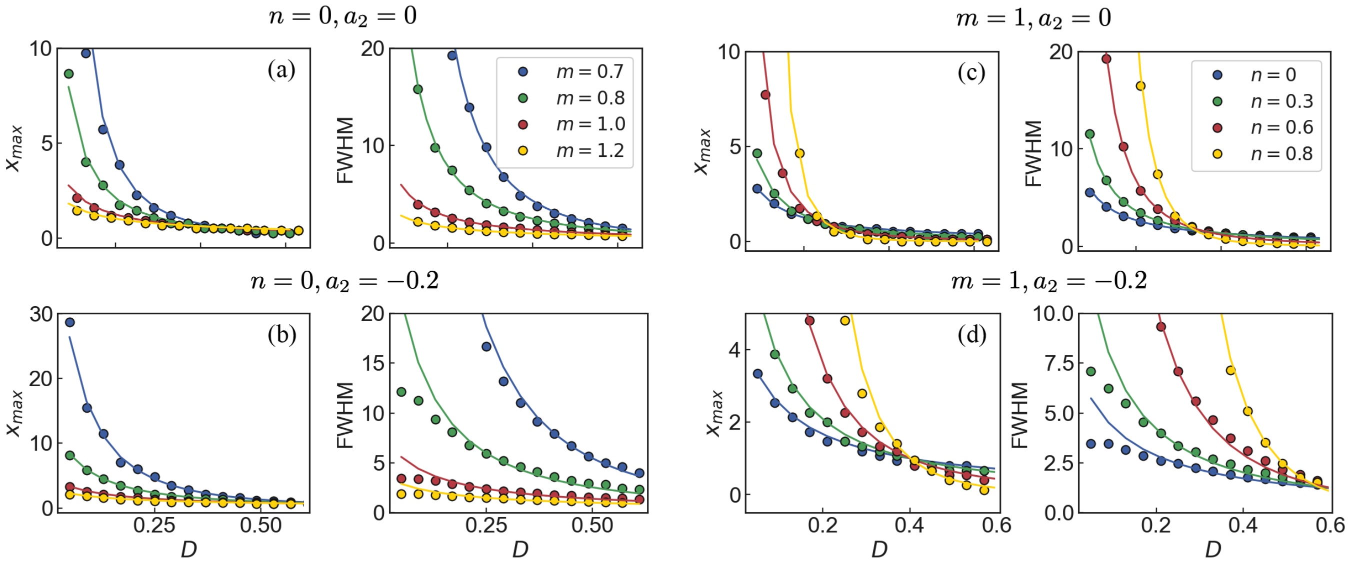

FIG. 5. $x_{\max}$ and FWHM of Eq. (23) (a) for fixed $n = 0, a_2 = 0$ ($a_1 = 0.5$), (b) $n = 0, a_2 = -0.2$ ($a_1 = 1$), (c) $m = 0, a_2 = 0$ ($a_1 = 0.5$), and (d) $m = 0, a_2 = -0.2$ ($a_1 = 1$). Lines of $x_{\max}$ and FWHM are fits using Eqs. (6) and (28), respectively.

is a constant independent of D). In this case we would expect the FWHM to vary with D as the maximum does according to

$$\text{FWHM} \propto \left(\frac{(2-\nu)mD - a_2}{a_1}\right)^{\frac{1}{n-2m+1}}. \tag{28}$$

This claim is explored in Figs. 5(a) and 5(b) for different m and Figs. 5(c) and 5(d) for different n. We see for $a_2 < 0$ in panels (b) and (d) that agreement between the conjecture and numerical results agrees increasingly well as D becomes larger. This is to be expected, since $\alpha \to 1$ and thus $y(x) \to x$ as $D \to \infty$. For $a_2 > 0$ the agreement seems to be good even for smaller D (results shown in the appendices). Numerical simulations suggest that a_2 is not a scale parameter.

B. Additive noise

We now consider the more general case where additive noise ξ_{add} is added to the SDE; i.e., we write Eq. (1) as

$$\begin{aligned}\dot{x} &= f(x) + \sqrt{2D_{\text{m}}}g(x)\xi_{\text{m}}(t) + \sqrt{2D_{\text{add}}}\xi_{\text{add}}(t)\\ &= f(x) + \sqrt{2D_{\text{m}}g(x) + 2D_{\text{add}}}\xi(t)\\ &= f(x) + \sqrt{2D_{\text{m}}}\tilde{g}(x)\xi(t),\end{aligned} \tag{29}$$

where $\tilde{g}(x) := \sqrt{g(x)^2 + D_{\text{add}}/D_{\text{m}}}$. The second line is a consequence that the two noises $\xi_{\text{add}}(t)$ and $\xi_{\text{m}}(t)$ are independent and Gaussian and thus the covariance is the sum of the individual covariances. It can easily be shown that $\tilde{g}(x)\tilde{g}'(x) = g(x)g'(x)$. For this reason this alteration has no affect on the maximum of the distribution, derived from Eq. (4). We therefore expect additive noise to also have no effect on the scale parameter, since $\beta \propto x_m$. We would for example expect the FWHM to vary with D_{m} in the presence of additive noise just as FWHM varies with D when no additive noise is present. This is indeed the case and will be explored more fully in a follow-up paper.

V. EXAMPLE

In this section we apply the results of this paper to elucidate the role played by multiplicative noise in the tipping points in a double-well potential. A classic model containing a double-well potential (and multiplicative noise) is that of the (noisy) Duffing oscillator described by the equation [51–53]

$$\frac{1}{\gamma}\ddot{x} + \dot{x} = [a + y(t)]x - bx^3, \tag{30}$$

where b is the (positive) friction coefficient and d is the mean frequency. $y(t)$ is Ornstein-Uhlenbeck noise with correlation function $\langle y(t)y(0)\rangle = qe^{-\Delta\omega|t|}$. This equation may for example be used to model the stability properties of the conductive state for Rayleigh-Bénard convection, if we consider the Rayleigh number to be a noisy function of time [31,32]. In the limits (i) $\gamma \gg \Delta\omega$ and (ii) $\Delta\omega \gg \gamma$ both may be represented by the same formal equation

