Most probable transition paths in piecewise-smooth stochastic differential equations

Kaitlin Hill∗†, Jessica Zanetell‡, and John A. Gemmer‡§

Abstract

We develop a path integral framework for determining most probable paths for a class of systems of stochastic differential equations with piecewise-smooth drift and additive noise. This approach extends the Freidlin-Wentzell theory of large deviations to cases where the system is piecewise-smooth and may be non-autonomous. In particular, we consider an \( n \)-dimensional system with a switching manifold in the drift that forms an \( (n - 1) \)-dimensional hyperplane and investigate noise-induced transitions between metastable states on either side of the switching manifold. To do this, we mollify the drift and use \( \Gamma \)-convergence to derive an appropriate rate functional for the system in the piecewise-smooth limit. The resulting functional consists of the standard Freidlin-Wentzell rate functional, with an additional contribution due to times when the most probable path slides in a crossing region of the switching manifold. We explore implications of the derived functional through two case studies, which exhibit notable phenomena such as non-unique most probable paths and noise-induced sliding in a crossing region.

Keywords: Piecewise smooth dynamical systems, Filippov systems, Freidlin-Wentzell rate functional, Gamma-convergence, noise induced tipping, rare events

2000 MSC: 37H10, 37J45

1 Introduction

In dynamical systems, a tipping event is loosely defined as occurring when a sudden or small change to a variable or parameter induces a large change to the state of the system. While tipping is often studied within the context of climate applications [1–7], it also has broad applications in ecology [8, 9], ecosystems [10, 11] and epidemiology [12–15], to name a few. While there is no precise mathematical definition of a tipping event, in [16] it was proposed that tipping events could be classified according to whether the underlying mathematical mechanism involves, predominantly, a bifurcation (B-tipping), noise-induced transitions (N-tipping), or transitions between basins of attraction induced by fast changes in parameters (R-tipping), i.e. rate-induced tipping; see also [17–20].

The term ‘noise-induced tipping’ encompasses a range of phenomena such as transitions between metastable states, bursting, stochastic resonance, stochastic coherence and stochastic synchronization [21] and can be analyzed from several points of view including the Freidlin-Wentzell (FW) theory of large deviations, [22–24], the path integral framework [25, 26], transition path theory [27, 28], and formal asymptotics [29, 30]. In particular, for smooth autonomous dynamical systems

∗Department of Mathematics, St. Mary’s University, San Antonio, TX 78228, USA
†Corresponding author, Email: khill5@stmarytx.edu
‡Department of Mathematics, Wake Forest University, Winston-Salem, NC 27109, USA
§Email: gemmerj@wfu.edu
additively perturbed by Gaussian white noise, the FW theory provides a framework for computing most probable transitions as minimizers of a rate functional in the asymptotic limit of vanishing noise strength. The benefit of the FW framework is that most probable transition paths can be numerically computed using iterative schemes such as the string method for gradient systems [31], the minimum action method [23], the geometric minimum action method [32] and explicit gradient descent [33]. Moreover, knowledge of the most probable transition path can then be coupled with statistical techniques such as importance sampling to compute quantities of interest such as the expected time of tipping for nonzero noise [24, 34].

In this paper we extend the path integral framework to a class of differential equations with additive noise and piecewise-smooth drift. There has been considerable recent interest in understanding the dynamics of piecewise-smooth stochastic dynamical systems, particularly in relation to rare events and tipping, or transitions between metastable states, but also in varying applications [35–37]. In [38, 39], Chen et al. use the backward Fokker-Planck technique to derive the distribution of paths in a model of dry friction. Notably, in [38], Chen et al. use a similar method to that of the present study, smoothing out the piecewise-smooth vector field to numerically analyze the desingularized SDE. In [40, 41], Baule et al. use the path integral approach, as in the present study, to investigate most probable paths in a piecewise-smooth model of stick-slip friction. Beyond mechanical systems, the study of noise-induced tipping in piecewise-smooth systems has applications in biology [42, 43] and climate models [44–46, 48, 49]. To our knowledge, most probable paths in general stochastic piecewise-smooth systems have not yet been addressed.

