

RESEARCH ARTICLE | JANUARY 10 2023

Rate and noise-induced tipping working in concert

Katherine Slyman   ; Christopher K. Jones 

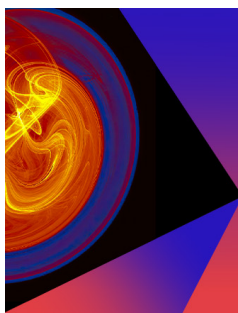


Chaos 33, 013119 (2023)

<https://doi.org/10.1063/5.0129341>



CrossMark



Chaos

Special Topic:
Nonautonomous Dynamical Systems:
Theory, Methods, and Applications

Submit Today

 AIP
Publishing

Rate and noise-induced tipping working in concert

Cite as: Chaos 33, 013119 (2023); doi: 10.1063/5.0129341

Submitted: 4 October 2022 · Accepted: 19 December 2022 ·

Published Online: 10 January 2023



View Online



Export Citation



CrossMark

Katherine Slyman^{1,a)} and Christopher K. Jones²

AFFILIATIONS

¹Department of Mathematics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27517, USA

²Renaissance Computing Institute, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27517, USA

^{a)}Author to whom correspondence should be addressed: kslyman@ad.unc.edu

ABSTRACT

Rate-induced tipping occurs when a ramp parameter changes rapidly enough to cause the system to tip between co-existing, attracting states. We show that the addition of noise to the system can cause it to tip well below the critical rate at which rate-induced tipping would occur. Moreover, it does so with significantly increased probability over the noise acting alone. We achieve this by finding a global minimizer in a canonical problem of the Freidlin–Wentzell action functional of large deviation theory that represents the most probable path for tipping. This is realized as a heteroclinic connection for the Euler–Lagrange system associated with the Freidlin–Wentzell action and we find it exists for all rates less than or equal to the critical rate. Its role as the most probable path is corroborated by direct Monte Carlo simulations.

Published under an exclusive license by AIP Publishing. <https://doi.org/10.1063/5.0129341>

The Intergovernmental Panel on Climate Change (IPCC)¹ defines a tipping point as “a level of change in system properties beyond which a system reorganizes, often in a nonlinear manner, and does not return to the initial state even if the drivers of the change are abated.” For climate systems, tipping points refer to critical thresholds when global or regional climates switch stable states.¹ Climate change is a rate-induced tipping problem. It is also a noisy system. These two components are true for many relevant climate subsystems. Since conceptual models of these climate systems contain multiple mechanisms that can induce tipping, there is a clear need for mathematical approaches, which synthesize techniques from these areas of research. To begin this analysis, we look at this question in its most simple form: in the context of a canonical problem. We study a system with a ramp parameter and impose additive noise on the dynamics to study to what extent the ramp parameter and noise interact to facilitate tipping. We approach the problem using a dynamical systems framework and prove the existence of a heteroclinic orbit between a stable base state and threshold boundary. We find that this heteroclinic orbit corresponds to the most probable path between these points. For rate values less than some critical rate, a ramp parameter alone does not allow tipping. The addition of noise to the system causes tipping well below the critical rate needed for rate-induced tipping to occur. However, the noise alone acting on the system induces tipping but only after a significantly longer time.

Therefore, a ramp parameter and noise conspire to cause tipping with increased probability over either acting alone.

I. INTRODUCTION

There are three main mechanisms for tipping in dynamical systems: bifurcation-induced, rate-induced, and noise-induced.² This work focuses on when there is a parameter shift (R-tipping) and the addition of random fluctuations (N-tipping), the schematics of which are shown in Fig. 1. The aim is to assess the extent rate and noise-induced tipping work together to facilitate tipping in cases where neither readily tip on their own.

We consider a canonical one-dimensional system with a ramp parameter and impose additive noise on its dynamics. We find that the addition of noise to the system causes the system to tip for all r values less than the critical rate needed for rate-induced tipping and does so with significantly increased probability over the noise or ramp acting alone. The most probable path to tip for all r values corresponds to the global minimizer of the Freidlin–Wentzell rate functional, which itself is a heteroclinic orbit. While we show these results in the context of a canonical problem, the phenomenon we find is suggestive of tipping between stable base states and threshold boundaries.

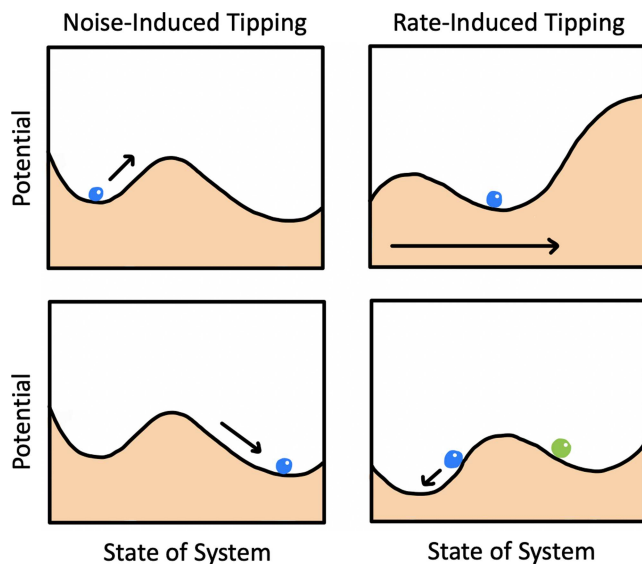


FIG. 1. Schematics for noise and rate-induced tipping in terms of a potential function and initializing with a particle at a minimum. N-tipping occurs when a noisy fluctuation is strong enough to push the particle out of the minimum and to some local maximum, where it can then fall to another minimum. R-tipping occurs when an external input varies too fast compared to the response rate of the system, resulting in a shift of the landscape and the deviation of a particle from its initialized stable state and the start of tracking a different stable state. This figure is inspired and recreated from van der Bolt and van Nes.⁴

Our methods are as follows. We compactify the system and derive the Euler–Lagrange equations associated with the Freidlin–Wentzell action functional.³ Using a dynamical systems framework, including tracking invariant manifolds, using the Wazewski principle, and applying shooting methods, we prove that there is always an intersection of the unstable manifold of the base state and stable manifold of the threshold state. Through numerical simulations, we find that this intersection is unique. The action values indicate that the heteroclinic connection through this intersection point is the global minimizer of the Freidlin–Wentzell functional. The fact that it does correspond to the most probable path at the appropriate noise levels is shown from Monte Carlo simulations.

As we consider nonzero rates within the ramp parameter, the ramp is a nontrivial component of the system. Consequently, this means that the additive noise should be of small levels, as otherwise, the noise effects would come after the ramp finishes, and we focus on the interplay of these phenomena. A drawback of the Freidlin–Wentzell theory is that it necessitates vanishingly small noise.³ In focusing on a small but not vanishingly small noise regime, the transient behavior of the underlying deterministic system will play an important role. We find that the Freidlin–Wentzell theory still holds in regard to the dynamical structure for small noise strengths, namely, the heteroclinic connection is the most probable path, but more discussion is needed when considering the expected time to tip. The extension from vanishingly small to small noise

levels is relevant for several applications of interest, especially in environmental, social, or biological contexts.

The addition of noise, regardless of size, will always result in the tipping of the system in infinite time. However, if we consider a finite time horizon, the addition of noise will cause the system to tip with a certain probability. The probability of tipping is dependent on both the noise strength and the speed of the ramp parameter, r , where the time horizon is chosen long enough to ensure the ramp function completes its transition. The size of small noise will change depending on the value of r being considered. Noise strengths are chosen so that the probability of tipping is less than 21%.

Our analysis builds off the work of Ashwin *et al.*⁵ and Ritchie and Sieber.⁶ Ashwin *et al.*⁵ introduced and studied the prototype model for rate-induced tipping

$$\dot{x} = (x + \lambda)^2 - 1, \quad (1)$$

with a monotonically increasing time-dependent parameter,

$$\lambda(t) = \frac{\lambda_{\max}}{2} \left(1 + \tanh \left(\frac{\lambda_{\max} r t}{2} \right) \right), \quad r > 0. \quad (2)$$

Using a compactification,⁷ they augment the system to an autonomous two-dimensional system containing equilibria and compact invariant sets, and in turn, the rate-induced tipping problem turns into a heteroclinic connection problem between two saddle equilibria. Perryman⁸ finds the critical rate needed for tipping within the system is $r_c = 4/3$. Ritchie and Sieber⁶ then considered this canonical problem with additive noise and found that an interplay between the noise and ramp parameter results in the tipping of the system before the critical rate, r_c , is reached. However, they only consider r values close to the critical rate. Ritchie and Sieber⁶ find solutions of the variational problem of determining the most likely tipping path using numerical continuation techniques. The majority of their work focuses on the most likely tipping time in the plane of two parameters: distance from tipping threshold and noise intensity.

