Saddle avoidance of noise-induced transitions in multiscale systems

Börner, R. ORCID: https://orcid.org/0000-0003-4593-7721, Deeley, R. ORCID: https://orcid.org/0009-0004-9653-8100, Römer, R. ORCID: https://orcid.org/0009-0006-4812-0150, Grafke, T. ORCID: https://orcid.org/0000-0003-0839-676X, Lucarini, V. ORCID: https://orcid.org/0000-0001-9392-1471 and Feudel, U. ORCID: https://orcid.org/0000-0002-6638-6095
(2024) Saddle avoidance of noise-induced transitions in multiscale systems. Physical Review Research, 6 (4). L042053. ISSN 2643-1564 doi: 10.1103/physrevresearch.6.l042053 Available at https://centaur.reading.ac.uk/119856/

Letter

Saddle avoidance of noise-induced transitions in multiscale systems

Reyk Börner [1,*,†], Ryan Deeley [2,*,‡], Raphael Römer [3], Tobias Grafke [4], Valerio Lucarini [5], and Ulrike Feudel [2]

[1]*Department of Mathematics and Statistics, University of Reading, Reading RG6 6AX, United Kingdom*
[2]*Theoretische Physik/Komplexe Systeme, ICBM, Carl von Ossietzky Universität Oldenburg, Oldenburg 26129, Germany*
[3]*Department of Mathematics and Statistics, University of Exeter, Exeter EX4 4QF, United Kingdom*
[4]*Mathematics Institute, University of Warwick, Coventry CV4 7AL, United Kingdom*
[5]*School of Computing and Mathematical Sciences, University of Leicester, Leicester LE1 7RH, United Kingdom*

(Received 16 November 2023; revised 11 June 2024; accepted 7 October 2024; published 2 December 2024)

In multistable dynamical systems driven by weak Gaussian noise, transitions between competing states are often assumed to pass via a saddle on the separating basin boundary. By contrast, we show that timescale separation can cause saddle avoidance in nongradient systems. Using toy models from neuroscience and ecology, we study cases where sample transitions deviate strongly from the instanton predicted by the Freidlin-Wentzell theory, even for weak finite noise. We attribute this to a flat quasipotential and present an approach based on the Onsager-Machlup action to aptly predict transition paths.

DOI: 10.1103/PhysRevResearch.6.L042053

Multistable systems, when randomly perturbed, may undergo transitions between their coexisting attracting states [1,2]. Examples range from brain activity [3] and gene regulation [4] to lasers [5], planetary atmospheres [6] and the earth system [7,8]. Transitions often represent high-impact low-probability events, e.g., financial crashes [9], ecosystem collapse [10], and climate tipping points [11–13]. Understanding critical transitions is crucial to assess a system's stability and resilience [14].

Noise-driven systems are commonly formulated as stochastic differential equations of the Itô type,

$$\mathrm{d}\boldsymbol{x} = \boldsymbol{b}(\boldsymbol{x})\,\mathrm{d}t + \sigma\,\boldsymbol{\Sigma}(\boldsymbol{x})\,\mathrm{d}\boldsymbol{W}_t, \quad \boldsymbol{x}(0) = \boldsymbol{x}_0, \quad t \geqslant 0, \qquad (1)$$

where $\boldsymbol{x}(t) \in \mathbb{R}^D$ evolves under the combined effect of the deterministic drift $\boldsymbol{b}(\boldsymbol{x}(t)) : \mathbb{R}^D \to \mathbb{R}^D$ and a stochastic forcing by a D-dimensional Wiener process $\boldsymbol{W}_t$ scaled with noise amplitude $\sigma > 0$. We assume that the system $\dot{\boldsymbol{x}} = \boldsymbol{b}(\boldsymbol{x})$ possesses multiple attractors (hence being multistable) and consider nondegenerate noise, ensuring the covariance matrix $\boldsymbol{Q}(\boldsymbol{x}) = \boldsymbol{\Sigma}(\boldsymbol{x})\boldsymbol{\Sigma}^\top(\boldsymbol{x}) \in \mathbb{R}^{D\times D}$ is invertible [15]. This modeling framework is fundamental to climate physics (Hasselmann's program [16–18]), chemical physics [19], theoretical biology [4,14], and further applications of complex physics [5,20–22].

The theory and intuition about systems described by Eq. (1) are often guided by the *gradient* case, where $\boldsymbol{b}(\boldsymbol{x}) = -\boldsymbol{\nabla} V(\boldsymbol{x})$ and isotropic noise is assumed. The potential $V : \mathbb{R}^D \to \mathbb{R}$ then describes an energy landscape that visualizes the system's global stability, with attractors located at local minima of the landscape and *saddles* marking "mountain passes" for noise-induced transitions—where trajectories are most likely to cross between different valleys. However, most systems of interest are out of equilibrium, meaning no potential V exists whose negative gradient is $\boldsymbol{b}$.

For this general *nongradient* case, the Freidlin-Wentzell (FW) theory [23] introduces the *quasipotential* as an energy-like scalar field [23–27]. Following a large deviation principle [28], the quasipotential is computed via a variational approach of minimizing the FW *action* functional S_T (see below). Intuitively, this functional measures the "energetic cost" of a trajectory in the limit of vanishing noise. As $\sigma \to 0$, transitions between competing attractors concentrate around a minimum action path, or *instanton*, and the probability of deviating from the instanton decays exponentially [23,29].

In most cases, the instanton from one attractor to another passes through a saddle of $\boldsymbol{b}$ where the quasipotential has a minimum along the boundary separating the different basins of attraction [15,30,31]. Transition rates can be computed in terms of the quasipotential value at the saddle [32]. These results of the FW theory have established the widespread view of saddles acting as gateways of noise-induced transitions.

However, the $\sigma \to 0$ limit is never attained in reality; various counterexamples attest that, if the noise is weak yet noninfinitesimal, noise-induced transitions do not necessarily pass near a saddle [21,33–39]. Here we show that saddle avoidance can occur when nongradient dynamics features fast and slow degrees of freedom. We demonstrate that sample transition paths may deviate significantly from the FW instanton even for noise so weak that transitions become extremely rare. We resolve the apparent disagreement between the FW theory and observations using the quasipotential and introducing a variational formulation based on the Onsager-Machlup (OM) action [40–45].

