


Minimum-Action Method for Nonequilibrium Phase Transitions

Ruben Zakine[✉] and Eric Vanden-Eijnden

Courant Institute, New York University, 251 Mercer Street, New York, New York 10012, USA

 (Received 6 June 2022; revised 7 March 2023; accepted 19 May 2023; published 7 December 2023)

First-order nonequilibrium phase transitions observed in active matter, fluid dynamics, biology, climate science, and other systems with irreversible dynamics are challenging to analyze, because they cannot be inferred from a simple free energy minimization principle. Rather, the mechanism of these transitions depends crucially on the system's dynamics, which requires us to analyze them in trajectory space rather than in phase space. Here, we consider situations where the path of these transitions can be characterized as the minimizer of an action, whose minimum value can be used in a nonequilibrium generalization of the Arrhenius law to calculate the system's phase diagram. We also develop efficient numerical tools for the minimization of this action. These tools are general enough to be transportable to many situations of interest, in particular, when the fluctuations present in the microscopic system are non-Gaussian and its dynamics is not governed by the standard Langevin equation. As an illustration, first-order phase transitions in two spatially extended nonequilibrium systems are analyzed: a modified Ginzburg-Landau equation with a chemical potential which is nongradient and a reaction-diffusion network based on the Schlögl model. The phase diagrams of both systems are calculated as a function of their control parameters, and the paths of the transitions, including their critical nuclei, are identified. These results clearly demonstrate the nonequilibrium nature of the transitions, with differing forward and backward paths.

DOI: [10.1103/PhysRevX.13.041044](https://doi.org/10.1103/PhysRevX.13.041044)

Subject Areas: Computational Physics,
Condensed Matter Physics, Soft Matter,
Statistical Physics

I. INTRODUCTION

Materials with identical microscopic constitutions can be found in very different macroscopic states when external conditions, such as temperature or pressure, vary. A major achievement of equilibrium statistical mechanics is to give a first-principles explanation of these phase transitions. The theory posits the existence of a distribution, for example, of Boltzmann-Gibbs type, that gives the probability of finding the microscopic system in any of its possible configurations. Macroscopic properties like the system's density, magnetization [1], population number in the ground state [2], etc., can then be deduced by enumerating all the microscopic configurations consistent with a given value of the chosen macroscopic observable and identifying which of these values is most likely. This information is typically encapsulated in a free energy, usually referred to as Landau's [3], whose local minima identify the system's metastable phases and whose global minimum is its thermodynamically preferred phase.

The statistical mechanics approach to phase transitions rests on the assumption that the probability distribution of the microscopic system is known. This information is available for equilibrium systems, whose microscopic dynamics is time reversible. In this work, however, we are primarily interested in nonequilibrium systems, whose dynamics is irreversible. Except for some special situations where it can be computed exactly [4,5] or asymptotically, e.g., via some thermodynamic mapping [6–9], the invariant distribution of these systems is not known, in general. Yet, these systems, too, can undergo phase transitions. Examples include driven systems arising from active matter [10–15], fluid dynamics [16,17], biology [18], neuroscience [19–21], climate science [22,23], etc. The description of such nonequilibrium phase transitions requires a generalization of the equilibrium statistical mechanics approach, in which we must consider the probability of trajectories rather than configurations, see Table 1 for a summary. Proposing such a generalization and exploring its consequences is the aim of this paper.

More specifically, our approach is based on a minimum-action principle in trajectory space that allows one to characterize the pathways that some suitably chosen macroscopic variables follow with high probability. This principle captures the effects of the microscopic fluctuations on the macroscopic quantities of interest and applies for systems

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

TABLE I. Correspondence of formalisms.

	Equilibrium approach in phase space	Nonequilibrium approach in trajectory space
Microscopic weights	Probability distribution (e.g., Boltzmann-Gibbs distribution) of the microscopic variables in state space	Probability distribution (e.g., path integral) of the microscopic trajectories
Macroscopic variable	Map from microscopic state space to coarse-grained macroscopic space (e.g., spin values to magnetization)	Map from microscopic trajectories in phase space to macroscopic trajectories
Macroscopic weights	Marginal distribution of the macroscopic variables and associated free energy	Marginal distribution of the macroscopic trajectories and associated action
Macroscopic predictions	Free energy minimization	Action minimization

whose dynamics do not need to satisfy detailed balance. Therefore, it offers a nonequilibrium generalization of Landau's free energy principle. In particular, the minimum-action principle can be leveraged to estimate probability ratios of state occupation as well as first-order phase transitions, identified as the parameter manifold where metastable phases are equally probable. The approach also lends itself naturally to numerical computations, using the tailored-designed algorithms developed here. The usefulness of these schemes is demonstrated below via the thorough treatment on two non-trivial examples of phase transitions with generic features.

A. Free energy principle at equilibrium

Let us begin with a short summary of Landau's free energy approach to equilibrium phase transitions, as this will be useful to contextualize the nonequilibrium approach introduced below. If we denote by ϕ the variable used to characterize the macroscopic state of the system, its statistical weight at equilibrium is obtained by summing the system's probability distribution over all microscopic states consistent with a given realization of ϕ . Performing this calculation typically require sophisticated tools such as renormalization group theory [24], the replica method, or the cavity method [25,26], along with tools from large deviation theory [27]. It generically shows that the statistical weight of ϕ is asymptotically given by $\exp[-V(\phi)/\epsilon]$, where $V(\phi)$ is some free energy to be calculated and ϵ is a small parameter that tends to zero in the thermodynamic limit when the number of microscopic constituents tends to infinity: As a result, the theory predicts that the system will be found in the macroscopic state ϕ of minimum free energy with probability one in this limit. This also explains equilibrium phase transitions: They take place when the topology of the free energy $V(\phi)$ changes as a control parameter, like the temperature or some applied external field, is varied. For example, if $V(\phi)$ has two wells whose relative depths change with the control parameter, a first-order phase transition occurs when the deepest well becomes more shallow than the other well, and, as a result, the macroscopic state of the system changes from the first to the second.

B. Minimum-action principle out of equilibrium

Even in situations where the stationary distribution of nonequilibrium systems is not known in general, we can often write down the probability distribution of their trajectories, using, e.g., path-integral approaches such as the Martin-Siggia-Rose-Janssen-De Dominicis [28] or the Doi-Peliti formalism [29] or Girsanov theorem. This offers the possibility to generalize the micro-to-macro mapping to trajectory space rather than phase space: That is, enumerate all the microscopic trajectories leading to the same evolution of a macroscopic variable and thereby deduce the probability weight of these macroscopic trajectories. While these calculations are again to be performed on a case-by-case basis, by analogy with the equilibrium setup we can deduce some of the generic features of the result. Let us discuss those next.

Assuming again that the macroscopic state of the system can be described by a variable or field ϕ in some differentiable manifold \mathcal{M} [for example \mathbb{R}^d , \mathbb{S}^d , or $L_2(\mathbb{R}^d)$], working in trajectory space amounts to calculating the probability weight of a macroscopic path $\{\phi(t)\}_{t \in [0, T]}$ by enumerating the microscopic trajectories consistent with $\{\phi(t)\}_{t \in [0, T]}$ and summing over their probability distribution. Generically, we expect the result of this sum to indicate that the weight of the macroscopic path $\{\phi(t)\}_{t \in [0, T]}$ is asymptotically given by the factor $\exp(-S_T[\phi]/\epsilon)$, where ϵ is again a small parameter that goes to zero as the number of microscopic constituents in the system goes to infinity and $S_T[\phi]$ is an action, that is, a functional of $\{\phi(t)\}_{t \in [0, T]}$. The specific form of this action depends on the problem under consideration (examples are given below), but it typically takes the form of an integral over a Lagrangian

$$S_T[\phi] = \int_0^T L(\phi, \dot{\phi}) dt, \quad (1)$$

where $\dot{\phi} = d\phi/dt \in T_\phi \mathcal{M}$. The action $S_T[\phi]$ is the non-equilibrium generalization of the free energy $V(\phi)$, and minimization of $S_T[\phi]$ allows us to quantify the probability

and mechanism of various macroscopic events in the limit as $\epsilon \rightarrow 0$. In particular,

- (i) The probability that the system started in state ϕ_a at time $t = 0$ ends up in state ϕ_b at time T is obtained by summing $\exp(-S_T[\phi]/\epsilon)$ over all paths with these end points. When $\epsilon \ll 1$, the path with minimum action dominates this sum, which means that the aforementioned probability is asymptotically given by

$$\mathbb{P}[\phi(T) = \phi_b | \phi(0) = \phi_a] \asymp \exp(-\inf S_T[\phi]/\epsilon), \quad (2)$$

where the minimization is taken over all paths $\{\phi(t)\}_{t \in [0, T]}$ such that $\phi(0) = \phi_a$ and $\phi(T) = \phi_b$ and \asymp means exponential asymptotics; i.e., the ratio of the logarithm of both sides in Eq. (2) tends to 1 as $\epsilon \rightarrow 0$. The minimizer of the action also gives the pathway by which the macroscopic transition event occurs with probability one in this limit.

- (ii) The nonequilibrium invariant distribution of the system can be characterized similarly via the quasipotential defined as

$$V_{\phi_a}(\phi_b) = \inf_{T > 0} \inf S_T[\phi] \quad (3)$$

where the inner minimization is again taken over all paths $\{\phi(t)\}_{t \in [0, T]}$ such that $\phi(0) = \phi_a$ and $\phi(T) = \phi_b$. The quasipotential $V_{\phi_a}(\phi_b)$ plays a role analogous to the free energy barrier from state ϕ_a to ϕ_b , and it can be used to identify the possible phases and formulate an equivalent of the Arrhenius law. More precisely, ϕ_a is a metastable phase if $V_{\phi}(\phi_a) \leq V_{\phi_a}(\phi)$ for all ϕ in a vicinity of ϕ_a (i.e., ϕ_a is the nonequilibrium equivalent of a local minimum on the free energy); and, if ϕ_a and ϕ_b are the only two metastable phases in the system, the asymptotic rates of transition from ϕ_a to ϕ_b and ϕ_b to ϕ_a are, respectively, given by

$$k_{a,b} \asymp e^{-V_{\phi_a}(\phi_b)/\epsilon}, \quad k_{b,a} \asymp e^{-V_{\phi_b}(\phi_a)/\epsilon}. \quad (4)$$

- (iii) Equation (4) is a nonequilibrium generalization of the Arrhenius law. It implies that the relative probability to find the system in states ϕ_a or ϕ_b on its nonequilibrium invariant distribution is asymptotically given by

$$\mathbb{P}(\phi_b)/\mathbb{P}(\phi_a) \asymp e^{-[V_{\phi_a}(\phi_b) - V_{\phi_b}(\phi_a)]/\epsilon}. \quad (5)$$

As a result, with probability 1 as $\epsilon \rightarrow 0$, the system is in state ϕ_a if $V_{\phi_a}(\phi_b) > V_{\phi_b}(\phi_a)$ and state ϕ_b if $V_{\phi_a}(\phi_b) < V_{\phi_b}(\phi_a)$. By analyzing how the quasipotential varies in terms of the system's control

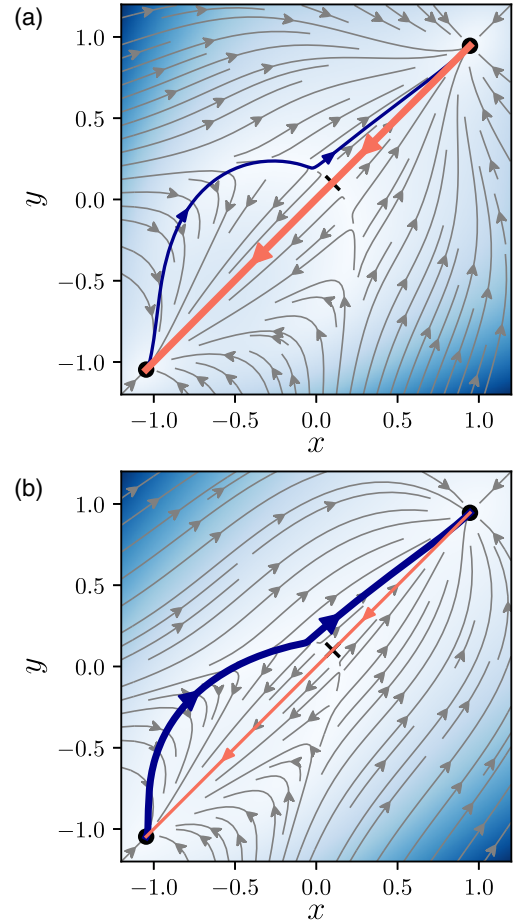


FIG. 1. Example of nonequilibrium phase transition in a toy two-dimensional system with two metastable states (black disks) and action $S_T[x, y] = \frac{1}{2} \int_0^T [|\dot{x} - f(x, y)|^2 + |\dot{y} - f(y, x)|^2] dt$, where $f(x, y) = D(y - x) + x - x^3 + h + \nu(x - y)^2$. The parameters D and h are fixed to 0.5 and -0.1 , respectively, while ν is used as a control parameter. The flow of the vector field f for $\nu = 0.5$ is shown in the (a) and $\nu = 1.5$ in the (b). This flow is nongradient (i.e., the deterministic dynamics is not steepest descent over an energy) and has two stable fixed points ϕ_a and ϕ_b , which solve $f(x, y) = f(y, x) = 0$ and which are the possible phases in this toy example. The blue line represents the most probable transition path (i.e., the minimizer of the action) from ϕ_a to ϕ_b , and the orange line is the most probable path from ϕ_b to ϕ_a . The thickest line indicates the path with the larger rate: That is, ϕ_a is the stable phase when $\nu = 0.5$ (a), while ϕ_b is the stable phase when $\nu = 1.5$ (b). The transition paths are calculated with the method developed in this paper and cross-checked using GMAM [30].

parameters, we can thereby identify nonequilibrium phase transitions that arise when $V_{\phi_b}(\phi_a) = V_{\phi_a}(\phi_b)$ and characterize their mechanism—the details of these calculations are given below. We also refer the reader to Fig. 1 for a graphical illustration in a toy system of nonequilibrium phase transition whose detection requires the formalism above. The figure shows the reaction paths for a particle subjected to a

Langevin additive noise in some nongradient force field with two stable fixed points. The reaction paths from ϕ_a to ϕ_b and from ϕ_b to ϕ_a do not coincide, confirming the nonequilibrium nature of the system. The relative stability of the fixed points is given by the numerical computation of the quasipotential.

C. Hamiltonian formalism

The minimum-action principle described in the last section offers a way to study transition events and phase transitions in nonequilibrium systems. Concrete predictions, however, rest on our ability to (i) derive the Lagrangian used in the action (1) and (ii) minimize this action as needed in Eqs. (2) and (3).

Like in the equilibrium case, resolving the first issue is again complicated, in general, and requires to be handled on a case-by-case basis. When these calculations can be done (see Sec. II for a list of examples), one often deduces that $L(\phi, \dot{\phi})$ is given as the Legendre-Fenchel transform of a Hamiltonian $H(\phi, \theta)$:

$$L(\phi, \dot{\phi}) = \sup_{\theta \in T_{\phi}\mathcal{M}} [\langle \dot{\phi}, \theta \rangle - H(\phi, \theta)], \quad (6)$$

where $\langle \cdot, \cdot \rangle$ denotes the scalar product in $T_{\phi}\mathcal{M}$ and θ is a field conjugate to ϕ whose physical meaning is explained below. The form of the Hamiltonian $H(\phi, \theta)$ is also problem dependent, but it is known in some instances; see Sec. II A.

Using Eq. (6), the minimization of the action can then be formulated as a min-max problem:

$$\inf_{\phi} S_T(\phi) = \inf_{\phi} \sup_{\theta} \int_0^T [\langle \dot{\phi}, \theta \rangle - H(\phi, \theta)] dt, \quad (7)$$

where the supremum is taken over all $\{\theta(t)\}_{t \in [0, T]}$ and the infimum over all $\{\phi(t)\}_{t \in [0, T]}$ such that $\phi(0) = \phi_a$ and $\phi(T) = \phi_b$. To get the quasipotential, we must also consider an extra minimization over all $T > 0$ to Eq. (7), while other applications may require adding terms to Eq. (7) or modifying the boundary conditions for this min-max problem. In most cases, these calculations must be performed numerically. One of the main goals of this paper is to develop robust numerical methods for these computations. These methods aim to be general enough to be applicable to a wide variety of systems that fit the framework above; here, we also use them to solve some nontrivial examples involving spatially extended systems undergoing nonequilibrium phase transitions.

D. Related works

The Euler-Lagrange equations associated with the min-max problem (7) are Hamilton's classical equations:

$$\dot{\phi} = \partial_{\theta} H, \quad \dot{\theta} = -\partial_{\phi} H. \quad (8)$$

What makes the problem nonstandard, however, are the boundary conditions imposed on $\phi(t)$ at $t = 0$ and $t = T$. The nature of these boundary conditions suggests to use shooting methods [31], as proposed, e.g., in Ref. [32], but such methods scale badly with dimension or can even be ill posed for the problems we are interested in, for which the equation for θ in Eq. (8) cannot be integrated forward in time. Shooting methods are also hard to use when $T = \infty$, which typically arises when we consider $\inf_{T > 0} \inf S_T[\phi]$.

To get around this difficulty, the minimum-action method (MAM) proposed in Ref. [33] evolves the whole trajectory $\{\phi(t)\}_{t \in [0, T]}$ while keeping $\phi(0) = \phi_a$ and $\phi(T) = \phi_b$ fixed. This amounts to performing gradient descent (GD) on the action in the landscape of all authorized paths satisfying these boundary conditions. Introducing the artificial optimization time τ , GD results in the following evolution equation for $\{\phi(\tau, t)\}_{\tau \geq 0, t \in [0, T]}$:

$$\partial_{\tau} \phi = -\frac{\delta S_T[\phi]}{\delta \phi(t)}, \quad \phi(\tau, 0) = \phi_a, \quad \phi(\tau, T) = \phi_b, \quad (9)$$

or using the Lagrangian formulation of the action

$$\partial_{\tau} \phi = -\left[\frac{\partial L}{\partial \phi} - \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\phi}} \right) \right], \quad (10)$$

with the same boundary conditions at $t = 0, T$. The main drawback of MAM is that it involves the Lagrangian $L(\phi, \dot{\phi})$ rather than the Hamiltonian $H(\phi, \theta)$, and there are many problems of interest where the latter is explicitly available but the former is not. Solving Eq. (10) then requires one to perform $\max_{\theta} [\langle \dot{\phi}, \theta \rangle - H(\phi, \theta)]$ for all $t \in [0, T]$ at each iteration step in τ to numerically get an estimate of the function $\vartheta(\phi, \dot{\phi})$ such that

$$L(\phi, \dot{\phi}) = \langle \dot{\phi}, \vartheta(\phi, \dot{\phi}) \rangle - H[\phi, \vartheta(\phi, \dot{\phi})]. \quad (11)$$

Proceeding similarly, we can also obtain numerical estimates for the derivative $\partial L / \partial \phi$ and $(d/dt) \partial L / \partial \dot{\phi}$ appearing on the right-hand side of Eq. (10). This approach is used in Ref. [34]. The downside is that Eq. (10) is a partial differential equation in physical time t , optimization time τ , and possibly space as well when ϕ and θ are fields; writing efficient numerical solvers for such equations typically requires one to use implicit schemes for numerical stability and/or efficiency, and such schemes are hard to design without explicit knowledge of $\vartheta(\phi, \dot{\phi})$. This is why here we want to bypass the computation of $\vartheta(\phi, \dot{\phi})$ and solve the min-max problem in Eq. (7) concurrently by treating the minimization over ϕ and the maximization over θ on equal footings.

If we now turn our attention to the quasipotential (3), an extra minimization over all $T > 0$ must be added to the min-max problem (7). The MAM can be generalized to handle this problem by using a reparametrization of the path $\{\phi(t)\}_{t \in [0, T]}$ by arclength rather than physical time.

This formulation leads to the geometric minimum-action method (GMAM), in which the minimization over T is performed explicitly beforehand [30]. GMAM is further developed and used in Ref. [34,35], and a variant of it is also recently proposed in Ref. [36] to compute the quasipotential by a spectral decomposition of paths and an optimization of basis coefficients. However, these methods are Lagrangian based, with the issues discussed before. We show below how to use the ideas behind GMAM in a Hamiltonian approach and, thereby, get an efficient method to solve $\inf_{T>0} \inf_{\phi} S_T[\phi]$.

On the analytical side, the quasipotential is related to the solution $V(\phi)$ of the Hamilton-Jacobi equation [37,38]

$$0 = H(\phi, \partial V / \partial \phi). \quad (12)$$

This equation can be used to deduce some structural properties of the quasipotential. It is a central object in problems that can be tackled through macroscopic fluctuation theory (MFT), and the review in Ref. [39] provides numerous examples and useful insights. Reference [40] also discusses how to perturbatively solve this equation in systems that are close to equilibrium. In general, however, Eq. (12) needs to be solved numerically, which is non-trivial, since it is a complicated partial differential equation (or even a functional equation when ϕ is a field). In dimension 2 or 3, this can be done globally using fast marching methods like the one discussed in Ref. [41]. In dimension higher than 3, these methods become inapplicable, and Eq. (12) must be solved locally by the method of characteristics using the variational formulation of this equation: This brings us back to solving $\inf_{T>0} \inf_{\phi} S_T[\phi]$.

E. Organization and main contributions

The remainder of this paper is organized into theory, numerics, and applications.

1. Theoretical results

In Sec. II, we give more details about the minimum-action principle that is at the core of our approach. To this end, in Sec. II A, we first present typical classes of stochastic dynamical systems that display metastability and nonequilibrium phase transitions and are amenable to analysis via action minimization. For illustration, we also use our method to compute the transition paths between metastable states in two low-dimensional benchmark models, namely, the Maier-Stein model [32] and the Schlögl model [42]. In Sec. II B, we discuss the features of the minimum-action framework in a general setup and summarize the main outputs of the approach.

2. Numerical contributions

In Sec. III, we present a numerical method for solving the min-max problem (7). The scheme is based on performing gradient descent-ascent (GDA) on the objective, which has

the advantage that it can be formulated directly in the Hamiltonian setup. The main issue we need to address is that, in our context, the GDA equations are partial differential equations of hyperbolic type in some optimization time τ and physical time t , to be solved with boundary conditions at $t = 0$ and $t = T$. In Sec. III A, we show that a simple linear change of variables in these equations allows us to reformulate them in a way that is convenient for numerical solution via Strang splitting. In Sec. III B, we then generalize this scheme to compute the quasipotential when an extra minimization over T is added to Eq. (7). This is done by using a geometric formulation in which the paths $\{\phi, \theta\}_{t \in [0, T]}$ are reparametrized using normalized arclength. This allows us to handle the minimization of T efficiently and calculate paths whose duration in physical time is infinite but whose length remains finite.