The motivation of this work relates to climate subsystems. The Earth’s climate is changing due to steadily warming temperatures caused by rising levels of greenhouse gases.⁹ Moreover, there are parts of the Earth system that have the potential for large, abrupt, and irreversible transitions in response to this warming and could lead to cascading effects.¹ These changes can be characterized as tipping points.¹ As presented in Lenton¹⁰ and Lenton *et al.*,¹¹ there are many such examples: Greenland ice sheet loss, break-off of Antarctic ice-sheets, boreal forest dieback, and permafrost loss, to name a few. Given the magnitude of the impacts of these phenomena, a comprehensive understanding of tipping phenomena is needed to predict and prevent these irreparable changes. Many conceptual models of climate systems contain multiple mechanisms that can induce tipping and there is a clear need for mathematical approaches that combine techniques from both rate-induced and noise-induced tipping.

The paper is structured as follows. Section II begins with the deterministic dynamics of the canonical problem. In Sec. III, we build the stochastic framework by introducing additive noise to the system. In Sec. IV, we derive and study the most probable path equations. These equations lead to a theorem about the existence of a heteroclinic orbit. Last, in Sec. V, we perform a numerical

investigation of this problem that includes finding the heteroclinic connections, path actions, and probability of tipping for different values of the rate and noise strength. We finish with a discussion and concluding remarks.

II. THE DETERMINISTIC DYNAMICS AND RATE-INDUCED TIPPING

Rate-induced tipping is where a sufficiently quick change to a parameter of a system may cause the system to move away from one attractor to another.² We consider

$$\dot{x} = (x + y)^2 - 1, \quad (3)$$

where $\dot{} = \frac{d}{dt}$, and a monotonically increasing time-dependent parameter, as proposed by Ashwin *et al.*,⁵

$$y(t) = \frac{3}{2} \left(1 + \tanh \left(\frac{3rt}{2} \right) \right), \quad r > 0. \quad (4)$$

Reformulating the nonautonomous system in (3) into the compactified system using the ramp function, (4), itself as the coordinate transformation, maps the real line onto the finite y -interval (0, 3). The y -interval is closed by including $y = 0, 3$, which come from the limits of (4) at $\pm\infty$. This leads to the autonomous compactified two-dimensional system

$$\begin{aligned} \dot{x} &= (x + y)^2 - 1, \\ \dot{y} &= ry(3 - y). \end{aligned} \quad (5)$$

The system given in (5) has four fixed points. We focus on the saddle equilibria $(-1, 0)$ and $(-2, 3)$. At a critical r , which we denote r_c , there is a heteroclinic connection between the two saddle points. Perryman⁸ found that $r_c = 4/3$ and the connecting orbit is the line given by $x = -\frac{y}{3} - 1$. However, for $0 < r < r_c$, the system end-point tracks the saddle equilibrium initialized at $(-1, 0)$ to $(-4, 3)$ and when $r > r_c$, the system tips to infinity. In Fig. 2, we show trajectories for different values of r for the system given in (5), demonstrating solution behaviors when initializing at the saddle $(-1, 0)$.

III. BUILDING THE STOCHASTIC FRAMEWORK

For the remainder of this work, we want to consider the effects of additive noise on the dynamics of x in (5). However, for the Freidlin–Wentzell theory, we have to consider noise on both the dynamics of x and y and take the limit as noise goes to zero in the y component. The stochastic version of the canonical problem is of the form

$$\begin{aligned} dx &= f(x, y)dt + \sigma_1 dW_1 = ((x + y)^2 - 1)dt + \sigma_1 dW_1, \\ dy &= g(x, y)dt + \sigma_2 dW_2 = (ry(3 - y))dt + \sigma_2 dW_2. \end{aligned} \quad (6)$$

Speaking generally of this form, x, y are stochastic processes parameterized by time, f, g are the deterministic pieces of the system often referred to as the drift, W_1, W_2 are standard Wiener processes, and σ_1, σ_2 denote the noise strength and often referred to as the diffusion coefficient.

With the addition of noise to the system, we will have, with probability equal to one, tipping in the system between the two

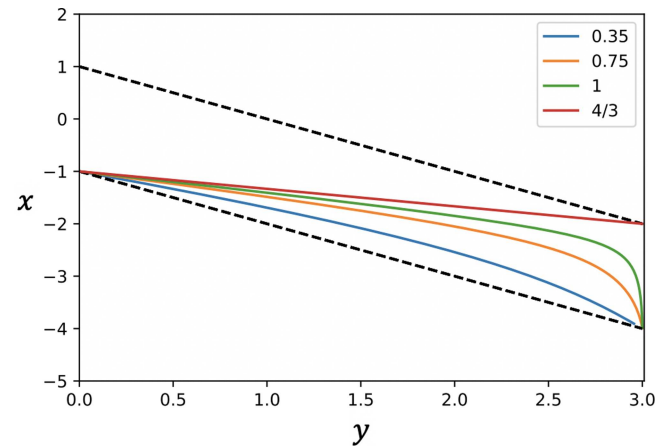


FIG. 2. Solutions of the compactified system given by (5). The black dashed curves track the fixed points $x = 1 - y$ and $x = -1 - y$ in the frozen system over time. The colored trajectories are solution curves initialized at $x = -1$, $y = 2.80729 \times 10^{-13}$ for different values of r . Solution curves with $0 < r < 4/3$ do not tip, whereas the solution curves with $r \geq 4/3$ tip. There is a heteroclinic connection between the two saddle equilibria at $r = 4/3$.

saddle equilibria. We want to find the most probable path to tip between these two points. The tool we use to study these transitions is the Freidlin–Wentzell theory of large deviations. This framework is fully presented in Freidlin and Wentzell’s monograph,³ Forgoston and Moore’s review article¹² and for gradient systems in Berglund’s review article.¹³

As presented in Freidlin and Wentzell,³ the most probable path between two points (x_0, y_0) and (x_f, y_f) is a curve of the form $(c_1(t), c_2(t))$ that minimizes the Freidlin–Wentzell functional, which is given by

$$I[c_1, c_2] = \int_{t_0}^{t_f} \left(\frac{(\dot{c}_1 - f)^2}{\sigma_1^2} + \frac{(\dot{c}_2 - g)^2}{\sigma_2^2} \right) dt, \quad (7)$$

where $(c_1(t_0), c_2(t_0)) = (x_0, y_0)$ and $(c_1(t_f), c_2(t_f)) = (x_f, y_f)$. $I[c_1, c_2]$ is non-negative and vanishes if and only if both $\dot{c}_1 = f$ and $\dot{c}_2 = g$ are solutions to the associated deterministic system. This functional represents the cost of straying from the deterministic dynamics. Minimizing this functional leads to the Euler–Lagrange equations, given by

$$\begin{aligned} \ddot{c}_1 &= f_y \dot{c}_2 + f f_x + \frac{\sigma_1^2}{\sigma_2^2} (g g_x - \dot{c}_2 g_x), \\ \ddot{c}_2 &= g_x \dot{c}_1 + g g_y + \frac{\sigma_2^2}{\sigma_1^2} (f f_y - \dot{c}_1 f_y), \end{aligned} \quad (8)$$

which are a condition that critical points, consequently minimizers, of the Freidlin–Wentzell functional must satisfy. These conditions are necessary but not sufficient for minimizers.³ We will use these conditions to derive the most probable path equations in Sec. IV.