$$\dot{x} = [a + \sqrt{2D}\xi(t)]x - bx^3 \tag{31}$$

if this equation is interpreted in the sense of Stratonovich in case (i) and in the sense of Itô in case (ii) [54]. This is the equation of a double-well potential with multiplicative noise. In the absence of noise this double-well potential has an unstable fixed point at $x = 0$ and two stable fixed points at $x = \pm\sqrt{a/b}$. The stationary probability distribution can be calculated as

$$p_s(x) = \frac{N}{2D}x^{\frac{a}{D}-2}\exp -\frac{b}{2D}x^2. \tag{32}$$

The FWHM is not simple to obtain from this equation, since by setting $p_s(x_H) = p_s(x_m)/2$ and calculating the maximum x_m we obtain the x_H at the point of the FWHM is a function of the nonanalytic Lambert W function. Nevertheless, considering the ideas of this paper [comparing $(b/2D)x^2$ and $(x/\beta)^2$] we see immediately that we have a scale parameter $\beta \propto b^{-1/2}$. Thus FWHM $\propto \beta \propto b^{-1/2}$. This agreement is illustrated in Fig. 6. In the case of multiple peaks for every single peak we

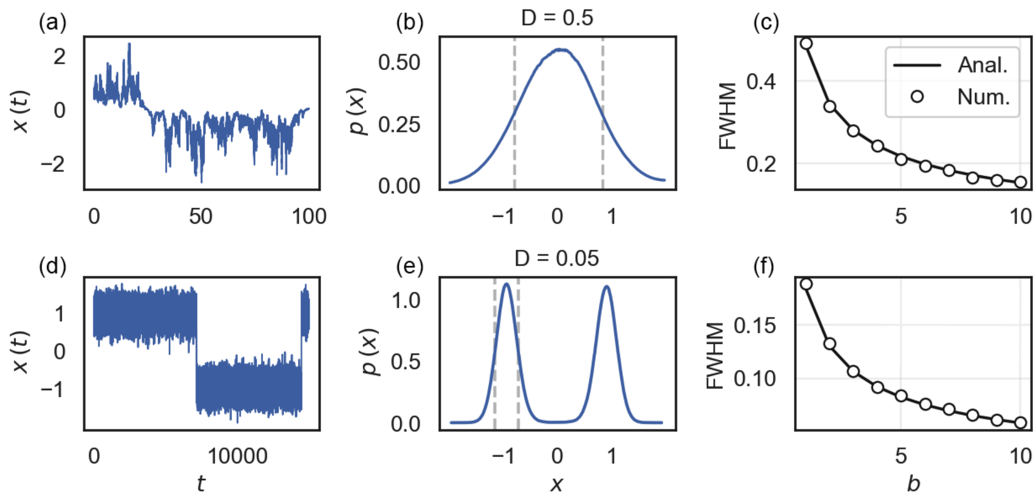

FIG. 6. Double-well potential of Eq. (31) with parameter $a = 1$, (a)–(c) $D = 0.5$, (a) time plot, $b = 1$, (b) stationary probability density, $b = 1$, (c) FWHM of (in deterministic case unstable) fixed point. (d)–(f) Same as (a)–(c) with $D = 0.05$, whereby a small amount of additive noise $D_{\rm add} = 0.01$ has been added to allow jumping across $x = 0$. Simulation time $T = 50\,000$ and time step $dt = 0.001$.

can define its individual FWHM as before in the case of the single peak. In the case of Fig. 6(e) both peaks are identical and so we may pick either peak for the analysis.

We now explore the behavior of the system in the vicinity of these fixed points. In the vicinity of $x = 0$ we argue that noise may stabilize the system. To show this we first calculate the maximum of the stationary probability distribution using Eq. (4) as

$$x_m = \pm\sqrt{(a - 2D)/b}. \tag{33}$$

As D is made larger ($\to a/2$) the maximum moves closer toward $x = 0$. In this region we make a simplifying approximation of $-bx^3 + ax \approx ax^{1-\epsilon}$ where $\epsilon > 0$ is a small number and $\epsilon \to 0$ as $D \to a/2$. We can see immediately that $n = 1 - \epsilon$ and $m = 1$ and thus $n - 2m < -1$. In other words the noise makes the fixed point at $x = 0$ M-stable. For $D > a/2$ the maximum of the stationary distribution is at $x = 0$. This again demonstrates the confining role of multiplicative noise, showing that white noise may stabilize the position of the potential well.

If D is small then we may linearize around $x = \sqrt{a/b}$ to obtain

$$\dot{x} \approx 2a\sqrt{a/b} - 2ax + \sqrt{2D}x\xi(t). \tag{34}$$

Using the substitution $y(x) = \alpha^{-1}x^{\alpha}$ where $\alpha = (D + 2a)/D$ we obtain

$$\dot{y} = 2a\sqrt{a/b}\left(\frac{D+2a}{D}y\right)^{\frac{2a}{D+2a}} + \sqrt{2D}\left(\frac{D+2a}{D}y\right)\xi(t). \tag{35}$$

We see that $n = 2a/(D + 2a) < 1$ and $m = 1$ and thus $n - 2m < -1$. Thus these minima are also M-stable. We also see as noted previously that $b^{-1/2}$ is a scale parameter. However the fact that noise of a similar double-well potential system causes the distribution to move toward zero at large intensities has been known from work from Graham *et al.* [54] that work focused on moments, such as $\langle x^2\rangle$ related to the variance. It was as such not able to account for the effect of the parameters on narrowing the peaks of the distribution.

The effect of multiplicative noise is argued in this paper to stabilize the system in the sense that it moves the body of the distribution to the point where the multiplicative noise term $|g(x)|$ is smallest, i.e., in this case at $x = 0$. It shows stabilization of a point in the phase space which is repelling without multiplicative noise. This is true of Eq. (35) both in the sense of the maximum [Eq. (33)] and the FWHM shown in Fig. 7. We distinguish again between the two cases $2D < a$ and $2D > a$. D is not a scale parameter of this system (even in limits of D). As mentioned when $2D < a$ there are two minima of the double well at $x = \sqrt{a/b}$ and when $2D \ll a$ we linearize around these points [see Fig. 6(e)]. We see from Eq. (33) that the maximum moves with increasing D away from these minima (toward $x = 0$) and from Fig. 7 that the FWHM becomes more diffused; i.e., with increasing D
the minima of the potential become less stable. As $2D > a$ the maximum is centered on $x = 0$ [see Fig. 6(b)], and here it is the linearization around $x = 0$, which is M-stable. In this domain increasing D causes the FWHM to become smaller; i.e., $x = 0$ becomes increasingly stabilized.