1.1 Most probable paths in smooth systems

During rare event transitions from one metastable state to another, a stochastic system will follow paths according to some distribution, which in the limit of vanishing noise strength is generally singly-peaked along a most probable transition path. Following [22], for a smooth vector field $\mathbf{F}$, the most probable transition path between $\mathbf{x}_0$ and $\mathbf{x}_f$ can be defined as follows. We first define an admissible set of transition paths $A(t_0,t_f)$ by

$$A(t_0,t_f) = \{ \alpha \in H^1([t_0,t_f]; \mathbb{R}^n) : \alpha(t_0) = \mathbf{x}_0 \text{ and } \alpha(t_f) = \mathbf{x}_f \},$$

where $H^1([t_0,t_f]; \mathbb{R}^n)$ is the Hilbert space of weakly differentiable curves $\alpha$ satisfying $\alpha \in H^1$ if and only if $\int_{t_0}^{t_f} |\alpha|^2 + |\dot{\alpha}|^2 dt < \infty$. The most probable transition path $\alpha^* \in A(t_0,t_f)$ is then defined to be the global minimizer of the Freidlin-Wentzell rate functional $I^{(t_0,t_f)} : A(t_0,t_f) \mapsto \mathbb{R}$, given by

$$I^{(t_0,t_f)}[\alpha] = \int_{t_0}^{t_f} \|\dot{\alpha}(t) - \mathbf{F}(\alpha(t),t)\|^2 dt,$$  \hspace{1cm} (1)

so that $\alpha^* = \arg \min_{\alpha \in A(t_0,t_f)} I^{(t_0,t_f)}[\alpha]$. For this functional, the necessary condition satisfied by second differentiable minimizers is the following Euler-Lagrange equations [24]:

$$\ddot{\alpha} = \mathbf{F}_t + \dot{\alpha} (\nabla \mathbf{F} - \nabla \mathbf{F}^\top) + \nabla \mathbf{F}^\top \mathbf{F}.$$

Here, the subscript $t$ refers to the partial derivative with respect to time.

1.2 Framing the piecewise-smooth problem

We determine the most probable transition path of a noise-induced tipping event when the drift is piecewise-smooth. In general, for systems with a piecewise-smooth drift, the appropriate rate functional to minimize is not known. Minimizers of the Freidlin-Wentzell rate functional may not
be well-defined when a region of the vector field is not continuous, or even Lipschitz continuous. It is not clear how the most probable path may traverse across a switching manifold, which may have an attracting or repelling sliding region that introduces more complex dynamics. Specifically, letting $F^\pm : \mathbb{R}^n \to \mathbb{R}^n$ be smooth vector fields, we consider a system of stochastic differential equations of the form
\[ dx_t = F(x_t) dt + \sigma dW_t, \]  
where $x = (x, y) \in \mathbb{R}^n$ such that $x \in \mathbb{R}$ and $y \in \mathbb{R}^{n-1}$, $\sigma \in \mathbb{R}$, $W = (W_1, \ldots, W_n)$ is an $n$-dimensional Wiener process, and $F : \mathbb{R}^n \setminus \{x = 0\} \to \mathbb{R}^n$ is defined by
\[ F(x) = \begin{cases} F^+(x), & x > 0, \\ F^-(x), & x < 0. \end{cases} \]  
We let $S_+ = \{x \in \mathbb{R}^n : x > 0\}$, $S_- = \{x \in \mathbb{R}^n : x < 0\}$, and $\Sigma = \{x \in \mathbb{R}^n : x = 0\}$. The set $\Sigma$ is often called the switching manifold or discontinuity boundary [50]. In general, $\Sigma$ may be defined as the zero level set of a smooth function $H : \mathbb{R}^n \to \mathbb{R}$, but we assume for simplicity that $\Sigma$ is a hyperplane in $\mathbb{R}^{n-1}$.

We define the deterministic skeleton of Equation (2) as the dynamical system
\[ \dot{x} = F(x). \]  
We assume that there exist asymptotically stable fixed points $x_0 \in S_+$ and $x_f \in S_-$ of Equation (4), such that $F^+(x_0) = 0$ and $F^-(x_f) = 0$. By noise-induced transitions, we mean realizations of Equation (2) that transition from the basin of attraction of $x_0$ to that of $x_f$. More precisely, we define a noise-induced transition on the interval $[t_0, t_f]$ as a solution to the stochastic boundary value problem given by realizations of Equation (2) that satisfy $x(t_0) = x_0$ and $x(t_f) = x_f$. Here the boundary conditions mean that the process is conditioned to transition from $x_0$ to $x_f$.