Predicting noise-induced transitions. Rare events tend to be predictable in the sense that they will most probably occur

*These authors contributed equally to this work.
†Contact author: reyk.boerner@reading.ac.uk
‡Contact author: ryan.deeley@uni-oldenburg.de

in the least unlikely way. In our context, noise-induced transitions between attractors of $\dot{\boldsymbol{x}} = \boldsymbol{b}(\boldsymbol{x})$ become increasingly rare as $\sigma \to 0$. The FW theory quantifies this via a large deviation principle: consider the set $\mathcal{C}_{IF}^{T}$ of paths leading from state $\boldsymbol{x}_I$ to state $\boldsymbol{x}_F$ in time T. For $\sigma \to 0$, the probability that a solution $\boldsymbol{\phi}_t$ to Eq. (1) initialized at $\boldsymbol{x}_I$ remains inside the δ tube of the path $\boldsymbol{\varphi}_t \in \mathcal{C}_{IF}^{T}$ follows

$$\mathbb{P}_\sigma\left(\sup_{0\leqslant t\leqslant T} ||\boldsymbol{\phi}_t - \boldsymbol{\varphi}_t|| < \delta\right) \overset{\sigma\downarrow 0}{\asymp} \exp\left(-\frac{S_T[\boldsymbol{\varphi}_t] + \varepsilon_\delta}{\sigma^2}\right), \quad (2)$$

where $\asymp$ denotes asymptotic logarithmic equivalence and ε_δ satisfies $\lim_{\delta\downarrow 0} \varepsilon_\delta = 0$ [23,46]. The rate function $S_T[\boldsymbol{\varphi}_t]$ is the FW action functional,

$$S_T[\boldsymbol{\varphi}_t] = \frac{1}{2}\int_0^T ||\dot{\boldsymbol{\varphi}}_t - \boldsymbol{b}(\boldsymbol{\varphi}_t)||^2_{\boldsymbol{Q}(\boldsymbol{\varphi}_t)} \mathrm{d}t, \quad (3)$$

which measures the "work" done against $\boldsymbol{b}$, as weighted via the $\boldsymbol{Q}$ metric $||\boldsymbol{v}||_{\boldsymbol{Q}} := \sqrt{\langle \boldsymbol{v}, \boldsymbol{Q}^{-1}\boldsymbol{v}\rangle}$. This implies that, in the limit $\sigma \to 0$, the most probable path connecting $\boldsymbol{x}_I$ to $\boldsymbol{x}_F$ is by far the instanton, $\boldsymbol{\varphi}^{*IF} := \arg\min_{\boldsymbol{\varphi}_t \in \mathcal{C}_{IF}^{T},\, T>0} S_T[\boldsymbol{\varphi}_t]$ [29]. Starting from a reference attractor $\boldsymbol{x}_R = \boldsymbol{x}_I$, the "difficulty" of reaching any state $\boldsymbol{x}_F$ is given by the FW quasipotential,

$$V_R(\boldsymbol{x}_F) = \inf_{T>0} \inf_{\boldsymbol{\varphi}_t \in \mathcal{C}_{RF}^{T}} S_T[\boldsymbol{\varphi}_t]. \quad (4)$$

By definition, $V_R(\boldsymbol{x}_R) = 0$ is the strict global and unique strict local minimum.

Now suppose Eq. (1) has two stable equilibria $\boldsymbol{x}_R$ and $\boldsymbol{x}_L$ with basins of attraction $\mathcal{B}_R, \mathcal{B}_L \subset \mathbb{R}^D$, separated by the basin boundary $\partial\mathcal{B}$. We define a transition path $\boldsymbol{\phi}_t^{RL}$ as a trajectory that, after exiting a small neighborhood R around $\boldsymbol{x}_R$, enters a small neighborhood L around $\boldsymbol{x}_L$ without re-entering R. Transitions can reach $\boldsymbol{x}_L$ at zero action after crossing the boundary $\partial\mathcal{B}$ (by following a flow line of $\boldsymbol{b}$); therefore, as $\sigma \to 0$, the most probable location to cross $\partial\mathcal{B}$ is at the global minimum of V_R when restricting to $\partial\mathcal{B}$. This minimum is typically a saddle of $\boldsymbol{b}$ and determines the Kramers-like scaling law of the *mean first-exit time* $\langle\tau_\sigma^{RL}\rangle$, i.e., the expected waiting time until a trajectory initialized at $\boldsymbol{x}_R$ leaves $\mathcal{B}_R$, through $\langle\tau_\sigma^{RL}\rangle \asymp \exp[\sigma^{-2}\min_{\boldsymbol{x}\in\partial\mathcal{B}} V_R(\boldsymbol{x})]$ [23,46–50]. This emphasizes the relevance of saddles for noise-induced transitions as $\sigma \to 0$.

Apart from the mean first-exit time, a second timescale characterizing the transition is the *mean transition time* $\langle t_\sigma^{RL}\rangle$, i.e., the average time $\boldsymbol{\phi}_t^{RL}$ takes to travel to L after last leaving R. Large values of $r_\sigma^{RL} := \langle\tau_\sigma^{RL}\rangle/\langle t_\sigma^{RL}\rangle$ indicate that individual transitions are *rare* and therefore occur as a memoryless Poisson process. In what follows we choose σ sufficiently small such that $r_\sigma^{RL} \gg 1$, in which case one might expect the FW theory to apply. We investigate the transition behavior in two paradigmatic two-dimensional multiscale models, one driven by additive noise and another by multiplicative noise.

FitzHugh-Nagumo model. Let us first consider the FitzHugh-Nagumo (FHN) model, which was originally conceived to describe a spiking neuron [51,52] and has been widely studied as a multiscale conceptual model in theoretical neuroscience. The model can be written as

$$\dot{\boldsymbol{x}} = \boldsymbol{b}_{\mathrm{FHN}}(u, v) = \begin{pmatrix} \varepsilon^{-1}(-u^3 + u - v) \\ -\beta v + u \end{pmatrix}. \quad (5)$$

Here $\boldsymbol{x} = (u, v) \in \mathbb{R}^2$ is the two-dimensional state vector, $\varepsilon > 0$ denotes the timescale parameter, and we set $\beta = 3$. We consider Eq. (1) with $\boldsymbol{b} = \boldsymbol{b}_{\mathrm{FHN}}$ and identity covariance matrix $\boldsymbol{Q} = \boldsymbol{I}_2$ (additive noise). The noise-free system is bistable, possessing stable equilibrium points at $\boldsymbol{x}_{R,L} = \pm(\sqrt{2/3}, \sqrt{2/3^3})$ and a saddle point at $\boldsymbol{x}_M = (0, 0)$.