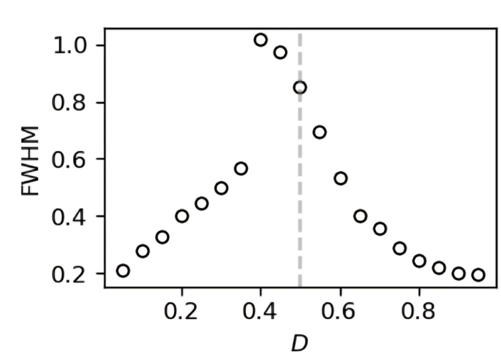

FIG. 7. FWHM of Eq. (31) where $b = 1$. The point where $2D = a$ is marked by a dashed line. When $2D$ is slightly less than a the two potentials still have separate maxima but are merged, and the FWHM is calculated from the merged distribution.

VI. CONCLUSIONS

We have shown that under quite general conditions (when $n - 2m < -1$) increasing multiplicative noise intensity shifts the mass of the stationary distribution closer to the minimum of the absolute value of $g(x)$ and behaves as a scale parameter. Intuitively $\dot{x} = g(x)\xi(t)$ is most stable at the minimum of $g(x)$, because at this point there is less noise acting to perturb the trajectory. We have shown that in this range there is exponentially distributed on-off intermittency, and that this goes hand in hand with the distribution having a scale parameter (a kind of self-similarity). The fact that both of these phenomena are connected is well known in turbulence. The boundary case $n - 2m = -1$ is scale-free power-law noise (with a power-law spectral density).

When exploring the analytically most simple case of natural boundaries we have defined the concept of a distribution being M-wide-sense stationary (having constant maximum and FWHM) and shown that the case of $n - 2m < -1$, even though it is not wide-sense stationary in the traditional sense (having diverging moments). We have shown importantly that the FWHM generally varies with the scale parameters in this regime in the same way as the maximum. The maximum and FWHM act as analogs to the first and second moments and collectively describe the body of the distribution, excluding the heavy tail. The bursts have been considered using first passage time theory.

Given essentially only the requirement of a significant multiplicative noise term we expect these results to be quite applicable to a range of different problems including extreme bursts, tipping points, and synchronization. We have applied the ideas of this paper to a double-well potential as a specific example. In particular these methods are important in the regime of large noise where standard small-noise perturbation methods are insufficient. In this regime we have shown that it is sufficient to study the first term of a Taylor expansion of both the deterministic and stochastic terms of the Langevin equation.

The authors have no conflict of interest to disclose.

APPENDIX A: SOME BASIC ANALYTICS

1. Delta function as $D \to \infty$

In this subsection we show that the stationary distribution of Eq. (5) with $n - 2m < -1$ becomes a delta function as $D \to \infty$. To prove these we calculate the integral of this function against a sufficiently good test function ϕ as $D \to \infty$. Rigorously, the Dirac function is only defined when acting on functions in the Schwartz space. Let ϕ be an infinitely smooth function on $\mathbb{R}$ with compact support $\phi \in C_c^\infty(\mathbb{R})$. We define

$$f_D := \frac{N}{x^{2m}} \exp - \frac{a}{Db^2} \frac{x^{n-2m+1}}{n-2m+1}. \tag{A1}$$

We now calculate the integral of this against our test function (using notation from functional analysis),

$$\begin{aligned} \langle f_D, \phi \rangle &= \int_0^\infty \frac{N}{x^{2m}} \exp - \frac{a}{Db^2} \frac{x^{n-2m+1}}{n-2m+1} \phi(x)\, dx \\ &= \int_0^\infty \frac{\widetilde{N}}{y^{2m}} \exp - \frac{a}{b^2} \frac{y^{n-2m+1}}{n-2m+1} \phi\big(yD^{\frac{1}{n-2m+1}}\big) dy, \end{aligned} \tag{A2}$$

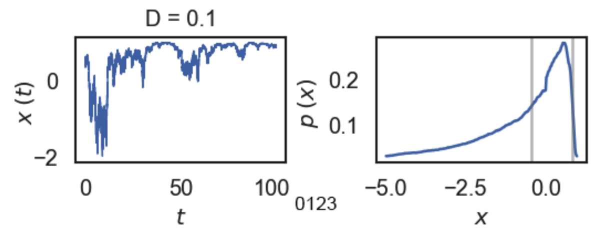

FIG. 8. Time plot and stationary solution of FPE of Eq. (A7). Gray lines indicate FWHM.

where we have used the fact the $p(x < 0) = 0$ and the substitution $y = xD^{-\frac{1}{n-2m+1}}$. We have defined $\widetilde{N} := ND^{\frac{1-2m}{n-2m+1}}$. We now look at the case where $n - 2m < -1$. As ϕ is continuous and compactly supported, this integrand is dominated by $(\widetilde{N}/y^{2m}) \exp(-\frac{a}{b^2}\frac{y^{n-2m+1}}{n-2m+1})\|\phi\|_\infty$, which integrates to $\|\phi\|_\infty$. Moreover, because ϕ is continuous the integrand converges pointwise to $\frac{\widetilde{N}}{y^{2m}} \exp - \frac{a}{b^2}\frac{y^{n-2m+1}}{n-2m+1}\phi(0)$ as $D \to \infty$. Applying the dominated convergences theorem yields

$$\begin{aligned} \lim_{D\to\infty} \langle f_D, \phi \rangle &= \frac{\widetilde{N}}{y^{2m}} \exp - \frac{a}{b^2} \frac{y^{n-2m+1}}{n-2m+1} \phi(0) \\ &= \phi(0) \end{aligned} \tag{A3}$$

since the stationary probability distribution is normalized to 1. Thus

$$\lim_{D\to\infty} \langle f_D, \phi \rangle = \langle \delta_0, \phi \rangle. \tag{A4}$$

This result is in a way very counterintuitive. As additive noise intensity increases distributions become wider and flatter (diffusion). Multiplicative noise in this sense seems to have exactly the opposite effect as additive noise, causing the distributions to become increasingly more peaked.