1.3 Background on piecewise-smooth systems

Dynamics that occur entirely within the smooth regions $S_\pm$ can be fully described by regular (smooth) dynamical systems theory. On the other hand, dynamics on a switching manifold $\Sigma$ may not be defined and are typically imposed a posteriori, e.g. using Filippov’s convex combination [50–52]. In general, the lack of smoothness requirements across the switching manifold allows for more diverse phenomena than in smooth systems of similar dimension, since the vector field across $\Sigma$ may be discontinuous or even point in opposite directions. Regions of $\Sigma$ where either occurs are called sliding regions. A solution that reaches $\Sigma$ at a sliding region may “slide” along it. On the other hand, crossing regions occur on $\Sigma$ where sliding is not possible. We differentiate these regions using the following notation:

1. **Attracting sliding regions**: $\Sigma_A = \{x \in \Sigma : F^+_1(0, y) \leq 0, F^-_1(0, y) \geq 0\}$,
2. **Repelling sliding regions**: $\Sigma_R = \{x \in \Sigma : F^+_1(0, y) \geq 0, F^-_1(0, y) \leq 0\}$,
3. **Positive crossing regions**: $\Sigma_+ = \{x \in \Sigma : F^+_1(0, y) > 0\}$,
4. **Negative crossing regions**: $\Sigma_- = \{x \in \Sigma : F^+_1(0, y) < 0\}$,

where $F^+_1$ is the first component of $F^\pm$.

When solutions may slide along the switching manifold $\Sigma$, we impose a flow using the Filippov convex method [50, 52]. That is, we define the sliding flow as a convex combination of $F^\pm$,
\[ F^s(0, y) = \lambda F^+(0, y) + (1 - \lambda)F^-(0, y) \]  
(5)
with
\[
\lambda(y) = \frac{F_1^-(0,y)}{F_1^-(0,y) - F_1^+(0,y)} \in [0,1].
\]

Notice that Equation (6) fixes \(\lambda(y)\). This flow naturally arises in the context of our main result without \textit{a priori} imposing it; see Theorem 3.8. Also, note that \(F^s(0,y)\) may vanish for some point \((0,y) \in \Sigma\). Since it is not an equilibrium of the smooth flow, we say that such a point \(x_s = (0,y_s) \in \Sigma\) is a \textit{pseudoequilibrium} if it is an equilibrium of the sliding flow; i.e., for some \(\lambda \in (0,1), F^s(0,y_s) = 0\).

Filippov’s convex combination interprets solutions of System (3),(4) as a continuous curve \(x(t)\) satisfying the following conditions:
\[
\begin{align*}
\dot{x}(t) &= F(x), \\
x(0) &= x_0,
\end{align*}
\]
where \(F: \mathbb{R}^n \rightarrow \mathbb{R}^n\) is defined as
\[
F(x) = \begin{cases} 
F^+(x), & x \in S_+ \cup \Sigma_+,
F^-(x), & x \in S_- \cup \Sigma_-,
F^s(x), & x \in \Sigma_A \cup \Sigma_R.
\end{cases}
\]

That is, for points in time in which a solution curve \(x(t)\) intersects \(\Sigma\) it will either cross \(\Sigma\), tracking the flow of either \(F^+\) or \(F^-\) depending on whether the curve is crossing into \(S_+\) or \(S_-\), or it will track \(\Sigma\) in the sliding regions until entering a crossing region. Note, implicit in this definition is that solution curves are differentiable everywhere except on \(\Sigma\).