For $\varepsilon \ll 1$, Eq. (5) describes a *fast-slow* system where u is *fast*, while v is *slow* [53]. In the u direction, deterministic trajectories rapidly approach the *critical manifold* $C_0 := \{(u, v) \in \mathbb{R}^2 : v = -u^3 + u\}$ to which the slow dynamics is confined as $\varepsilon \to 0$ [53]. The cubic form of C_0 yields two saddle-node bifurcations with respect to v. Two stable branches of C_0 are separated by an unstable branch between the fold points $v_f^{\pm} = \pm\sqrt{4/3^3}$.

For such stochastic fast-slow systems with $0 < \varepsilon \ll 1$, one anticipates sample trajectories to closely track the stable part of C_0 until they approach a fold point, where they may abruptly transition to the opposite stable branch [54,55]. By contrast, the FW theory predicts that sample transition paths pass arbitrarily near the saddle for sufficiently small σ [23]. We thus face two competing limits, $\sigma \to 0$ and $\varepsilon \to 0$, when studying weakly noise-driven multiscale systems [56].

Using Monte Carlo simulations, we sample 100 transitions $\boldsymbol{\phi}_t^{RL}$ for each $\varepsilon \in \{0.01, 0.1, 1, 10\}$, fixing $\sigma(\varepsilon)$ to maintain $r_\sigma^{RL} \approx 10^5$ (details can be found in the Supplemental Material [57]). The ensembles of transition paths concentrate within a tube around a *mean sample transition path* (MTP), which we compute by spatial averaging over the ensemble [Fig. 1(a)]. Additionally, using the geometric minimum action method (gMAM) [58], we compute the corresponding instantons $\boldsymbol{\varphi}^{*RL}$ which minimize the FW action functional S_T [Eq. (3)].

The computed instantons always pass through the saddle. They approach it at an angle relative to $\partial\mathcal{B}$ that narrows as ε decreases, becoming tangential to $\partial\mathcal{B}$ when $\varepsilon < 1/\beta$ [36] [Fig. 1(a)]. For $\varepsilon \in \{1, 10\}$, the MTP closely matches the instanton and all sample paths cross the basin boundary within $\sim\sigma$ distance on either side of the saddle. Contrarily, for $\varepsilon \in \{0.01, 0.1\}$, sample transitions avoid the saddle: the MTP diverges from the instanton after getting close to the basin boundary [34], which happens before reaching the saddle. Once the noise kicks the trajectory into the competing basin, it is repelled from the basin boundary stronger than it is attracted towards the saddle.

This multiple timescale effect manifests itself in the ratio of the negative and positive eigenvalues $\lambda_\mp$ of the Jacobian of $\boldsymbol{b}$ at the saddle, $\mu := |\lambda_-|/\lambda_+$. If $\mu < 1$, Ref. [36] has shown that sample transition paths avoid the saddle on an extended lengthscale $\mathcal{O}(\sigma^\mu)$ [59,60]: their exit locations (where they cross the basin boundary) follow a one-sided Weibull distribution [61], whose mode approaches the saddle only logarithmically as $\sigma \to 0$ [35,36]. This contrasts with the case $\mu > 1$ where the distribution of exit locations is centered around the saddle on the lengthscale $\mathcal{O}(\sigma)$ as $\sigma \to 0$ [36]. In the FHN model, $\mu < 1$ if and only if $0 < \varepsilon < 1/\beta$, and we have $\mu = (\beta - 1)\varepsilon + \mathcal{O}(\varepsilon^2)$ as $\varepsilon \to 0$, which highlights the link between saddle avoidance and timescale separation (see Fig. S1 in the Supplemental Material [57]). The case $\varepsilon = 0.01$ exemplifies the fast-slow behavior anticipated for $\varepsilon \to 0$: the MTP escapes from the basin boundary near the