3. Applications

In Secs. IV and V, we then use the methods we propose to analyze two spatially extended nonequilibrium systems that display first-order phase transitions. The method is relevant in regimes far from critical points but close to the first-order transition line. In Sec. IV, we study a modified Ginzburg-Landau (GL) dynamics subject to an additive Gaussian white noise. The noiseless evolution of the field is nongradient with two stable fixed points; the noise makes these points metastable, and we must resort to minimum-action algorithms to compute the nonequilibrium transition pathways between them. The action along these paths allows us to estimate the relative probability of the metastable states. We use this procedure to compute the phase diagram of the system in function of two control parameters. In Sec. V, we study a spatially extended version of the Schlögl model, in which the fluctuations are driven by both diffusion of the microscopic molecules and reactions between them. This system displays a first-order nonequilibrium phase transition in terms of the diffusivity of the molecules, which we characterize. We also show that the predictions of the minimum-action approach explain the transition events observed in the microscopic system.

Some other possible applications of our approach are discussed in Sec. VI, and concluding remarks are given in Sec. VII. Technical developments are deferred to several Appendixes.

II. PROBLEM SETUP AND INTERPRETATION

The aim of this section is to provide a better motivation of the minimum-action principle introduced in the introduction and pinpoint some of its key predictive features. For the reader's convenience, we begin by listing a collection of motivating problems where the formalism applies. In Sec. II B, we then put the approach in a broader context and explain how to use it to analyze metastability.

A. A collection of motivating problems

The first two examples (a and b) involve no coarse-graining from micro-to-macro and are included because they are simple and transparent; the last three examples (c–e) require one to define proper macroscopic variables to derive the minimum-action principle and its Hamiltonian.

- (a) *Diffusion in detailed balance.*—Consider the motion of a particle $x(t) \in \mathbb{R}^d$ whose evolution is governed by the overdamped Langevin equation

$$\dot{x} = -\nabla U(x) + \sqrt{2kT}\eta(t), \quad (13)$$

where $U(x)$ is some potential, kT is the product of Boltzmann constant k and the temperature T , and $\eta(t)$ is a white noise. The dynamics (13) is in detailed balance with respect to the Boltzmann-Gibbs probability density function $\rho(x) = Z^{-1}e^{-U(x)/kT}$, where $Z = \int_{\mathbb{R}^d} e^{-U(x)/kT} dx$. If the potential $U(x)$ has multiple local minima and the temperature kT is much smaller than the barriers between them, Eq. (13) displays metastability: The system stays confined for a long time in the well around a minimum of $U(x)$ before finally hopping to another well where the process repeats. In this example, we can use a WKB expansion to analyze the Fokker-Planck equation associated with Eq. (13). The eikonal equation obtained at leading order in kT is a Hamilton-Jacobi equation whose Hamiltonian is given by

$$H(x, \theta) = -\langle \theta, \nabla U(x) \rangle + |\theta|^2. \quad (14)$$

This Hamiltonian is the one to be used in the min-max problem (7), as can also be proven rigorously using Freidlin-Wentzell large deviation theory (LDT) [37]. In this example, the quasipotential $V_{x_a}(x_b)$ can be calculated explicitly. If x_a and x_b are the locations of two local minima of $U(x)$ with adjacent wells, the path minimizing the action (3) is the minimum energy path between these two points, and $V_{x_a}(x_b)$ is given by

$$V_{x_a}(x_b) = U(x_s) - U(x_a), \quad (15)$$

where $U(x_s)$ is the energy of the saddle point of minimum height (also known as mountain pass) between x_a and x_b . Thus, we recover the Arrhenius law for the rate of transition from x_a to x_b :

$$k_{a,b} \asymp e^{-V_{x_a}(x_b)/kT} = e^{-[U(x_s) - U(x_a)]/kT}. \quad (16)$$

- (b) *Diffusion out of equilibrium.*—The picture above can be generalized to systems whose evolution is described by the stochastic differential equation (SDE)

$$\dot{x} = b(x) + \sqrt{\epsilon}\sigma(x)\eta(t), \quad (17)$$

even if this equation is not in detailed balance; i.e., it is not possible to write the drift $b(x)$ as $-D(x)\nabla U(x) + kT\nabla \cdot D(x)$ for $D(x) = (\sigma\sigma^T)(x)$ and some potential $U(x)$. Metastability is observed with Eq. (17) in situations where the noiseless deterministic system $\dot{x} = b(x)$ has multiple stable fixed points and the noise amplitude is small but finite: The system hovers for long times in the basin around one of these fixed points, but a noise-driven transition to another basin eventually occurs. These transitions can be described by the minimum-action principle using the Hamiltonian which can again be obtained via WKB analysis of the Fokker-Planck equation associated with Eq. (17). It is given by

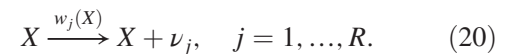
$$H(x, \theta) = \langle b(x), \theta \rangle + \frac{1}{2}|\sigma(x)\theta|^2. \quad (18)$$

Given two stable fixed points x_a and x_b of $\dot{x} = b(x)$ with adjacent basins of attraction, it is no longer possible, in general, to solve Eq. (3) analytically and calculate $V_{x_a}(x_b)$ —to do so requires numerical tools of the type developed below. Still, we know that the rates of transition from x_a to x_b and x_b to x_a satisfy, respectively,

$$k_{a,b} \asymp e^{-V_{x_a}(x_b)/\epsilon}, \quad k_{b,a} \asymp e^{-V_{x_b}(x_a)/\epsilon}, \quad (19)$$

and, with probability 1 as $\epsilon \rightarrow 0$, the system performs the transition by following the optimal path minimizing Eq. (3)—for an illustration in the context of Maier-Stein model [32], see Fig. 2. The results in SDE with small noise of this type can be made rigorous using the Freidlin-Wentzell theory of large deviations (LDT) [37].

- (c) *Reaction networks.*—Consider a well-stirred chemical network between M chemical species, where the quantity of species i is denoted X_i . Define the population vector $X = (X_1, \dots, X_M)^T$, and assume that there are R reaction channels with rates $w_j(x)$ and change (stoichiometric) vectors $\nu_j \in \mathbb{Z}^M$, i.e.,



When the typical number of agents, Ω , tends to infinity, the dynamics of $x = X/\Omega$ is captured by the mass-action law

$$\dot{x} = \sum_{j=1}^R w_j(x)\nu_j, \quad (21)$$

and the minimum-action principle is useful to quantify the effects of fluctuations when the number of agents is large but finite. In particular, metastability arises if Eq. (21) has multiple stable fixed points

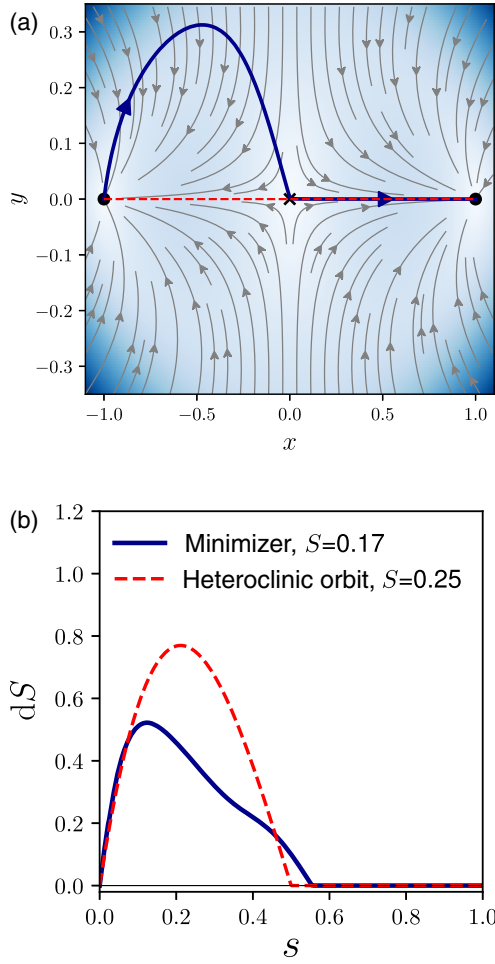


FIG. 2. Optimal reaction path in the Maier-Stein model [32]: The evolution of $(x, y) \in \mathbb{R}^2$ is governed by the SDE $\dot{x} = x - x^3 - \beta xy^2 + \sqrt{2\epsilon}\eta_1$, $\dot{y} = -(1+x^2)y + \sqrt{2\epsilon}\eta_2$, where $\eta_{1,2}$ are independent Gaussian white noises and β some parameter (here, $\beta = 10$). The flow lines of the noiseless model ($\epsilon = 0$) are shown as gray lines in (a), and the background color indicates the magnitude of the drift (darker means larger): There are two stable fixed points $x_a = (-1, 0)$ and $x_b = (1, 0)$ (black disks) and one unstable fixed point $x_c = (0, 0)$ (black cross). When ϵ is small but finite, these two fixed points become metastable, and the noise induces transitions between them: The most likely path from x_a to x_b is the minimum-action path, which is shown as the blue line. Also shown as a red dashed line is the heteroclinic orbit between x_a and x_b : This orbit is different from the minimum-action path, indicative of a nonequilibrium transition. (b) shows the increment of the action along the minimum-action path and the heteroclinic orbit, confirming that the former is more likely.

(see Refs. [43,44]), and it can be analyzed using the Hamiltonian

$$H(x, \theta) = \sum_{j=1}^R w_j(x) (e^{\nu_j \theta} - 1). \quad (22)$$

Here, too, this Hamiltonian can be obtained rigorously via Freidlin-Wentzell LDT, formally via WKB

analysis of the system's master equation [43,45] (see also the appendix in Ref. [13] for a pedagogical derivation), or via a Doi-Peliti field theory computation [29,46–48]. If x_a and x_b denote two stable fixed points of Eq. (21) with adjacent basins of attraction, the rates of transition from x_a to x_b and x_b to x_a are, respectively, given by

$$k_{a,b} \asymp e^{-\Omega V_{x_a}(x_b)}, \quad k_{b,a} \asymp e^{-\Omega V_{x_b}(x_a)}, \quad (23)$$

and with probability 1 as $\Omega \rightarrow \infty$, when the network performs the transition, X/Ω follows the optimal path minimizing Eq. (3). This minimization needs again to be performed numerically, in general. As an illustrating example, Fig. 3 displays reaction paths in the bistable Schlögl model with two reactive compartments: This model belongs to the class of reaction-diffusion networks that are properly introduced in Sec. V. Note that if we reduce the number of compartments to one, the quasipotential of this model can be explicitly obtained from the Hamiltonian; see Fig. 4.

(d) *Interacting particle systems.*—Consider N particles $x_i \in \Lambda \subset \mathbb{R}^d$, $i = 1, \dots, N$, that evolve according to

$$\dot{x}_i = b(x_i) + \frac{1}{N} \sum_{j=1}^N k(x_i, x_j) + \sigma(x_i) \eta_i(t), \quad (24)$$

where $b(x)$ is a drift as in Eq. (17), $k(x, y)$ is some interaction kernel, and $\eta_i(t)$ are independent white noises. To analyze such interacting particle systems, it is convenient to introduce the empirical density of the particles: $\rho_N(t, x) = N^{-1} \sum_{i=1}^N \delta[x - x_i(t)]$. As $N \rightarrow \infty$, this empirical density converges toward the density $\rho(t, x)$ that satisfies McKean-Vlasov equation

$$\partial_t \rho = -\nabla \cdot [B(x, [\rho])\rho] + \frac{1}{2} \nabla \nabla : [D(x)\rho], \quad (25)$$

where $B(x, [\rho]) = b(x) + \int_{\mathbb{R}^d} k(x, y)\rho(y)dy$ and $D(x) = (\sigma\sigma^T)(x)$. Large fluctuations away from the mean-field dynamics (25) can be captured by the minimum-action principle, by using the Hamiltonian

$$\begin{aligned} H(\rho, \theta) = & \int_{\mathbb{R}^d} \nabla \theta(x) \cdot b(x) \rho(x) dx \\ & + \int_{\mathbb{R}^d \times \mathbb{R}^d} \nabla \theta(x) \cdot k(x, y) \rho(x) \rho(y) dx dy \\ & + \frac{1}{2} \int_{\mathbb{R}^d} |\sigma(x) \nabla \theta(x)|^2 \rho(x) dx. \end{aligned} \quad (26)$$

This Hamiltonian can again be derived rigorously using LDT [49], and it formally follows from a WKB analysis, now performed on the functional master equation for the empirical particle density.

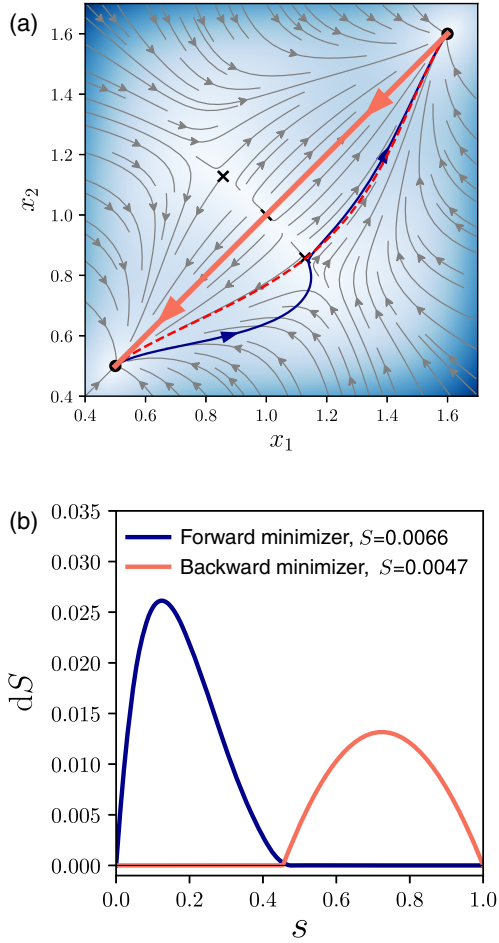


FIG. 3. Optimal reactions paths in the bistable Schlögl model presented in Sec. V, with two reactive compartments and where the particles are allowed to jump from one compartment to another at rate γ . This system belongs to the class of reaction networks. Here, the Hamiltonian is given by $H = H^R + H^D$ with $H^R = \sum_{i=1}^2 w^+(x_i)(e^{\theta_i} - 1) + w^-(x_i)(e^{-\theta_i} - 1)$ and $H^D = \sum_{i=1}^2 \gamma x_i (e^{\theta_{i-1} - \theta_i} + e^{\theta_{i+1} - \theta_i} - 2)$. The flow lines of the mass-action law are shown in (a), along with its stable fixed points (black dots) and unstable fixed points (black crosses). When the number of agents is large but finite, these stable fixed points become metastable states, and the most likely paths between them are shown as full lines in blue and orange. Also shown as a dashed red line is the heteroclinic orbit. All three paths differ, indicative of a nonequilibrium transition. (b) shows the increment along the action of the two minimum-action paths, indicating that the bottom left state is most stable under random fluctuations.

(e) *Fast-slow systems.*—Consider a system made of a pair of variables $(x, y) \in \mathbb{R}^{d \times D}$ whose evolution is governed by

$$\begin{aligned} \dot{x} &= f(x, y), \\ \dot{y} &= \alpha^{-1} b(x, y) + \alpha^{-1/2} \sigma(x, y) \eta(t), \end{aligned} \quad (27)$$

where $\alpha > 0$ measure the separation of timescale between x and y . For small α , this separation is large, and y

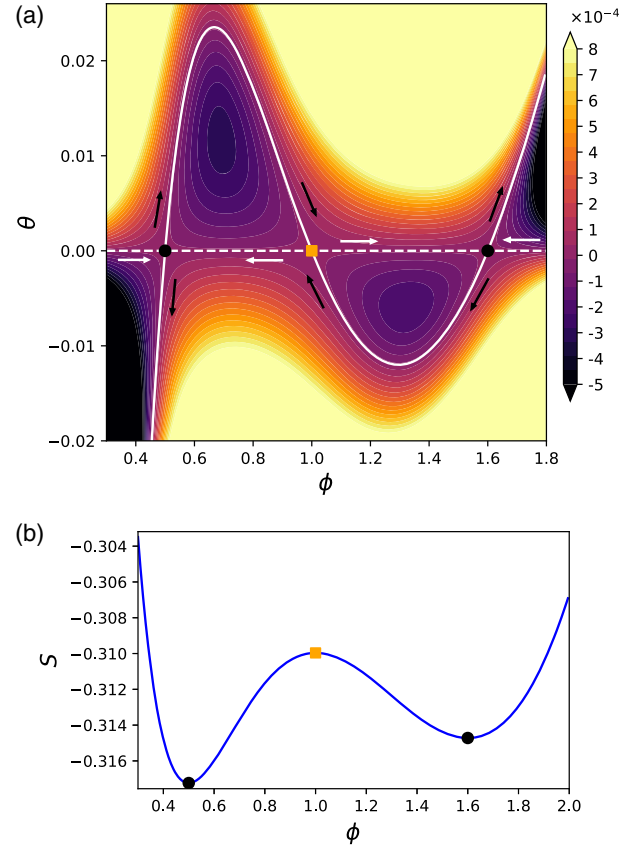


FIG. 4. (a) Contour plot of the Hamiltonian $H(\phi, \theta)$ of the bistable Schlögl model introduced in Sec. V [Eq. (71)] with a single reactive compartment containing a large number of particles Ω . The two black disks show the stable fixed points ϕ_a (left disk) and ϕ_b (right disk) of the law of mass action valid when $\Omega \rightarrow \infty$, and the orange square shows its unstable fixed point at $\phi_s = 1$. The solid and dashed white lines show the level set $H(\phi, \theta) = 0$ on which the solutions to Hamilton's equations (8) evolve in the limit when $T \rightarrow \infty$. The horizontal dashed line at $\theta = 0$ corresponds to the noiseless dynamics of the mass-action law by which the system relaxes to one of the two stable fixed points ϕ_a or ϕ_b by moving in the direction of the white arrows. The solid white line, to follow in the direction of the black arrows, shows the solutions of Hamilton's equations on $H(\phi, \theta) = 0$ at $\theta \neq 0$: These trajectories indicate the most probable path by which the fluctuations can drive the system away from ϕ_a and ϕ_b and induce transitions when reaching ϕ_s . In this example, the action required to drive the system away from ϕ_a to some $\phi \leq \phi_s$ in its basin of attraction is the area between the white and the dashed solid lines, going from ϕ_a to ϕ : This area is also the quasipotential $V_{\phi_a}(\phi)$ for $\phi \leq \phi_s$, and $V_{\phi_b}(\phi)$ for $\phi \geq \phi_s$ can be defined similarly. (b) Global quasipotential $V(\phi)$ obtained by gluing the two quasipotentials $V_{\phi_a}(\phi)$ and $V_{\phi_b}(\phi)$ at $\phi = \phi_s$ after vertical shifting; $V(\phi)$ is also the viscosity solution of the Hamilton-Jacobi equation (12).

evolves much faster than x . In particular, when $\alpha \rightarrow 0$, the dynamics of x is effectively captured by the deterministic limiting equation

$$\dot{x} = F(x), \quad (28)$$

where $F(x)$ is obtained by averaging $f(x, y)$ over the stationary distribution of the equation for y at x fixed, assuming that it exists. This equation defines the so-called virtual fast process, which on the fast timescale $\tau = \alpha t$ reads

$$dy_x/d\tau = b(x, y_x) + \sigma(x, y_x)\eta(\tau) \quad (x \text{ frozen}). \quad (29)$$

Denoting by \mathbb{E}^x the expectation over the stationary distribution of $y_x(\tau)$, the function F entering Eq. (28) is given by

$$F(x) = \mathbb{E}^x f(x, y_x). \quad (30)$$

If we want to analyze the effect of fluctuations on the dynamics of x when α is small but finite, we can use the minimum-action principle with a Hamiltonian that can be derived from LDT [50–53]:

$$H(x, \theta) = \log \mathbb{E}^x \exp [\langle \theta, f(x, y_x) \rangle]. \quad (31)$$

In general, this Hamiltonian needs to be calculated numerically, which makes slow-fast systems of the type above more difficult to treat than the models previously discussed in this section. Still, provided that we can design some numerical routine to estimate H as well as its derivatives $\partial_x H$ and $\partial_\theta H$, the numerical methods presented below are applicable in the context of slow-fast systems, too. In models of this type, $\epsilon = \alpha$.

B. General setup

The problems listed in Sec. II A all share a Hamiltonian with the following features:

- (A1) $H(\phi, 0) = 0$ for all $\phi \in \mathcal{M}$;
- (A2) $H(\phi, \theta)$ is strictly convex in θ for all $\phi \in \mathcal{M}$;
- (A3) $H(\phi, \theta)$ is twice differentiable in both its arguments.

Assumptions A1 and A2 follow from the fact that, generically, the Hamiltonian can be expressed as a cumulant generating function, i.e., in the form of an expectation generalizing Eq. (31):

$$H(\phi, \theta) = \log \mathbb{E}^\phi \exp [\langle \theta, F(\phi, y_\phi) \rangle], \quad (32)$$

where F is problem dependent and the expectation is taken over the statistics of some underlying process y_ϕ conditional on ϕ being fixed. Assumption A3 is added for simplicity, as it guarantees that Hamilton's equations (8) are well posed.

The aim of this section is to discuss at a generic level the meaning we can give to the min-max problem (7) assuming that the Hamiltonian satisfies these assumptions. In particular, we show that the conjugate field θ appearing in the min-max problem (7) can be generically interpreted as measuring the mean effects of the fluctuations in the system dynamics needed to achieve some rare event and the minimum of the action as the total cost of these fluctuations from which the probability of the event can be estimated as

well as its mechanism and rate (if we add the minimization over $T > 0$).

1. Mean behavior

To begin, notice that, for the interpretation that θ measures the effects of the fluctuations to be consistent, the stochastic system under consideration should, in some appropriate limit in which the fluctuations disappear, satisfy the deterministic evolution equation obtained by setting $\theta = 0$ in Hamilton's equations (8):

$$\dot{\phi} = \partial_\theta H(\phi, \theta = 0). \quad (33)$$

For example, returning to the problems mentioned in Sec. II A, Eq. (33) reduces to the ordinary differential equation $\dot{x} = b(x)$ as $\epsilon \rightarrow 0$ in the SDE (17) with Hamiltonian (18); to the law of mass action (21) as $\Omega \rightarrow \infty$ for the reaction network (20) with Hamiltonian (22); to the McKean-Vlasov equation as $N \rightarrow \infty$ for the interacting particle system (24) with Hamiltonian (26); and to Eq. (28) as $\alpha \rightarrow 0$ in the slow-fast system (27) with Hamiltonian (31). More generally, under our assumptions, the solution of Eq. (33) is indeed a special solution with $\theta(t) = 0$ of Hamilton's equations (8), since assumptions A1 and A3 imply that $\partial_\phi H(\phi, 0) = 0$ for all ϕ .