IV. A DYNAMICAL SYSTEMS PERSPECTIVE ON THE CANONICAL PROBLEM

Using the Euler–Lagrange equations given by (8), we use a Legendre transform¹⁴ to create a degree four Hamiltonian system of the form

$$\begin{aligned}\dot{x} &= f + \sigma_1^2 p, \\ \dot{p} &= -f_x p, \\ \dot{y} &= g + \sigma_2^2 q, \\ \dot{q} &= -g_y q.\end{aligned}\quad (9)$$

The Hamiltonian function itself is

$$H(x, p, y, q) = fp + gq + \frac{\sigma_1^2}{2} p^2 + \frac{\sigma_2^2}{2} q^2. \quad (10)$$

As mentioned earlier, we want to only consider noise on the dynamics of x , as y is a time parameterization, and, thus, we send σ_2 to zero. It follows that $\dot{x}, \dot{p}, \dot{y}$ are all independent of q and we are able to project onto our equations into x, p, y space. Using this independence of q and substituting f and g as they are defined in (6) results in (9) becoming

$$\begin{aligned}\dot{x} &= (x + y)^2 - 1 + \sigma_1^2 p, \\ \dot{p} &= -2(x + y)p, \\ \dot{y} &= ry(3 - y).\end{aligned}\quad (11)$$

In addition, notice that $p = 0$ is invariant and carries the deterministic flow given by (5). These equations in (11) are the most probable path equations. Throughout this work, x is the original state variable, y is a time reparameterization, and p is the extra variable representing the work a trajectory has to do against the vector field.

We note that alternatively we could have used the Freidlin–Wentzell functional on the nonautonomous system (3) to derive the Euler–Lagrange equations, use a Legendre transform to create a degree two Hamiltonian system, and finish by compactifying the system. The compactification process and the Euler–Lagrange and Legendre transform procedures commute, and we would have the same resulting equations as shown in (11). This alternative method is useful when we perform numerical experiments in Sec. V B.

Performing a phase portrait analysis on (11), we have six equilibria: three on $y = 0$ and three on $y = 3$. We are interested in the heteroclinic connection between the saddle points $(-1, 0, 0)$ and $(-2, 0, 3)$, as these correspond to the saddles $(-1, 0)$ and $(-2, 3)$ in our two-dimensional phase space. For notation purposes, we refer to $(-1, 0, 0)$ as s_1 and $(-2, 0, 3)$ as s_2 . A quick check of the eigenvalues of (11) linearized at s_1 show s_1 has a 1D stable manifold and a 2D unstable manifold. Similar methods show s_2 has a 1D unstable manifold and a 2D stable manifold. We denote unstable and stable manifolds of a point p by $W^u(p)$ and $W^s(p)$ respectively. Using this notation, the desired heteroclinic will lie on $W^u(s_1)$ and also on $W^s(s_2)$. See Fig. 3 for what the phase space looks like on $y = 0$. We note that asymptotically the phase space dynamics are identical on $y = 3$.

Using the Hamiltonian structure in the invariant planes $y = 0$ and $y = 3$ creates two possible tipping paths between the two saddles

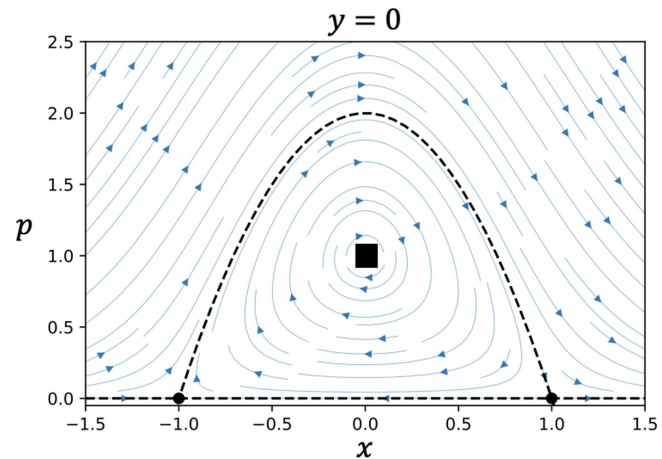


FIG. 3. The phase space for (11) on the plane $y = 0$. We have two saddles (black circles) and one center (black square). The black dashed lines represent where $H = 0$. The blue arrows show the direction of the vector field. The phase space is asymptotically identical on the plane $y = 3$.

of interest. The first possible path is to tip from s_1 to $(1, 0, 0)$ in $y = 0$ and then end-point track from $(1, 0, 0)$ to s_2 in $p = 0$. The second possible path is to end-point track in $p = 0$ from s_1 to $(-4, 0, 3)$ and then tip to s_2 in $y = 3$. However, as we will see in Secs. V C and V D, these paths have essentially an infinite time until tipping occurs and have a high action value.

We claim that there is always a third heteroclinic connection that is the most probable path and is the path of least action. We first show the existence of a heteroclinic orbit between the two saddles s_1 and s_2 for all $0 < r \leq r_c$ by showing $W^u(s_1)$ is continuous on the plane $y = -x$ for $y \leq \frac{3}{2}$ for $0 < r \leq r_c$, and that $W^u(s_1)$ and $W^s(s_2)$ are symmetric.

Proposition 1: *The primary intersection of $W^u(s_1)$ with the plane $y = -x$ is continuous on the plane $y = -x$ for $y < \frac{3}{2}$ and $r > 0$.*

Proof. The Wazewski principle¹⁵ states the following: Let W^- be the immediate exit set of W and let W^0 be the eventual exit set of W . If W^- is closed relative to W^0 , then W is a Wazewski set and the map $K: W^0 \rightarrow W^-$, which takes each point to the first where it exits W is continuous.

We define the primary intersection of $W^u(s_1)$ with the plane $y = -x$ to be the first crossing of this plane from trajectories initialized in the unstable subspace of s_1 coming from $-\infty$. For the system given in (11), we say the Wazewski set, W , is the space bounded by the following planes: $y = -1 - x$, $y = -x$, $y = 0$, $y = 3/2$, and $p = 0$, and extends infinitely in the positive p direction. Based on flow of the vector field, the following are true about the boundaries of W : $y = -1 - x$ is an entrance set, $y = 0$ and $p = 0$ are neither entrance nor exit sets, as they are invariant planes, and $y = 3/2$ is an exit set. On $y = -x$, below the curve $p = \frac{1}{\sigma_1^2}(1 - ry(3 - y))$ is an entrance set and above it, an exit set. Refer to Fig. 4 for a visual of W .

We have to determine what happens on the curve $p = \frac{1}{\sigma_1^2}(1 - ry(3 - y))$ itself, which is the boundary between an entrance set and an immediate exit set. Consider x and p as functions of y .

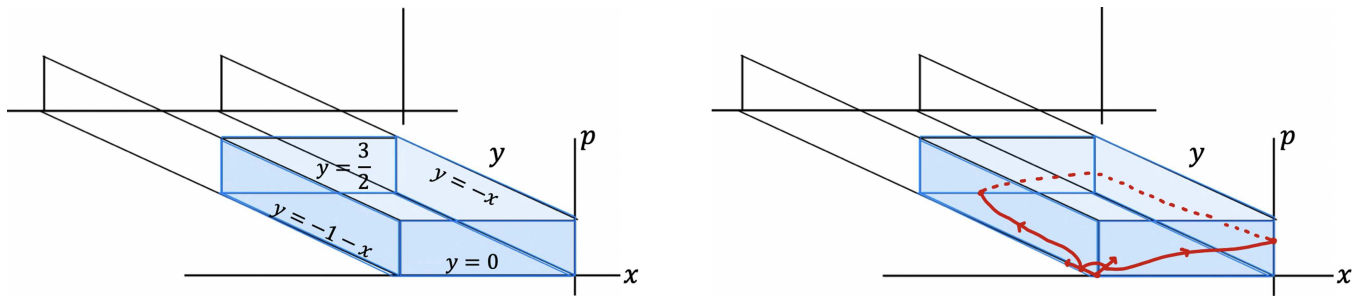


FIG. 4. The boundary of the Wazewski set is in blue and extends infinitely in the positive p direction. It is the space bounded by $y = -1 - x$, $y = 0$, $y = -x$, $y = \frac{3}{2}$, and $p = 0$. Taking a quarter circle of radius ϵ about s_1 intersected with the piece of $W^u(s_1)$ lying in W , and applying map K , results in the dotted red curve.