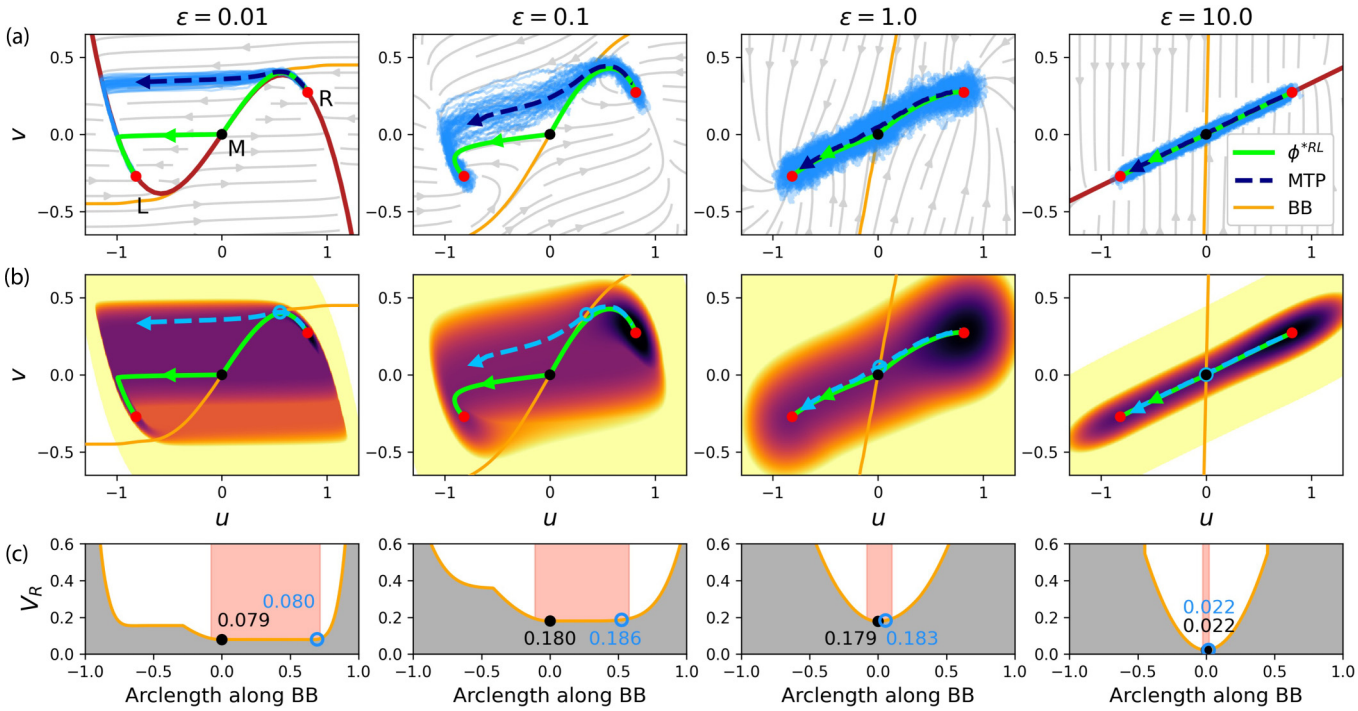

FIG. 1. FHN model for $\varepsilon = (0.01, 0.1, 1, 10)$ at $\sigma(\varepsilon) \approx (0.08, 0.12, 0.12, 0.04)$. (a) Phase space with equilibria $\boldsymbol{x}_{L,R}$ (red), saddle $\boldsymbol{x}_M$ (black), drift $\boldsymbol{b}$ (gray flow lines), instanton $\boldsymbol{\varphi}^{*RL}$ (green), 50 sample transition paths (light blue), MTP (blue dashed), and basin boundary (BB, orange); critical manifold C_0 (red) for $\varepsilon \to 0$ (left) and $1/\varepsilon \to 0$ (right panel). (b) Like panel (a) but with $V_R(\boldsymbol{x})$ on a logarithmic colormap (black: $V_R = 0$; brighter = larger V_R), showing $\boldsymbol{x}_c$ (blue circle) where MTP crosses BB. (c) $V_R(\boldsymbol{x})$ along BB, indicating values at $\boldsymbol{x}_M$ (black) and $\boldsymbol{x}_c$ (blue), and the region $\mathcal{Z}$ (red shading).

bifurcation point v_f^+ where C_0 becomes unstable [Fig. 1(a)]. In the inverse limit $1/\varepsilon \to 0$ (where $\mu \gg 1$), the corresponding critical manifold given by $\beta v = u$ is globally stable; the limiting behavior of both σ and $1/\varepsilon$ aligns by forcing the instanton and sample transition paths onto C_0, as seen for $\varepsilon = 10$.

Can we understand the observed transition behavior by means of the quasipotential V_R [Eq. (4)], which measures the difficulty of reaching a point $\boldsymbol{x}$ from $\boldsymbol{x}_R$? As $\sigma \to 0$, the probability of passing through the global minimum $\boldsymbol{x}^*$ of V_R along the basin boundary $\partial\mathcal{B}$ approaches 1 [23]. For finite σ, however, the large deviation principle underlying the FW theory [Eq. (2)] suggests that transitions may cross with similar probability in regions $\mathcal{Z}(\sigma) = \{\boldsymbol{z} \in \partial\mathcal{B} : V_R(\boldsymbol{z}) \leqslant V_R(\boldsymbol{x}^*) + \sigma^2\}$.

We compute V_R for the four ε cases considered using the OLIM4VAD algorithm [62] (see the Supplemental Material [57]). For $\varepsilon = 10$, a steep and narrow trench of low quasipotential connects $\boldsymbol{x}_R$ with $\boldsymbol{x}_L$ [Fig. 1(b)]. As ε decreases, however, V_R increasingly flattens along the v direction, leading to an extended *quasipotential plateau* around the saddle $\boldsymbol{x}_M$. Along $\partial\mathcal{B}$, V_R indeed assumes a global minimum at $\boldsymbol{x}_M$ for all ε, but its curvature around $\boldsymbol{x}_M$ becomes small for $\varepsilon \ll 1$ [Fig. 1(c)]. The low curvature implies the lack of a clear "mountain pass" and a widened "danger zone" $\mathcal{Z}(\sigma)$ around $\boldsymbol{x}_M$, whose arc length converges more slowly to 0 as $\sigma \to 0$. This large deviation theoretical argument aligns with the occurrence of saddle avoidance on the extended lengthscale $\mathcal{O}(\sigma^\mu)$ (see above and Ref. [36]).

Sample transition paths do not distribute proportionally to $\exp(-V_R/\sigma^2)$ across the flat quasipotential region but are skewed away from the saddle: the boundary-crossing location $\boldsymbol{x}_c(\varepsilon, \sigma)$ of the MTP lies at the end of $\mathcal{Z}$, far from $\boldsymbol{x}_M$. This is likely related to the instanton approaching the basin boundary tangentially. Yet, Kramers' scaling of the mean first-exit time $\langle \tau_\sigma^{RL} \rangle \asymp \exp[\sigma^{-2} V_R(\boldsymbol{x}_M)]$ is empirically observed for all transition path ensembles (see Fig. S2 of the Supplemental Material [57]). We attribute this to the quasipotential barrier height at the mean crossing point $\boldsymbol{x}_c(\varepsilon, \sigma)$ being $V_R(\boldsymbol{x}_c) \approx V_R(\boldsymbol{x}_M) + \sigma^2$, yielding a comparable exponential scaling as long as $\sigma^2 \ll V_R(\boldsymbol{x}_M)$. The subexponential prefactor in Kramers' formula is approximately constant in the examined range of σ values (not shown) and hence does not break the exponential scaling. This supports the suitability of the saddle for predicting transitions rates, even if the corresponding transition paths avoid the saddle.