2. Impact of the fluctuations

At the same time, the solution to the deterministic evolution (33) with $\phi(0) = \phi_a$ does not satisfy $\phi(T) = \phi_b$, in general. Therefore, for general boundary conditions $\phi(0) = \phi_a$ and $\phi(T) = \phi_b$, the solution of Eq. (8) must have $\theta(t) \neq 0$ —in the minimum-action framework, this is a reflection that in the original system fluctuations are needed to drive the system's trajectory away from the solution of Eq. (33), and the value of $\theta(t) \neq 0$ allows us to quantify the cost or probability of observing the event $\phi(T) = \phi_b$ given that $\phi(0) = \phi_a$. Specifically, under the strict convexity assumption A2, we have

$$\langle \theta, \partial_\theta H(\theta, \phi) \rangle - H(\phi, \theta) \geq -H(\phi, 0) = 0, \quad (34)$$

with equality if and only if $\theta = 0$. Since $\dot{\phi} = \partial_\theta H(\phi, \theta)$ along the solution to Hamilton's equations (8), we deduce that along this solution

$$\inf_\phi S_T[\phi] = \inf_\phi \int_0^T [\langle \theta, \partial_\theta H(\phi, \theta) \rangle - H(\phi, \theta)] dt \geq 0. \quad (35)$$

Thus, the action $S_T[\phi]$ can be indeed interpreted as a cost, which is zero only if $\theta(t) = 0$ (i.e., when the event can occur without fluctuations) and is strictly positive otherwise (i.e., when the event requires fluctuations). For the problems listed in Sec. II A, the form of the integrand $\langle \theta, \partial_\theta H \rangle - H$ is $|\sigma(x)\theta|^2 \geq 0$ for the SDE (17); $\sum_{j=1}^R a_j(x) (\langle \nu_j, \theta \rangle e^{\langle \nu_j, \theta \rangle} - e^{\langle \nu_j, \theta \rangle} + 1) \geq 0$ for the reaction network (20); $\frac{1}{2} \int_{\mathbb{R}^d} |\sigma(x)\nabla\theta(x)|^2 \rho(x) dx \geq 0$ for the

interacting particle system (24); and $\mathbb{E}^x f(x, y_x) e^{\langle \theta, f(x, y_x) \rangle} / \mathbb{E}^x e^{\langle \theta, f(x, y_x) \rangle} - H$ for the slow-fast system (27).

In this interpretation, minimizing $S_T[\phi]$ amounts to minimizing the cost of the fluctuations or, equivalently, finding the most likely fluctuation that drives the event $\phi(T) = \phi_b$ given that $\phi(0) = \phi_a$, leading to the asymptotic estimate (2) for the probability of the event.

3. Long-time limit

Turning now our attention to the problem $\inf_{T \geq 0} \inf_{\phi} S_T[\phi]$, its interpretation is easiest if we assume that the noiseless equation (33) has N stable fixed points $\phi_1, \phi_2, \dots, \phi_N$, and the basins of attraction of these fixed points under Eq. (33), denoted respectively as B_1, \dots, B_N , partition $\mathcal{M} = \cup_{i=1}^N \bar{B}_i$ (note that $B_i \cap B_j = \emptyset$ if $i \neq j$ by definition). In this case, we can calculate the quasipotentials $V_{\phi_i}(\phi_j)$ of every ϕ_i and ϕ_j with $i \neq j$ such that these points have adjacent basins: These are defined as $V_{\phi_i}(\phi_j) = \inf_{T > 0} \inf_{\phi} S_T[\phi]$ for all paths $\{\phi(t)\}_{T \in [0, T]}$ such that $\phi(0) = \phi_i$, $\phi(T) = \phi_j$, and $\phi(t) \in \bar{B}_i \cup \bar{B}_j$ for all $t \in [0, T]$. Consistently, we also set $V_{\phi_i}(\phi_j) = +\infty$ if B_i and B_j have no common boundary. These quasipotentials quantify the cost of the fluctuations needed to escape ϕ_i conditional on entering ϕ_j next, and these costs can be used to deduce the asymptotic rate of these transition events. More precisely, as the parameter ϵ measuring the amplitude of the macroscopic fluctuations tends to zero, the system dynamics can be approximated by a Markov jump process (MJP) between the metastable states ϕ_1, \dots, ϕ_N , with rates given asymptotically by

$$k_{i,j} = e^{-V_{\phi_i}(\phi_j)/\epsilon}, \quad i, j = 1, \dots, N \quad (i \neq j). \quad (36)$$

Questions about the asymptotic behavior of the system's dynamics can be answered by analyzing this MJP. Since the rates are all vanishing exponentially at different rates as $\epsilon \rightarrow 0$, this process is quite singular and can be analyzed by the method of decomposition into cycles developed by Freidlin and Wentzell [37] (see also Ref. [38]). This method is, however, rather intricate, and in practice it is often simpler to solve specific questions in the MJP directly (e.g., what is its invariant distribution or what is the mean first passage time from state i to state j) and take the limit as $\epsilon \rightarrow 0$ afterward. Note also that the quasipotentials $V_{\phi_i}(\phi)$ for $i = 1, \dots, N$ can be used to construct a global non-equilibrium potential $V(\phi)$, solution of the Hamilton-Jacobi equation (12) [37,38]; since this construction is not doable in practice for the high-dimensional examples we are interested in, we do not dwell upon it here. We do, however, look at the minimizers of $\inf_{T \geq 0} \inf_{\phi} S_T[\phi]$, which can exist if the path $\{\phi(t)\}_{t \in [0, T]}$ is reparametrized using arclength rather than physical time on any minimizing sequence: These geometric paths give the mechanism of the transition between ϕ_i and ϕ_j .

The statements made in this section summarize what can be deduced from the minimum-action principle when it applies. As stated earlier, the starting point is always an action giving the weights of trajectories of the stochastic system, whether in or out of equilibrium. This action can be obtained rigorously in specific cases using, e.g., tools from LDT, in the small noise [37] and weak interaction limit [49], or in more complicated setups involving hydrodynamic limits [54,55]. This is notably how the starting point of MFT can be rigorously derived [56]. These approaches may directly yield the Lagrangian of the dynamics, when the noise structure allows for it. A second route to obtain the action can be to start from the master equation and write the probability density with a WKB approximation. This is typically the route followed to study chemical networks and population dynamics [45]. It leads to the Hamilton-Jacobi equation, which is a convenient starting point to derive analytical results on the quasipotential [7,39]. The Lagrangian is then implicitly defined as the Legendre transform of the Hamiltonian. Finally, one must mention derivation based on the path-integral formulation. In particular, the Martin-Siggia-Rose-Jensen-De Dominicis formalism [28] or the Doi-Peliti formalism [29] lead both to an action that can be interpreted as above (with H satisfying assumptions A1, A2, and A3). The search of the saddle points of the action is similar to the min-max problem (7) and gives again the most probable path between two system configurations.

The numerical methods that we propose below are of general purpose and are designed with such general situations in mind.

III. COMPUTATIONAL ASPECTS

A. Min-max on the action as a wave propagation problem

In this section, we discuss how to solve the min-max problem stated in Eq. (7), i.e.,

$$\inf_{\phi} S_T[\phi] = \min_{\phi} \max_{\theta} I_T[\phi, \theta], \quad (37)$$

where we define the functional

$$I_T[\phi, \theta] = \int_0^T [\langle \dot{\phi}, \theta \rangle - H(\phi, \theta)] dt \quad (38)$$

and the optimization is to be performed over trajectories $\{\phi(t), \theta(t)\}_{t \in [0, T]}$ subject to the boundary conditions $\phi(0) = \phi_a$ and $\phi(T) = \phi_b$. By assumption A2, the functional I_T is strictly concave in θ ; assuming that it is also locally convex in ϕ around the local minimizers of $S_T[\phi]$, it is established [57,58] that this min-max problem can be solved by amortizing the minimization and maximization over small alternating steps of steepest descent in ϕ and steepest ascent in θ . If these alternating steps are infinitesimal in some artificial optimization time τ , this GDA method leads to the evolution equation

$$\partial_\tau \phi = -\alpha \delta I_T / \delta \phi, \quad \alpha \partial_\tau \theta = \delta I_T / \delta \theta, \quad (39)$$

where for convenience we introduce a parameter $\alpha > 0$ that sets the relative timescales over which ϕ and θ evolve—in the jargon of GDA, this is referred to as two-timescale GDA [59]. Calculating the functional derivatives, the system (39) is explicitly given by

$$\partial_\tau \phi = \alpha \partial_t \theta + \alpha \partial_\phi H, \quad \alpha \partial_\tau \theta = \partial_t \phi - \partial_\theta H. \quad (40)$$

These equations for $\{\phi(\tau, t), \theta(\tau, t)\}$ are to be solved with the boundary conditions (in physical time t)

$$\phi(\tau, t=0) = \phi_a, \quad \phi(\tau, t=T) = \phi_b \quad (41)$$

for some initial conditions (in optimization time τ)

$$\phi(\tau=0, t) = \phi^0(t), \quad \theta(\tau=0, t) = \theta^0(t), \quad (42)$$

with $\phi^0(t)$ such that $\phi^0(0) = \phi_a$ and $\phi^0(T) = \phi_b$.

It is easy to see that the fixed points (in τ) of Eq. (40) are a solution to Hamilton's equations (8) that satisfy $\phi(0) = \phi_a$ and $\phi(T) = \phi_b$. In Appendix A, we show that (i) there is a one-to-one correspondence between the fixed points of Eq. (40) and the critical points of the action $S_T[\phi]$, and (ii) if α is small enough, these fixed points are stable if and only if they are local minimizers of the action. Thus, solving Eq. (40) is indeed a way to perform $\inf_\phi S_T[\phi]$. For illustrative purposes, we derive in Appendix B how the GDA converges to the instanton for an Ornstein-Uhlenbeck process.

Let us now show how to put Eq. (40) in a form that is convenient for numerical integration. Since Eq. (40) is a hyperbolic system of partial differential equations (PDEs), it is useful to diagonalize the problem, introduce the fields $u = \phi + \alpha\theta$ and $v = \phi - \alpha\theta$ that propagate along characteristics, and verify

$$\partial_\tau u = \partial_t u + f(u, v), \quad (43)$$

$$\partial_\tau v = -\partial_t v + g(u, v), \quad (44)$$

where we define

$$f(u, v) = \alpha \partial_\phi H - \partial_\theta H, \quad (45)$$

$$g(u, v) = \alpha \partial_\phi H + \partial_\theta H. \quad (46)$$

The boundary conditions now involve only the propagating fields

$$v(\tau, t=0) = -u(\tau, t=0) + 2\phi_a, \quad (47)$$

$$u(\tau, t=T) = -v(\tau, t=T) + 2\phi_b. \quad (48)$$

This formulation shows that the system made of Eqs. (43) and (44) is well posed under these boundary conditions (see Ref. [60]), since the fields v and u propagate, respectively, forward and backward in physical time t as the optimization time τ increases. It also immediately suggests an algorithm to solve Eqs. (43) and (44) based on Strang splitting [61]: To update the fields at every iteration step in τ , first update u at v fixed by propagating the final condition at $t = T$ for u in Eq. (48) toward $t = 0$ using Eq. (43) with forward differentiation in t , and then update v at u fixed by propagating the initial condition at $t = 0$ for v in Eq. (47) toward $t = T$ using Eq. (44) with backward differentiation in t . Of course, the convergence does not change if the algorithm starts by updating v before updating u .

In practice, the continuous paths $\phi(\tau, t)$ and $\theta(\tau, t)$ are discretized in physical time on $M + 1$ points with index $i \in I = \{0, \dots, M\}$ such that $T = M\Delta t$, and we use index $n \in \mathbb{N}_0$ to encode the evolution of the paths in optimization time τ using steps of size $\Delta\tau$, so that for any field $\psi(\tau, t)$, $\psi_i^n \equiv \psi^n(i\Delta t)$. The details are given in Algorithm 1. The stability of the code relies on two important features.

Algorithm 1. Action minimization by gradient descent-ascent.

-
- 1: **Inputs:** $M \in \mathbb{N}$; a path $\{\phi_i^0\}_{i \in I}$ with $\phi_0^0 = \phi_a$ and $\phi_M^0 = \phi_b$; the functions $f(u, v)$ and $g(u, v)$; $T > 0$, $\Delta\tau > 0$, $\alpha > 0$.
 - 2: **Initialization:** For every $i \in I$, take $\theta_i^0 = 0$, and set $u_i^0 = \phi_i^0 + \alpha\theta_i^0$ and $v_i^0 = \phi_i^0 - \alpha\theta_i^0$; set $\Delta t = T/M$.
 - 3: **for** $n \geq 0$ **do**
 - 4: Update u with an implicit upwind scheme; namely, solve $\{u_i^{n+1}\}_{i \in I}$ sequentially from $i = M$ to $i = 0$ using

$$u_M^{n+1} = -v_M^n + 2\phi_b,$$

$$\frac{u_i^{n+1} - u_i^n}{\Delta\tau} = \frac{u_{i+1}^{n+1} - u_i^{n+1}}{\Delta t} + f(u_{i+1}^n, v_{i+1}^n), \quad i = M-1, \dots, 0.$$

- 5: Update v with an implicit upwind scheme; namely, solve $\{v_i^{n+1}\}_{i \in I}$ sequentially from $i = 0$ to $i = M$ using

$$v_0^{n+1} = -u_0^{n+1} + 2\phi_a,$$

$$\frac{v_i^{n+1} - v_i^n}{\Delta\tau} = -\frac{v_i^{n+1} - v_{i-1}^{n+1}}{\Delta t} + g(u_{i-1}^{n+1}, v_{i-1}^n), \quad i = 1, \dots, M.$$

- 6: Compute $\{\phi_i^{n+1} = \frac{1}{2}(u_i^{n+1} + v_i^{n+1})\}_{i \in I}$ and $\{\theta_i^{n+1} = \frac{1}{2}\alpha^{-1}(u_i^{n+1} - v_i^{n+1})\}_{i \in I}$ (if needed).
-

First, advection of the fields is treated with an implicit upwind scheme for both v and u . Second, reaction terms g and f are also evaluated on an upwind grid point with respect to the direction of advection. Reaction terms could also be evaluated on site i , but we empirically find that the upwind implementation strongly stabilizes the code when dealing with spatially extended diffusive fields. The stability analysis of the numerical scheme is detailed in Appendix C.

We should also emphasize that our interest resides in the fixed point of the dynamics that solves Hamilton's equation in physical time t rather than in the details of the dynamics in algorithmic time τ , which has no physical relevance. This consideration enjoins us to look for the largest time step $\Delta\tau$ that still provides a converging algorithm. The time step Δt , however, crucially needs to remain smaller than some characteristic time t_c needed to correctly resolve the dynamics of the instanton. This issue is discussed in more detail in Appendix C.

Finally, it is important to mention that a higher-order finite-difference stencil for the advection of the fields can be implemented while keeping the same algorithmic complexity. Such a scheme can significantly improve computation time, since we need a smaller number of grid points to get the same accuracy as the first-order scheme, introduced in the text for the purpose of simplicity. The second-order scheme is presented in Appendix D. The various implementations of the algorithm are accessible in [62].

B. Geometric formulation on unbounded time intervals

Let us now turn to the min-max problem

$$V_{\phi_a}(\phi_b) = \inf_{T>0} \inf_{\phi} S_T[\phi] = \inf_{T>0} \inf_{\phi} \sup_{\theta} I_T[\phi, \theta], \quad (49)$$

Algorithm 2. Geometric action minimization by gradient descent-ascent.

-
- 1: **Inputs:** $M \in \mathbb{N}$; two stable fixed points ϕ_a and ϕ_b of the noiseless dynamics where $H(\phi_{a,b}, 0) = \partial_\theta H(\phi_{a,b}, 0) = 0$; a path $\{\hat{\phi}_i^0\}_{i \in I}$ with $\hat{\phi}_0^0 = \phi_a$ and $\hat{\phi}_M^0 = \phi_b$, such that $|\hat{\phi}_{i+1}^0 - \hat{\phi}_i^0|$ is constant in i ; the functions $f(u, v)$, $g(u, v)$, and $\lambda(u, v)$; $\Delta\tau > 0$, $\alpha > 0$.
 - 2: **Initialization:** For every $i \in I$, take $\hat{\theta}_i^0 = 0$, and set $u_i^0 = \hat{\phi}_i^0 + \alpha \hat{\theta}_i^0$ and $v_i^0 = \hat{\phi}_i^0 - \alpha \hat{\theta}_i^0$; set $\Delta s = 1/M$.
 - 3: **for** $n \geq 0$ **do**
 - 4: Update u with an implicit upwind scheme, namely, solve $\{u_i^{n+1}\}_{i \in I}$ sequentially from $i = M$ to $i = 0$ using

$$\begin{aligned} u_M^{n+1} &= -v_M^n + 2\phi_b, \\ \frac{u_i^{n+1} - u_i^n}{\Delta\tau} &= \lambda_i(u^n, v^n) \frac{u_{i+1}^{n+1} - u_i^{n+1}}{\Delta s} + f(u_i^n, v_i^n), \quad i = M-1, \dots, 0. \end{aligned}$$

- 5: Update v with an implicit upwind scheme, namely, solve $\{v_i^{n+1}\}_{i \in I}$ sequentially from $i = 0$ to $i = M$ using

$$\begin{aligned} v_0^{n+1} &= -u_0^{n+1} + 2\phi_a, \\ \frac{v_i^{n+1} - v_i^n}{\Delta\tau} &= -\lambda_i(u^{n+1}, v^n) \frac{v_i^{n+1} - v_{i-1}^{n+1}}{\Delta s} + g(u_i^{n+1}, v_i^n), \quad i = 1, \dots, M. \end{aligned}$$

- 6: Compute $\{\bar{\phi}^{n+1} = \frac{1}{2}(u^{n+1} + v^{n+1})\}_{i \in I}$ and $\{\bar{\theta}^{n+1} = \frac{1}{2}\alpha^{-1}(u^{n+1} - v^{n+1})\}_{i \in I}$.
 - 7: Interpolate $\{(\bar{\phi}_i^{n+1}, \bar{\theta}_i^{n+1})\}_{i \in I}$ onto a path $\{(\hat{\phi}_i^{n+1}, \hat{\theta}_i^{n+1})\}_{i \in I}$ such that $|\hat{\phi}_{i+1}^{n+1} - \hat{\phi}_i^{n+1}|$ is constant in i , as in the string method.
 - 8: Set $\{u_i^{n+1} = \hat{\phi}_i^{n+1} + \alpha \hat{\theta}_i^{n+1}\}_{i \in I}$ and $\{v_i^{n+1} = \hat{\phi}_i^{n+1} - \alpha \hat{\theta}_i^{n+1}\}_{i \in I}$.
-

where $I_T[\phi, \theta]$ is the functional defined in Eq. (38) and the inner min-max is again to be performed over trajectories $\{\phi(t), \theta(t)\}_{t \in [0, T]}$ subject to the boundary conditions $\phi(0) = \phi_a$ and $\phi(T) = \phi_b$.

In general, a trajectory $\{\phi(t), \theta(t)\}_{t \in [0, T]}$ of finite time duration T cannot solve Eq. (49); i.e., we can always reduce the value of the inner min-max by increasing T . This lack of optimizer complicates the solution of Eq. (49). To proceed, it is useful to follow the strategy of the GMAM in Ref. [30] and parametrize the physical time as $t(s)$ for $s \in [0, 1]$. Writing $dt/ds = \lambda^{-1}(s)$ and introducing $[\hat{\phi}(s), \hat{\theta}(s)] = \{\phi[t(s)], \theta[t(s)]\}$, the minimization over T in Eq. (49) can be turned into a minimization over λ :

$$V_{\phi_a}(\phi_b) = \min_{\lambda \geq 0} \min_{\hat{\phi}} \max_{\hat{\theta}} \int_0^1 [(\hat{\phi}', \hat{\theta}) - \lambda^{-1} H(\hat{\phi}, \hat{\theta})] ds, \quad (50)$$

where $\hat{\phi}' = d\hat{\phi}/ds$ and the min-max is to be performed over the triple $\{\hat{\phi}(s), \hat{\theta}(s), \lambda(s)\}_{s \in [0, 1]}$ subject to the boundary conditions $\hat{\phi}(0) = \phi_a$ and $\hat{\phi}(1) = \phi_b$. Since we add degrees of freedom by representing T by the function $\lambda(s)$, we can add a constraint on the parametrization of $\hat{\phi}(s)$, e.g., by imposing that $|\hat{\phi}'| = c s t$ (in s), in which case s is the normalized arclength along the path $\hat{\phi}$. This choice is convenient numerically, as it guarantees that discretization points in s are uniformly distributed along the path.

The min-max problem (50) is now in a form that can be solved using GDA, similar to what we do for Eq. (37). As shown in Appendix E, it is, however, convenient to treat λ separately, and this leads to the following GDA equations [compare Eq. (40)]:

$$\partial_\tau \hat{\phi} = \alpha \lambda \partial_s \hat{\theta} + \alpha \partial_\phi H, \quad \alpha \partial_\tau \hat{\theta} = \lambda \partial_s \hat{\phi} - \partial_\theta H, \quad (51)$$

with λ given by

$$\lambda = \frac{|\langle \partial_\theta H, \partial_s \hat{\phi} \rangle| + \sqrt{\max(0, \langle \partial_\theta H, \partial_s \hat{\phi} \rangle^2 - 4H|\partial_s \hat{\phi}|^2)}}{2|\partial_s \hat{\phi}|^2}. \quad (52)$$

Equations (51) are to be solved under the constraint that $|\partial_s \hat{\phi}| = cst$ (in s , not τ), with the boundary conditions (in s)

$$\hat{\phi}(\tau, s = 0) = \phi_a, \quad \hat{\phi}(\tau, s = 1) = \phi_b, \quad (53)$$

for some initial conditions (in optimization time τ)

$$\hat{\phi}(\tau = 0, s) = \hat{\phi}^0(s), \quad \hat{\theta}(\tau = 0, s) = \hat{\theta}^0(s), \quad (54)$$

with $\hat{\phi}^0(s)$ such that $\hat{\phi}^0(0) = \phi_a$ and $\hat{\phi}^0(1) = \phi_b$.

Similar to what we do with Eq. (40), the system of equations (51) as well as the expression (52) for λ can be put in a form suitable for numerical solution by rewriting them in terms of the fields $u = \hat{\phi} + \alpha \hat{\theta}$ and $v = \hat{\phi} - \alpha \hat{\theta}$. For brevity, we do not write these equations explicitly and refer the reader to Algorithm 2 for their discretized version. At convergence, $\hat{\phi}$ and $\hat{\theta}$ satisfy

$$\lambda \hat{\phi}' = \partial_\theta H, \quad \lambda \hat{\theta}' = -\partial_\phi H. \quad (55)$$

If we insert the first of these equations in Eq. (52), we can reorganize this equation into

$$\lambda = \frac{\lambda |\hat{\phi}'|^2 + \sqrt{\lambda^2 |\hat{\phi}'|^4 - 4H |\hat{\phi}'|^2}}{2 |\hat{\phi}'|^2}. \quad (56)$$

This equation shows that at convergence we must also have

$$\forall s \in [0, 1], \quad H[\hat{\phi}(s), \hat{\theta}(s)] = 0. \quad (57)$$

This is consistent with the fact that the term involving $\lambda^{-1}H$ in Eq. (50) can also be interpreted as a Lagrangian multiplier term added to the objective function to enforce the constraint that $H = 0$; see Ref. [30] for details.