Looking at the first and second derivatives at the point $z = \left(-y, \frac{1}{\sigma_1^2}, (1 - ry(3 - y)), y\right)$, representing any point on this curve, we have

$$\begin{aligned} \frac{dx}{dy} \Big|_z &= -1, \\ \frac{d^2x}{dy^2} \Big|_z &= \frac{3 - 2y}{y(3 - y)} > 0 \quad \text{for } y < 3/2, \\ \frac{dp}{dy} \Big|_z &= 0, \\ \frac{d^2p}{dy^2} \Big|_z &= 0. \end{aligned} \quad (12)$$

We see in (12) that $\frac{dx}{dy} = -1$ and $\frac{d^2x}{dy^2} > 0$. By the second derivative test, we know a trajectory would be concave up at this point, forcing any points to leave and consequently, not enter W . Therefore, we have shown that the boundary of the immediate exit set is contained in the immediate exit set. We conclude the following about the immediate exit set and eventual exit set of W :

$$\begin{aligned} W^- &= \left\{ (x, p, y) \mid y = 3/2, y = -x \text{ for } p \geq \frac{1}{\sigma_1^2}(1 - ry(3 - y)) \right\}, \\ W^0 &= \{(x, p, y) \mid W \setminus \{y = 0, y = -1 - x, p = 0, (-1, 0, 0), (0, 1, 0)\}\}. \end{aligned}$$

The boundary of the immediate exit set is in the immediate exit set, and it easily follows that W^- is closed relative to W^0 . Therefore, W is a Wazewski set and the map $K: W^0 \rightarrow W^-$, is continuous for $y < \frac{3}{2}$. This implies that $W^u(s_1)$ intersected with the plane $y = -x$ is continuous for $y < \frac{3}{2}$ and $r > 0$. $\square$

Proposition 2: $W^u(s_1)$ intersected with W^- crosses the plane $y = \frac{3}{2}$ for $0 < r \leq r_c$.

Proof. Notice $y = \frac{3}{2}$ separates W^- into two pieces. Take the quarter circle of radius ϵ around the fixed point s_1 intersected with $W^u(s_1)$ that lies in W , and call this curve C_W . Applying the map K to C_W results in a curve in $\mathbb{R}^3$, specifically a curve lying in W^- by the definition of Wazewski map K .

Since C_W is a closed curve, we track where the two end points of C_W map to under K . The first end point of C_W has $y = 0, p \neq 0$ and

second end point of C_W has $p = 0, y \neq 0$. Take the end point of C_W that lies in $y = 0$. Since the $y = 0$ plane is invariant, when we apply K , the trajectory must stay in this plane and eventually exit through $y = -x$ and above $p = \frac{1}{\sigma_1^2}(1 - ry(3 - y))$. Take the end point of C_W that lies in $p = 0$. Since the $p = 0$ plane is invariant, when we apply K , the trajectory must stay in this plane. Since $\dot{y} > 0$, this trajectory will eventually exit through $y = \frac{3}{2}$, when $0 < r \leq r_c$.

$W^u(s_1)$ intersected with W^- actually intersects $y = \frac{3}{2}$ by the intermediate value theorem, as K is a continuous map, and one end point of C_W maps to the plane $y = -x$ in $y = 0$ while the other end point of C_W maps to the plane $y = 3/2$ in $p = 0$. See Fig. 4 for an illustration of this shooting argument.

Therefore, the intersection of $W^u(s_1)$ and the plane $y = -x$ is continuous for $y \leq \frac{3}{2}$ for $0 < r \leq r_c$. $\square$

Proposition 3: $W^u(s_1)$ and $W^s(s_2)$ are symmetric.

Proof. Recall our system given in (11). Making the change of variables $\tau = -t$, we get the time reversed system given by

$$\begin{aligned} x' &= -(x + y)^2 + 1 - \sigma_1^2 p, \\ p' &= 2(x + y)p, \\ y' &= ry(y - 3). \end{aligned} \quad (13)$$

We transform the variables x, p, y by

$$\hat{x} = -x - 3, \quad \hat{p} = p, \quad \hat{y} = 3 - y$$

and substitute them into the time reversed system we found in (13). The equations simplify to

$$\begin{aligned} \hat{x}' &= (\hat{x} + \hat{y})^2 - 1 + \sigma_1^2 \hat{p}, \\ \hat{p}' &= -2(\hat{x} + \hat{y})\hat{p}, \\ \hat{y}' &= r\hat{y}(3 - \hat{y}). \end{aligned} \quad (14)$$

We see that (14) is in the original form, as given in (11), and it follows that $W^u(s_1)$ and $W^s(s_2)$ are symmetric. $\square$

Theorem 1: There exists a heteroclinic connection between the saddle points s_1 and s_2 that goes through the plane $y = -x$ at $y = \frac{3}{2}$ for $0 < r \leq r_c$.

Proof. We found that the intersection of $W^u(s_1)$ and the plane $y = -x$ was continuous for $y \leq \frac{3}{2}$ for $0 < r \leq r_c$ using Propositions 1 and 2. The symmetry of $W^u(s_1)$ and $W^s(s_2)$, proven in Proposition 3, implies the intersection of $W^s(s_2)$ intersected with the plane

$y = -x$ is continuous for $y \geq \frac{3}{2}$ for $0 < r \leq r_c$. Therefore, $W^u(s_1)$ and $W^s(s_2)$ will always intersect once in the plane $y = -x$ at $y = \frac{3}{2}$, implying a heteroclinic connection between s_1 and s_2 for $0 < r \leq r_c$, and concluding our proof of Theorem 1. $\square$

We have, thus, proven the existence of a heteroclinic connection between s_1 and s_2 for all $0 < r \leq r_c$ and demonstrated how to find this heteroclinic using the intersection of the invariant manifolds. In the deterministic system, for $0 < r < r_c$, we would not have tipping or a heteroclinic connection between the saddles. The presence of noise, regardless of the size, allows direct tipping between these saddles within the system.

V. COMPUTATIONAL METHODS AND NUMERICAL RESULTS

A. Visualization of invariant manifolds and the heteroclinic connection

We proved in Sec. IV the existence of the intersection of $W^u(s_1)$ and $W^s(s_2)$ at $y = 3/2$, giving rise to a heteroclinic connection between the two saddle points through that specific point. We numerically compute these manifolds, plot them in $y = -x$, and observe their intersection at $y = 3/2$. This enables us to visualize their intersection as well as compute the trajectory through the intersection point. The trajectory is then projected into the xy plane to find the heteroclinic connection in the two-dimensional extended phase space.

The local unstable subspace of s_1 is spanned by the two vectors

$$\left(\frac{\sigma_1^2}{4}, 1, 0\right)^T \text{ and } \left(\frac{-2}{3r+2}, 0, 1\right)^T, \quad (15)$$

which span the plane

$$4(x+1) - (\sigma_1^2 + 4)p - \frac{12r}{3r+2}y = 0. \quad (16)$$

Intersecting this plane with the sphere

$$(x+1)^2 + p^2 + y^2 = (0.001)^2 \quad (17)$$

and taking points such that $y, p > 0$ result in a curve of points that lie in the unstable subspace, as seen in Fig. 5. We discretize this curve and use the tuples as a set of initial conditions. We numerically run system (11) forward in time, for each initial condition, until the trajectory first hits the plane $y = -x$. Similarly, you can perform this

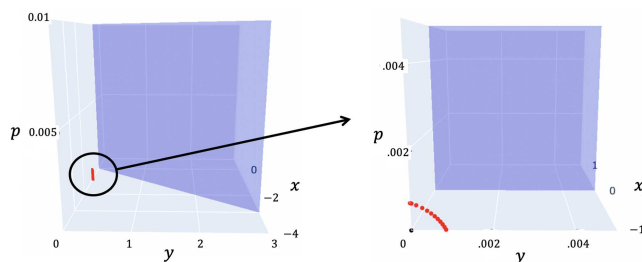


FIG. 5. Parameters are set at $r = 1$ and $\sigma_1 = 0.15$. The red curve is the intersection of the sphere (17) and the plane (16) spanned by (15).