A finite-noise variational formulation. While the quasipotential can explain the occurrence of saddle avoidance, the question remains how to predict the location distribution of transition paths for finite noise. If the noise is additive, there is increasing confidence [41,43,44,63] that the appropriate variational problem entails minimizing the OM action [40–45],

$$\tilde{S}_T^\sigma[\boldsymbol{\varphi}_t] = S_T[\boldsymbol{\varphi}_t] + \frac{\sigma^2}{2} \int_0^T \nabla \cdot \boldsymbol{b}(\boldsymbol{\varphi}_t)\, \mathrm{d}t, \tag{6}$$

which adds a σ-dependent correction term to the FW action. This term acts as a "penalty" proportional to the divergence of $\boldsymbol{b}$ along the path. As we show in the Supplemental

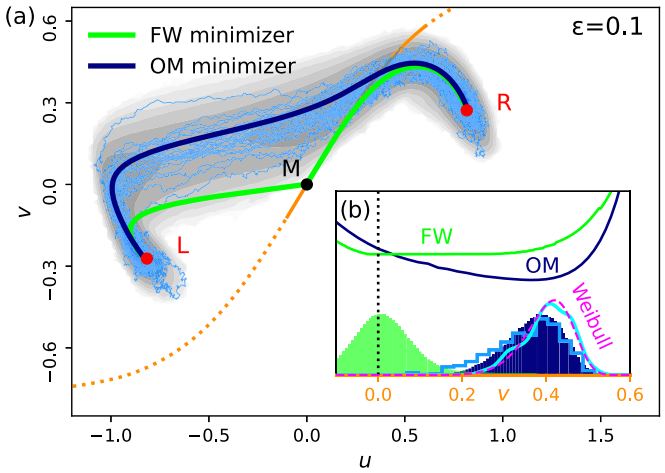

FIG. 2. FHN model. (a) FW (green) and OM (dark blue) minimizer compared with sample paths (light blue) for the transition $\boldsymbol{\phi}_t^{RL}$, showing the path-space sampling density (gray) and the basin boundary (orange). (b) Cross section along the basin boundary (projected onto v). Top lines: quasipotential from minimizing FW/OM; histograms: crossing point distributions for direct sampling (light blue) vs path-space sampling using FW (green) and OM (dark blue) functionals; sampled first-exit point distribution (cyan) compared to rescaled Weibull distribution (magenta).

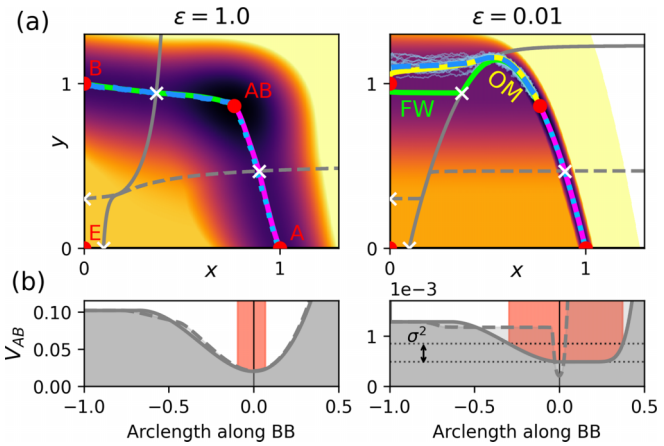

FIG. 3. COMP model. (a) V_{AB} with respect to $\boldsymbol{x}_{AB}$ (brighter indicates larger V_{AB}) for $\varepsilon \in \{1,\ 0.01\}$, FW instantons to $\boldsymbol{x}_A$ (magenta) and $\boldsymbol{x}_B$ (green), and corresponding MTPs for $\sigma(\varepsilon) = (0.0425, 0.006)$ (dashed blue); the OM minimizer (yellow) is shown for the saddle-avoidant scenario. Basin boundaries (BB, gray lines), attractors (red points), and saddles (white crosses) are shown. (b) V_{AB} along the solid and dashed BB displayed in panel (a), as a function of arc length from the respective saddle, highlighting $\mathcal{Z}$ (red shading) for the solid BB.

Material [57], saddle avoidance is directly related to a positive divergence at the saddle.

Minimizing the OM action allows us to derive a candidate for the most probable transition path. Although the minimization over all $T > 0$ is ill-defined in the OM case [45], one can often select a characteristic path travel time T; here we fix $T = \langle t_\sigma^{RL} \rangle$ to match the kinetics of the transition process. Numerically, we use the OM action formulation to perform *path-space sampling* (see the Supplemental Material [57]), which yields a transition path density identical to that of sampling Eq. (1) via Monte Carlo simulation conditioned on the start and end points [64,65].

We focus on the FHN model for $\varepsilon = 0.1$ and $\sigma \approx 0.119$. The minimizer of the OM action over $\mathcal{C}_{RL}^T$ avoids the saddle and yields an appropriate approximation of the MTP [Fig. 2(a)]. Further, path-space sampling using the OM action indeed recovers the observed transition path density [unlike using the FW action, see Fig. 2(b)], and the first-exit points are accurately described by the Weibull distribution $\rho(\sigma, \varepsilon, A)$ with $A \approx 1.23$, corroborating Ref. [36]. The difference between the first-exit point and boundary crossing point distributions in Fig. 2(b) results from transition paths sometimes crossing the basin boundary multiple times before reaching the competing attractor.

Our results show that the OM minimizer gives better predictions of sample transition paths. This motivates constructing a finite-noise, finite-time quasipotential landscape in the spirit of Eq. (4) but with T fixed and $\tilde{S}_T^\sigma$ replacing S_T. This quantity exhibits a minimum on the basin boundary where sample transition paths are observed to cross [Fig. 2(b)].