In practice, the continuous paths $\hat{\phi}(\tau, s)$ and $\hat{\theta}(\tau, s)$ are discretized in s on a grid of $M + 1$ points with index $i \in I = \{0, \dots, M\}$ such that $1 = M\Delta s$, and we use the index $n \in \mathbb{N}_0$ to encode their evolution with step size $\Delta\tau$ in artificial time τ so that for any field $\psi(\tau, s)$, $\psi_i^n \equiv \psi^n(i\Delta s)$. The details are given in Algorithm 2, which is a modified version of Algorithm 1 that includes the step of reparametrization of the path. Note that we keep the implicit upwind discretization for the advection term. Also, the reaction terms f and g and the coefficient λ should be evaluated at

the same grid point as originating from the same term $\lambda^{-1}H(\hat{\phi}, \hat{\theta})$ in Eq. (50). We find that evaluating these terms at the target grid point i is better for stability, in general.

Algorithms 1 and 2 are simple to implement and require only evaluating the first derivative of the Hamiltonian $H(\phi, \theta)$ with respect to its arguments, just as we would have to do to solve Hamilton's equations (8). It is used to calculate the paths shown in Figs. 1–3. In the more complicated examples we treat below, because the state variable is a field that depends on space as well as time, and as a result the PDEs (43) and (44) involve spatial derivatives, some additional consideration must be given to the way we discretize space and evaluate these derivatives to ensure that the resulting scheme is numerically stable. As usual with PDEs, this issue needs to be addressed on a case-by-case basis, depending on the nature of the PDE.

IV. FIRST-ORDER PHASE TRANSITIONS IN A NONEQUILIBRIUM GL SYSTEM

A possible starting point of the method borrows from Landau's theory of phase transitions where, relying on the symmetries of the system, we postulate the free energy of the macroscopic variable of interest ϕ rather than deriving it from a microscopic distribution [3,63–65]. When the system is in equilibrium, this procedure is well understood [66,67]. For nonequilibrium systems, MFT [39,68,69] offers a generalization of Landau's approach where we directly start from the action. In what follows, we apply a similar approach to study phase transitions in a modified Ginzburg-Landau system.

A. Nonequilibrium GL dynamics

Following the general framework introduced in Refs. [70,71], the stochastic evolution of a nonconservative field $\rho(t, x)$ can formally be described by the Langevin equation

$$\partial_t \rho = -\mu([\rho], x) + \sqrt{2\epsilon} \eta, \quad (58)$$

where the drift $\mu([\rho], x)$ can be interpreted as a chemical potential, $\eta(t, x)$ is a standard Gaussian white noise in space and time whose amplitude is measured by $\epsilon > 0$, and for simplicity we set the mobility to 1. For a field ρ defined on a one-dimensional domain, say, $[0, 1]$, the action corresponding to this equation is

$$S_T[\rho] = \frac{1}{2} \int_0^T \int_0^1 |\partial_t \rho + \mu([\rho], x)|^2 dx dt, \quad (59)$$

and it is the subject of our investigations.

The dynamics (58) is in detailed balance when $\mu = \mu_E$, with μ_E such that it can be written as a derivative of a free energy functional $\mathcal{F}[\rho]$:

$$\mu_E([\rho], x) = \frac{\delta \mathcal{F}[\rho]}{\delta \rho(x)}. \quad (60)$$

In this case, at equilibrium, the field configurations are distributed according to the Gibbs measure associated with the free energy \mathcal{F} , at temperature ϵ .

Here, we are mostly interested in the active version of dynamics (58), when $\mu = \mu_E + \mu_A$ with μ_A that cannot be cast into the form (60) or, equivalently, does not satisfy the functional Schwarz relation [9,13]:

$$\frac{\delta \mu_A([\rho], x)}{\delta \rho(y)} - \frac{\delta \mu_A([\rho], y)}{\delta \rho(x)} \neq 0. \quad (61)$$

When $\mu_A \neq 0$, the stationary distribution of the field configurations, if it exists, is a nonequilibrium distribution which is not available in closed form.

For concreteness, we focus on the following example that displays a nonequilibrium first-order phase transition: We assume that the field $\rho(t, x)$ is one-dimensional, with $x \in [0, 1]$ and periodic boundary conditions, and we use chemical potentials given by

$$\mu_E([\rho], x) = -(D\partial_x^2 \rho + \rho - \rho^3 + h), \quad (62)$$

$$\mu_A([\rho]) = -\kappa \int_0^1 \rho^2(y) dy, \quad (63)$$

such that Eq. (58) becomes

$$\partial_t \rho = D\partial_x^2 \rho + \rho - \rho^3 + h + \kappa \int_0^1 \rho^2(y) dy + \sqrt{2\epsilon} \eta. \quad (64)$$

Here, $D > 0$ is a diffusion constant which effectively depends on the system size in the dimensionless variables we use (reducing D is equivalent to enlarging the domain size), h the strength of an externally applied field, and κ is the strength of the nonequilibrium coupling. The chemical potential μ_E is borrowed from the Ginzburg-Landau ϕ^4 theory. For $\kappa = 0$, Eq. (64) is referred to as the time-dependent-Ginzburg-Landau model and is the space-continuous version of the Ising model studied by Glauber [72]. This dynamical equation is also referred to as model A [66] and was studied by Allen and Cahn in the context of front propagation in first-order phase transitions in metallic alloys [73]. A direct functional integration shows that $\mu_E = \delta \mathcal{F} / \delta \rho$ with

$$\mathcal{F}[\rho] = \int_0^1 \left(\frac{1}{2} D |\partial_x \rho|^2 + \frac{1}{4} [1 - \rho^2(x)]^2 - h \rho(x) \right) dx. \quad (65)$$

The additional force μ_A we consider in Eq. (63) is nonlinear and strongly nonlocal. Such terms can be found in different contexts (nonlinear optical systems [74], population dynamics [75], and species competition [76]). We opt for

this particular choice of μ_A as it produces nonequilibrium effects that are nontrivial. Specifically, μ_A acts as an additional uniform applied field that is active and depends nonlinearly on the value of ρ rather than being externally imposed. While μ_E drives the system toward the minimizers of the energy (65), which are homogeneous state solutions of $\rho - \rho^3 + h$, μ_A homogeneously pushes the field upward when $\kappa > 0$ and downward when $\kappa < 0$. Therefore, the applied field h and μ_A have competing effects when h and κ have opposite signs. Since there is a region in the (κ, h) space where the noiseless dynamics has two stable fixed points (see Sec. IV B), this means that the system can undergo a nonequilibrium first-order phase transition for critical values of h and κ which we determine in Sec. IV C.

B. Phase boundaries for coexisting homogeneous fixed points

Numerical evidence indicates that the stable fixed points of the noiseless dynamics [i.e., Eq. (64) with $\epsilon = 0$] are homogeneous states. As a result, they are solutions to $\rho - \rho^3 + h + \kappa \rho^2 = 0$. In the domain where this equation has three real roots, ρ_- , ρ_c , and ρ_+ with $\rho_- < \rho_c < \rho_+$, ρ_- and ρ_+ are stable fixed points of the noiseless dynamics, whereas ρ_c is an unstable point. The coexistence region in the parameter space (κ, h) where both ρ_- and ρ_+ are present is marked as region (ii) in Fig. 5; it is where $h \in [h_c^-, h_c^+]$, with h_c^- and h_c^+ given by

$$h_c^- = -\frac{1}{27} \left(\sqrt{\kappa^2 + 3} + \kappa \right)^2 \left(2\sqrt{\kappa^2 + 3} - \kappa \right), \quad (66)$$

shown as a blue line in Fig. 5, and

$$h_c^+ = \frac{1}{27} \left(\kappa - \sqrt{\kappa^2 + 3} \right)^2 \left(2\sqrt{\kappa^2 + 3} + \kappa \right), \quad (67)$$

shown as a yellow line in Fig. 5. Exactly on these boundaries, only two real roots coexist, and one state is, thus, marginally stable. In regions (i) and (iii), only one stable state exists.

Our next goal is to analyze the relative stability of ρ_- and ρ_+ under the effect of the noise, i.e., derive the phase diagram of the system. Even though we lack a free energy that yields the stationary measure, we expect ρ_+ to be the stable phase for $h, \kappa > 0$ and ρ_- for $h, \kappa < 0$. However, when h and κ have opposite signs (and, thus, opposite effects on ρ), determining the most likely phase becomes nontrivial.

C. Extracting the minimum-action paths

To assess whether ρ_+ or ρ_- is the stable phase in the coexistence region, we compute the difference of the minimal actions $\Delta S \equiv V_{\rho_-}(\rho_+) - V_{\rho_+}(\rho_-)$, obtained from the minimum-action paths from ρ_- to ρ_+ and vice versa. The Hamiltonian entering the action is the one associated with Eq. (58):

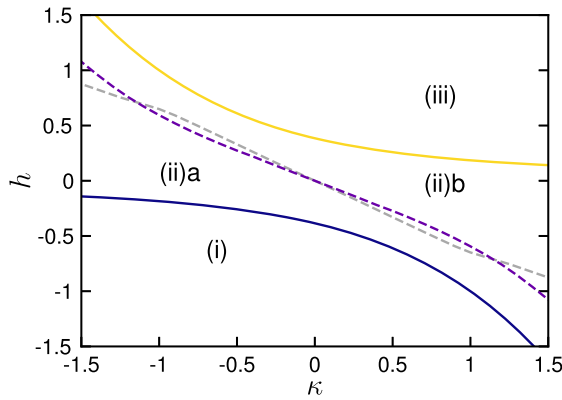


FIG. 5. Phase diagram of the modified Ginzburg-Landau system. The phase diagram is divided in four regions. Region (i): ρ_- is the only phase. Region (ii): coexistence region where the noiseless dynamics has two stable homogeneous fixed points ρ_{\pm} and one unstable homogeneous fixed point ρ_c . Region (ii)(a): ρ_- is the stable phase. Region (ii)(b): ρ_+ is the stable phase. Region (iii): ρ_+ is the only phase. Yellow solid line: $h_c^+(\kappa)$. Purple solid line: $h_c^-(\kappa)$. Purple dashed line: phase transition curve between regions (ii)(a) and (ii)(b) obtained by the minimum-action method. Gray dashed line: phase transition curve obtained by treating the transition as if the system were in equilibrium (wrong prediction).

$$H(\rho, \theta) = \langle -\mu[\rho], \theta \rangle_{L^2} + \langle \theta, \theta \rangle_{L^2}, \quad (68)$$

where the scalar product of two functions f and g is given by $\langle f, g \rangle_{L^2} = \int_0^1 f(x)g(x)dx$. To obtain the phase diagram, these calculations must be repeated for a set of values (κ, h) in the coexistence region to compute the minimal action difference as a function of these parameters, $\Delta S(\kappa, h)$: From Eq. (5), the line of phase transition is then the curve where $\Delta S(\kappa, h) = 0$.

In practice, we use Algorithm 2 with $M = 400$ copies along the path, $N_x = 64$ points of space discretization, $\Delta\tau = 10^{-3}$, and $\alpha = 0.33$, and we monitor convergence by looking at the decay of the action. The result of these computations is shown in Fig. 5, where the purple dashed line frontier in the phase diagram corresponds to $\Delta S(\kappa, h) = 0$. The stable phase is ρ_- below the line [region (ii)(a)], and ρ_+ above it [region (ii)(b)].

For comparison, we also compute the phase diagram under the (wrong) assumption that the escape paths are given by the heteroclinic orbits followed in a time-reversed way. This would have to be the case in equilibrium by time-reversal symmetry. While these escape paths are incorrect, in general, in nonequilibrium systems, their respective cost in the action gives an upper bound on the actual minima $V_{\rho_-}(\rho_+)$ and $V_{\rho_+}(\rho_-)$. Denoting by $S_{\rho_-}^{\text{het}}(\rho_+)$ and $S_{\rho_+}^{\text{het}}(\rho_-)$ the actions along the heteroclinic orbit, we compute, for instance, $S_{\rho_-}^{\text{het}}(\rho_+)$ as

$$S_{\rho_-}^{\text{het}}(\rho_+) = \int_0^1 [\langle \partial_s \hat{\rho}, \hat{\theta} \rangle - \lambda^{-1} H(\hat{\rho}, \hat{\theta})] ds, \quad (69)$$

where the path $\hat{\rho}(s)$ and the parametrization $\lambda(s)$ ($s \in [0, 1]$) are obtained by the string method [77,78] that identifies the heteroclinic orbit between $\hat{\rho}(0) = \rho_-$ and $\hat{\rho}(1) = \rho_+$ and θ solves

$$\lambda \partial_s \hat{\rho} = \partial_{\theta} H(\hat{\rho}, \hat{\theta}). \quad (70)$$

As a sanity check, we verify that one always has $V_{\rho_-}(\rho_+) \leq S_{\rho_-}^{\text{het}}(\rho_+)$ and $V_{\rho_+}(\rho_-) \leq S_{\rho_+}^{\text{het}}(\rho_-)$, namely, that the minimizer of the action is always smaller than the action along the heteroclinic orbit. It is also worth noticing that these bonds offer no information about the location of the phase transition line: The line where $\Delta S^{\text{het}}(\kappa, h) = 0$ is plotted as the gray dashed line in Fig. 5, and it is different from the actual transition line $\Delta S(\kappa, h) = 0$.

The minimum-action paths also give physical insights about the mechanism of the transition: The contour plot in (s, x) space of these paths is shown in Fig. 6 for the specific value $(\kappa, h) = (1, -0.5)$ [which is in region (ii)(a)]: Fig. 6(a) shows the forward path from ρ_- to ρ_+ and Fig. 6(b) the reversed path from ρ_+ to ρ_- . Also shown in Fig. 6(c) is the heteroclinic orbit. The actual path $\{\hat{\rho}(s)\}_{s \in [0,1]}$ is shown in the first row, while the second row displays the conjugate momentum $\{\hat{\theta}(s)\}_{s \in [0,1]}$, and the third the action increment (i.e., the Lagrangian) $\langle \hat{\theta}(s), \hat{\rho}'(s) \rangle$ along the path.

As can be seen in Fig. 6, the forward and the backward minimum-action paths are different, and they cross the separatrix (marked as a dashed black line in the figure) at different places: That is, the critical nuclei for the forward and backward transitions are different (for the backward path, this ‘‘critical nucleus’’ is actually flat). This is a signature of time-symmetry breaking that can be intuitively explained as follows: When $\kappa > 0$, as in Fig. 6, the nonequilibrium term $\kappa \int \rho^2 dx$ favors the movement from ρ_- to ρ_+ but opposes the one from ρ_+ to ρ_- . In the forward path from ρ_- to ρ_+ , it is better to have $\int \rho^2 dx$ large, which favors nucleation; conversely, in the backward path from ρ_+ to ρ_- , it is better to have $\int \rho^2 dx$ small, and the way to minimize this quantity given the value of its changing mean is to have ρ spatially uniform. Notice, however, that this is a finite size effect: If D were decreased to even smaller values, we would observe ‘‘classical’’ inhomogeneous nucleation events in both directions (albeit different ones in each). To further emphasize the impact of the $\int \rho^2 dx$ term on the transition path, let us also analyze the behavior of the conjugate field θ . The forward path of θ whose contour plot is displayed in Fig. 6(a) shows the coexistence of a region $\theta(x) < 0$ for $x \gtrsim 0.45$ with a region $\theta(x) > 0$ for $x \lesssim 0.45$: This sign structure of θ enhances the inhomogeneities in ρ , which, in turn, helps the term $\int \rho^2 dx$ drive the system to the separatrix. This effect is absent in an equilibrium Allen-Cahn dynamics where the noise field θ acts positively on ρ to drive it out of the state ρ_- . Conversely, along the backward path of θ displayed in

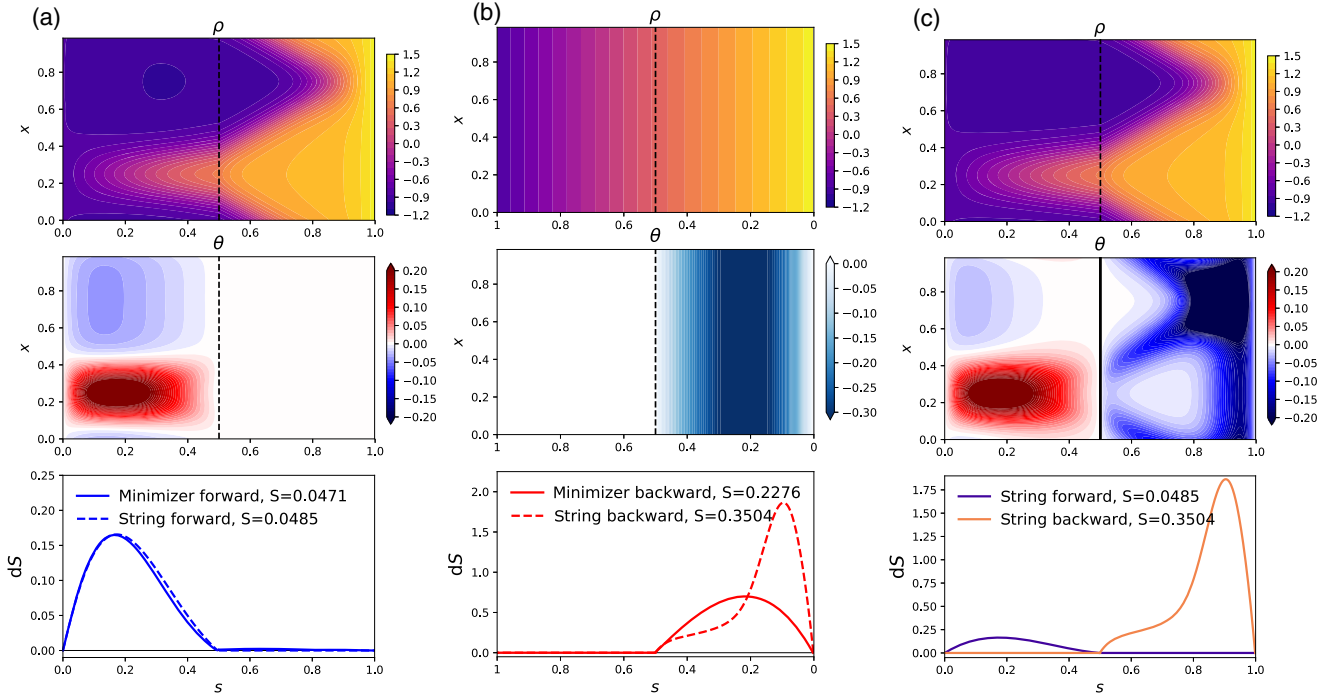


FIG. 6. Nonequilibrium GL system for $D = 5 \times 10^{-3}$, $\kappa = 1$, $h = -0.5$, and, hence, $\rho_+ = 1.45161$ and $\rho_- = -0.854638$. For these values, we have $V_{\rho_-}(\rho_+) < V_{\rho_+}(\rho_-)$, indicating that ρ_+ is the stable phase. Upper: contour plots of the paths $\hat{\rho}(s, x)$; middle: contour plots of the conjugate momentum $\hat{\theta}(s, x)$; lower: Lagrangian along the paths. (a) Minimum-action path from ρ_- to ρ_+ ; (b) minimum-action path from ρ_+ to ρ_- . Notice the strong difference between the forward and backward paths. The black dashed line at $s = 0.5$ marks the critical nucleus, past which θ and L are both zero as they should be. (c) Heteroclinic orbit joining ρ_- and ρ_+ calculated by the string method. Since the ascent is set as the reverse descent here, we plot on one graph the conjugate momentum and the Lagrangian in both the forward and reversed directions.

Fig. 6(b), the noise remains homogeneous and negative, which homogeneously decreases ρ without breaking the space-translational invariance.

Note also that the interesting part of these minimum-action paths is their escape half, where the noise is needed and the action increment is, therefore, positive: It is the first half for the forward path in Fig. 6(a) and the second half for the backward path in Fig. 6(b); for the heteroclinic orbit shown in Fig. 6(c), we display both halves at once, since the forward and reversed paths are symmetric. In all situations, past the critical nucleus, the paths simply follow the noiseless dynamics, and this half of the path may not be unique if the critical nucleus has more than one unstable direction. This nonuniqueness has no impact on the action, however, since the Lagrangian is zero along the solution of the noiseless dynamics. Finally, we check that the reaction paths that we obtain do not depend on the initial paths: Starting from the heteroclinic orbit or from a (strongly) perturbed linear interpolation of the critical states, the algorithm always finds the same instantons, up to translational symmetry (since periodic boundary conditions are taken here). This result increases the confidence that the path we find is the global minimizer of the action.

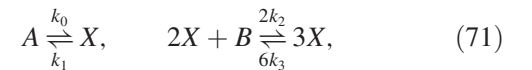
The effective nonequilibrium Ginzburg-Landau-like dynamics we consider in this section can be modified to

include different nonequilibrium terms, like, for example, $\mu_A = \kappa |\partial_x \rho|^2$, which naturally appears in the coarse-grained field description of interface growth phenomena [79] and active matter systems [12,70,80,81]. We could also use our approach to compute the phase diagram of modified Cahn-Hilliard systems, which naturally emerge in active matter field theories [8,34,70,71,81,82].

V. PHASE TRANSITIONS IN A BISTABLE REACTION-DIFFUSION SYSTEM

A. The Schlögl model

In 1972, Schlögl introduced the following chemical reaction network [42]:



with microscopic rates $k_i > 0$ and where the concentration of A and B is held constant. In a certain regime of the reaction rates, this system displays metastability between a low-density and a high-density phase. Here, we consider a spatially extended variant of this model introduced by Tănase-Nicola and Lubensky [44], in which a one-dimensional domain is split into $L \in \mathbb{N}$ well-stirred

compartments: The molecules react only within their compartment and randomly jump to neighboring ones with rate $\gamma > 0$. We also impose periodic boundary conditions.

When the number n_i of molecules in compartment $i \in \{1, \dots, L\}$ is large, it is convenient to introduce the rescaled $\rho_i = n_i/\Omega$, where $\Omega \gg 1$ is the typical number of molecules per compartments. In the limit as $\Omega \rightarrow \infty$, the law of mass action for ρ_i is a (discrete) reaction-diffusion equation

$$\dot{\rho}_i = \gamma(\rho_{i+1} + \rho_{i-1} - 2\rho_i) + w_+(\rho_i) - w_-(\rho_i), \quad (72)$$

where $w_+(\rho_i) = \lambda_0 + \lambda_2\rho_i^2$ and $w_-(\rho_i) = \lambda_1\rho_i + \lambda_3\rho_i^3$ are the rescaled reaction rates with $\lambda_i = k_i\Omega^{i-1}$ (see Appendix G). Equation (72) can be written as a gradient flow:

$$\dot{\rho}_i = -\partial_{\rho_i} E(\rho), \quad (73)$$

where $\rho = (\rho_1, \dots, \rho_L)$ and we introduce

$$E(\rho) = \sum_{i=1}^L \left(\frac{1}{2} \gamma (\rho_{i+1} - \rho_i)^2 + U(\rho_i) \right) \quad (74)$$

with

$$U(\rho_i) = -\lambda_0\rho_i - \frac{1}{3}\lambda_2\rho_i^3 + \frac{1}{2}\lambda_1\rho_i^2 + \frac{1}{4}\lambda_3\rho_i^4. \quad (75)$$

We recognize a (discrete) Ginzburg-Landau free energy, from which we conclude that the stable fixed points of Eq. (72) are homogeneous states as long as γ is not too small. For the value of $(\lambda_0, \lambda_1, \lambda_2, \lambda_3)$ that we consider here, there are two such fixed points: $\rho_i = \rho_{\pm}$ for all $i \in \{1, \dots, L\}$, where ρ_- and ρ_+ are, respectively, the smallest and the largest root of $w_+(z) - w_-(z) = 0$.