2. Generalized minima of $g(x)$

If $g(x)$ has a different zero $g(x_u) = 0$ to $x = 0$ or if the deterministic drift is instead going in the negative direction and the noise creates a gradient in the positive direction such as

$$\dot{x} = -a|x|^n + \sqrt{2D}c|x_u - x|^m \xi(t), \tag{A5}$$

then we may use a transformation $y = x_u - x$ to bring the equation into the form

$$\dot{y} = a(y - x_u)^n + \sqrt{2D}y^m \xi(t), \tag{A6}$$

which in the case of $n = 1$ takes the form of Eq. (5). In the case that the drift and diffusion are acting in the same direction the variable will simply drift into the zero $g(x) = 0$ at which point the noise will switch off. An example is shown in Fig. 8 for the equation

$$\dot{x} = -ax^n + b(x-1)^m \xi(t). \tag{A7}$$

The trajectory has laminar lengths in the vicinity of $x = 1$ and experiences bursts in the negative direction due to the negative drift term.

3. Noise-induced drift

In this section we discuss the drift induced by multiplicative noise. In the Stratonovich case noise induces a well known "Stratonovich" drift $Dg(x)g'(x)$. In the Itô interpretation this drift term is also present. To justify that claim we use a substitution, which obeys $h'(x) := \frac{1}{g(x)}$. Thus the general substitution must be $y = h(x) = \int \frac{dx}{g(x)}$ and $h''(x) = -g'(x)/g^2(x)$. In this way we obtain (the well known equation)

$$\begin{aligned} dy &= h'(x)\,dx + \frac{1}{2}h''(x)\,dx^2 \\ &= \frac{1}{g(x)}[f(x) - (2-\nu)Dg(x)g'(x)]\,dt + \sqrt{2D}dW_t. \end{aligned} \tag{A8}$$

The transformation $y = h(x)$ is known as the Lamperti transformation [70,71]. When $f(x)$ is more strongly nonlinear than $g(x)g'(x)$ then the drift is dominant. When it is the other way around then it is the multiplicative noise, which is dominant. When $f(x) \propto g(x)g'(x)$ then there is balance. These three cases correspond to our observed, nonstationarity [Fig. 1(e)], on-off intermittency [Fig. 1(c)], and power-law noise [Fig. 1(a)] cases, respectively. This transformation does however not shed light on properties such as the FWHM. These results highlight how the multiplicative noise creates a drift $Dg(x)g'(x)$, which holds in both the Itô and Stratonovich interpretations. The potential of y is given by $V = -\int (f - Dgg')/g\,dx$.

Due to the symmetry of the white noise the function $g(x)$ can be replaced with the function $|g(x)|$. In this way it can be seen that $y = h(x)$ is a monotonic function of x. This is in line with the behavior of the maximum [Eq. (4)].

It is clear that when $g(x)g'(x)$ varies more strongly x over the domain than $f(x)$ then the overall drift direction of the particle (at large x) is toward the zero $g(x) = 0$ (this is a generalization of the case $n - 2m < -1$). On the other hand if $f(x)$ varies more strongly x over the domain than $g(x)g'(x)$ then the dominant force on the particle (at large x) is forcing it toward the fixed point $f(x) = 0$ (this is a generalization of $n - 2m > -1$).

4. FPE

The following theory is well known (see for example [1,68]). The Fokker-Planck equation is given by

$$\begin{aligned} \frac{\partial}{\partial t}p(\phi,t) &= -\frac{\partial}{\partial \phi}\Big[\Big\{f(\phi) + \frac{\nu}{2}g'(\phi)g(\phi)\Big\}p(\phi,t)\Big] \\ &\quad + \frac{1}{2}\frac{\partial^2}{\partial \phi^2}[g^2(\phi)p(\phi,t)] \qquad \text{(A9)} \\ &= -\frac{\partial}{\partial \phi}J(\phi,t). \qquad \text{(A10)} \end{aligned}$$

We wish to obtain the stationary distribution $\partial_t p_s(\phi,t) = 0$. The solution to the homogeneous equation, known as the fundamental solution, is found to be

$$\psi(\phi) := \exp\left[2\int^{\phi}\frac{f(\phi') + \frac{1}{2}g'(\phi')g(\phi')}{g^2(\phi')}\,d\phi'\right]. \tag{A11}$$

The complete solution is found to be

$$p_s(\phi) = \frac{\psi(\phi)}{g^2(\phi)}\left(N - 2J\int^{\phi}\frac{d\phi'}{\psi(\phi')}\right). \tag{A12}$$

In the case of absorbing or natural boundary conditions $J = 0$. Thus the stationary solution becomes

$$p_s(x) = \frac{N}{g^2(x)}\exp 2\int^{x}\frac{f(u)}{g^2(u)}du. \tag{A13}$$

5. Proof-of-scale parameter

To show this we prove that $p_s(x;\beta) \propto p_s(x/\beta;1)$. From this Eq. (9) follows automatically, due to normalization. We now prove the first statement. Since $g(x)g'(x) = 2Dmx^{2m-1}$, the left-hand side of Eq. (9) is equal to

$$\begin{aligned} p_s(x;\beta) &= \frac{N}{2Dg^2(x)}e^{\frac{1}{D}\int_A^x \frac{f(u)+(\nu/2)g'(u)g(u)}{g^2(u)}\,du} \\ &\propto x^{(\nu-2)m}\exp\frac{a}{D}\frac{x^{n-2m+1}}{n-2m+1}, \end{aligned} \tag{A14}$$

where A is the left boundary, and $\nu = 1$ in the Stratonovich interpretation and $\nu = 0$ in the Itô interpretation. On the other hand we also have

$$\begin{aligned} p_s(x/\beta;1) &= N(x/\beta)^{-2m}e^{\int^{x/\beta}\frac{u^n}{u^{2m}}+\frac{m}{u}\,du} \\ &\propto x^{(\nu-2)m}\exp\frac{(x/\beta)^{n-2m+1}}{n-2m+1}. \end{aligned} \tag{A15}$$