Two competing species. In multiscale systems with more than two attractors, the FW theory may accurately predict the path of one transition scenario while failing for another. We show this in a system of two competing species A and B perturbed by *multiplicative* noise (COMP) [66]. The growth of each population is modeled with an Allee effect [67] according to

$$\boldsymbol{b}_{\text{comp}}(x, y) = \begin{pmatrix} x(x-\alpha_A)(1-x) - \beta_A xy \\ \varepsilon[y(y-\alpha_B)(1-y) - \beta_B xy] \end{pmatrix}. \tag{7}$$

Here x and y denote the population densities of species A and B, respectively, which each go extinct below their critical densities $\alpha_A = 0.1$ and $\alpha_B = 0.3$. The parameter ε represents the ratio of net growth rates of the two species. The competition term is controlled by $\beta_A = 0.18$ and $\beta_B = 0.1$. This choice of parameters yields four stable equilibrium points: a state $\boldsymbol{x}_{AB}$ where both species coexist, two states $\boldsymbol{x}_A$ and $\boldsymbol{x}_B$ where only one species survives, and a full extinction state $\boldsymbol{x}_E$. Additionally, four saddle points and one repeller exist in the non-negative quadrant [Fig. 3(a)]. As with the FHN model, we investigate the dynamics of Eq. (1), now for $\boldsymbol{b} = \boldsymbol{b}_{\text{comp}}$ and a state-dependent covariance matrix $\boldsymbol{Q} = \text{diag}(x, y)$ mimicking population fluctuations [68,69].

Several transition scenarios are possible. We focus on the extinction of one species, realized by the two scenarios $AB \to A$ and $AB \to B$. If both species have similar net growth rates ($\varepsilon \approx 1$), both scenarios are characterized by a well-defined transition channel and distinct quasipotential minimum on the relevant basin boundary (Fig. 3). The MTPs track the FW instanton and cross near the corresponding saddle. Contrarily, the two scenarios differ when one species grows faster than the other (e.g., $\varepsilon = 0.01$): transitions $AB \to A$ follow the instanton in a narrow quasipotential channel, whereas for $AB \to B$ the MTP detaches from the instanton and avoids the saddle in a region of flat quasipotential V_{AB} induced by the fast-slow dynamics.

For the COMP system, it is possible to obtain the OM minimizer as before by performing a coordinate transform, $x \to 2\sqrt{x}$ and $y \to 2\sqrt{y}$, that effectively leads to an additive

noise problem (see Supplmental Material [57]). Again, we find that the OM minimizer avoids the saddle and agrees with the MTP obtained from direct simulation samples [Fig. 3(a)].

Discussion. Multiscale dynamics can cause noise-induced transitions to bypass the saddle point between competing attracting states. Since physical systems are typically nongradient, noisy, and multiscale, this phenomenon can appear in various applications, even where transitions classify as rare events. Transition path ensembles may deviate strongly from the minimizer of the FW action for weak yet finite noise, while Kramers' law remains valid. These properties are possible due to a flat quasipotential along the basin boundary, which may occur due to timescale separation in the drift term for additive and multiplicative noise alike.

Despite avoiding the instanton, sample transition paths still tend to bundle within a tube around a typical transition path, manifesting the general notion of dynamical typicality [70,71]. For additive noise of finite amplitude, minimizing the OM action (for an appropriate path travel time T) yields an apt prediction of this most probable transition path and the transition path distribution, in contrast to minimizing the FW action. For multiplicative noise, computing finite-noise most probable transition paths is a topic of ongoing research [41,63,72]. Interestingly, estimating the FW minimizer with machine learning using deep gMAM [73] yields a path that more closely resembles the OM minimizer instead of the true FW instanton [74]. Addressing the mathematical challenges of multiplicative and degenerate noise is of great interest for future work.

Reference [36] presented a path-geometric study of saddle avoidance in two dimensions based on the local stability at the saddle. Our work complements this by (a) providing a global stability viewpoint based on the quasipotential, (b) considering the stability of the critical manifold around the saddle, and (c) clarifying the link between saddle avoidance, positive divergence at the saddle, and multiscale dynamics. These concepts may enable anticipating saddle avoidance also in nongradient fast-slow systems of higher dimension: if the saddle exists on an unstable branch of the critical manifold, we conjecture that repulsion away from the saddle outweighing attraction towards it will generally cause the instanton to approach the basin boundary tangentially, resulting in a quasipotential plateau along the saddle's stable set. Other physical interpretations of saddle avoidance, such as anisotropic friction [37], may be interpreted as a multiscale feature. Linking our findings to recent literature on coarse-graining [75,76] could provide further insights into multiscale stochastic systems.

While transitions driven by Lévy noise are known to avoid the saddle [77], our results challenge the generic role of saddles as gateways of noise-induced transitions also under Gaussian noise. Saddle avoidance limits the classical applicability of the FW theory for predicting most probable transition paths in multiscale systems, since the regime of weak but finite noise will often apply to rare events of relevance. Especially for high-impact low-likelihood events, understanding where in state space transitions will likely occur is crucial to assess a system's resilience to random fluctuations.

The data shown in Figs. 1–3 are publicly available at [78].

Acknowledgments. The authors wish to thank P. Ditlevsen for suggesting the FHN model; M. Harsh for suggesting to transform the COMP system; P. Ashwin, M. Cameron, K. Lux, C. Nesbitt, T. Tél, and J. Wouters for valuable discussions; E. Simonnet for checking our results on the FHN model using deep gMAM; and the anonymous reviewers for their helpful comments. R.B., R.D., R.R., V.L. and U.F. gratefully acknowledge funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skodowska-Curie Grant Agreement No. 956170 (CriticalEarth). V.L. acknowledges support received from the EPSRC Project No. EP/T018178/1, the EU Horizon 2020 project TiPES (Grant No. 820970) and the EU Horizon Europe project ClimTIP (Grant No. 101137601). T.G. acknowledges support from EPSRC Projects No. EP/T011866/1 and No. EP/V013319/1. The simulations were performed at the HPC Cluster ROSA, located at the University of Oldenburg (Germany) and funded by the DFG through its Major Research Instrumentation Programme (INST 184/225-1 FUGG) and the Ministry of Science and Culture (MWK) of the Lower Saxony State.

[1] U. Feudel, Complex dynamics in multistable systems, Int. J. Bifurcation Chaos **18**, 1607 (2008).

[2] A. N. Pisarchik and U. Feudel, Control of multistability, Phys. Rep. **540**, 167 (2014).

[3] B. A. W. Brinkman, H. Yan, A. Maffei, I. M. Park, A. Fontanini, J. Wang, and G. La Camera, Metastable dynamics of neural circuits and networks, Appl. Phys. Rev. **9**, 011313 (2022).

[4] P. C. Bressloff, Stochastic switching in biology: From genotype to phenotype, J. Phys. A: Math. Theor. **50**, 133001 (2017).