To provide a brief illustration, Figure 1(a)-(c) shows some possible dynamics across \(\Sigma\) in a two-dimensional system, with (a) \(\Sigma_+\), (b) \(\Sigma_{\pm}\) and \(\Sigma_A\), and (c) \(\Sigma_{\pm}\) and \(\Sigma_R\). In (a) the flow traverses across \(\Sigma\) continuously but not smoothly; in (b) and (c) the flow may also remain on \(\Sigma\), in intervals indicated by the blue line (for \(\Sigma_A\)) and the green line (for \(\Sigma_R\)). Note that the flow leaving a repelling sliding region is non-unique in forward time, while the flow leaving an attracting sliding region is non-unique in backward time. In addition, complex dynamics may appear on \(\Sigma\) itself, depending on the manner in which the imposed vector field on \(\Sigma\) transitions between \(S_+\) and \(S_-\). Figure 1(d) shows one example of a path traversing from a stable equilibrium in \(S_+\), through the switching manifold in an attracting sliding region using dynamics imposed using Filippov’s convex combination [52], then traversing to a stable equilibrium in \(S_-\). The system in this illustration is a piecewise-smooth version of the Lorenz 63 model,
\[
\begin{pmatrix}
\dot{x} \\
\dot{y} \\
\dot{z}
\end{pmatrix} = \begin{pmatrix}
\sigma_-(y-y_0-x+x_0) \\
(x-x_0)(\rho_--z+z_0)-y+y_0 \\
(x-x_0)(y-y_0)-\beta_-(z-z_0)
\end{pmatrix}, \quad x < 0, \quad \text{and} \quad \begin{pmatrix}
\dot{x} \\
\dot{y} \\
\dot{z}
\end{pmatrix} = \begin{pmatrix}
\sigma_+(y-x) \\
x(\rho_+--z)-y \\
yx-\beta_+z
\end{pmatrix}, \quad x > 0,
\]
originally studied in [55]. Here, \(\sigma_{\pm}, \rho_{\pm}, \beta_{\pm} > 0\) are parameters and \(x_0, y_0, z_0 \in \mathbb{R}\). For this example, \(S_{\pm}\) and \(\Sigma\) are defined as previously, with \(n = 3\).

### 1.4 Outline of analysis

For our problem, noise-induced transitions must cross \(\Sigma\) and the Freidlin-Wentzell rate functional, Equation (1), must be modified to account for the discontinuity in \(F\) and possible discontinuity-induced dynamics. For example, a most probable path that reaches a repelling sliding region \(\Sigma_R\) leaves it non-uniquely due to the non-uniqueness of the flow, as illustrated in the case studies in Sections 4 and 5. Additionally, if one were to naively piece together the most probable paths of
Most probable paths in piecewise-smooth SDEs

Figure 1: (a)-(c) Phase plane of a generic $F$ with $n = 2$, with (a) a crossing region, (b) crossing and attracting sliding, and (c) crossing and repelling sliding. (d) Phase diagram of a piecewise-smooth version of the Lorenz 63 model, System (9). Red points are equilibria of $S_{\pm}$, and the black point is a pseudoequilibrium of the imposed flow on $\Sigma$. The sliding region, $\Sigma_A$, is shaded blue. The black curve connects the two equilibria, going through the pseudoequilibrium at the origin and sliding in $\Sigma_A$. The blue curve is the sliding trajectory. Parameters used are $\sigma_+ = 10$, $\beta_+ = 2$, $\rho_\pm = 2$, $\sigma_- = 11$, $\beta_- = 3$, $x_0 = 1$, $y_0 = -1$, and $z_0 = 0$.

the smooth regions and neglect possible contribution from the manner in which the path crosses the switching manifold, this may lead to paths that do not reflect a global minimum of the rate functional.

We resolve these issues by smoothing out $F$ via mollification with a compactly supported, smooth, radially symmetric kernel of characteristic width $\varepsilon$, and consider the sequence of minimizers to the rate functional Equation (1) as $\varepsilon \to 0$. We show for any piecewise-smooth system of the form given by System (2),(3) that the most probable path minimizes Equation (1) in the smooth regions $S_{\pm}$, plus an additional functional whose contribution represents time spent sliding in a crossing region of the switching manifold. For $F = (F_1, G)$ such that $F_1 : \mathbb{R}^n \setminus \Sigma \to \mathbb{R}$ and $G : \mathbb{R}^n \setminus \Sigma \to \mathbb{R}^{n-1}$, and $\alpha = (\alpha, \beta) \in A^{(t_0, t_f)}$ and $0 < \varepsilon \ll 1$, the appropriate rate functional with both contributions is

$$
I^{(t_0, t_f)}[\alpha] = \int_{\{t \in S_+ \cup S_-\}} \|\dot{\alpha}(t) - F(\alpha(t))\|^2 dt + \int_{\{t \in \Sigma\}} \min_{\lambda \in [0, 1]} \left\{ \|\hat{\beta} - \lambda G^+(0, \beta) - (1 - \lambda)G^-(0, \beta)\|^2 \right\} dt.
$$

We make the following observations about this result:

1. The form of the above rate functional is independent of the chosen mollifier.

2. When there is a sliding region, the most probable path may track the Filippov dynamics; that is, no additional contribution is made during times where the most probable path slides via the imposed flow on the switching manifold.