From simple comparison β is obtained. Thus Eqs. (9) and (10) have been proven. The normalization constant is calculated as

$$\begin{aligned} N^{-1} &= \frac{1}{2D}\int_0^{\infty}x^{(\nu-2)m}e^{\frac{a}{D}\frac{x^{n-2m+1}}{n-2m+1}} \\ &= \frac{1}{2a}x^{\nu m-n}\left(\frac{a}{D\delta}x^{-\delta}\right)^{\frac{\nu m-n}{\delta}} \\ &\quad \times \Gamma\left(\frac{(2-\nu)m-1}{\delta},\frac{a}{D\delta}x^{-\delta}\right)\Bigg|_0^{\infty}, \end{aligned} \tag{A16}$$

where $\delta := 2m - n - 1 > 0$. In the Itô interpretation $N^{-1} < \infty$ for all $\delta > 0$. In the Stratonovich interpretation this is only guaranteed if we impose the additional requirement that $m > 1$. In the case that $n = 0$ and $m = 1$ in the Stratonovich interpretation we have

$$N^{-1} \propto \lim_{b\to\infty}\int_0^b \frac{e^{-\frac{a}{Dx}}}{x}\,dx = \Gamma(0), \tag{A17}$$

which is undefined. A similar result holds for all $m < 1$. In the special case $n = 0$ and $m > 0.5$ in the Itô interpretation we obtain the simple result

$$N^{-1} = \frac{1}{2D}\int_0^{\infty}x^{-2m}\exp\frac{a}{D}\frac{x^{1-2m}}{1-2m}\,dx = \frac{1}{2a}. \tag{A18}$$

In fact more generally we notice in the Itô interpretation for $m = 1$ ($n < 1$ follows from $\delta > 0$) that

$$N^{-1} \propto \frac{1}{D}\left(\frac{D}{a}\right)^{-\frac{1}{n-1}}. \tag{A19}$$

The fact that $P_s(x)$ is normalizable when $m \leqslant 1$ only in the Itô interpretation, not in the Stratonovich interpretation, requires some discussion. It is generally thought that both the Itô and Stratonovich interpretations should lead to qualitatively similar results. It is suggested in [1] (page 112) that if an SDE admits a stationary solution in the Itô interpretation, but not in the Stratonovich interpretation (or vice versa), then such a discrepancy should be interpreted as a "red warning light" and may indicate pathological features.

Our model does not seem to have any pathological features in its domain $x \in (0, \infty)$ such as not being differentiable at a point. Beyond this, in this paper we are interested in the maximum and FWHM, which are both independent of the normalization and are in fact qualitatively similar. Both distributions are M-stable. For this reason we are not concerned with the fact that one is normalizable while the other is not, which is a reflection only of the heaviness of the tails. The tails are however nevertheless also qualitatively the same, both decaying according to a power law.

We now ask in the more general case of Eq. (23) if there exists a scale parameter analogous to $D^{\frac{1}{n-2m+1}}$ when $a_2 \neq 0$:

$$p_s(x;\beta) = \frac{N}{2D} x^{\frac{a_2}{D}-2m} e^{\frac{a_1}{D}\frac{x^{n-2m+1}}{n-2m+1}} \propto x^{\frac{a_2}{D}-2m} e^{\frac{a_1}{D}\frac{x^{n-2m+1}}{n-2m+1}}, \tag{A20}$$

$$p_s(x/\beta;1) \propto (x/\beta)^{\frac{a_2}{1}-2m} e^{\frac{(x/\beta)^{n-2m+1}}{n-2m+1}} \propto x^{a_2-2m} e^{\frac{(x/\beta)^{n-2m+1}}{n-2m+1}}. \tag{A21}$$

These results seem to show from the mismatches in the power-law terms in general that D is not a scale parameter unless $a_2 = 0$ or $a_2 \propto D$ (equivalent to $a_2 = 0$ in the Stratonovich interpretation). a_1 is in general a scale parameter. We see that x scales according to $\beta_x \propto \text{FWHM} \propto (1/a_1)^{\frac{1}{n-2m+1}}$. Although the parameters a_2 and D are inextricably linked, and a_2 is not a scale parameter, we do still see in the limit of large D or small a_2 (more precisely $D \gg a_2$) that $\beta_x \propto D^{\frac{1}{n-2m+1}}$.

6. Diverging moments

We show diverging moments of Eq. (A20) when $\delta = 2m - n - 1 > 0$. We first note that due to the average of the noise being zero the expectation of the trajectory is identical to the deterministic trajectory, i.e., $\langle \dot{x} \rangle = f(x)$. Since the potential is chosen to be nonconfining the $\langle x \rangle$ drifts to infinity.

To calculate the moments of the stationary distribution more generally we start with Eq. (A20) and then

$$p_s(x) = \frac{N}{2D} x^{\frac{a_2}{D}-2m} \exp -\frac{a_1}{D}\frac{x^{-\delta}}{\delta}. \tag{A22}$$

The first moment is calculated as

$$\langle x \rangle = \frac{N}{2D}\int_0^\infty x^{1+\frac{a_2}{D}-2m} \exp -\frac{a_1}{D}\frac{x^{-\delta}}{\delta}\, dx \tag{A23}$$

$$= \frac{1}{\delta} x^\nu \left(\frac{a_1 x^{-\delta}}{D\delta}\right)^{\frac{\nu}{\delta}} \Gamma\left(-\frac{\nu}{\delta}, \frac{a_1}{D\delta} x^{-\delta}\right)\Big|_0^\infty \tag{A24}$$

$$= \infty, \tag{A25}$$

where $\nu := 2 + (a_2/D) - 2m$. Since the first moment diverges so too do all of the higher moments.