The gradient structure of Eq. (73) may suggest that ρ_- is the stable phase if $E(\rho_-) < E(\rho_+)$, whereas ρ_+ is if $E(\rho_-) > E(\rho_+)$. This conclusion is, however, incorrect, as already observed by Tănase-Nicola and Lubensky [44], who show that the system undergoes a nonequilibrium first-order phase transition when γ changes. Since $E(\rho_{\pm}) = U(\rho_{\pm})$ and, therefore, is independent of γ , the phase transition cannot be predicted by analyzing $E(\rho)$ only: The system is not in detailed balance with respect to the Gibbs measure associated to $E(\rho)$. What is not determined in Ref. [44] is the critical value γ_c at which the phase transition occurs. This is the question we solve next, using the approach introduced above.

B. Change of relative stability with increasing jumping rate

The spatially extended Schlögl model is a reaction network of the type considered in Sec. II A, and its phase diagram can be analyzed by minimizing the action associated with a Hamiltonian similar to Eq. (22). Using the structure of the model, it is natural to decompose

$H = H^R + H^D$, with H^R accounting for the reaction and H^D for the jumps:

$$H^R(\rho, \theta) = \sum_{i=1}^L w_+(\rho_i)(e^{\theta_i} - 1) + w_-(\rho_i)(e^{-\theta_i} - 1), \quad (76)$$

$$H^D(\rho, \theta) = \gamma \sum_{i=1}^L \rho_i (e^{\theta_{i-1} - \theta_i} + e^{\theta_{i+1} - \theta_i} - 2). \quad (77)$$

We use this Hamiltonian in the action that we minimize using Algorithm 2 to calculate the quasipotentials $V_{\rho_-}(\rho_+)$ and $V_{\rho_+}(\rho_-)$. We repeat these calculations for different values of the jump rate γ while keeping the rates fixed at $(\lambda_0, \lambda_1, \lambda_2, \lambda_3) = (0.8, 2.9, 3.1, 1)$, for which $\rho_- = 0.5$ and $\rho_+ = 1.6$. We use $L = 40$ compartments, and in Algorithm 2 we set $M = 400$ ($\Delta s = 2.5 \times 10^{-3}$), $\alpha = 1$, and $\Delta \tau = 0.01$. The graphs of $V_{\rho_-}(\rho_+)$ and $V_{\rho_+}(\rho_-)$ versus γ are shown in Fig. 7. These results indicate that the nonequilibrium first-order phase transition occurs at $\gamma_c \simeq 5$: ρ_+ is the stable phase for $\gamma > \gamma_c$, while ρ_- is the stable one for $\gamma < \gamma_c$. We stress again that this result cannot be deduced by looking at $E(\rho)$ even though the law of mass action can be written as the gradient flow (73). This is, of course, not a contradiction: The Schlögl model is not in detailed balance and lacks time-reversal symmetry, because this property also depends on the nature of the noise.

The nonequilibrium nature of the phase transition can be confirmed by looking at the transition paths from ρ_- and ρ_+ and vice versa. They are shown in Fig. 8, in which we use the same plotting conventions as in Fig. 6. These results are for $\gamma = 12$, when ρ_- is the stable phase. We can see that the forward [Fig. 8(a)] and the backward [Fig. 8(b)] paths are different and go through different critical nuclei (marked as a dashed vertical black line on the graphs).

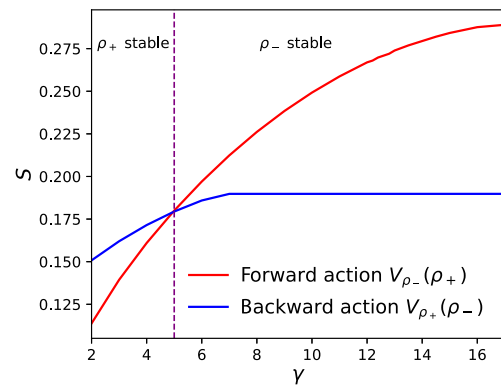


FIG. 7. Phase diagram of the Schlögl model in the limit of a large number of particles jumping between well-stirred reactive boxes. The parameters λ_i are fixed, and we vary the jump rate γ . For $\gamma < \gamma_c = 5 \pm 0.1$, we have $V_{\rho_+}(\rho_-) > V_{\rho_-}(\rho_+)$, which indicates that ρ_+ is a stable state phase, whereas ρ_- is for $\gamma > \gamma_c$. We use the parameters from Ref. [34]— $\lambda_0 = 0.8$, $\lambda_1 = 2.9$, $\lambda_2 = 3.1$, and $\lambda_3 = 1$ —and take $L = 40$.

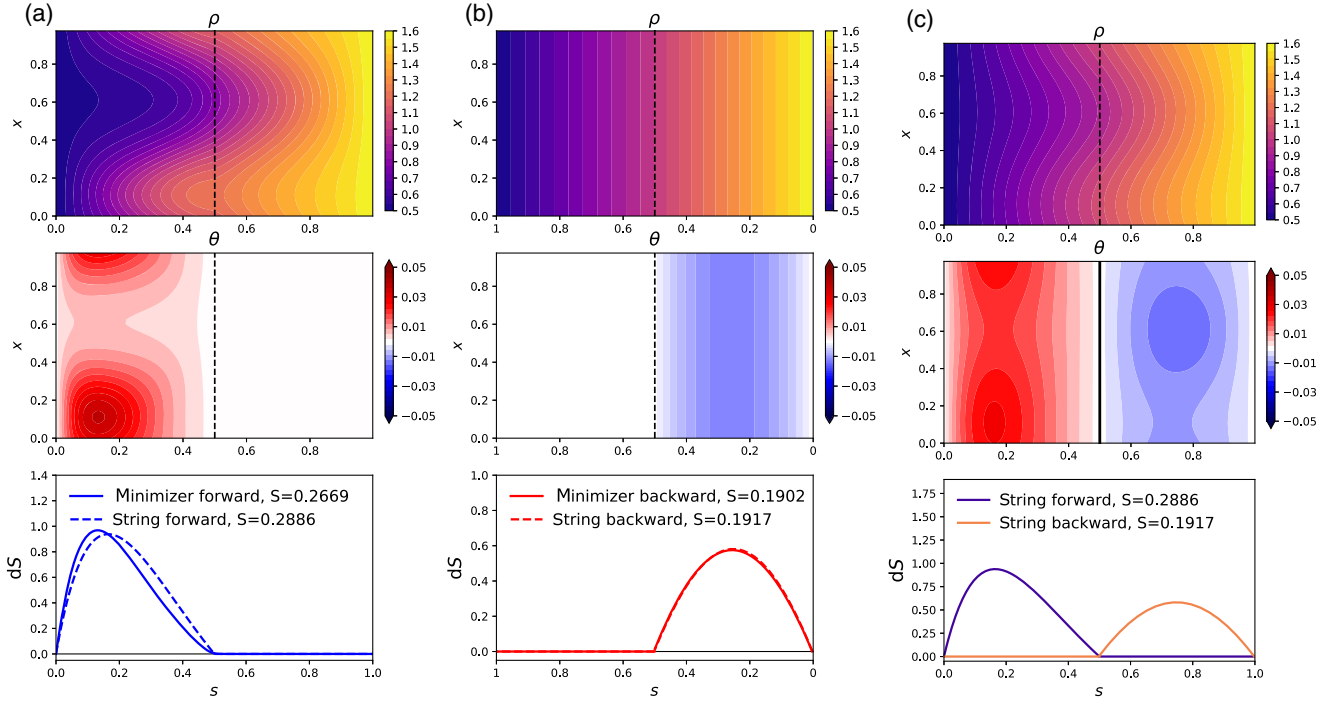


FIG. 8. Spatially extended Schlögl model for $L = 40$ reactive boxes, $(\lambda_0, \lambda_1, \lambda_2, \lambda_3) = (0.8, 2.9, 3.1, 1)$, and $\gamma = 12$. For these values, we have $\rho_- = 0.5$ and $\rho_+ = 1.6$. The figure is organized as Fig. 6. As spatial coordinate, we use $x = i/L$ and plot the transitions paths and associated momenta as if they were continuous in space.

These paths are also different from the heteroclinic orbit [Fig. 8(c)]. If we increase the value of γ , the forward path eventually becomes homogeneous (results not shown): This is consistent with the fact that, at high γ , the system behaves essentially as one single well-stirred compartment. In that limit, we can calculate the nonequilibrium steady distribution of the system and use it to calculate $V_{\rho_-}(\rho_+)$ and $V_{\rho_+}(\rho_-)$: This is done in Appendix G, and it gives the same values as the ones obtained by Algorithm 2 when γ is large and the transition paths are both homogeneous. Conversely, if we decrease the value of γ , the backward path becomes inhomogeneous (results not shown). This transition from homogeneous to inhomogeneous backward path occurs around $\gamma = 7$. How to gain intuition about these changes of behavior is harder in this model than in the GL model with additive noise in Sec. IV, because the noise is non-Gaussian (and somehow multiplicative), and ultimately the shape of the minimum-action paths depends on a complex interplay between many effects in the dynamics. Yet, once the instanton is known, the effect of the noise θ on the density ρ can still be partially interpreted recalling that the dynamics of the instanton is given by $\partial_t \rho_i = \partial_{\theta_i} H$, which here reads

$$\begin{aligned} \partial_t \rho_i &= w_+(\rho_i) e^{\theta_i} - w_-(\rho_i) e^{-\theta_i} \\ &+ \gamma [\rho_{i+1} e^{\theta_i - \theta_{i+1}} - \rho_i e^{\theta_{i-1} - \theta_i} \\ &+ \rho_{i-1} e^{\theta_i - \theta_{i-1}} - \rho_i e^{\theta_{i+1} - \theta_i}]. \end{aligned} \quad (78)$$

While the contribution from diffusion has an effect which is not easily amenable to interpretation, the reaction terms clearly show that $\theta_i > 0$ leads to particle creation through w_+ , while $\theta_i < 0$ leads to particle destruction through w_- . This effect is displayed in Fig. 8, where the spatial maxima of θ_i and ρ_i are both attained at the same site (corresponding to maximal particle creation). The drive coming from the diffusion noise is more difficult to explain. Clues can be obtained by measuring the mean value of ρ along the path (not shown), but elucidating the effect of the diffusion-related noise is hard.

Finally, as with the previous example treated in Sec. IV, we check that the reaction path we find is independent from the initial condition, which again supports the claim that the global minimizer of the action has been identified.

C. Comparison with microscopic simulations

In this section, we corroborate the conclusions in Sec. VB by performing Markov chain Monte Carlo (MCMC) simulations of the microscopic system.

For the parameters value used in Sec. VB, it is not possible to calculate this way the phase diagram shown in Fig. 7: This is because Ω needs to be large in order for the phases ρ_- and ρ_+ to be (meta)stable under the noise, and the timescales of the forward or backward transition between the phases ρ_- and ρ_+ is of the order of the nucleation times $e^{\Omega V_{\rho_-}(\rho_+)}$ and $e^{\Omega V_{\rho_+}(\rho_-)}$. These timescales are too big to be accessible with MCMC. We can, however,

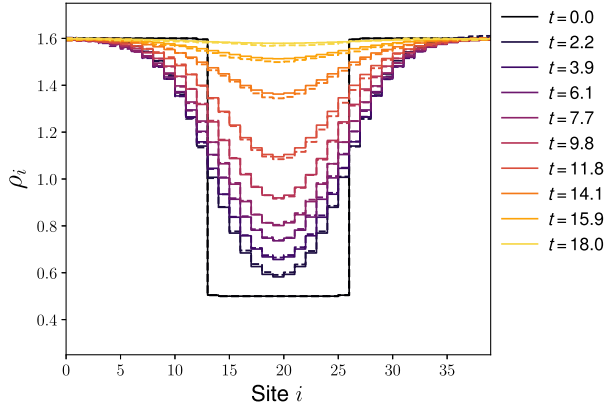


FIG. 9. Spatially extended Schlögl model: relaxation of the microscopic system with a large number of particles per site starting from a step profile at $t = 0$. The microscopic dynamics (dashed lines) closely follows the solution of Eq. (72) (solid lines) with the same initial condition and converges to the stationary fixed point $\rho = \rho^+$ within the same physical time $t \simeq 18$. Parameters: $\lambda_0 = 0.8$, $\lambda_1 = 2.9$, $\lambda_2 = 3.1$, $\lambda_3 = 1$, $\gamma = 4$, $L = 40$, and $\Omega = 16 \times 10^4$.

perform two types of experiments to test the results of the minimum-action principle.

First, we can check that MCMC simulations of the microscopic system initiated with some nonuniform profile in the compartments behave as predicted by Eq. (72). The results of these simulations are shown in Fig. 9, where we compare the evolution of a step profile: The microscopic system follows the deterministic dynamics (72), as expected. To do these calculations, we fix $\Omega = 1.6 \times 10^5$, set the number of molecules in each compartment $i \in \{1, \dots, L\}$ to be $n_i = \lfloor \rho_i \Omega \rfloor$ for all i , and simulate the microscopic dynamics exactly with the Gillespie algorithm [83,84] using the microscopic rates $k_i = \lambda_i \Omega^{1-i}$.

Second, to check the prediction of the minimum-action principle in terms of nucleation times, we can change the rates λ_i to make one of two phases (say, ρ_+) only very weakly metastable, i.e., such that $V_{\rho_+}(\rho_-) \ll 1$. We can then calculate the mean escape time $\tau_{+,-}$ from this state toward ρ_- , repeat this calculation for different values of Ω , and check that $\tau_{+,-} \simeq e^{\Omega V_{\rho_+}(\rho_-)}$. The result is shown in Fig. 10, which confirms that the prediction from the minimum-action framework explains the microscopic simulations.

D. Continuous limit

Finally, let us consider the continuous-space limit of the model by sending the number of compartments $L \rightarrow \infty$. To this end, let us set $\rho_i = \tilde{\rho}(x_i)/L$ and $\theta_i = \tilde{\theta}(x_i)$, with $x_i = i/L$ and where $\tilde{\rho}(x)$ and $\tilde{\theta}(x)$ are fields on $x \in [0, 1]$. Let us also set $\gamma = DL^2$ for some diffusion coefficient $D > 0$ and $\lambda_i = \tilde{\lambda}_i L^{i-1}$ for some rescaled rates $\tilde{\lambda}_i$. Assuming that D and $\tilde{\lambda}_i$ are $O(1)$ in L , in the limit as $L \rightarrow \infty$, it is easy to verify

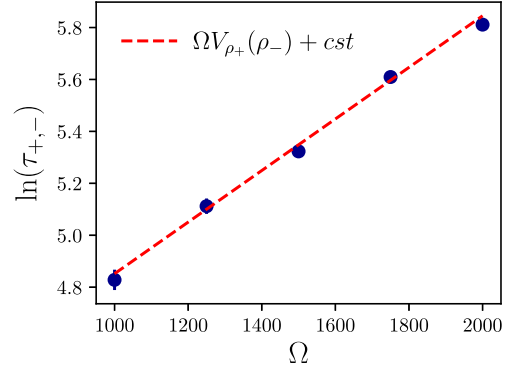


FIG. 10. Logarithm of the average nucleation time $\tau_{+,-}$ needed to reach the separatrix starting from a state metastable state ρ_+ close to linear instability, as a function of the typical number of particles Ω per site. In this regime, large deviation theory can be tested with Monte Carlo simulations in reasonable time. LDT does not predict the prefactor in front of the exponential scaling. Blue dots: results of MC simulations. Dashed line: slope predicted by theory $V_{\rho_+}(\rho_-) = 9.931 \times 10^{-4}$. Parameters: $\lambda_0 = 0.728$, $\lambda_1 = 2.9$, $\lambda_2 = 3.1$, $\lambda_3 = 1$, $\gamma = 20$, and $L = 40$.

that the Hamiltonians H^R and H^D in Eqs. (76) and (77) now become

$$\begin{aligned} \tilde{H}^R(\tilde{\rho}, \tilde{\theta}) &= \int_0^1 [\tilde{w}_+(\tilde{\rho})(e^{\tilde{\theta}} - 1) + \tilde{w}_-(\tilde{\rho})(e^{-\tilde{\theta}} - 1)] dx, \\ \tilde{H}^D(\tilde{\rho}, \tilde{\theta}) &= D \int_0^1 [\tilde{\rho} \partial_x^2 \tilde{\theta} + \tilde{\rho} (\partial_x \tilde{\theta})^2] dx, \end{aligned} \quad (79)$$

where we define $\tilde{w}_+(\tilde{\rho}) = \tilde{\lambda}_0 + \tilde{\lambda}_2 \tilde{\rho}^2$ and $\tilde{w}_-(\tilde{\rho}) = \tilde{\lambda}_1 \tilde{\rho} + \tilde{\lambda}_3 \tilde{\rho}^3$. Note that in this limit the discrete Poisson jumps of the molecules are approximated as a Gaussian noise on the density: We recover the multiplicative Gaussian noise that appears in the Dean-Kawasaki equation [48,85]. The structure of the Poisson noise of the reaction is, however, left unaffected by the limit.

We use Algorithm 2 to calculate transition pathways at continuous level by minimizing the action associated with $\tilde{H} = \tilde{H}^R + \tilde{H}^D$. The pathways (not shown) are not significantly different from those shown in Fig. 8 when $L = 40$: This indicates that the system is already close to its continuous limit at that value of L .

VI. PROSPECTS AND APPLICATIONS

The modified Ginzburg-Landau dynamics discussed in Sec. IV and the reaction-diffusion system based on the Schlögl model discussed in Sec. V are two nontrivial examples involving spatially extended systems undergoing nonequilibrium phase transitions, whose phase diagrams, paths of the transitions, and critical nuclei can be calculated by our approach. Below, we discuss other systems that can be revisited numerically with the use of our method.

First, the algorithm applies to many types of Langevin dynamics in low dimension. For instance, the algorithm is suited to compute escape rates of active particles in mechanical traps, in two dimensions or more [86]. Because of the Hamiltonian formulation, the algorithm offers a simple way to treat the common situation of hypoelliptic noise, which naturally appears in underdamped systems, or in active particle dynamics [87–89]. More generally, the Hamiltonian formulation avoids the computation of the invert of the correlation matrix when the noise is Gaussian, whether additive or multiplicative. More than that, the Hamiltonian formulation allows one to compute reaction paths in systems that are subject to noise of non-Gaussian nature, as illustrated with the Schlögl model on two or numerous sites.

Considering next infinite-dimensional systems, i.e., fields evolving according to some stochastic PDEs, the two examples we treat in Secs. IV and V show how the method can be implemented with finite-difference schemes in space and time. In the Schlögl model notably, the algorithm allows us to simultaneously treat the Poisson noise coming from the chemical reactions and the multiplicative Gaussian noise coming from the diffusion when the continuous-space limit is taken. Besides, the diffusive noise is also subject to a mass conservation constraint that the algorithm takes naturally into account. The method can be exported to compute nucleation paths in various setups, whenever the system may reach a stationary steady state, like found in motility-induced phase separation [12] or in the laning transition [90,91]. For active matter systems, in particular, described by exact coarse-grainings [81,92] or by effective field theories [12,71,93], the question of the nucleation of a phase-separated state in the binodal is the subject of strong investigations [94].

It is also worth stressing that our algorithm gives similar roles to the boundary conditions, whether set for the variable ϕ or for the conjugate variable θ , since the problem is reformulated in terms of the new variables u and v . Typically, this symmetry can be used to easily enforce a final time condition on θ , which appears when computing extreme realizations of an observable A_T , a function of the stochastic process $\{\phi_t\}_{t \in [0, T]}$ [95–98]. Typical choices for A_T can be some final time observable $F(\phi_T)$, with F some function, or one can take $A_T = \int_0^T f(\phi_t) dt$, which, for instance, computes the residence time in state ϕ_i when $f(\cdot) \equiv \delta(\cdot - \phi_i)$, or also $A_T = \int_0^T \langle g(\phi_t), d\phi_t \rangle$, which serves to compute dynamical currents or entropy production for well-chosen function g .

VII. CONCLUSION

In summary, the analysis of first-order phase transitions and other activated processes in nonequilibrium systems can be reduced to the minimization of an action in situations where these processes are rare and occur via

reproducible pathways. The approach can be justified rigorously within the framework of LDT, and minimum-action principles can also be derived formally in other instances using, e.g., the Martin-Siggia-Rose-Janssen-De Dominicis [28] or the Doi-Peliti [29] formalisms. The minimum of the action can be used to generalize Arrhenius law, and its minimizer to explain the mechanism of the transitions, including the shape of the critical nucleus that serves as transition state. Concrete predictions, however, rest on our ability to solve this minimization problem, which often needs to be done numerically.

Here, we develop algorithms to perform these calculations, in both finite and infinite times. These algorithms are designed to be used directly within the Hamiltonian formulation of the action, which leads to a min-max problem, and do not require the user to calculate the Lagrangian beforehand. In particular, it can be used for systems where the fluctuations are non-Gaussian and the Lagrangian is typically unavailable in closed form.

We hope that our work will pave the way for a systematic approach to study activated processes. These applications will depend on the possibility to derive an appropriate minimum-action principle, potentially through coarse-graining, which is a nontrivial question on its own.

ACKNOWLEDGMENTS

We thank Jasna Brujic, Giovanni Ciccotti, Tobias Grafke, Tobias Kühn, Cesare Nardini, and Frédéric van Wijland for useful comments. This work was supported by the Materials Research Science and Engineering Center (MRSEC) program of the National Science Foundation under Grant No. NSF DMR-1420073 and in part by Grant No. NSF DMR-1710163. R. Z. thanks Laboratoire MSC Paris for hospitality. R. Z. and E. V.-E. also thank the Center for Data Science ENS Paris for hospitality.

APPENDIX A: CONVERGENCE OF THE MIN-MAX FOR α SMALL BUT FINITE

We would like to show that if a path is stable with respect to the Lagrangian minimization, then the path is also stable with respect to the Hamiltonian min-max algorithm, for α small enough. We start from Eq. (39), but in this section, we conveniently rescale artificial time $\tau = \tilde{\tau}/\alpha$, set $\tilde{\alpha} = \alpha^2$, and drop the tilde such that the evolution equations read

$$\begin{aligned} \partial_\tau x &= \partial_t \theta + \partial_x H, \\ \alpha \partial_\tau \theta &= \partial_t x - \partial_\theta H. \end{aligned} \quad (\text{A1})$$

We focus on the dynamical system subjected to an additive Gaussian white noise, as presented in Eqs. (17) and (18); i.e., the Hamiltonian takes the form

$$H(x, \theta) = \langle b(x), \theta \rangle + \frac{1}{2} |\theta|^2. \quad (\text{A2})$$

We assume that a minimum-action path $\{(x^*, \theta^*)\}_{t \in [0, T]}$ is obtained. We look at a perturbed path $(x^* + X, \theta^* + \Theta)$ with $X = (x_1, \dots, x_p)^T$ and $\Theta = (\theta_1, \dots, \theta_p)^T$ the perturbations, and we assess the conditions for path relaxation to the minimum-action path. The evolution of the perturbation reads

$$\begin{aligned} \partial_\tau x_p &= \dot{\theta}_p + \theta_j^* \frac{\partial^2 b_j^*}{\partial x_k \partial x_p} x_k + \theta_j \frac{\partial b_j^*}{\partial x_p}, \\ \alpha \partial_\tau \theta_p &= \dot{x}_p - \frac{\partial b_p^*}{\partial x_k} x_k - \theta_p, \end{aligned} \quad (\text{A3})$$

where we use the Einstein convention for the sum on repeated indices, and the shorthand notations $(\partial b_p^* / \partial x_k) = (\partial b_p / \partial x_k)|_{x^*}$ and $(\partial^2 b_j^* / \partial x_k \partial x_p) = (\partial^2 b_j / \partial x_k \partial x_p)|_{x^*}$, for any indices j, k , and p .