[5] C. Masoller, Noise-induced resonance in delayed feedback systems, Phys. Rev. Lett. **88**, 034102 (2002).

[6] F. Bouchet, J. Rolland, and E. Simonnet, Rare event algorithm links transitions in turbulent flows with activated nucleations, Phys. Rev. Lett. **122**, 074502 (2019).

[7] N. Boers, M. Ghil, and T. F. Stocker, Theoretical and paleoclimatic evidence for abrupt transitions in the Earth system, Environ. Res. Lett. **17**, 093006 (2022).

[8] D.-D. Rousseau, W. Bagniewski, and V. Lucarini, A punctuated equilibrium analysis of the climate evolution of cenozoic exhibits a hierarchy of abrupt transitions, Sci. Rep. **13**, 11290 (2023).

[9] J.-P. Bouchaud and R. Cont, A Langevin approach to stock market fluctuations and crashes, Eur. Phys. J. B **6**, 543 (1998).

[10] I. Bashkirtseva and L. Ryashko, Sensitivity analysis of stochastic attractors and noise-induced transitions for population model with Allee effect, Chaos **21**, 047514 (2011).

[11] 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. London A **370**, 1166 (2012).
[12] V. Lucarini and T. Bódai, Transitions across melancholia states in a climate model: Reconciling the deterministic and stochastic points of view, Phys. Rev. Lett. **122**, 158701 (2019).
[13] P. D. Ditlevsen and S. J. Johnsen, Tipping points: Early warning and wishful thinking, Geophys. Res. Lett. **37**, 2010GL044486 (2010).
[14] C. S. Holling, Resilience and stability of ecological systems, Annu. Rev. Ecol. Syst. **4**, 1 (1973).
[15] V. Lucarini and T. Bódai, Global stability properties of the climate: Melancholia states, invariant measures, and phase transitions, Nonlinearity **33**, R59 (2020).
[16] K. Hasselmann, Stochastic climate models Part I. Theory, Tellus **28**, 473 (1976).
[17] L. Arnold, Hasselmanns program revisited: The analysis of stochasticity in deterministic climate models, in *Stochastic Climate Models* (Springer, Berlin, 2001), pp. 141–157.
[18] V. Lucarini and M. D. Chekroun, Theoretical tools for understanding the climate crisis from Hasselmanns programme and beyond, Nat. Rev. Phys. **5**, 744 (2023).
[19] M. I. Dykman, E. Mori, J. Ross, and P. M. Hunt, Large fluctuations and optimal paths in chemical kinetics, J. Chem. Phys. **100**, 5735 (1994).
[20] S. L. T. 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).
[21] B. Schäfer, M. Matthiae, X. Zhang, M. Rohden, M. Timme, and D. Witthaut, Escape routes, weak links, and desynchronization in fluctuation-driven networks, Phys. Rev. E **95**, 060203 (2017).
[22] M. I. Dykman and M. A. Krivoglaz, Theory of fluctuational transitions between stable states of a nonlinear oscillator, Sov. Phys. JETP **50**, 30 (1979).
[23] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems* (Springer, Berlin, 1998).
[24] M. K. Cameron, Finding the quasipotential for nongradient SDEs, Phys. D (Amsterdam, Neth.) **241**, 1532 (2012).
[25] J. X. Zhou, M. D. S. Aliyu, E. Aurell, and S. Huang, Quasipotential landscape in complex multi-stable systems, J. R. Soc. Interface **9**, 3539 (2012).
[26] P. Zhou and T. Li, Construction of the landscape for multi-stable systems: Potential landscape, quasi-potential, A-type integral and beyond, J. Chem. Phys. **144**, 094109 (2016).
[27] R. Graham and T. Tél, Nonequilibrium potential for coexisting attractors, Phys. Rev. A **33**, 1322 (1986).
[28] H. Touchette, The large deviation approach to statistical mechanics, Phys. Rep. **478**, 1 (2009).
[29] T. Grafke and E. Vanden-Eijnden, Numerical computation of rare events via large deviation theory, Chaos **29**, 063118 (2019).
[30] G. Margazoglou, T. Grafke, A. Laio, and V. Lucarini, Dynamical landscape and multistability of a climate model, Proc. R. Soc. London A **477**, 20210019 (2021).
[31] V. Lucarini and T. Bódai, Edge states in the climate system: Exploring global instabilities and critical transitions, Nonlinearity **30**, R32 (2017).
[32] F. Bouchet and J. Reygner, Generalisation of the Eyring–Kramers transition rate formula to irreversible diffusion processes, Ann. Henri Poincaré **17**, 3499 (2016).
[33] R. S. Maier and D. L. Stein, Escape problem for irreversible systems, Phys. Rev. E **48**, 931 (1993).
[34] R. S. Maier and D. L. Stein, Transition-rate theory for nongradient drift fields, Phys. Rev. Lett. **69**, 3691 (1992).
[35] D. G. Luchinsky, R. S. Maier, R. Mannella, P. V. E. McClintock, and D. L. Stein, Observation of saddle-point avoidance in noise-induced escape, Phys. Rev. Lett. **82**, 1806 (1999).
[36] R. S. Maier and D. L. Stein, Limiting exit location distributions in the stochastic exit problem, SIAM J. Appl. Math. **57**, 752 (1997).
[37] A. Berezhkovskii, L. Berezhkovskii, and V. Zitzerman, The rate constant in the Kramers multidimensional theory and the saddle-point avoidance, Chem. Phys. **130**, 55 (1989).
[38] S. H. Northrup and J. A. McCammon, Saddle-point avoidance in diffusional reactions, J. Chem. Phys. **78**, 987 (1983).
[39] N. Agmon and R. Kosloff, Dynamics of two-dimensional diffusional barrier crossing, J. Phys. Chem. **91**, 1988 (1987).
[40] R. L. Stratonovich, On the probability functional of diffusion processes, Selected Trans. Math. Stat. Probab. **10**, 273 (1971).
[41] D. Dürr and A. Bach, The Onsager-Machlup function as Lagrangian for the most probable path of a diffusion process, Commun. Math. Phys. **60**, 153 (1978).
[42] W. Horsthemke and A. Bach, Onsager-Machlup function for one dimensional nonlinear diffusion processes, Z. Phys. B: Condens. Matter Quanta **22**, 189 (1975).
[43] F. J. Pinski and A. M. Stuart, Transition paths in molecules at finite temperature, J. Chem. Phys. **132**, 184104 (2010).
[44] J. Gladrow, U. F. Keyser, R. Adhikari, and J. Kappler, Experimental measurement of relative path probabilities and stochastic actions, Phys. Rev. X **11**, 031022 (2021).
[45] T. Li and X. Li, Gamma-limit of the Onsager–Machlup functional on the space of curves, SIAM J. Math. Anal. **53**, 1 (2021).
[46] A. Dembo and O. Zeitouni, *Large Deviations Techniques and Applications*, Stochastic Modelling and Applied Probability, Vol. 38 (Springer, Berlin, 2010).
[47] S. Arrhenius, ber die Reaktionsgeschwindigkeit bei der Inversion von Rohrzucker durch Suren, Z. Phys. Chem. **4U**, 226 (1889).
[48] H. A. Kramers, Brownian motion in a field of force and the diffusion model of chemical reactions, Physica **7**, 284 (1940).
[49] N. Berglund, Kramers' law: Validity, derivations and generalisations, Markov Processes Relat. Fields **19**, 459 (2013).
[50] M. V. Day, Large deviations results for the exit problem with characteristic boundary, J. Math. Anal. Appl. **147**, 134 (1990).
[51] R. FitzHugh, Impulses and physiological states in theoretical models of nerve membrane, Biophys. J. **1**, 445 (1961).
[52] J. Nagumo, S. Arimoto, and S. Yoshizawa, An active pulse transmission line simulating nerve axon, Proc. IRE **50**, 2061 (1962).
[53] C. Kuehn, *Multiple Time Scale Dynamics*, Applied Mathematical Sciences, Vol. 191 (Springer, Berlin, 2015).
[54] C. Kuehn, A mathematical framework for critical transitions: Bifurcations, fast-slow systems and stochastic dynamics, Phys. D (Amsterdam, Neth.) **240**, 1020 (2011).
[55] X. Zhang, C. Kuehn, and S. Hallerberg, Predictability of critical transitions, Phys. Rev. E **92**, 052905 (2015).
[56] C. Kuehn, N. Berglund, C. Bick, M. Engel, T. Hurth, A. Iuorio, and C. Soresina, A general view on double limits in