7. MFPT

We now look at intermittency without resetting. For the case that $D_2 = 0$ we can assume reflecting boundaries as $x = 0$ and the above equation for the MFPT can be simplified dramatically to

$$T(x) = \langle T \rangle = \int_0^\infty \int_0^b \rho(x', t|x, 0)\, dx'\, dt, \tag{A26}$$

which has a general solution

$$T(x) = 2\int_x^b \frac{dy}{\psi(y)} \int_a^y \frac{\psi(z)}{g^2(z)}\, dz. \tag{A27}$$

We now look at Eq. (5) ($n = 0$, $m = 1$) and calculate the fundamental solution

$$\psi(x) = \exp \int_a^x dx' \frac{\omega}{Dx^2} = Ce^{-\frac{\omega}{Dx}}, \tag{A28}$$

where C is a constant. When the left-hand side has a reflecting boundary this becomes

$$T(x) = 2\int_x^b dy e^{\frac{\omega}{Dy}} \int_0^y \frac{e^{-\frac{\omega}{Dz}}}{2Dz^2}\, dz = \int_x^b \frac{dy}{\omega} = \frac{b - x}{\omega}. \tag{A29}$$

From this we can obtain the mean laminar length $\langle l \rangle$ as

$$\langle l \rangle = |T(b) - T(x)| = \frac{|b - x|}{\omega}, \tag{A30}$$

and thus the number of bursts in a time period L, which we denote as χ, is

$$\chi(L) = \frac{L}{\langle l \rangle} = \frac{2\pi}{\langle l \rangle}. \tag{A31}$$

It is known that the absolute values of the outfluxes of probability are related to the conditional escape time densities [72].

8. Additive noise

The stationary probability distribution is given by

$$p_s(x) = \frac{N}{2D_1 x^{2m} + 2D_2} e^{2\int^x \frac{au^n}{2D_1 u^{2m} + 2D_2} du} = \frac{N}{2D_1 x^{2m} + 2D_2} \times \exp \frac{ax^{n+1}}{D_2(n+1)} {}_2F_1\left(1, \frac{n+1}{2m}; \frac{\chi}{2m}; \frac{-x^{2m}}{D_2/D_1}\right), \tag{A32}$$

where $\chi = 2m + n + 1$. For the case of $n = 0$ and $m = 1$ this reduces to

$$p_s(x) = \frac{N}{2D_1 x^{2m} + 2D_2} \times \exp \frac{2a}{\sqrt{D_1 D_2}}\left[n\pi + \tan^{-1}\left(\sqrt{\frac{D_1}{D_2}} x\right)\right], \tag{A33}$$

where $n \in \mathbb{Z}$ is defined such that $\pi(n-\frac{1}{2}) \leqslant \theta < \pi(n+\frac{1}{2})$. Simulations show that the FWHM decreases as $\propto D^{\frac{1}{n-2m+1}}$. We assert that additional additive noise D_2 makes no difference to this behavior. To be more explicit the noise does broaden the distribution, making the FWHM wider. The FWHM is nevertheless made smaller again by increasing multiplicative noise according to $\propto D^{\frac{1}{n-2m+1}}$.

9. Stability of paths

The conditions required for the existence and uniqueness in a time interval $[t_0, T]$ are the Lipschitz condition and the growth condition. Almost every stochastic differential equation encountered in practice satisfies the Lipschitz condition since it is essentially a smoothness condition. The growth condition is that there exists a K such that for all t in the range $[t_0, T]$

$$|f(x,t)|^2 + |g(x,t)|^2 \leqslant K^2(1+|x|^2). \tag{A34}$$

The growth condition in contrast to the Lipschitz condition is often violated. If these conditions hold then it is guaranteed that the solution will not become infinite in a finite time. For our case of $n = 0$ and $m = 1$ it is guaranteed that the solution will not explode. This is however not guaranteed if either $n > 1$ or $m > 1$. Simulations suggest that the trajectory does not explode over not too long time frames for n or m slightly greater than 1. Nevertheless it is for this reason that we generally restrict ourselves to the cases $n \leqslant 1$ and $m \leqslant 1$.

APPENDIX B: GENERAL

1. Results for $a_2 > 0$

Results for $a_2 > 0$ are shown in Fig. 9. Agreement seems to exist between conjecture and simulations even for smaller D, although simulations become increasingly more challenging in this domain, requiring a longer simulation time and a smaller time step.

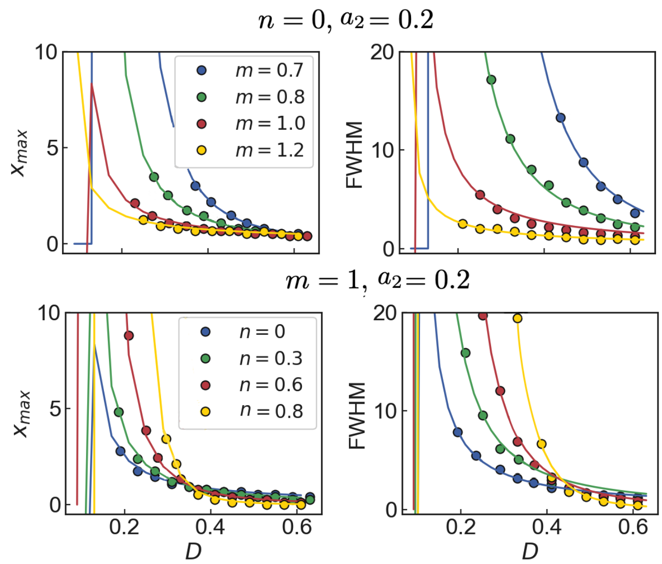

FIG. 9. $x_{\max}$ and FWHM of Eq. (23) for $a_2 = 0.2$ and $a_1 = 0.5$. Lines of $x_{\max}$ and FWHM are fits using Eqs. (6) and (28), respectively.

2. Higher-order "M-moments"

We also speculate that instead of higher-order Gaussian moments such as the kurtosis, quantifying a heavy-tailed distribution using the power law of the tail is more appropriate. The link between power laws and multiplicative noise has already been the work of recent literature [73–75].