The evolution of the fields can now be cast into the following form:

$$\begin{aligned} \partial_\tau X &= \mathcal{L}_1 X + \mathcal{L}_2 \Theta, \\ \partial_\tau \Theta &= \alpha^{-1} (\mathcal{L}_3 X - \Theta), \end{aligned} \quad (\text{A4})$$

where

$$(\mathcal{L}_1)^{(pk)} = \theta_j^* \frac{\partial^2 b_j^*}{\partial x_k \partial x_p}, \quad (\text{A5})$$

$$(\mathcal{L}_2)^{(pk)} = \left(\delta_{pk} \partial_t + \frac{\partial b_k^*}{\partial x_p} \right), \quad (\text{A6})$$

$$(\mathcal{L}_3)^{(pk)} = \left(\delta_{pk} \partial_t - \frac{\partial b_p^*}{\partial x_k} \right). \quad (\text{A7})$$

In the Lagrangian algorithm, the equation $\Theta = \mathcal{L}_3 X$ is always verified, so the evolution of the perturbation X is simply given by

$$\partial_\tau X = (\mathcal{L}_1 + \mathcal{L}_2 \mathcal{L}_3) X. \quad (\text{A8})$$

Any perturbation X vanishes if the eigenvalues of the operator $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2 \mathcal{L}_3$ are all of the negative real part. The operator \mathcal{L} is self-adjoint (since \mathcal{L} is the Hessian of the action S), and one can, thus, extract a basis of normalized orthogonal eigenvectors $X_n^\mathcal{L}$. Let us denote by μ_n the n th eigenvalue and $X_n^\mathcal{L}$ the corresponding eigenvector. We have $\mathcal{L} X_n^\mathcal{L} = \mu_n X_n^\mathcal{L}$, with $\mu_n < 0$.

Now, in the Hamiltonian algorithm, we would like to find the conditions under which any perturbation (X, Θ) close to a path of minimum action vanishes when evolving in artificial time τ . The perturbation vanishes if and only if the eigenvalues λ_n of the linear operator given in Eq. (A4) have a negative real part. The eigenvalue λ_n associated to the eigenvector $(\Theta_n, X_n)^T$ should verify

$$\begin{aligned} \mathcal{L}_1 X_n + \mathcal{L}_2 \Theta_n &= \lambda_n X_n, \\ \alpha^{-1} (\mathcal{L}_3 X_n - \Theta_n) &= \lambda_n \Theta_n. \end{aligned} \quad (\text{A9})$$

The second equation in Eqs. (A9) yields $\Theta_n = (1 + \lambda_n \alpha)^{-1} \mathcal{L}_3 X_n$, assuming that there exists $\alpha > 0$ such that $(1 + \lambda_n \alpha) \neq 0$ for every n . Injecting this result into the first equation of Eqs. (A9) yields a closed equation for X_n and λ_n :

$$(1 + \lambda_n \alpha) (\mathcal{L}_1 X_n - \lambda_n X_n) + \mathcal{L}_2 \mathcal{L}_3 X_n = 0. \quad (\text{A10})$$

For $\alpha \ll 1$, the system (A4) can be seen as a perturbation of the Lagrangian problem, where an additional degree of freedom Θ relaxes to $\mathcal{L}_3 X$ on a fast timescale $1/\alpha$. This suggests to look for eigenvalues with a specific form: (i) A first set of eigenvalues $\lambda_n^{(1)}$ should be the perturbed eigenvalues μ_n with perturbed $X_n^\mathcal{L}$ as corresponding eigenvectors; (ii) a second set of eigenvalues $\lambda_n^{(2)}$ is expected to scale as $O(\alpha^{-1})$ and encodes the fast relaxation of the variable Θ to $\mathcal{L}_3 X$.

Therefore, we look for eigenvalues of a general form:

$$\lambda_n = \frac{\eta_n}{\alpha} + \nu_n + O(\alpha) \quad (\text{A11})$$

with η_n and ν_n of $O(1)$. Expanding at leading order $O(\alpha^{-1})$ in Eq. (A10) yields

$$\eta_n (1 + \eta_n) X_n = 0, \quad (\text{A12})$$

implying that $\eta_n = 0$ or $\eta_n = -1$. The case $\eta_n = 0$ leads us to consider the first case (i) mentioned above, in which we look for eigenvalues $\lambda_n^{(1)} = \mu_n + \alpha \xi_n$, with $\xi_n = O(1)$, and we expand Eq. (A10) to leading order. At order 0 in α , we find that X_n must solve

$$(\mathcal{L}_1 + \mathcal{L}_2 \mathcal{L}_3) X_n - \mu_n X_n = 0, \quad (\text{A13})$$

where we recognize the operator $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2 \mathcal{L}_3$, which confirms that we expand around the eigenvalues $X_n^\mathcal{L}$ of the Lagrangian system. We write $X_n = X_n^\mathcal{L} + \alpha Y_n$ with $\|Y_n\| = O(1)$. Now expanding Eq. (A10) at order 1 in α and using the relation $\mathcal{L} X_n^\mathcal{L} = \mu_n X_n^\mathcal{L}$, we get

$$\mathcal{L} Y_n - \mu_n Y_n = \xi_n X_n^\mathcal{L} + \mu_n \mathcal{L}_2 \mathcal{L}_3 X_n^\mathcal{L}. \quad (\text{A14})$$

By taking the scalar product with the eigenvector $X_n^\mathcal{L}$ on both sides of Eq. (A14), we obtain

$$\xi_n = -\mu_n \langle X_n^\mathcal{L}, \mathcal{L}_2 \mathcal{L}_3 X_n^\mathcal{L} \rangle. \quad (\text{A15})$$

The brackets $\langle \cdot, \cdot \rangle$ stand for the scalar product in $L^2([0, T]; \mathbb{R}^p)$, which is given by

$$\langle \phi, \psi \rangle = \int_0^T \phi_p(t) \psi_p(t) dt. \quad (\text{A16})$$

Using the fact that $\mathcal{L}_2^* = -\mathcal{L}_3$, we have

$$\xi_n = \mu_n \|\mathcal{L}_3 X_n^\mathcal{L}\|^2, \quad (\text{A17})$$

where $\|\cdot\|$ is the norm associated with $\langle \cdot, \cdot \rangle$. Inserting this result as well as $\eta_n = 0$ in Eq. (A11), we deduce that

$$\lambda_n^{(1)} = \mu_n [1 + \alpha \|\mathcal{L}_3 X_n^\mathcal{L}\|^2 + O(\alpha^2)]. \quad (\text{A18})$$

This shows that, for α small enough, the signs of the real part of λ_n and μ_n are the same.

Now, in case (ii), where $\eta_n = -1$, at leading order 1 in α , Eq. (A10) becomes

$$-\nu_n X_n = \mathcal{L}_2 \mathcal{L}_3 X_n. \quad (\text{A19})$$

Hence, using again the fact that $\mathcal{L}_2^* = -\mathcal{L}_3$, ν_n verifies

$$\nu_n = \frac{\|\mathcal{L}_3 X_n\|^2}{\|X_n\|^2}, \quad (\text{A20})$$

where X_n is an eigenvector of $\mathcal{L}_2 \mathcal{L}_3$. Therefore, we have another set of eigenvalues given by

$$\lambda_n^{(2)} = -\alpha^{-1} + \|\mathcal{L}_3 X_n^{\mathcal{L}_2 \mathcal{L}_3}\|^2 + O(\alpha), \quad (\text{A21})$$

with $X_n^{\mathcal{L}_2 \mathcal{L}_3}$ an eigenvector of $\mathcal{L}_2 \mathcal{L}_3$. Since $\lambda_n^{(2)} < 0$ if α is small enough, the associated eigenvectors are always stable. Therefore, the stability of the fixed points of the Hamiltonian system is determined by the sign of $\lambda_n^{(1)}$, which is the same as the sign of the eigenvalues μ_n of the Lagrangian system for α small enough.

APPENDIX B: CONVERGENCE OF THE GDA IN AN ANALYTICALLY SOLVABLE CASE

To gain insight about the convergence of the GDA algorithm, we consider an Ornstein-Uhlenbeck (OU) process in one dimension for which the evolution equations of the GDA are amenable to analytic solution. The example is also relevant since it is the dynamics verified by each Fourier mode of a freely diffusive field in a one-dimensional periodic box. The stability of the numerical scheme associated with this example is analyzed in Appendix C.

For an OU process, the Hamiltonian is given by

$$H(x, \theta) = -\zeta x \theta + \frac{1}{2} \theta^2, \quad (\text{B1})$$

where $\zeta > 0$ is the stiffness of the confining harmonic potential. We look for the instanton joining x_a to $x_b \neq x_a$,

with x_a possibly nonzero (i.e., not the stable point of the deterministic dynamics). The GDA equations read

$$\begin{aligned} \partial_\tau x &= \alpha(\partial_t \theta - \zeta \theta), \\ \alpha \partial_\tau \theta &= \partial_t x + \zeta x - \theta, \end{aligned} \quad (\text{B2})$$

subject to the boundary conditions

$$x(\tau, 0) = x_a, \quad x(\tau, T) = x_b \quad (\text{B3})$$

and to the initial conditions

$$x(0, t) = x_0(t), \quad \theta(0, t) = \theta_0(t). \quad (\text{B4})$$

Again, as in the body of text, we introduce a scale $\alpha > 0$ that controls the relative evolution of x and θ in time τ . By taking the derivative with respect to τ of the first equation in Eqs. (B2) and applying the operator $\partial_t - \zeta$ to the second equation, a closed PDE for the function $x(\tau, t)$ can be obtained:

$$\partial_\tau^2 x = -\frac{1}{\alpha} \partial_\tau x + (\partial_t^2 x - \zeta^2 x). \quad (\text{B5})$$

For $\alpha > 0$, we set $\gamma(t) = x_a(1 - t/T) + x_b t/T$ and define

$$\varphi(\tau, t) \equiv x(\tau, t) - \gamma(t) \quad (\text{B6})$$

such that $\varphi(\tau, 0) = \varphi(\tau, T) = 0$. To proceed, let us use the Fourier decomposition of φ :

$$\varphi(\tau, t) = \sum_{q=1}^{\infty} \varphi_q(\tau) \sin(q\pi t/T), \quad (\text{B7})$$

where we take the convention

$$\varphi_q(\tau) = \frac{2}{T} \int_0^T \varphi(\tau, t) \sin(q\pi t/T) dt. \quad (\text{B8})$$

Inserting Eqs. (B6) and (B7) into Eq. (B5) and projecting, we arrive at the equation verified by the modes φ_q ($q \in \mathbb{N}$):

$$\alpha \partial_\tau^2 \varphi_q + \partial_\tau \varphi_q + \alpha \omega_q^2 \varphi_q = \alpha \zeta^2 \frac{2[x_a - x_b(-1)^q]}{\pi q}, \quad (\text{B9})$$

with $\omega_q^2 = \zeta^2 + (\pi q/T)^2$. We recognize the equation for the damped harmonic oscillator, implying that the modes eventually decay exponentially in τ to their stationary value. Setting

$$\Delta_q \equiv 1 - 4\alpha^2 \omega_q^2 \quad (\text{B10})$$

and defining

$$\bar{\varphi}_q = \frac{\zeta^2 [x_a - x_b (-1)^q]}{\pi q \omega_q^2}, \quad (\text{B11})$$

we have for $\Delta_q > 0$,

$$\begin{aligned} \varphi_q(\tau) &= \bar{\varphi}_q + C_{1q} e^{-\tau/(2\alpha)} e^{\sqrt{\Delta_q} \tau / (2\alpha)} \\ &\quad + C_{2q} e^{-\tau/(2\alpha)} e^{-\sqrt{\Delta_q} \tau / (2\alpha)}, \end{aligned} \quad (\text{B12})$$

for $\Delta_q < 0$,

$$\begin{aligned} \varphi_q(\tau) &= \bar{\varphi}_q + C_{1q} e^{-\tau/(2\alpha)} \cos \left[\sqrt{|\Delta_q|} \tau / (2\alpha) \right] \\ &\quad + C_{2q} e^{-\tau/(2\alpha)} \sin \left[\sqrt{|\Delta_q|} \tau / (2\alpha) \right]; \end{aligned} \quad (\text{B13})$$

and, for $\Delta_q = 0$,

$$\varphi_q(\tau) = \bar{\varphi}_q + (C_{1q} + C_{2q} \tau) e^{-\tau/(2\alpha)}. \quad (\text{B14})$$

In each case, the constants C_{1q} and C_{2q} are then determined by the initial conditions $x_0(t)$ and $\theta_0(t)$.

This computation addresses the question of the convergence rate, which is different for each mode q . For a given $q \geq 1$, the best decay rate λ_q is obtained when choosing $\alpha = 1/(2\omega_q)$, which corresponds to the damped critical regime. In practice, we need to use the same α for all modes q . The convergence to the final path is then determined by the smallest rate, which must be chosen as large as possible. Since the decay rates λ_q are ordered according to $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_q$, this prescribes the choice of $\alpha = 1/(2\omega_1)$ that maximizes the convergence rate of the mode $q = 1$, for which we have $\lambda_1 = \omega_1$. This also yields identical rates $\lambda_q = \omega_1$ for all modes q , and the modes $q > 1$ display damped oscillations.

APPENDIX C: STABILITY OF THE NUMERICAL SCHEME

In this section, we analyze the stability of Algorithm 1 for the Ornstein-Uhlenbeck system introduced in Appendix B. Starting from Eq. (B2), where x is subject to the boundary conditions (B3), we set $u = x + \alpha\theta$ and $v = x - \alpha\theta$. The GDA equations in these variables become

$$\begin{aligned} \partial_\tau u &= \partial_t u + f(u, v), \\ \partial_\tau v &= -\partial_t v + g(u, v) \end{aligned} \quad (\text{C1})$$

with

$$f(u, v) = \left(\zeta + \frac{1}{\alpha} \right) v - \frac{1}{\alpha} u, \quad (\text{C2})$$

$$g(u, v) = \left(-\zeta + \frac{1}{\alpha} \right) u - \frac{1}{\alpha} v \quad (\text{C3})$$

and with the boundary conditions

$$v(\tau, t = 0) = -u(\tau, t = 0) + 2x_a, \quad (\text{C4})$$

$$u(\tau, t = 0) = -v(\tau, t = 0) + 2x_b. \quad (\text{C5})$$

Since Eqs. (C1) are linear equations, Algorithm 1 can be expressed in terms of matrix multiplications. To this end, let us define $r = \Delta\tau/\Delta t$, the ratio between the algorithm evolution time step and the physical time step, and set $U^n = (u_0^n, \dots, u_M^n)^T$ and $V^n = (v_0^n, \dots, v_M^n)^T$, such that $(U^n, V^n)^T = (u_0^n, \dots, u_M^n, v_0^n, \dots, v_M^n)^T$. The scheme we prescribe in Algorithm 1, thus, writes in a block-matrix form, where each block $K, L, \{G_i\}_{1 \leq i \leq 4}$ is a $(M+1) \times (M+1)$ matrix and where $\mathbf{1}$ is the identity matrix: We solve $(U^{n+1}, V^{n+1})^T$ such that

$$\begin{pmatrix} K & 0 \\ 0 & \mathbf{1} \end{pmatrix} \begin{pmatrix} U^{n+1} \\ V^{n+1} \end{pmatrix} = \begin{pmatrix} G_1 & G_2 \\ 0 & \mathbf{1} \end{pmatrix} \begin{pmatrix} U^n \\ V^n \end{pmatrix} + R_1, \quad (\text{C6})$$

$$\begin{pmatrix} \mathbf{1} & 0 \\ 0 & L \end{pmatrix} \begin{pmatrix} U^{n+1} \\ V^{n+1} \end{pmatrix} = \begin{pmatrix} \mathbf{1} & 0 \\ G_3 & G_4 \end{pmatrix} \begin{pmatrix} U^{n+1} \\ V^n \end{pmatrix} + R_2, \quad (\text{C7})$$

with

$$K = \begin{pmatrix} 1+r & -r & 0 & \dots & 0 \\ 0 & 1+r & -r & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & 0 & 1+r & -r \\ 0 & & & 0 & 1 \end{pmatrix}; \quad (\text{C8})$$

$$L = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -r & 1+r & 0 & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & & -r & 1+r & 0 \\ 0 & & 0 & -r & 1+r \end{pmatrix}, \quad (\text{C9})$$

and, setting $z \equiv \Delta\tau[\zeta + (1/\alpha)]$ and $\bar{z} \equiv \Delta\tau[-\zeta + (1/\alpha)]$,

$$G_1 = \begin{pmatrix} 1 & -\frac{\Delta\tau}{\alpha} & 0 & \dots & 0 \\ 0 & 1 & -\frac{\Delta\tau}{\alpha} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & 0 & 1 & -\frac{\Delta\tau}{\alpha} \\ 0 & & 0 & 0 & 0 \end{pmatrix}; \quad (\text{C10})$$

$$G_2 = \begin{pmatrix} 0 & z & 0 & \cdots & 0 \\ 0 & 0 & z & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & 0 & 0 & z \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}; \quad (\text{C11})$$

$$G_3 = \begin{pmatrix} -1 & & & & \\ \bar{z} & 0 & & & \mathbf{0} \\ & \ddots & \ddots & & \\ \vdots & & \bar{z} & 0 & \\ 0 & 0 & \bar{z} & 0 & \end{pmatrix}; \quad (\text{C12})$$

$$G_4 = \begin{pmatrix} 0 & & & & \\ -\frac{\Delta\tau}{\alpha} & 1 & & & \mathbf{0} \\ 0 & \ddots & \ddots & & \\ \vdots & & -\frac{\Delta\tau}{\alpha} & 1 & \\ 0 & 0 & -\frac{\Delta\tau}{\alpha} & 1 & \end{pmatrix}, \quad (\text{C13})$$

and with $R_1 = (\mathbf{0}_M, 2x_b, \mathbf{0}_{M+1})^T$ and $R_2 = (\mathbf{0}_{M+1}, 2x_a, \mathbf{0}_M)^T$ and where $\mathbf{0}_M$ indicates a list with M zeros.

The scheme is stable if the eigenvalues of the matrix Q

$$Q \equiv \begin{pmatrix} \mathbf{1} & 0 \\ 0 & L^{-1} \end{pmatrix} \begin{pmatrix} \mathbf{1} & 0 \\ G_3 & G_4 \end{pmatrix} \begin{pmatrix} K^{-1} & 0 \\ 0 & \mathbf{1} \end{pmatrix} \begin{pmatrix} G_1 & G_2 \\ 0 & \mathbf{1} \end{pmatrix} \quad (\text{C14})$$

are of module ≤ 1 . To check the stability for different values of $\Delta\tau$ and Δt , we prescribe a value of ζ , which defines a typical relaxation timescale $t_\zeta = 1/\zeta$ of the noiseless dynamics. We then either prescribe the final time T and vary the number of points M along the path, or we fix the number of points and vary Δt , which changes the final time T . In particular, it is important to check the stability for a final time $T \gg t_\zeta$, such that a transition from x_a or x_b to the critical point $x = 0$ can follow the deterministic flow on most of the path (such transitions need noise only close to the critical point that must be reached exactly at time $t = T$). To correctly resolve the instanton, we should then typically take $\Delta t < t_\zeta$.

In Fig. 11, we display the stability region of the scheme in space $(\Delta t, \Delta\tau)$ for some values of the parameters. The result shows that the stability region depends on ζ and α . Since the fixed point (the instanton) is independent of the dynamical parameters $\Delta\tau$ and α , we should take the best values that bring stability and convergence. Interestingly, sending $\alpha \rightarrow 0$ is not the best choice for stability, even if it looks appealing at first sight, since it corresponds to solving

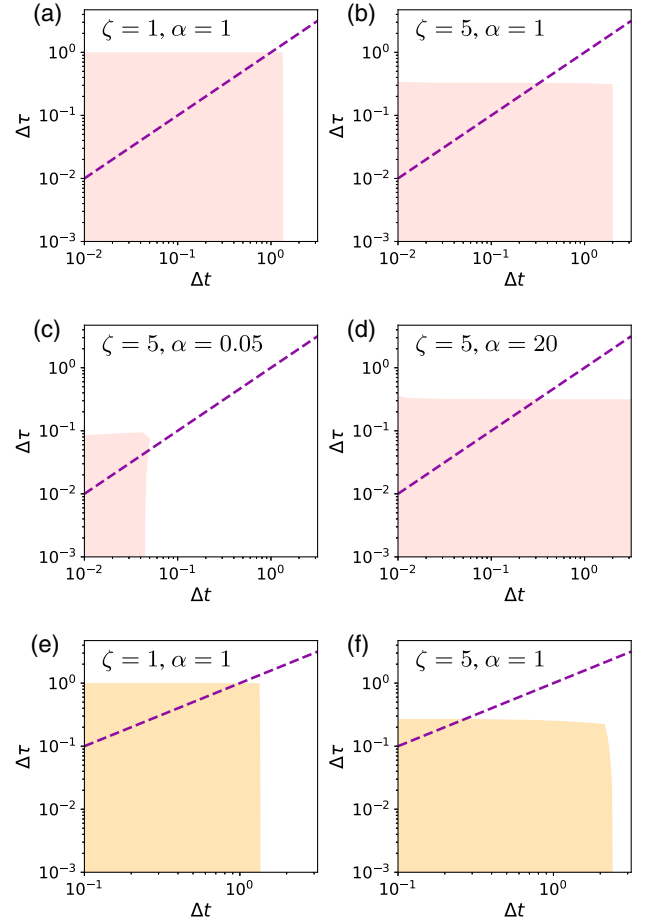


FIG. 11. Stability of the numerical scheme represented as the shaded zone in the space $(\Delta t, \Delta\tau)$. The dashed line materializes the usual CFL conditions for explicit advection scheme $\Delta\tau = \Delta t$, which is relaxed here since the scheme is semi-implicit. For (a)–(d), we fix the number of discretization points $M = 100$ and T varies according to $T = M\Delta t$. For (e) and (f), we fix $T = 10t_k$, and M is then given by $M = T/\Delta t$.

the Legendre-Fenchel transform $\operatorname{argmax}_\theta[(\dot{x}, \theta) - H(x, \theta)]$ for a given x . It is instead preferable to take $\alpha = O(1)$ to update x and θ on similar timescales. It is also worth noticing that the use of an implicit scheme for advection relaxes the Courant-Friedrichs-Lewy (CFL) stability condition; i.e., one can take $\Delta\tau > \Delta t$ and still have a stable scheme. Unfortunately, it is clear from the graphs that the limit $\Delta\tau \rightarrow \infty$ is not stable. This limit is, however, interesting, since it corresponds to an infinitely fast update of u^{n+1} at v^n fixed, i.e., that u^{n+1} solves in the continuous limit

$$0 = \partial_t u^{n+1} + f(u^n, v^n) \quad (\text{C15})$$

at v^n fixed, and, following, v^{n+1} solves

$$0 = -\partial_t v^{n+1} + g(u^{n+1}, v^{n+1}) \quad (\text{C16})$$

at u^{n+1} fixed. This greedy procedure would bring the convergence to the fixed point in a few steps, when it converges. Here, instead, $\Delta\tau$ must remain finite to guarantee convergence.