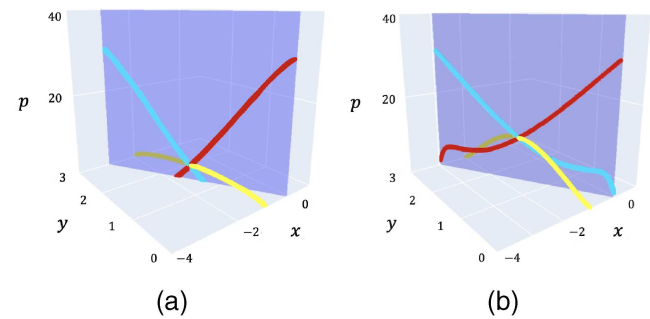


FIG. 6. $W^u(s_1)$ (red) and $W^s(s_2)$ (cyan) in the plane $y = -x$ (purple) for $y \in (0, 3)$. The trajectory through their intersection point is the heteroclinic orbit connecting s_1 and s_2 (yellow). (a) Parameters are set at $r = 1$ and $\sigma_1 = 0.25$. (b) Parameters are set at $r = 0.5$ and $\sigma_1 = 0.25$.

process when looking at the stable subspace of s_2 and running system (11) backward in time. We find the intersection of these two curves in the plane $y = -x$. Through these simulations, we find the intersection point of these two curves in $y = -x$ is unique. Running the system both forward and backward in time from the intersection point supplies the full heteroclinic trajectory. Refer to Fig. 6 to see a visualization of $W^u(s_1)$ and $W^s(s_2)$ intersecting in the plane $y = -x$ as well as the trajectory through the intersection point for two different parameter pairs of r and σ_1 , corresponding to the heteroclinic orbit between s_1 and s_2 . Projecting this heteroclinic orbit into xy space is the connecting orbit between $(-1, 0)$ and $(-2, 3)$, and we show in Sec. VB that this orbit is, in fact, the most probable path between these points.

In the compactified system, the heteroclinic connections do not look identical when varying r . However, for fixed r and varying σ_1 , the heteroclinic connections will look identical when projected into the xy plane as this variation of σ_1 only scales the p variable. We note that in Fig. 8(b) while the heteroclinic orbit connecting the two saddle equilibria looks linear, we have verified that it is not a line. As r approaches the critical rate of the deterministic system, the heteroclinic orbit becomes deceptively linear.

In addition to the existence of the heteroclinic connection between s_1 and s_2 proven in Sec. IV, this first set of numerical simulations now verifies the uniqueness of the heteroclinic connection between s_1 and s_2 . However, we still need to determine if this heteroclinic connection is the most probable path between these two points, implying that we need to show that it is the global minimizer of the Freidlin–Wentzell action functional. We perform these calculations in Sec. VB.

B. Monte Carlo simulations and the most probable path

We corroborate that the heteroclinic connection constructed in Sec. IV is in fact the most probable path using Monte Carlo simulations. Recall that our original problem was a one-dimensional differential equation. Consider its stochastic version, given by

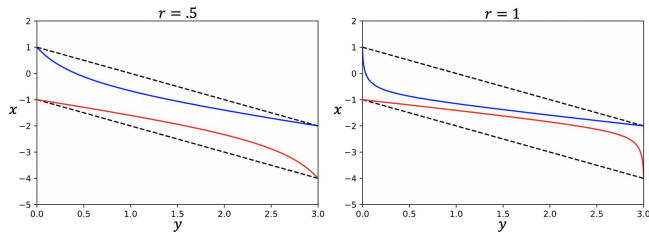


FIG. 7. $W^u(-1, 0)$ (red) and $W^s(-2, 3)$ (blue) for $r = 0.5$ and $r = 1$. Tipping occurs when a realization of (19) crosses $W^s(-2, 3)$ and $\lim_{t \rightarrow \infty} \neq -4$.

$$\begin{aligned} dx &= ((x + y)^2 - 1) dt + \sigma_1 dW, \\ y(t) &= \frac{3}{2} \left(1 + \tanh \left(\frac{3rt}{2} \right) \right), \quad r > 0. \end{aligned} \quad (18)$$

As we said in Sec. IV, the order of the compactification process and the Euler–Lagrange and Legendre procedures commute, and, therefore, we use (18) for running simulations as it is computationally less expensive.

We numerically approximate the solutions of (18) by using the Euler–Maruyama method to create a discretized Markov process¹⁶ over a time interval of length 30. To apply the Euler–Maruyama method, we partition the time interval into sub-intervals of width $\Delta t = 0.001$ and initialize the solution at $x = -1$ and $y = 2.80729 \times 10^{-13}$. This y initial condition is used to find the corresponding t_0 by solving the rescaled time $rt = -10$. This explains why the exact time intervals of length 30 vary for each case. We note that changing the initial y value, corresponding to changing the starting time, only shortens or extends the time for a realization to tip.⁶ To create the discretized Markov process, recursively define x as

$$x_{n+1} = x_n + ((x_n + y(t_n))^2 - 1) \Delta t + \sigma_1 \Delta W_n. \quad (19)$$

A standard Wiener process, W , satisfies the property that Brownian increments are independent and normally distributed with mean zero and variance Δt . Therefore, it follows that $\Delta W = W_n - W_{n-1}$

can be numerically simulated using $\sqrt{\Delta t} \cdot N(0, 1)$. This can be shown by manipulating the probability density function of $N(0, \Delta t)$.

We simulate M realizations of (18) using the Euler–Maruyama method given in (19). We map these realizations to two-dimensional phase space by plotting $(y(t), x)$. We define tipping to be when a realization of (18) crosses $W^s(-2, 3)$, and $\lim_{t \rightarrow \infty} \neq -4$. $W^s(-2, 3)$ can be seen in Fig. 7.

Of the M realizations, we define M_{tip} to be the number of realizations that tip on the finite time interval of our choosing. Thus, $M - M_{tip}$ does not tip, an example of which is shown in Fig. 8(a) for $r = 1, \sigma_1 = 0.15$. There are M_{tip} points within the M_{tip} realizations that tipped for every discretized time. We use Python to get the kernel smoothing density estimation of the M_{tip} points at each time. This finds the “most probable point” at every time step, which is determined by the peak of the kernel density estimation (KDE). This peak corresponds to the mode of the M_{tip} points at that time. Plotting the mode at each time step, we have an approximation for the most probable path. Overlaying the numerically simulated most probable path, with what we found using the projection of the trajectory through the intersection of $W^u(s_1)$ and $W^s(s_2)$ in $y = -x$ in Sec. V A, we see that the approximation matches the actual path extremely well, an example of which is shown in Fig. 8(b) for $r = 1, \sigma_1 = 0.15$. Therefore, we can say the trajectories that tipped followed the heteroclinic connection and that the heteroclinic connection between the two saddles is the most probable path.

As mentioned above, we see via the Monte Carlo simulations that trajectories either tip to infinity or end-point track the stable path to $(-4, 3)$ on the given finite time horizon. The trajectories that end-point track the stable path follow the pullback attractor⁵ of $(-1, 0)$. Performing another kernel smoothing density estimation on the realizations that did not tip, we see these trajectories actually peak along this pullback attractor, an example of which is shown in Fig. 8(c) for $r = 1$ and $\sigma_1 = 0.15$.

The heteroclinic orbit and the pullback attractor are objects that can be used to separate trajectories of the system. These computations show that the addition of noise allows the system to tip when its deterministic equivalent would not tip, as the trajectory would be the pullback attractor. For the specific parameter regime

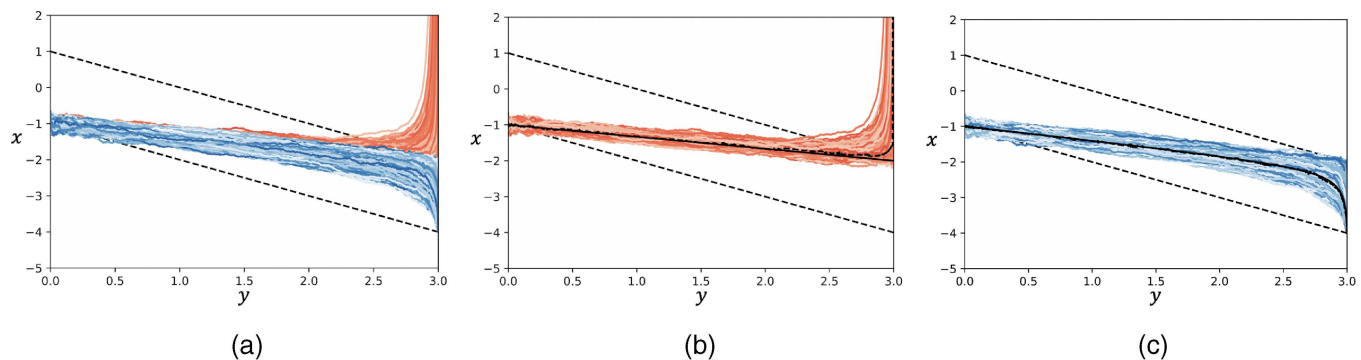


FIG. 8. (a) 3000 Monte Carlo simulations of (18) with $r = 1$ and $\sigma_1 = 0.15$ on the interval $[0, 30]$ with $dt = 0.001$. 2807 realizations do not tip (blue) and 193 tip (red). (b) The realizations that tipped overlaid with the heteroclinic orbit found (solid black) and a kernel density estimate of the realizations that tipped (dashed black). (c) The realizations that did not tip, overlaid with the pullback attractor of $(-1, 0)$ (solid black) and the kernel density estimate (dashed black).