differential equations, Phys. D (Amsterdam, Neth.) **431**, 133105 (2022).
[57] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevResearch.6.L042053 for methodological details, further explanation, and supplemental figures.
[58] M. Heymann and E. Vanden-Eijnden, The geometric minimum action method: A least action principle on the space of curves, Commun. Pure Appl. Math. **61**, 1052 (2008).
[59] B.-Z. Bobrovsky and O. Zeitouni, Some results on the problem of exit from a domain, Stochastic Processes Their Appl. **41**, 241 (1992).
[60] M. V. Day, Cycling and skewing of exit measures for planar systems, Stochastics Stochastic Rep. **48**, 227 (1994).
[61] S. Coles, *An Introduction to Statistical Modeling of Extreme Values*, Springer Series in Statistics, Vol. 208 (Springer, Berlin, 2001).
[62] D. Dahiya and M. Cameron, An Ordered Line Integral Method for computing the quasi-potential in the case of variable anisotropic diffusion, Phys. D **382-383**, 33 (2018).
[63] A. L. Thorneywork, J. Gladrow, U. F. Keyser, M. E. Cates, R. Adhikari, and J. Kappler, Resolution dependence of most probable pathways with state-dependent diffusivity, arXiv:2402.01559.
[64] A. M. Stuart, J. Voss, and P. Wilberg, Conditional path sampling of SDEs and the Langevin MCMC method, Commun. Math. Sci. **2**, 685 (2004).
[65] M. Hairer, A. M. Stuart, and J. Voss, Analysis of SPDEs arising in path sampling part II: The nonlinear case, Ann. Appl. Probab. **17**, 1657 (2007).
[66] A. D. Bazykin, Dynamics of Isolated Populations, in *World Scientific Series on Nonlinear Science Series A* (World Scientific, Singapore, 1998), Vol. 11, pp. 7–17.
[67] P. A. Stephens, W. J. Sutherland, and R. P. Freckleton, What is the Allee effect? Oikos **87**, 185 (1999).
[68] H. Weissmann, N. M. Shnerb, and D. A. Kessler, Simulation of spatial systems with demographic noise, Phys. Rev. E **98**, 022131 (2018).
[69] Y. Meng, Y.-C. Lai, and C. Grebogi, Tipping point and noise-induced transients in ecological networks, J. R. Soc. Interface **17**, 20200645 (2020).
[70] V. M. Galfi and V. Lucarini, Fingerprinting heatwaves and cold spells and assessing their response to climate change using large deviation theory, Phys. Rev. Lett. **127**, 058701 (2021).
[71] V. Lucarini, V. M. Galfi, J. Riboldi, and G. Messori, Typicality of the 2021 Western North America summer heatwave, Environ. Res. Lett. **18**, 015004 (2023).
[72] J. Kappler, M. E. Cates, and R. Adhikari, Sojourn probabilities in tubes and pathwise irreversibility for Itô processes, arXiv:2009.04250.
[73] E. Simonnet, Computing non-equilibrium trajectories by a deep learning approach, J. Comput. Phys. **491**, 112349 (2023).
[74] E. Simonnet (personal communication).
[75] F. Legoll, T. Leliévre, and U. Sharma, Effective dynamics for non-reversible stochastic differential equations: A quantitative study, Nonlinearity **32**, 4779 (2019).
[76] C. Hartmann, L. Neureither, and U. Sharma, Coarse graining of nonreversible stochastic differential equations: Quantitative results and connections to averaging, SIAM J. Math. Anal. **52**, 2689 (2020).
[77] V. Lucarini, L. Serdukova, and G. Margazoglou, Lévy noise versus Gaussian-noise-induced transitions in the Ghil–Sellers energy balance model, Nonlinear Processes Geophys. **29**, 183 (2022).
[78] https://doi.org/10.6084/m9.figshare.27844776.