[1] W. Horsthemke, Noise induced transitions, in *Non-Equilibrium Dynamics in Chemical Systems: Proceedings of the International Symposium, Bordeaux, France* (Springer, Heidelberg, Germany, 1984), pp. 150–160.
[2] U. Feudel and C. Grebogi, Multistability and the control of complexity, Chaos **7**, 597 (1997).
[3] A. N. Pisarchik and U. Feudel, Control of multistability, Phys. Rep. **540**, 167 (2014).
[4] U. Feudel, Complex dynamics in multistable systems, Int. J. Bifurcation Chaos **18**, 1607 (2008).
[5] S. Kraut and U. Feudel, Multistability, noise, and attractor hopping: The crucial role of chaotic saddles, Phys. Rev. E **66**, 015207(R) (2002).
[6] S. Kraut, U. Feudel, and C. Grebogi, Preference of attractors in noisy multistable systems, Phys. Rev. E **59**, 5253 (1999).
[7] G. Huerta-Cuellar, A. N. Pisarchik, and Y. O. Barmenkov, Experimental characterization of hopping dynamics in a multistable fiber laser, Phys. Rev. E **78**, 035202(R) (2008).
[8] S. L. de Souza, A. M. Batista, I. L. Caldas, R. L. Viana, and T. Kapitaniak, Noise-induced basin hopping in a vibro-impact system, Chaos, Solitons Fractals **32**, 758 (2007).
[9] P. S. Landa and P. V. E. McClintock, Changes in the dynamical behavior of nonlinear systems induced by noise, Phys. Rep. **323**, 1 (2000).
[10] E. Forgoston and R. O. Moore, A primer on noise-induced transitions in applied dynamical systems, SIAM Rev. **60**, 969 (2018).
[11] C. Van den Broeck, J. M. R. Parrondo, and R. Toral, Noise-induced nonequilibrium phase transition, Phys. Rev. Lett. **73**, 3395 (1994).
[12] L. Hodgkinson and M. Mahoney, Multiplicative noise and heavy tails in stochastic optimization, in *International Conference on Machine Learning* (PMLR, 2021), pp. 4262–4274.
[13] S. Kim, S. H. Park, and C. S. Ryu, Noise-enhanced multistability in coupled oscillator systems, Phys. Rev. Lett. **78**, 1616 (1997).
[14] S. Kim, S. H. Park, C. R. Doering, and C. S. Ryu, Reentrant transitions in globally coupled active rotators with multiplicative and additive noises, Phys. Lett. A **224**, 147 (1997).
[15] K. J. Lee, Y. Kwak, and T. K. Lim, Phase jumps near a phase synchronization transition in systems of two coupled chaotic oscillators, Phys. Rev. Lett. **81**, 321 (1998).

[16] S. Boccaletti, J. Kurths, G. Osipov, D. Valladares, and C. Zhou, The synchronization of chaotic systems, Phys. Rep. **366**, 1 (2002).
[17] J. L. Cabrera and J. G. Milton, On-off intermittency in a human balancing task, Phys. Rev. Lett. **89**, 158702 (2002).
[18] A. Krawiecki, J. A. Hołyst, and D. Helbing, Volatility clustering and scaling for financial time series due to attractor bubbling, Phys. Rev. Lett. **89**, 158701 (2002).
[19] B. B. Mandelbrot and B. B. Mandelbrot, *The Variation of Certain Speculative Prices* (Springer, Heidelberg, Germany, 1997).
[20] P. Imkeller and J.-S. Von Storch, *Stochastic Climate Models* (Springer Science & Business Media, 2001), Vol. 49.
[21] IPCC, Technical summary, in *Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by V. Masson-Delmotte, P. Zhai, A. Pirani, S. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J. Matthews, T. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (Cambridge University Press, Cambridge, 2021), pp. 33–144.
[22] A. J. Majda, I. Timofeyev, and E. Vanden Eijnden, Models for stochastic climate prediction, Proc. Natl. Acad. Sci. USA **96**, 14687 (1999).
[23] A. J. Majda, C. Franzke, and D. Crommelin, Normal forms for reduced stochastic climate models, Proc. Natl. Acad. Sci. USA **106**, 3649 (2009).
[24] P. D. Sardeshmukh and P. Sura, Reconciling non-Gaussian climate statistics with linear dynamics, J. Clim. **22**, 1193 (2009).
[25] J. Berner, U. Achatz, L. Batte, L. Bengtsson, A. d. l. Cámara, H. M. Christensen, M. Colangeli, D. R. Coleman, D. Crommelin, S. I. Dolaptchiev *et al.*, Stochastic parameterization: Toward a new view of weather and climate models, Bull. Am. Meteorol. Soc. **98**, 565 (2017).
[26] P. Sura, M. Newman, C. Penland, and P. Sardeshmukh, Multiplicative noise and non-Gaussianity: A paradigm for atmospheric regimes? J. Atmos. Sci. **62**, 1391 (2005).
[27] R. Buizza, M. Milleer, and T. N. Palmer, Stochastic representation of model uncertainties in the ECMWF ensemble prediction system, Q. J. R. Meteorol. Soc. **125**, 2887 (1999).
[28] B. Wang and J. Aljadeff, Multiplicative shot-noise: A new route to stability of plastic networks, Phys. Rev. Lett. **129**, 068101 (2022).
[29] J. Bauermann and B. Lindner, Multiplicative noise is beneficial for the transmission of sensory signals in simple neuron models, Biosystems **178**, 25 (2019).
[30] I. Pavlyukevich, Lévy flights, non-local search and simulated annealing, J. Comput. Phys. **226**, 1830 (2007).
[31] G. Ahlers, P. C. Hohenberg, and M. Lücke, Externally modulated Rayleigh-Bénard convection: Experiment and theory, Phys. Rev. Lett. **53**, 48 (1984).
[32] M. Lücke and F. Schank, Response to parametric modulation near an instability, Phys. Rev. Lett. **54**, 1465 (1985).
[33] M. Das and H. Kantz, Stochastic resonance and hysteresis in climate with state-dependent fluctuations, Phys. Rev. E **101**, 062145 (2020).
[34] N. Platt, E. A. Spiegel, and C. Tresser, On-off intermittency: A mechanism for bursting, Phys. Rev. Lett. **70**, 279 (1993).
[35] J. F. Heagy, N. Platt, and S. M. Hammel, Characterization of on-off intermittency, Phys. Rev. E **49**, 1140 (1994).
[36] M. Ding and W. Yang, Stability of synchronous chaos and on-off intermittency in coupled map lattices, Phys. Rev. E **56**, 4009 (1997).
[37] S. Elaskar and E. del Río, Review of chaotic intermittency, Symmetry **15**, 1195 (2023).
[38] P. Ashwin, J. Buescu, and I. Stewart, Bubbling of attractors and synchronisation of chaotic oscillators, Phys. Lett. A **193**, 126 (1994).
[39] A. E. Hramov, A. A. Koronovskii, M. K. Kurovskaya, and S. Boccaletti, Ring intermittency in coupled chaotic oscillators at the boundary of phase synchronization, Phys. Rev. Lett. **97**, 114101 (2006).
[40] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Van Nes, M. Rietkerk, and G. Sugihara, Early-warning signals for critical transitions, Nature (London) **461**, 53 (2009).
[41] M. R. Young and S. Singh, Statistical properties of a laser with multiplicative noise, Opt. Lett. **13**, 21 (1988).
[42] S. Zhu, Steady-state analysis of a single-mode laser with correlations between additive and multiplicative noise, Phys. Rev. A **47**, 2405 (1993).
[43] S. N. Chowdhury, A. Ray, S. K. Dana, and D. Ghosh, Extreme events in dynamical systems and random walkers: A review, Phys. Rep. **966**, 1 (2022).
[44] D. Schertzer and S. Lovejoy, Physical modeling and analysis of rain and clouds by anisotropic scaling multiplicative processes, J. Geophys. Res.: Atmos. **92**, 9693 (1987).
[45] S. Lovejoy, Spectra, intermittency, and extremes of weather, macroweather and climate, Sci. Rep. **8**, 12697 (2018).
[46] P. Ashwin, S. Wieczorek, R. Vitolo, and P. Cox, Tipping points in open systems: Bifurcation, noise-induced and rate-dependent examples in the climate system, Philos. Trans. R. Soc. A **370**, 1166 (2012).
[47] P. Sura, Noise-induced transitions in a barotropic β-plane channel, J. Atmos. Sci. **59**, 97 (2002).
[48] E. T. Phillips, The synchronizing role of multiplexing noise: Exploring Kuramoto oscillators and breathing chimeras, Chaos **33**, 073140 (2023).
[49] W. Gilpin, Desynchronization of jammed oscillators by avalanches, Phys. Rev. Res. **3**, 023206 (2021).
[50] A. Schenzle and H. Brand, Multiplicative stochastic processes in statistical physics, Phys. Rev. A **20**, 1628 (1979).
[51] S. Aumaître, K. Mallick, and F. Pétrélis, Noise-induced bifurcations, multiscaling and on-off intermittency, J. Stat. Mech: (2007) P07016.
[52] R. Bourret, U. Frisch, and A. Pouquet, Brownian motion of harmonic oscillator with stochastic frequency, Physica **65**, 303 (1973).
[53] K. Lindenberg, V. Seshadri, and B. J. West, Brownian motion of harmonic systems with fluctuating parameters: III Scaling and moment instabilities, Physica A **105**, 445 (1981).
[54] R. Graham and A. Schenzle, Stabilization by multiplicative noise, Phys. Rev. A **26**, 1676 (1982).
[55] L. Arnold, H. Crauel, and V. Wihstutz, Stabilization of linear systems by noise, SIAM J. Control Optim. **21**, 451 (1983).
[56] L. Arnold, W. Horsthemke, and J. Stucki, The influence of external real and white noise on the Lotka-Volterra model, Biom. J. **21**, 451 (1979).