$r = 1, \sigma_1 = 0.15$ as depicted above, even with r being $3/4$ of the critical rate, we are able to get tipping within the system.

C. Time to tip

There is concern that the influence of noise on (3) is the sole reason the system exhibits tipping. However, we verify in this section that the tipping occurs due to the interplay of both the ramp parameter and noise. Additionally, the frequency of tipping largely increases with this addition of small noise strengths interacting with the ramp parameter.

Recall that the original goal is to tip from $(-1, 0)$ to $(-2, 3)$, which corresponds to s_1 and s_2 in the three-dimensional system in (11). Using (11), based on the stable and unstable directions of these saddles, and the direction of the vector field, we initially had two possible ways to tip due to the Hamiltonian structure in the invariant planes $y = 0$ and $y = 3$. We proved in Sec. IV that we had a third way of tipping via a heteroclinic orbit between the two saddles, which we now know is the most probable path from the Monte Carlo simulations. Refer to Fig. 9 for a visual of these possible tipping paths.

Notice that if tipping occurs in either $y = 0$ or $y = 3$, there is no interplay with the ramp parameter, as it would be before or after the ramping occurs. In planes $y = 0, 3$, we have a one-dimensional stochastic differential equation. We can find the approximate the expected time to tip as we have asymptotic formulas for gradient systems,¹³ given by

$$\mathbb{E}[\tau] \approx e^{\frac{2\Delta V}{\sigma_1^2}}, \quad (20)$$

where V is the associated potential function and ΔV gives the height of the potential barrier. We calculate the expected time to tip for each tipping path. In our numerical analysis, $0.08 \leq \sigma_1 \leq 0.3$ for all experiments, as we consider a small noise regime.

Case 1. Assume we tip from s_1 to $(1, 0, 0)$ in $y = 0$ and end-point track from $(1, 0, 0)$ to s_2 along $W^s(-2, 3)$. We can find the expected time to tip between the fixed points in $y = 0$ as the system is a gradient system in this plane. The associated form of the gradient system and potential function, V , is

$$\begin{aligned} dx &= (x^2 - 1)dt + \sigma_1 dW \\ &= -\nabla V dt + \sigma_1 dW \\ &= -\nabla \left(x - \frac{1}{3}x^3 \right) dt + \sigma_1 dW. \end{aligned} \quad (21)$$

The extrema of V correspond to the fixed points of the problem when $y = 0 : x = -1, 1$. Solving for the expected time to tip, we find that without the ramp parameter,

$$\mathbb{E}[\tau] \approx e^{\frac{8}{3\sigma_1^2}} > 10^{12}. \quad (22)$$

Therefore, the time to tip from s_1 to s_2 along this path will be some time greater than 10^{12} . See a depiction of Case 1 in Fig. 9(a).

Case 2. Similarly, assume we end-point track the path from s_1 to $(-4, 0, 3)$ along $W^u(-1, 0)$ and then tip from $(-4, 0, 3)$ to s_2 in $y = 3$. We first find the expected time to tip in $y = 3$. Again, we have a gradient system and can rewrite the system in terms of the

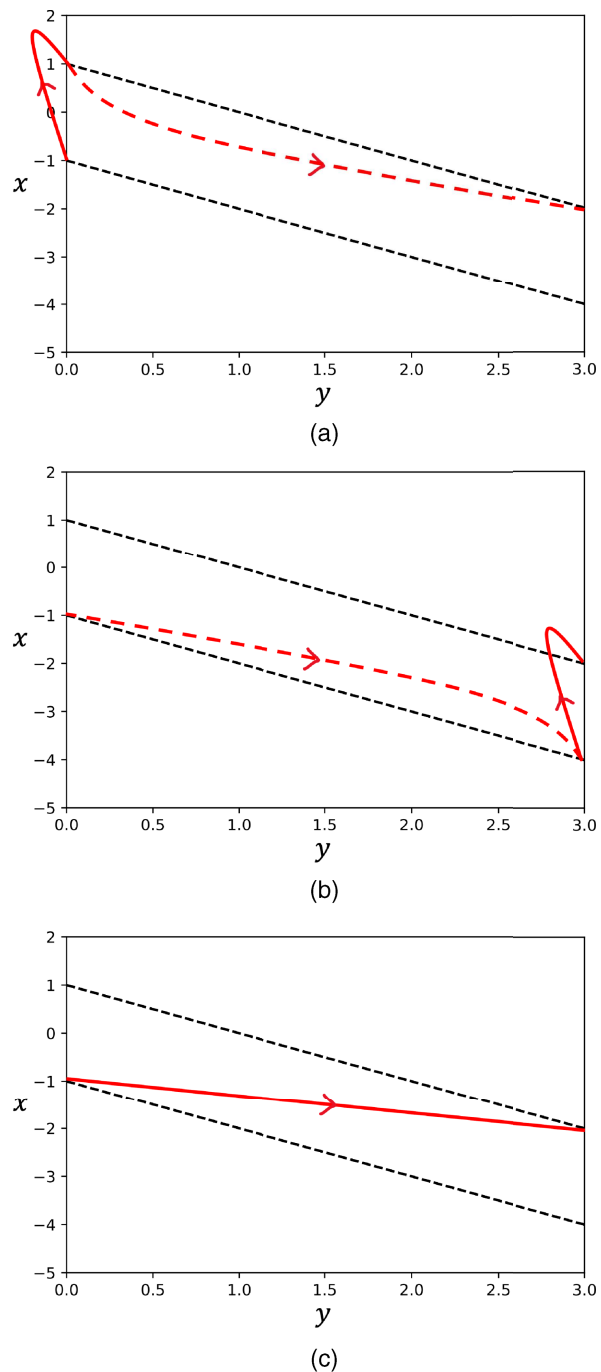


FIG. 9. Schematic diagrams for the three possible ways to tip from $(-1, 0)$ to $(-2, 3)$. (a) Case 1. (b) Case 2. (c) Case 3. In Cases 1 and 2, the end-point tracking curves (dashed red) will be dependent on $W^u(-1, 0)$ and $W^s(-2, 3)$. Here, they are sketched for $r = 0.5$. Additionally, in Cases 1 and 2, the solid red curves are to represent tipping via the heteroclinic orbit in $y = 0$ and $y = 3$. These are for presentation purposes as one would only see vertical lines instead of these parabolic curves if projected into the xy plane.

potential function V , written as

$$\begin{aligned} dx &= ((x+3)^2 - 1)dt + \sigma_1 dW \\ &= -\nabla \left(-8x - 3x^2 - \frac{1}{3}x^3 \right) dt + \sigma_1 dW. \end{aligned} \quad (23)$$

Solving for the expected time to tip, we find that without the ramp parameter,

$$\mathbb{E}[\tau] \approx e^{\frac{8}{3\sigma_1^2}} > 10^{12}. \quad (24)$$

The time to tip from s_1 to s_2 along this path will also be some time greater than 10^{12} . See a depiction of Case 2 in Fig. 9(b).