Finally, let us discuss the issue of preconditioning the scheme for large values of ζ , which is relevant, e.g., if ζ refers to the different relaxation rates of the different Fourier modes of a diffusive field [we typically have $\zeta(k) = Dk^2$ with D the diffusion coefficient of the field]. Specifically, we would like to choose time step $\Delta\tau$ for which the numerical scheme remains stable when evolving separately each Fourier component in a semispectral method for solving PDEs. Preconditioning the evolution (B2) aims at keeping the same $\Delta\tau$ for each mode and allowing higher modes [large $\zeta(k)$] to relax slower than lower modes [small $\zeta(k)$]. This procedure does not change the final path, which still solves Hamilton's equations. Such preconditioning for the GDA reads

$$\begin{aligned}\partial_\tau u &= \left(1 + \frac{1}{\alpha} + \zeta\right)^{-1} [\partial_t u + f(u, v)], \\ \partial_\tau v &= \left(1 + \frac{1}{\alpha} + \zeta\right)^{-1} [-\partial_t v + g(u, v)].\end{aligned}\quad (\text{C17})$$

The shape of the matrices K , L , and $\{G_i\}_{1 \leq i \leq 4}$ is unchanged, but $\Delta\tau$ should be modified following the substitution rule:

$$\Delta\tau \rightarrow \Delta\tau \left(1 + \frac{1}{\alpha} + \zeta\right)^{-1}, \quad (\text{C18})$$

which also modifies r , z , and \bar{z} . We find that this procedure ensures the stability of the scheme up to $\Delta\tau = 1$, for all ζ , assuming Δt is small enough.

The semianalytical proof presented here in the Ornstein-Uhlenbeck setup provides insights on the behavior of the system, but the problems we are usually interested in display strong nonlinearities, and their stability cannot be analyzed through the spectrum of a linear operator. We keep in mind, however, that a few ingredients should be reused and that they stabilize the code, in general: The advection should be treated with an implicit upwind scheme, the reaction terms can be evaluated upwind, and preconditioning the dynamics allows us to take the same time step $\Delta\tau$ for each Fourier mode in a semispectral scheme.

APPENDIX D: HIGHER-ORDER SCHEME

In Algorithms 1 and 2 that we present in the text, the derivative of u and v with respect to physical time t (or with respect to parametrization s) is approximated with a first-order finite-difference upwind derivative. This finite-difference scheme is straightforward to implement, but it is only first-order accurate; i.e., the error with respect to the

analytical solution decreases as $O(M^{-1})$ when M increases, M being the number of points used to discretize the interval $[0, T]$ (or interval $[0, 1]$ if working with reparametrized time s). Yet, it is possible to implement a higher-order stable finite-difference scheme for the advection while keeping large steps $\Delta\tau$. This can be done by keeping the implicit and upwind features of the scheme while approximating the derivatives with the second-order difference stencil. The second-order stencil involves two upwind grid points, which means that it can be used for points $i = 2, \dots, M$ in forward advection for v_i and for $i = M - 2, \dots, 0$ for backward advection for u_i . The values v_1 and u_{M-1} are still computed with a first-order finite-difference stencil. We checked that the higher-order scheme indeed significantly improves the accuracy.

The higher-order implementation of Algorithms 1 and 2 is detailed in Algorithms 3 and 4, respectively. The algorithmic complexity of the higher-order schemes is the same as the algorithmic complexity of lower-order ones. Indeed, in both cases, the implicit fields u^{n+1} and v^{n+1} simply solve a triangular system.

APPENDIX E: DETAILS ON THE CALCULATION OF λ IN THE GEOMETRIC FORMULATION

We detail here the steps that lead us to consider the geometric formulation of the GDA scheme with a specific treatment to λ . We start from Eq. (50), and we notice that the minimization of λ can be performed before the minimization over $\hat{\phi}$. This reads

$$\min_{\hat{\phi}} \min_{\lambda \geq 0} \max_{\hat{\theta}} \int_0^1 [\langle \hat{\phi}', \hat{\theta} \rangle - \lambda^{-1} H(\hat{\phi}, \hat{\theta})] ds. \quad (\text{E1})$$

The max on $\hat{\theta}$ can be performed pointwise in s , and this gives the following relation that the minimum-action path should satisfy:

$$\lambda \hat{\phi}' = \partial_{\hat{\theta}} H, \quad (\text{E2})$$

which implicitly defines $\hat{\theta}(s) \equiv \vartheta[\hat{\phi}(s), \hat{\phi}'(s), \lambda(s)]$. Since $\theta(s)$ can be understood as a function of $\lambda(s)$, the minimization in λ of Eq. (E1) gives the equation

$$\langle \hat{\phi}', \partial_\lambda \vartheta \rangle + \lambda^{-2} H - \lambda^{-1} \langle \partial_{\hat{\theta}} H, \partial_\lambda \vartheta \rangle = 0, \quad (\text{E3})$$

which, using Eq. (E2), simplifies into $H(\hat{\phi}, \vartheta(\hat{\phi}, \hat{\phi}', \lambda)) = 0$. This last equation, however, cannot be inverted, in general, to obtain $\lambda(s)$ explicitly. Instead, as an update rule for $\lambda(s)$, we start back from Eq. (E3), and we replace $\partial_\lambda \vartheta$ by its value at convergence. Indeed, by taking the derivative of Eq. (E2) with respect to λ , we get $\hat{\phi}' = (\partial_{\hat{\theta}}^2 H) \partial_\lambda \vartheta$, or $\partial_\lambda \vartheta = (\partial_{\hat{\theta}}^2 H)^{-1} \hat{\phi}'$, since $C^{-1} \equiv (\partial_{\hat{\theta}}^2 H)$ is invertible under assumption A3 that H strictly convex in θ . Inserting this

Algorithm 3. Action minimization by higher-order gradient descent-ascent.

- 1: Follow steps 1 and 2 in Algorithm 1.
- 2: **for** $n \geq 0$ **do**
- 3: Update u with an implicit upwind scheme, namely, solve $\{u_i^{n+1}\}_{i \in I}$ sequentially from $i = M$ to $i = 0$ using

$$\begin{aligned} u_M^{n+1} &= -v_M^n + 2\phi_b, \\ \frac{u_{M-1}^{n+1} - u_{M-1}^n}{\Delta\tau} &= \frac{u_M^{n+1} - u_{M-1}^{n+1}}{\Delta t} + f(u_M^n, v_M^n), \\ \frac{u_i^{n+1} - u_i^n}{\Delta\tau} &= \frac{-u_{i+2}^{n+1} + 4u_{i+1}^{n+1} - 3u_i^{n+1}}{2\Delta t} + f(u_{i+1}^n, v_{i+1}^n), \quad i = M-2, \dots, 0. \end{aligned}$$

- 4: Update v with an implicit upwind scheme, namely, solve $\{v_i^{n+1}\}_{i \in I}$ sequentially from $i = 0$ to $i = M$ using

$$\begin{aligned} v_0^{n+1} &= -u_0^{n+1} + 2\phi_a, \\ \frac{v_1^{n+1} - v_1^n}{\Delta\tau} &= -\frac{v_1^{n+1} - v_0^{n+1}}{\Delta t} + g(u_0^{n+1}, v_0^n), \\ \frac{v_i^{n+1} - v_i^n}{\Delta\tau} &= -\frac{3v_i^{n+1} - 4v_{i-1}^{n+1} + v_{i-2}^{n+1}}{2\Delta t} + g(u_{i-1}^{n+1}, v_{i-1}^n), \quad i = 2, \dots, M. \end{aligned}$$

- 5: Compute $\{\phi_i^{n+1} = \frac{1}{2}(u_i^{n+1} + v_i^{n+1})\}_{i \in I}$ and $\{\theta_i^{n+1} = \frac{1}{2}\alpha^{-1}(u_i^{n+1} - v_i^{n+1})\}_{i \in I}$ (if needed).
-
-

expression for $\partial_\lambda \vartheta$ in Eq. (E3), and extracting the root, we get

$$\lambda = \frac{\langle \partial_\theta H, C\hat{\phi}' \rangle + \sqrt{\langle \partial_\theta H, C\hat{\phi}' \rangle^2 - 4H\langle \hat{\phi}', C\hat{\phi}' \rangle}}{2\langle \hat{\phi}', C\hat{\phi}' \rangle}. \quad (\text{E4})$$

Equation (52) follows from this equation if we modify a few terms to guarantee that $\lambda \geq 0$ [which may not always

be satisfied, since Eq. (E2) holds only at convergence and not during the optimization]. Note that Eqs. (E2) and (E4) impose $H = 0$ along the trajectory, for any definite positive C . This enjoins us to consider replacing C by the identity in our numerical algorithms in order to avoid computing $(\partial_\theta^2 H)$ and its inverse. Note also that the value $H = 0$ is the only possible value, since $H(0) = H(1) = 0$ (end points are critical points) and that $H(s) = H(0)$ for every $s \in [0, 1]$ (Hamiltonian system). In this sense, the

Algorithm 4. Geometric action minimization by higher-order gradient descent-ascent.

- 1: Follow steps 1 and 2 in Algorithm 2.
- 2: **for** $n \geq 0$ **do**
- 3: Update u with an implicit upwind scheme, namely, solve $\{u_i^{n+1}\}_{i \in I}$ sequentially from $i = M$ to $i = 0$ using

$$\begin{aligned} u_M^{n+1} &= -v_M^n + 2\phi_b, \\ \frac{u_{M-1}^{n+1} - u_{M-1}^n}{\Delta\tau} &= \lambda_{M-1}(u^n, v^n) \frac{u_M^{n+1} - u_{M-1}^{n+1}}{\Delta s} + f(u_{M-1}^n, v_{M-1}^n), \\ \frac{u_i^{n+1} - u_i^n}{\Delta\tau} &= \lambda_i(u^n, v^n) \frac{-u_{i+2}^{n+1} + 4u_{i+1}^{n+1} - 3u_i^{n+1}}{2\Delta s} + f(u_i^n, v_i^n), \quad i = M-2, \dots, 0. \end{aligned}$$

- 4: Update v with an implicit upwind scheme, namely, solve $\{v_i^{n+1}\}_{i \in I}$ sequentially from $i = 0$ to $i = M$ using

$$\begin{aligned} v_0^{n+1} &= -u_0^{n+1} + 2\phi_a, \\ \frac{v_1^{n+1} - v_1^n}{\Delta\tau} &= -\lambda_1(u^{n+1}, v^n) \frac{v_1^{n+1} - v_0^{n+1}}{\Delta s} + g(u_1^{n+1}, v_1^n), \\ \frac{v_i^{n+1} - v_i^n}{\Delta\tau} &= -\lambda_i(u^{n+1}, v^n) \frac{3v_i^{n+1} - 4v_{i-1}^{n+1} + v_{i-2}^{n+1}}{2\Delta s} + g(u_i^{n+1}, v_i^n), \quad i = 2, \dots, M. \end{aligned}$$

- 5: Follow steps 5–7 in Algorithm 2.
-
-

coefficient λ^{-1} can also be seen as a Lagrange multiplier enforcing $H = 0$. Finally, the min on $\hat{\phi}$ brings the second Hamilton equation

$$\lambda \hat{\theta}' = -\partial_{\hat{\phi}} H. \quad (\text{E5})$$

APPENDIX F: MODIFIED GINZBURG-LANDAU DYNAMICS DISCRETIZED ON TWO SITES

To work in low dimensions, we discretize Eq. (64) on two sites with periodic boundary conditions. Defining $x(t)$ and $y(t)$ the values of the field on site 1 and site 2, respectively, the system is now equivalent to studying a particle at position $(x, y)^T$ subjected to a nonequilibrium force. The dynamics read

$$\begin{aligned} \dot{x} &= b_x(x, y) + \sqrt{2\epsilon}\eta_x, \\ \dot{y} &= b_y(x, y) + \sqrt{2\epsilon}\eta_y, \end{aligned} \quad (\text{F1})$$

where, following Eq. (64), we set

$$b_x(x, y) = 8D(y - x) + x - x^3 + \frac{\kappa}{2}(x^2 + y^2), \quad (\text{F2})$$

$$b_y(x, y) = 8D(x - y) + y - y^3 + \frac{\kappa}{2}(x^2 + y^2), \quad (\text{F3})$$

and where η_x and η_y are Gaussian white noise of variance unity. The term $\kappa(x^2 + y^2)/2$ is the two-site analog of $\kappa \int \rho^2$. For simplicity, we take $h = 0$, such that the noiseless system has always two stable fixed points in $(x_1, y_1) = (q_1, q_1)$ and $(x_2, y_2) = (q_2, q_2)$, with $q_1 = (-\kappa - \sqrt{\kappa^2 + 4})/2$ and $q_2 = (-\kappa + \sqrt{\kappa^2 + 4})/2$, and one unstable fixed point at $(x_u, y_u) = (0, 0)$. The system displays time-reversal

symmetry for $\kappa = 0$, and we, therefore, know that the dynamics follows a stochastic gradient descent on the landscape $V(x, y) = -x^2/2 + x^4/4 - y^2/2 + y^4/4 + 4D(x - y)^2$ and that the minimum-action path in this case corresponds to the minimum energy path, which simply follows the gradient ascent to escape the basin of attraction of one or the other stable fixed point. For $\kappa \neq 0$, however, we need to resort to numerical algorithms to find the new minimum-action paths between the stable fixed points. In this diffusive system, as presented in Sec. II A, the Hamiltonian is explicitly given by

$$H = b_x(x, y)\theta_x + \theta_x^2 + b_y(x, y)\theta_y + \theta_y^2, \quad (\text{F4})$$

where θ_x and θ_y are the conjugated fields of x and y , respectively. The solution to Hamilton's equation of motion is found via iteration of the gradient ascent-descent procedure introduced above. The results can be found in Fig. 12. We checked that the scheme perfectly recovers the prediction of previous methods [36,37,39]. As expected, the action decreases when the final time T increases; see Fig. 12(b). Also, the Hamiltonian H is conserved along the path, since the final path solves Hamilton's equation of motion. In addition, we notice that $H \rightarrow 0$ when T increases.

Now, we know that the minimal action is reached when we allow $T \rightarrow \infty$; see Sec. III B. However, since the number of points along the trajectory must remain finite, increasing T translates into an increase of the physical time step Δt and decreases the resolution of the path. To overcome this issue, we use the geometric parametrization of the path presented in Sec. III B. In Fig. 13, we see that the geometric parametrization allows us to reach the minimum of the action. We also check that $H = 0$ along the trajectory.

This example illustrates the approach on low-dimensional systems subjected to an additive Gaussian

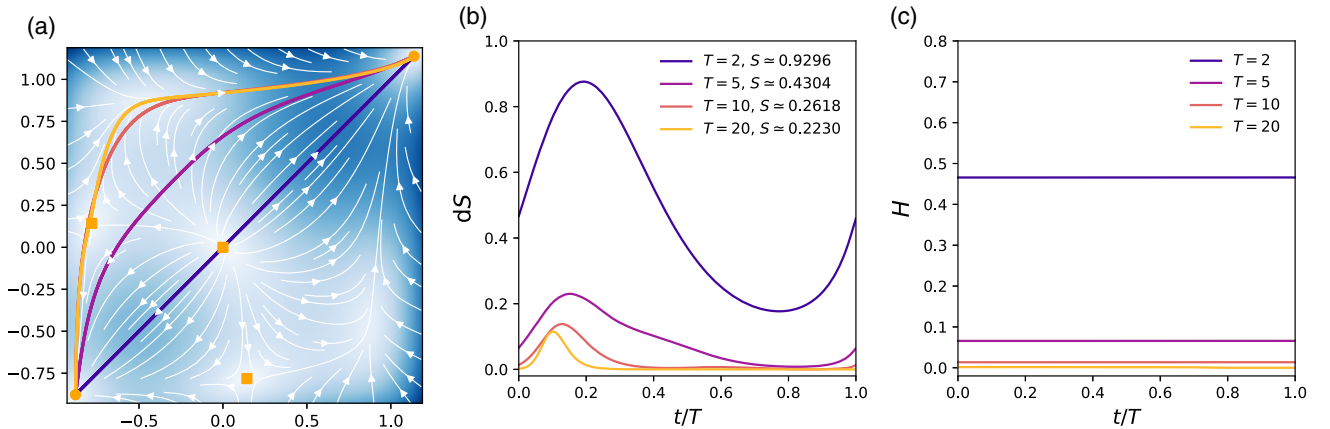


FIG. 12. Comparison between paths in the low-dimensional modified Ginzburg-Landau system for different values of the final time $T = 2, 5, 10$, and 20 . Minimum-action paths (left), Lagrangian (middle), and Hamiltonian (right) along the paths. Orange squares (left): unstable fixed points. Orange disks: stable fixed points. Background color: intensity of the force field; darker encodes stronger force. As T increases, the action decreases. For large T , the Lagrangian plateaus in the vicinity of critical points. We notice that the final Hamiltonian is constant along the trajectory. Parameters: $N = 400$, $D = 0.03$, $\kappa = 0.26$, and $\Delta\tau = 0.002$.

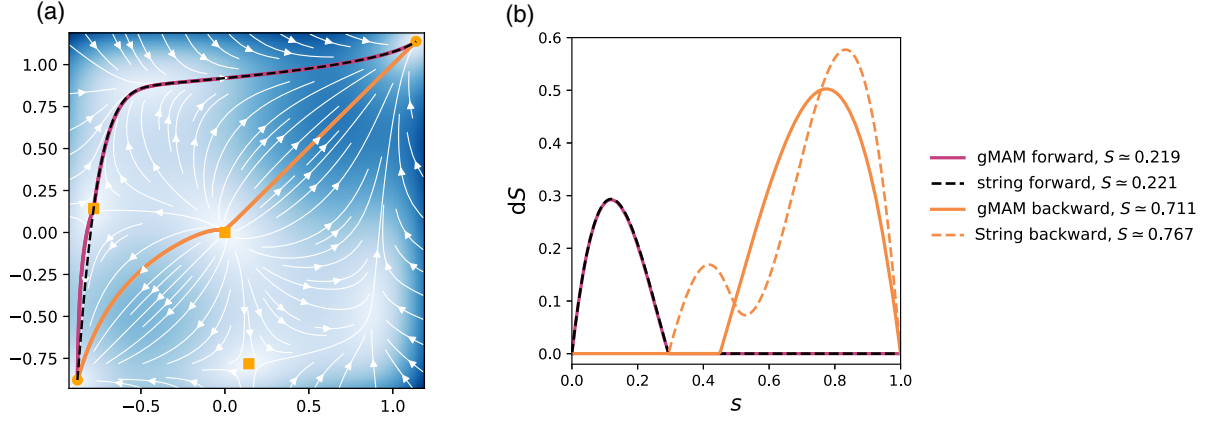


FIG. 13. Comparison between paths in the low-dimensional modified Ginzburg-Landau system. Left: flow lines and paths. Right: action increment along the paths parametrized by normalized arclength. (a) Dashed line: heteroclinic orbit obtained with the string method. Purple line: forward minimizer obtained with the geometric algorithm. Orange line: backward minimizer obtained with the geometric algorithm. Squares: unstable fixed points. Disks: stable fixed points. Background color: intensity of the force field; darker encodes stronger force. Parameters: $N = 600$, $D = 0.03$, $\kappa = 0.26$, and $\Delta\tau = 0.01$.

noise. The method yields the optimal paths of finite duration but also the paths of infinite time length if the geometric formulation is used. Note that the numerical scheme can be adapted to determine transition paths for spatially extended fields: This is what we do in Secs. IV and V.

APPENDIX G: LARGE Ω EXPANSION AND PROBABILITY DENSITY

In this section, we focus on the case of a unique well-stirred compartment of fixed volume. We work in the limit of a large number of particles, and we choose the microscopic rates k_i such that the mean-field equation always displays two stable fixed points n_- and n_+ that solve $0 = k_0 - k_1 n + k_2 n(n-1) - k_3 n(n-1)(n-2)$ and where n is the number of particles in the compartment. When all the particles coexist in the same compartment, the evolution of the probability $P(n)$ to find a number n of particles X at time t is given by

$$\partial_t P(n) = W_+(n-1)P(n-1) + W_-(n+1)P(n+1) - [W_+(n) + W_-(n)]P(n), \quad (\text{G1})$$

where the rates are

$$W_+(n) = k_0 + k_2 n(n-1), \quad (\text{G2})$$

$$W_-(n) = k_1 n + k_3 n(n-1)(n-2). \quad (\text{G3})$$

The stationary probability P_{eq} can be obtained explicitly:

$$P_{\text{eq}}(n) = K \prod_{i=1}^n \frac{W_+(i-1)}{W_-(i)} = K \exp\left(\sum_{i=1}^n \ln \frac{W_+(i-1)}{W_-(i)}\right), \quad (\text{G4})$$

where K is a normalization constant. We are interested in the case of a large number of particles per compartment, such that we can extract a large deviation principle. We denote $\Omega \gg 1$ the typical number of particles in the compartment, we define the rescaled number of particles $\rho = n/\Omega$, and we write the rates as

$$W_+(n) = \Omega[w_+(\rho) + O(\Omega^{-1})], \quad (\text{G5})$$

$$W_-(n) = \Omega[w_-(\rho) + O(\Omega^{-1})], \quad (\text{G6})$$

with $w^+(\rho) = \lambda_0 + \lambda_2 \rho^2$ and $w^-(\rho) = \lambda_1 \rho + \lambda_3 \rho^3$ and where the λ_i are now rescaled reaction rates verifying $\lambda_i = k_i \Omega^{i-1}$. This rescaling ensures that the deterministic mean-field dynamics

$$\partial_t \rho = \lambda_0 + \lambda_2 \rho^2 - \lambda_1 \rho - \lambda_3 \rho^3 \quad (\text{G7})$$

keeps the same fixed points ρ_- , ρ_s , and ρ_+ when $\Omega \rightarrow \infty$, with ρ_- and ρ_+ stable, while ρ_s is unstable. Following Refs. [43,44,97,99], using the WKB (or eikonal) approximation and the continuum limit, the probability now becomes a probability density and can be cast into the following form:

$$P_{\text{eq}}(\rho) = K(\rho, \Omega) e^{-\Omega V_{\text{eq}}(\rho)}, \quad (\text{G8})$$

where the equilibrium potential is given by

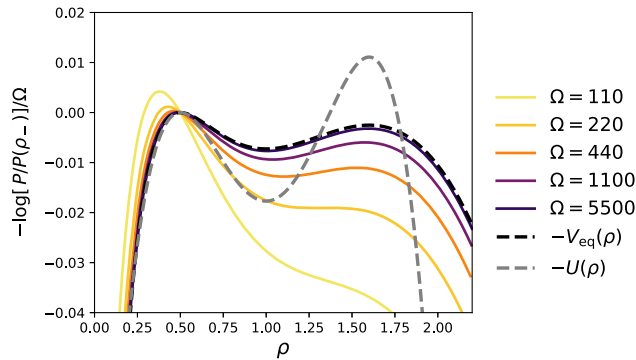


FIG. 14. Rescaled and shifted logarithms of the probability density on a unique site in the Schlögl model when the local number of particle Ω increases. The probability is explicitly known for this equilibrium case, and we show the convergence to the function $V_{\text{eq}}(\rho)$ as the number of particles in the box increases. For the parameters chosen here, ρ_- is the stable phase, while looking at the naive Ginzburg-Landau potential $U(\rho) = -\lambda_0\rho - \lambda_2\rho^3/3 + \lambda_1\rho^2/2 + \lambda_3\rho^4/4$ wrongly predicts that ρ_+ is the stable phase. Parameters: $\lambda_0 = 0.8$, $\lambda_1 = 2.9$, $\lambda_2 = 3.1$, and $\lambda_3 = 1$.

$$V_{\text{eq}}(\rho) = \int^{\rho} dy \ln \frac{w_+(y)}{w_-(y)} \quad (\text{G9})$$

and where $K(\rho, \Omega)$ is a function with the property that, for any ρ_1 and ρ_2 , the fraction $K(\rho_1, \Omega)/K(\rho_2, \Omega)$ is bounded. As such, the ratio of the probabilities $P(\rho_1)/P(\rho_2)$ is completely determined by the difference $\Delta V \equiv V_{\text{eq}}(\rho_1) - V_{\text{eq}}(\rho_2)$, in the large Ω limit. We show in Fig. 14 that the potential V_{eq} and the naive Ginzburg-Landau potential $U(\rho) = -\lambda_0\rho - \lambda_2\rho^3/3 + \lambda_1\rho^2/2 + \lambda_3\rho^4/4$ display different maxima and, thus, different predictions for the relative stability of the states. In the large Ω limit, we check that the probability converges as expected to the probability density given by the function $V_{\text{eq}}(\rho)$. We also notice that many particles per box may be needed to observe the convergence to the large deviation function.