[57] R. Bobrik and L. Stettner, The stabilizing effect of a random parametric perturbation on an unstable linear system, J. Math. Sci. **96**, 3038 (1999).
[58] C. Gardiner, *Stochastic Methods* (Springer, Berlin, 2009).
[59] A. W. Lau and T. C. Lubensky, State-dependent diffusion: Thermodynamic consistency and its path integral formulation, Phys. Rev. E **76**, 011123 (2007).
[60] B. Kaulakys, J. Ruseckas, V. Gontis, and M. Alaburda, Nonlinear stochastic models of $1/f$ noise and power-law distributions, Physica A **365**, 217 (2006).
[61] B. Kaulakys and M. Alaburda, Modeling scaled processes and $1/f^{\beta}$ noise using nonlinear stochastic differential equations, J. Stat. Mech. (2009) P02051.
[62] J. Ruseckas and B. Kaulakys, $1/f$ Noise from nonlinear stochastic differential equations, Phys. Rev. E **81**, 031105 (2010).
[63] F. Arecchi and F. Lisi, Hopping mechanism generating $1/f$ noise in nonlinear systems, Phys. Rev. Lett. **49**, 94 (1982).
[64] A. Ben-Mizrachi, I. Procaccia, N. Rosenberg, A. Schmidt, and H. Schuster, Real and apparent divergencies in low-frequency spectra of nonlinear dynamical systems, Phys. Rev. A **31**, 1830 (1985).
[65] I. Procaccia and H. Schuster, Functional renormalization-group theory of universal $1/f$ noise in dynamical systems, Phys. Rev. A **28**, 1210 (1983)
[66] H. G. Schuster and W. Just, *Deterministic Chaos: An Introduction* (John Wiley & Sons, 2006).
[67] A. E. Hramov, A. A. Koronovskii, M. K. Kurovskaya, A. A. Ovchinnikov, and S. Boccaletti, Length distribution of laminar phases for type-I intermittency in the presence of noise, Phys. Rev. E **76**, 026206 (2007).
[68] J. E. Hirsch, B. A. Huberman, and D. J. Scalapino, Theory of intermittency, Phys. Rev. A **25**, 519 (1982).
[69] W.-H. Kye and C.-M. Kim, Characteristic relations of type-I intermittency in the presence of noise, Phys. Rev. E **62**, 6304 (2000).
[70] J. K. Møller and H. Madsen, *From State Dependent Diffusion to Constant Diffusion in Stochastic Differential Equations by the Lamperti Transform* (2010).
[71] B. Oksendal, *Stochastic Differential Equations: An Introduction with Applications* (Springer Science & Business Media, 2013).
[72] B. Lindner, M. Kostur, and L. Schimansky-Geier, Optimal diffusive transport in a tilted periodic potential, Fluct. Noise Lett. **1**, R25 (2001).
[73] M. Levy and S. Solomon, Power laws are logarithmic Boltzmann laws, Int. J. Mod. Phys. C **7**, 595 (1996).
[74] D. Sornette and R. Cont, Convergent multiplicative processes repelled from zero: Power laws and truncated power laws, J. Phys. I **7**, 431 (1997).
[75] D. Sornette, Multiplicative processes and power laws, Phys. Rev. E **57**, 4811 (1998).