Case 3. To determine the expected time to tip of the most probable path found in Sec. IV, we run a sufficient number of Monte Carlo simulations so that the expected time to tip distribution converges. We use the Euler–Maruyama method to simulate M realizations of (18) on a time interval of length 30, initialized at $x = -1, y = 2.80729 \times 10^{-13}$, with a step size of $dt = 0.001$. Recall we use this initial condition for y to find the corresponding t_0 for each value of r . We want to find the realizations that have tipped to infinity and capture when the mapped versions, $(y(t), x)$, have crossed $W^s(-2, 3)$.

Let τ_i denote the first time a path, X_i of the form $(y(t), x)$, crosses $W^s(-2, 3)$. We define escape events to be the paths X_i that have $\tau_i \leq 30$ and component $x \rightarrow \infty$. Assume for M realizations there are K escape events. We construct the distribution for the K crossing times of $W^s(-2, 3)$. To verify we have a converged result for the distribution of the time of escape events, we use the following process.

1. Bin the crossing times of the K escape events by the Freedman Diaconis¹⁷ rule. This separates the K escape events into B bins of equal length.
2. Run another M realizations of (18) on the same time interval and with the same step size. Assume there are J escape events. We bin the J escape events by the same number of bins B found in Step 1.
3. There are two vectors D_1, D_2 of the same length, where each component of the vector represents the amount of paths that tipped in that time interval. Calculate $Err = \frac{\|D_1 - D_2\|_2}{\|D_1\|_2}$, which is the relative error between the two data sets.
4. If $Err < 0.1$, we say we have found the converged distribution. However, if $Err \geq 0.1$, we iterate this process with larger N until the relative error of D_1 and D_2 is small enough. In addition, we use the Kolmogorov–Smirnov two sample test¹⁸ as a final verification that we have a converged distribution.

We conduct this experiment for different values of r, σ_1 pairs. In Table I, we see ranges of some of the expected times to tip. Notice that unlike Cases 1 and 2, the expected time to tip is now finite. The different times to tip between s_1 and s_2 , depending on which path taken, demonstrates that tipping without the ramp is extremely rare to the point of almost never tipping. In addition, if we just had a ramp parameter and no stochastic component, there is no tipping for when $0 < r < 4/3$. Thus, there is an interplay of additive noise and a ramp parameter, and together they facilitate tipping on a finite timescale.

TABLE I. Monte Carlo simulation results for the expected time to tip, τ , for $r = 0.75, 0.85, 1, 1.1$ for different values of σ_1 . These times come from converged results of the Monte Carlo simulations using the method described above.

r	σ_1 range	t_0	τ range
0.75	0.15–0.30	−13.33	$\in(0.07, 0.67)$
0.85	0.10–0.30	−11.76	$\in(0.04, 0.74)$
1	0.08–0.25	−10.00	$\in(-0.20, 0.48)$
1.1	0.08–0.25	−9.09	$\in(-0.29, 0.41)$

Notice in Table I that the expected time to tip for the majority r, σ_1 pairs considered correspond approximately to $t = 0$, implying $y \approx 3/2$. This might be seen as a validation that the symmetric extremal calculated in Sec. IV is indeed close to where the first passages are taking place.

D. Path actions

In addition to using Monte Carlo simulations to see how realizations behave and to calculate the expected time to tip, we can compute the action along the different path options. The most probable path should be the path of the least action. Due to the choice for the p variable in the Legendre transform, for a fixed r , the variation of σ_1 results in a scaling in p . Scaling by some constant does not change the actual minimizer of the functional given by (7). Therefore, we want to consider the normalized action when calculating the path actions to make the action more numerically tractable. The normalized action is given by

$$I_N[x] = \int_{t_0}^{t_f} (\dot{x} - f)^2 dt = \int_{t_0}^{t_f} \sigma_1^4 p^2 dt. \quad (25)$$

Using (25), we find that the heteroclinic constructed in Sec. IV, Case 3, has the least action compared to the other two paths of tipping, Cases 1 and 2. We see that if we tip before the ramp starts or after the ramp finishes, the action value is 5.333. However, tipping along the most probable path gives the least action value, by multiple orders of magnitude. Refer to Table II for the comparison of the action size for each of Cases 1–3 for different r values.

In regard to the action values for Cases 1 and 2 being the same, we first want to stress that both of these paths for tipping are extremely unlikely on reasonable timescales. Practically speaking, the action values are the same when looking along the heteroclinic

TABLE II. Action values for Cases 1–3. Cases 1 and 2 do not depend on r as they tip either before or after the ramp. For Case 3, which is dependent on r , we see that for different r values, the action is much less than the action of the other paths to tip.

Case No.	r value	Action
1	...	5.333
2	...	5.333
3	1.1	0.023
3	1	0.054
3	0.75	0.226
3	0.5	0.684

orbits in the planes $y = 0, 3$ as the p values are exactly the same when considering the curves where $H = 0$. The rest of each path lies in $p = 0$ and so there is no other contribution to the action. While one might expect that it would be easier to track a stable quasi-steady state and then transition as opposed to transitioning and then tracking an unstable quasi-steady state, we believe it has to do with the compactification, as this brought $\pm\infty$ to the finite world.

VI. DISCUSSION AND CONCLUSIONS

A. The heteroclinic orbit as the most probable path

A key observation is the proximity of the heteroclinic orbit to the Monte Carlo KDE most probable path. The KDE most probable path estimate tracks the heteroclinic connection quite closely far beyond the first half of the ramping phase. A consequence of this fact is that we can conclude that the heteroclinic connection does carry the realizations that tip through the ramping phase. The heteroclinic connection provides a surprising amount of guidance on how the trajectories that tip do so, especially as we are dealing with not vanishingly small noise; a regime we were forced into because otherwise there would be no interaction of the ramp parameter and additive noise.

Another interpretation for how realizations are tipping would involve considering trajectories following the corresponding deterministic pullback attractor, and tipping would occur when $W^u(-1, 0)$ and $W^s(-2, 3)$ were closest. However, running simulations of the noisy system for different values of r and overlaying the tipped realizations with $W^u(-1, 0)$ and $W^s(-2, 3)$, we see these realizations do not track $W^u(-1, 0)$ (when you consider where the tipped realizations are most densely concentrated). As we decrease r , we see the significance of the heteroclinic orbit more clearly as the heteroclinic orbit becomes more distinguishable from the pullback attractor. The realizations that tip closely follow the heteroclinic orbit from the start. However, as r increases, $W^u(-1, 0)$ and $W^s(-2, 3)$ grow closer together, which is why it looks as the tipped realizations track the pullback attractor for some time.

B. A scaling law for the expected time to tip

For vanishingly small noise, the Freidlin–Wentzell theory of large deviations, which gives the probability of a specific trajectory in a stochastic dynamical system, aids in finding the most probable path between two points. This is obtained by minimizing the Freidlin–Wentzell action functional. Additionally, the Freidlin–Wentzell theory gives the expected time to tip.³ We saw in this work that the Freidlin–Wentzell theory holds in regard to the dynamical structure of the most probable path for small noise strengths. It is still an open question if the expected time to tip aligns between the vanishingly small noise case and the small noise case.

We discovered a power scaling law for the expected time to tip via Monte Carlo simulations, τ , and $1/\sigma_1^2$, for set $r > 0$ and varying values of σ_1 . The log–log plot of these coordinate pairs results in a linear relationship, examples of which are shown in Fig. 10. This linear relationship in the log–log space corresponds to a power law of the form $a\left(\frac{1}{\sigma_1^2}\right)^b$ between $1/\sigma_1^2$ and the expected time to tip. While the scaling laws in Fig. 10 are for $r = 1$ and $r = 0.75$, the linear