In summary, for a well-stirred and unique compartment, the analytical expression of the large deviation function can be obtained explicitly. This is no longer the case when spatial diffusion is taken into account, and this is why one has to resort to other techniques to access the quasipotential.

[1] P. Weiss, *L'Hypothèse du Champ Moléculaire et la Propriété Ferromagnétique*, *J. Phys. Theor. Appl.* **6**, 661 (1907).
 [2] A. Einstein, *Quantum Theory of a Monoatomic Ideal Gas, a translation of Quantentheorie des Einatomigen Idealen Gases (1924)* (Princeton University, Princeton, NJ, 1925).

[3] L. D. Landau and E. M. Lifshitz, in *Course of Theoretical Physics*, third ed. (Pergamon, New York, 1980), Chap. XIV, pp. 446–516.
 [4] B. Derrida, E. Domany, and D. Mukamel, *An Exact Solution of a One-Dimensional Asymmetric Exclusion Model with Open Boundaries*, *J. Stat. Phys.* **69**, 667 (1992).
 [5] B. Derrida, M. R. Evans, V. Hakim, and V. Pasquier, *Exact Solution of a 1D Asymmetric Exclusion Model Using a Matrix Formulation*, *J. Phys. A* **26**, 1493 (1993).
 [6] J. Tailleur, J. Kurchan, and V. Lecomte, *Mapping Non-equilibrium onto Equilibrium: The Macroscopic Fluctuations of Simple Transport Models*, *Phys. Rev. Lett.* **99**, 150602 (2007).
 [7] J. Tailleur, J. Kurchan, and V. Lecomte, *Mapping Out-of-Equilibrium into Equilibrium in One-Dimensional Transport Models*, *J. Phys. A* **41**, 505001 (2008).
 [8] J. Tailleur and M. E. Cates, *Statistical Mechanics of Interacting Run-and-Tumble Bacteria*, *Phys. Rev. Lett.* **100**, 218103 (2008).
 [9] J. O'Byrne and J. Tailleur, *Lamellar to Micellar Phases and beyond: When Tactic Active Systems Admit Free Energy Functionals*, *Phys. Rev. Lett.* **125**, 208003 (2020).
 [10] A. P. Solon and J. Tailleur, *Revisiting the Flocking Transition Using Active Spins*, *Phys. Rev. Lett.* **111**, 078101 (2013).
 [11] I. Buttinoni, J. Bialké, F. Kümmel, H. Löwen, C. Bechinger, and T. Speck, *Dynamical Clustering and Phase Separation in Suspensions of Self-Propelled Colloidal Particles*, *Phys. Rev. Lett.* **110**, 238301 (2013).
 [12] M. E. Cates and J. Tailleur, *Motility-Induced Phase Separation*, *Annu. Rev. Condens. Matter Phys.* **6**, 219 (2015).
 [13] T. Grafke, M. E. Cates, and E. Vanden-Eijnden, *Spatiotemporal Self-Organization of Fluctuating Bacterial Colonies*, *Phys. Rev. Lett.* **119**, 188003 (2017).
 [14] D. Geyer, D. Martin, J. Tailleur, and D. Bartolo, *Freezing a Flock: Motility-Induced Phase Separation in Polar Active Liquids*, *Phys. Rev. X* **9**, 031043 (2019).
 [15] D. Martin, H. Chaté, C. Nardini, A. Solon, J. Tailleur, and F. Van Wijland, *Fluctuation-Induced Phase Separation in Metric and Topological Models of Collective Motion*, *Phys. Rev. Lett.* **126**, 148001 (2021).
 [16] T. Grafke, R. Grauer, and T. Schäfer, *Instanton Filtering for the Stochastic Burgers Equation*, *J. Phys. A* **46**, 062002 (2013).
 [17] T. Grafke, R. Grauer, and T. Schäfer, *The Instanton Method and Its Numerical Implementation in Fluid Mechanics*, *J. Phys. A* **48**, 333001 (2015).
 [18] G. Liu, A. Patch, F. Bahar, D. Yllanes, R. D. Welch, M. C. Marchetti, S. Thutupalli, and J. W. Shaevitz, *Self-Driven Phase Transitions Drive Myxococcus xanthus Fruiting Body Formation*, *Phys. Rev. Lett.* **122**, 248102 (2019).
 [19] D. Amit, *Model of Global Spontaneous Activity and Local Structured Activity during Delay Periods in the Cerebral Cortex*, *Cereb. Cortex* **7**, 237 (1997).
 [20] L. Mazzucato, A. Fontanini, and G. La Camera, *Dynamics of Multistable States during Ongoing and Evoked Cortical Activity*, *J. Neurosci.* **35**, 8214 (2015).
 [21] D. Jercog, A. Roxin, P. Barthó, A. Luczak, A. Compte, and J. de la Rocha, *Up-Down Cortical Dynamics Reflect State Transitions in a Bistable Network*, *eLife* **6**, e22425 (2017).

- [22] F. Ragone, J. Wouters, and F. Bouchet, *Computation of Extreme Heat Waves in Climate Models Using a Large Deviation Algorithm*, *Proc. Natl. Acad. Sci. U.S.A.* **115**, 24 (2018).
- [23] E. Simonnet, J. Rolland, and F. Bouchet, *Multistability and Rare Spontaneous Transitions in Barotropic β -Plane Turbulence*, *J. Atmos. Sci.* **78**, 1889 (2021).
- [24] K. G. Wilson, *The Renormalization Group: Critical Phenomena and the Kondo Problem*, *Rev. Mod. Phys.* **47**, 773 (1975).
- [25] S. F. Edwards and P. W. Anderson, *Theory of Spin Glasses*, *J. Phys. F* **5**, 965 (1975).
- [26] M. Mézard, G. Parisi, and M. A. Virasoro, *Spin Glass Theory and beyond: An Introduction to the Replica Method and Its Applications* (World Scientific, Singapore, 1987), Vol. 9.
- [27] R. S. Ellis, *Entropy, Large Deviations, and Statistical Mechanics* (Taylor & Francis, London, 2006), Vol. 1431.
- [28] P. C. Martin, E. D. Siggia, and H. A. Rose, *Statistical Dynamics of Classical Systems*, *Phys. Rev. A* **8**, 423 (1973); H.-K. Janssen, *On a Lagrangean for Classical Field Dynamics and Renormalization Group Calculations of Dynamical Critical Properties*, *Z. Phys. B* **23**, 377 (1976); C. De Dominicis, *Techniques de Renormalisation de la Théorie des Champs et Dynamique des Phénomènes Critiques*, *J. Phys. (Paris)*, Colloq. **37**, C1-247 (1976).
- [29] M. Doi, *Second Quantization Representation for Classical Many-Particle System*, *J. Phys. A* **9**, 1465 (1976); L. Peliti, *Path Integral Approach to Birth-Death Processes on a Lattice*, *J. Phys. (Paris)* **46**, 1469 (1985).
- [30] E. Vanden-Eijnden and M. Heymann, *The Geometric Minimum Action Method for Computing Minimum Energy Paths*, *J. Chem. Phys.* **128**, 061103 (2008).
- [31] H. B. Keller, *Numerical Methods for Two-Point Boundary-Value Problems* (Courier Dover, New York, 2018).
- [32] R. S. Maier and D. L. Stein, *A Scaling Theory of Bifurcations in the Symmetric Weak-Noise Escape Problem*, *J. Stat. Phys.* **83**, 291 (1996).
- [33] W. E, W. Ren, and E. Vanden-Eijnden, *Minimum Action Method for the Study of Rare Events*, *Commun. Pure Appl. Math.* **57**, 637 (2004).
- [34] T. Grafke, T. Schäfer, and E. Vanden-Eijnden, *Long Term Effects of Small Random Perturbations on Dynamical Systems: Theoretical and Computational Tools*, in *Recent Progress and Modern Challenges in Applied Mathematics, Modeling and Computational Science* (Springer, New York, 2017), pp. 17–55.
- [35] M. Heymann and E. Vanden-Eijnden, *Pathways of Maximum Likelihood for Rare Events in Nonequilibrium Systems: Application to Nucleation in the Presence of Shear*, *Phys. Rev. Lett.* **100**, 140601 (2008).
- [36] L. Kikuchi, R. Singh, M. E. Cates, and R. Adhikari, *Ritz Method for Transition Paths and Quasipotentials of Rare Diffusive Events*, *Phys. Rev. Res.* **2**, 033208 (2020).
- [37] M. I. Freidlin and A. D. Wentzell, *Random Perturbations of Dynamical Systems*, Grundlehren der mathematischen Wissenschaften Vol. 260 (Springer, New York, 1998).
- [38] R. Graham and T. Tél, *Nonequilibrium Potential for Coexisting Attractors*, *Phys. Rev. A* **33**, 1322 (1986).
- [39] L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, *Macroscopic Fluctuation Theory*, *Rev. Mod. Phys.* **87**, 593 (2015).
- [40] F. Bouchet, K. Gawędzki, and C. Nardini, *Perturbative Calculation of Quasi-Potential in Non-equilibrium Diffusions: A Mean-Field Example*, *J. Stat. Phys.* **163**, 1157 (2016).
- [41] M. Cameron, *Finding the Quasipotential for Nongradient SDEs*, *Physica (Amsterdam)* **241D**, 1532 (2012); T. Gan and M. Cameron, *A Graph-Algorithmic Approach for the Study of Metastability in Markov Chains*, *J. Nonlinear Sci.* **27**, 927 (2017); D. Dahiya and M. Cameron, *Ordered Line Integral Methods for Computing the Quasi-Potential*, *J. Sci. Comput.* **75**, 1351 (2018).
- [42] F. Schlögl, *Chemical Reaction Models for Non-equilibrium Phase Transitions*, *Z. Phys.* **253**, 147 (1972).
- [43] 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).
- [44] S. Tănase-Nicola and D. K. Lubensky, *Exchange of Stability as a Function of System Size in a Nonequilibrium System*, *Phys. Rev. E* **86**, 040103(R) (2012).
- [45] M. Assaf and B. Meerson, *WKB Theory of Large Deviations in Stochastic Populations*, *J. Phys. A* **50**, 263001 (2017).
- [46] J. L. Cardy and P. Grassberger, *Epidemic Models and Percolation*, *J. Phys. A* **18**, L267 (1985).
- [47] F. van Wijland, K. Oerding, and H. Hilhorst, *Wilson Renormalization of a Reaction-Diffusion Process*, *Physica (Amsterdam)* **251A**, 179 (1998).
- [48] A. Lefèvre and G. Biroli, *Dynamics of Interacting Particle Systems: Stochastic Process and Field Theory*, *J. Stat. Mech.* (2007) P07024.
- [49] D. A. Dawson and J. Gärtner, *Large Deviations from the McKean-Vlasov Limit for Weakly Interacting Diffusions*, *Stochastics* **20**, 247 (1987).
- [50] M. I. Freidlin, *The Averaging Principle and Theorems on Large Deviations*, *Russ. Math. Surv.* **33**, 117 (1978).
- [51] A. Y. Veretennikov, *On the Averaging Principle for Systems of Stochastic Differential Equations*, *Math. USSR-Sbornik* **69**, 271 (1991).
- [52] Y. Kifer, *Averaging Principle for Fully Coupled Dynamical Systems and Large Deviations*, *Ergod. Theory Dyn. Syst.* **24**, 847 (2004).
- [53] F. Bouchet, T. Grafke, T. Tangarife, and E. Vanden-Eijnden, *Large Deviations in Fast-Slow Systems*, *J. Stat. Phys.* **162**, 793 (2016).
- [54] H. Spohn, *Large Scale Dynamics of Interacting Particles* (Springer, Berlin, 1991).
- [55] C. Kipnis and C. Landim, *Scaling Limits of Interacting Particle Systems*, Grundlehren der mathematischen Wissenschaften Vol. 320 (Springer, Berlin, 1999).
- [56] C. Kipnis, S. Olla, and S. R. S. Varadhan, *Hydrodynamics and Large Deviation for Simple Exclusion Processes*, *Commun. Pure Appl. Math.* **42**, 115 (1989).
- [57] T. Kose, *Solutions of Saddle Value Problems by Differential Equations*, *Econometrica* **24**, 59 (1956).
- [58] A. Cherukuri, B. Ghahesifard, and J. Cortés, *Saddle-Point Dynamics: Conditions for Asymptotic Stability of Saddle Points*, *SIAM J. Control Optim.* **55**, 486 (2017).

- [59] T. Lin, C. Jin, and M. Jordan, *On Gradient Descent Ascent for Nonconvex-Concave Minimax Problems*, Proc. Mach. Learn. Res. **119**, 6083 (2020).
- [60] J. C. Strikwerda, *I. Hyperbolic Partial Differential Equations*, in *Finite Difference Schemes and Partial Differential Equations, Second Edition*, Other Titles in Applied Mathematics (Society for Industrial and Applied Mathematics, Philadelphia, 2004), pp. 1–36.
- [61] S. MacNamara and G. Strang, *Operator Splitting, in Splitting Methods in Communication, Imaging, Science, and Engineering*, edited by R. Glowinski, S. J. Osher, and W. Yin (Springer International, Cham, 2016), pp. 95–114.
- [62] <https://github.com/rzakine/MAM-gradient-descent-ascent>
- [63] L. D. Landau, *On the Theory of Phase Transitions. I.*, Zh. Eksp. Teor. Fiz. **11**, 19 (1937).
- [64] V. L. Ginzburg and L. D. Landau, *On the Theory of Superconductivity*, in *On Superconductivity and Superfluidity: A Scientific Autobiography* (Springer, Berlin, 2009), pp. 113–137.
- [65] P. G. De Gennes, *Short Range Order Effects in the Isotropic Phase of Nematics and Cholesterics*, Mol. Cryst. Liq. Cryst. **12**, 193 (1971).
- [66] P. C. Hohenberg and B. I. Halperin, *Theory of Dynamic Critical Phenomena*, Rev. Mod. Phys. **49**, 435 (1977).
- [67] P. M. Chaikin, T. C. Lubensky, and T. A. Witten, *Principles of Condensed Matter Physics* (Cambridge University Press, Cambridge, England, 1995), Vol. 10.
- [68] L. Bertini, A. D. Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim, *Non Equilibrium Current Fluctuations in Stochastic Lattice Gases*, J. Stat. Phys. **123**, 237 (2006).
- [69] Y. Baek, Y. Kafri, and V. Lecomte, *Dynamical Phase Transitions in the Current Distribution of Driven Diffusive Channels*, J. Phys. A **51**, 105001 (2018).
- [70] C. Nardini, E. Fodor, E. Tjhung, F. van Wijland, J. Tailleur, and M. E. Cates, *Entropy Production in Field Theories without Time-Reversal Symmetry: Quantifying the Non-equilibrium Character of Active Matter*, Phys. Rev. X **7**, 021007 (2017).
- [71] M. E. Cates, *Active Field Theories*, arXiv:1904.01330.
- [72] R. J. Glauber, *Time-Dependent Statistics of the Ising Model*, J. Math. Phys. (N.Y.) **4**, 294 (1963).
- [73] S. M. Allen and J. W. Cahn, *A Microscopic Theory for Antiphase Boundary Motion and Its Application to Antiphase Domain Coarsening*, Acta Metall. **27**, 1085 (1979).
- [74] N. I. Nikolov, D. Neshev, O. Bang, and W. Z. Królikowski, *Quadratic Solitons as Nonlocal Solitons*, Phys. Rev. E **68**, 036614 (2003).
- [75] M. A. Fuentes, M. N. Kuperman, and V. M. Kenkre, *Non-local Interaction Effects on Pattern Formation in Population Dynamics*, Phys. Rev. Lett. **91**, 158104 (2003); M. G. Clerc, D. Escaff, and V. M. Kenkre, *Patterns and Localized Structures in Population Dynamics*, Phys. Rev. E **72**, 056217 (2005).
- [76] E. Hernández-García, C. López, S. Pigolotti, and K. H. Andersen, *Species Competition: Coexistence, Exclusion and Clustering*, Phil. Trans. R. Soc. A **367**, 3183 (2009).
- [77] W. E, W. Ren, and E. Vanden-Eijnden, *String Method for the Study of Rare Events*, Phys. Rev. B **66**, 052301 (2002).
- [78] W. E, W. Ren, and E. Vanden-Eijnden, *Simplified and Improved String Method for Computing the Minimum Energy Paths in Barrier-Crossing Events*, J. Chem. Phys. **126**, 164103 (2007).
- [79] M. Kardar, G. Parisi, and Y.-C. Zhang, *Dynamic Scaling of Growing Interfaces*, Phys. Rev. Lett. **56**, 889 (1986).
- [80] R. Wittkowski, A. Tiribocchi, J. Stenhammar, R. J. Allen, D. Marenduzzo, and M. E. Cates, *Scalar ϕ^4 Field Theory for Active-Particle Phase Separation*, Nat. Commun. **5**, 4351 (2014).
- [81] A. P. Solon, J. Stenhammar, M. E. Cates, Y. Kafri, and J. Tailleur, *Generalized Thermodynamics of Motility-Induced Phase Separation: Phase Equilibria, Laplace Pressure, and Change of Ensembles*, New J. Phys. **20**, 075001 (2018).
- [82] J. O’Byrne, *Nonequilibrium Currents in Stochastic Field Theories: A Geometric Insight*, Phys. Rev. E **107**, 054105 (2023).
- [83] D. T. Gillespie, *A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions*, J. Comput. Phys. **22**, 403 (1976).
- [84] J. Elf and M. Ehrenberg, *Spontaneous Separation of Bi-stable Biochemical Systems into Spatial Domains of Opposite Phases*, Syst. Biol. **1**, 230 (2004).
- [85] D. S. Dean, *Langevin Equation for the Density of a System of Interacting Langevin Processes*, J. Phys. A **29**, L613 (1996).
- [86] O. Dauchot and V. Démery, *Dynamics of a Self-Propelled Particle in a Harmonic Trap*, Phys. Rev. Lett. **122**, 068002 (2019).
- [87] T. J. Newman, A. J. Bray, and A. J. McKane, *Inertial Effects on the Escape Rate of a Particle Driven by Colored Noise: An Instanton Approach*, J. Stat. Phys. **59**, 357 (1990).
- [88] E. Woillez, Y. Zhao, Y. Kafri, V. Lecomte, and J. Tailleur, *Activated Escape of a Self-Propelled Particle from a Metastable State*, Phys. Rev. Lett. **122**, 258001 (2019).
- [89] E. Woillez, Y. Kafri, and V. Lecomte, *Nonlocal Stationary Probability Distributions and Escape Rates for an Active Ornstein-Uhlenbeck Particle*, J. Stat. Mech. **2020**, 063204 (2020).
- [90] J. Dzubiella, G. P. Hoffmann, and H. Löwen, *Lane Formation in Colloidal Mixtures Driven by an External Field*, Phys. Rev. E **65**, 021402 (2002).
- [91] C. Reichhardt, J. Thibault, S. Papanikolaou, and C. J. O. Reichhardt, *Laning and Clustering Transitions in Driven Binary Active Matter Systems*, Phys. Rev. E **98**, 022603 (2018).
- [92] M. Kourbane-Houssene, C. Erignoux, T. Bodineau, and J. Tailleur, *Exact Hydrodynamic Description of Active Lattice Gases*, Phys. Rev. Lett. **120**, 268003 (2018).
- [93] T. Speck, *Collective Behavior of Active Brownian Particles: From Microscopic Clustering to Macroscopic Phase Separation*, Eur. Phys. J. Special Topics **225**, 2287 (2016).
- [94] M. E. Cates and C. Nardini, *Classical Nucleation Theory for Active Fluid Phase Separation*, Phys. Rev. Lett. **130**, 098203 (2023).
- [95] J. L. Lebowitz and H. Spohn, *A Gallavotti-Cohen-Type Symmetry in the Large Deviation Functional for Stochastic Dynamics*, J. Stat. Phys. **95**, 333 (1999).

- [96] H. Touchette, *Introduction to Dynamical Large Deviations of Markov Processes*, *Physica (Amsterdam)* **504A**, 5 (2018).
- [97] T. Grafke and E. Vanden-Eijnden, *Numerical Computation of Rare Events via Large Deviation Theory*, *Chaos* **29**, 063118 (2019).
- [98] J. Yan, H. Touchette, and G. M. Rotskoff, *Learning Non-equilibrium Control Forces to Characterize Dynamical Phase Transitions*, *Phys. Rev. E* **105**, 024115 (2022).
- [99] G. Nicolis and J.W. Turner, *Effect of Fluctuations on Bifurcation Phenomena*, *Ann. N.Y. Acad. Sci.* **316**, 251 (1979).