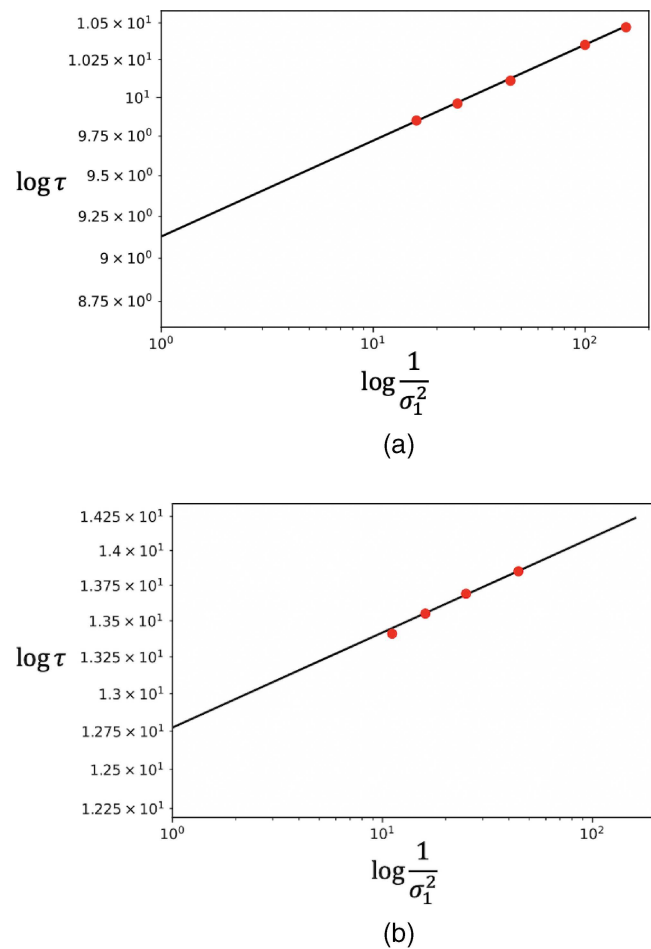


FIG. 10. (a) Log–log plot of τ vs $\frac{1}{\sigma_1^2}$ for $r = 1, \sigma_1 = 0.25, 0.2, 0.15, 0.1, 0.08$. The linear relationship in the log–log space corresponds to a power law of the form $9.14\left(\frac{1}{\sigma_1^2}\right)^{0.027}$. (b) Log–log plot of τ vs $\frac{1}{\sigma_1^2}$ for $r = 0.75, \sigma_1 = 0.3, 0.25, 0.2, 0.15$. The linear relationship in the log–log space corresponds to a power law of the form $12.42\left(\frac{1}{\sigma_1^2}\right)^{0.021}$.

relationship in the log–log space held true for multiple r values we studied. An interesting observation is the slope of the line in log–log space for $r = 1$ in Fig. 10(a) is the same as $\frac{1}{2}$ (action value $r = 1$), the value of which can be seen in Table II.

This scaling law is different from the asymptotic formula given by the Freidlin–Wentzell theory. However, there are various explanations for this mismatch. The most likely is that we are considering a small noise regime and not $\sigma_1 \rightarrow 0$, and so it is not necessarily surprising the known scaling law does not hold. Alternatively, we have yet to find the leading coefficient, c , which could be dependent on r . We hypothesize that you can find the leading coefficient, by finding more expected times to tip and switching perspectives to that of inverse problems.

We believe that this task needs to implement importance sampling¹⁹ to aid in speeding up the time required to gather the converged data sets. Importance sampling is commonly used to speed up Monte Carlo simulations of rare events by biasing realizations to those rare events.¹²

We would like to point out that Ritchie and Sieber⁶ found that for rate values between $r = 1.05$ and $r = 1.25$ that as the noise is decreased, the time to tip increases slowly. They find a similar relationship for the delay in the rate-induced tipping as that of Bakhtin²⁰ for rare escapes of an autonomous system.

C. Final conclusions

Using compactification⁷ with a coordinate transformation of the ramp parameter (4) allows us to frame the canonical problem as a two-dimensional autonomous system with fixed points and invariant objects as well as study the heteroclinic connection. We have shown that the addition of additive noise causes the system to tip well below the critical rate needed for rate-induced tipping to occur. The system will always have a heteroclinic connection directly between the two saddle equilibria for all $0 < r \leq r_c$. Moreover, the heteroclinic orbit found using the intersection of invariant manifolds matches the kernel density estimate of the noisy realizations found by Monte Carlo simulations, corroborating it as the most probable path of tipping between these two points. Calculating the action over all the possible paths between the two saddles, we find that the heteroclinic connection we constructed has the least action by multiple orders of magnitude, verifying we have truly found the most probable path between these two points. Additionally, we find that rate and noise-induced tipping conspire to facilitate tipping with increased probability, when neither tip on their own when considering a finite time horizon and $0 < r < r_c$.

We have pushed on the levels of noise to a size where the Freidlin–Wentzell theory may no longer hold as the noise strength was not vanishingly small. However, we find that the Freidlin–Wentzell theory actually is still relevant in the extent of the most probable path.

This paper has considered a one-dimensional canonical problem, but we believe this work can extend to understanding the tipping between a base state and a threshold state of similar forms. Our method made use of the symmetry within the system. If that symmetry does not exist, other implementations of the Wazewski principle will need to be used to prove that an intersection of the invariant manifolds exists. Thus, an extension to this case is still required.

ACKNOWLEDGMENTS

Both investigators were supported by the Office of Naval Research under Grant No. N000141812204.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

Katherine Slyman: Visualization (lead); Writing – original draft (lead). **Christopher K. Jones:** Writing – review & editing (lead).

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

- ¹M. Collins, M. Sutherland, L. Bouwer, S.-M. Cheong, T. Frölicher, H. Jacot Des Combes, M. Koll Roxy, I. Losada, K. McInnes, B. Ratter, E. Rivera-Arriaga, R. D. Susanto, D. Swingedouw, and L. Tibig, “Extremes, abrupt changes and managing risk,” in *IPCC Special Report on the Ocean and Cryosphere in a Changing Climate* (Cambridge University Press, 2019), pp. 589–655.
- ²P. Ashwin, S. Wiczeorek, R. Vitolo, and P. Cox, *Philos. Trans. R. Soc. A* **370**, 1166 (2012).
- ³M. I. Freidlin and A. D. Wentzell, “Random perturbations of dynamical systems,” in *Grundlehren der Mathematischen Wissenschaften* (Springer, Berlin, 2012), Vol. 260.
- ⁴B. van der Bolt and E. H. van Nes, *PLoS One* **16**, e0253003 (2021).
- ⁵P. Ashwin, C. Perryman, and S. Wiczeorek, *Nonlinearity* **30**, 2185 (2017).
- ⁶P. Ritchie and J. Sieber, *Chaos* **26**, 093116 (2016).
- ⁷S. Wiczeorek, C. Xie, and C. K. R. T. Jones, *Nonlinearity* **34**, 2970 (2021).
- ⁸C. G. Perryman, Ph.D. thesis (University of Exeter, 2015).
- ⁹IPCC (Intergovernmental Panel on Climate Change), “Climate change 2021: The physical science basis. Working Group I contribution to the IPCC Sixth Assessment Report,” Cambridge University Press, 2021, <https://www.ipcc.ch/assessment-report/ar6>.
- ¹⁰T. M. Lenton, *Nat. Clim. Change* **1**, 201 (2011).
- ¹¹T. M. Lenton, H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber, *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1786 (2008).
- ¹²E. Forgoston and R. O. Moore, *SIAM Rev.* **60**, 969 (2018).
- ¹³N. Berglund, “Kramers’ law: Validity, derivations and generalisations,” *Markov Processes Relat. Fields* **19**, 459–490 (2013); [arXiv:1106.5799](https://arxiv.org/abs/1106.5799) (2013).
- ¹⁴V. I. Arnold, “Mathematical methods of classical mechanics,” in *Graduate Texts in Mathematics*, 2nd ed. (Springer, New York, 1997), Vol. 60.
- ¹⁵R. Szrednicki, in *Handbook of Differential Equations: Ordinary Differential Equations* (Elsevier, 2004), Vol. 1, pp. 591–684.
- ¹⁶D. J. Higham, *SIAM Rev.* **43**, 525 (2001).
- ¹⁷D. Freedman and P. Diaconis, *Z. Wahrschein. Verwandte Gebiete* **57**, 453 (1981).
- ¹⁸Y. Dodge, in *The Concise Encyclopedia of Statistics* (Springer, New York, NY, 2008), pp. 283–287.
- ¹⁹Y. Yu, C. B. Muratov, and R. O. Moore, *IEEE Trans. Magn.* **55**, 7205011 (2019).
- ²⁰Y. Bakhtin, “On Gumbel limit for the length of reactive paths,” [arXiv:1312.1939](https://arxiv.org/abs/1312.1939) (2013).