We rigorously prove the above result using a technique from the calculus of variations called $\Gamma$-convergence to show that $I^{(t_0, t_f)}$ is the $\Gamma$-limit of the sequence of rate functionals for the mollified system [56, 57]. The $\Gamma$-limit is the natural notion of a limiting functional in this context in the sense
that minimizers of the rate functional with the mollified drift converge (weakly) to a minimizer of \( I^{(t_0, t_f)} \) in the limit as \( \varepsilon \to 0 \). While \( \Gamma \)-convergence has been used to study the convergence of minimizers of the Onsager-Machlup functional in the small noise limit [58–62], to our knowledge we are the first to use \( \Gamma \)-convergence to compute an appropriate extension of the Freidlin-Wentzell rate functional for piecewise smooth systems.

The utility of the derived rate functional is demonstrated through two case studies. Each case study also presents distinct phenomena in their most probable transition paths that are not possible in systems with smooth drift. The first case study analyzes the most probable transition path in a two-dimensional piecewise-linear system, a simple demonstrative case in which the typical rate functional formulation of the most probable path breaks down. Notably, in this case the most probable path follows the switching manifold along an attracting sliding region, but it does not follow the switching manifold along a repelling sliding region.

The second case study analyzes the most probable path in a simple one-dimensional periodically-forced piecewise-smooth system, constructed similarly to some conceptual models for Arctic energy balance [4, 63]. A significant portion of the analysis of the deterministic dynamics and the Monte Carlo simulations were first derived in Zanetell’s Master’s thesis [64]. Here we again observe the breakdown of the typical formulation of the most probable path. Additionally, we observe an emergent dynamical phenomenon, which we describe as noise-induced sliding, in which the most probable path follows the switching manifold in a crossing region, despite the energy cost of the associated rate functional.

In both case studies, we note in some regimes the most probable path(s) predicted by our result do not match observed Monte Carlo simulations. Several factors may contribute to this discrepancy, including needing a smaller noise strength \( \sigma \), a higher-order functional in the smoothing parameter \( \varepsilon \), or including the correction term in the Onsager-Machlup functional. Such possibilities will be explored in future work.

The paper is organized as follows: in Section 2 we briefly highlight notation we use and some technical assumptions on \( F \). Then in Section 3 we derive the mollified vector field and use \( \Gamma \)-convergence to determine the appropriate limiting functional for calculating the most probable path for stochastic differential equations of the form given by System (2),(3). Next, in Section 4 we provide a simple case study of a planar linear piecewise-smooth system as an illustration of the most probable path calculated in Section 3. Finally, in Section 5 we provide a case study that extends the most probable path derivation to the case of a one-dimensional non-autonomous system.

## 2 Notation and Assumptions

In this paper we use the following conventions for notation:

1. We use the convention that \( \| \cdot \| \) represents the \( L^2 \) norm for functions and \( | \cdot | \) represents the \( \ell^2 \) norm for vectors in \( \mathbb{R}^n \). Furthermore, we use \( \langle \cdot, \cdot \rangle \) to denote the inner product in \( \ell^2 \).

2. Subscripts refer to the index of a vector unless otherwise indicated. E.g., for \( F = (F_1, F_2, \ldots, F_n) \), we write the components as \( F_j, j = 1, \ldots, n \).

3. Unless noted otherwise, \( F \) may be non-autonomous. Although our analysis applies to non-autonomous systems, for simplicity of presentation we suppress the time dependence.

4. Since we assume that the underlying drift is only piecewise-smooth in the first component, we use a convenient shorthand that separates this first component from the remaining ones.
Namely, we write $x = (x, y) \in \mathbb{R}^n$, where $x \in \mathbb{R}$ and $y \in \mathbb{R}^{n-1}$, and $F = (F_1, G)$, where $F_1 : \mathbb{R}^n \setminus \{x = 0\} \to \mathbb{R}$ and $G : \mathbb{R}^n \setminus \{x = 0\} \to \mathbb{R}^{n-1}$.

5. For the path $\alpha \in A^{(t_0, t_f)}$, we separate the first component from the remaining ones as $\alpha = (\alpha, \beta)$, so that $\alpha \in H^1([t_0, t_f]; \mathbb{R})$, $\alpha(t_0) = x_0$, and $\alpha(t_f) = x_f$, where $x_{0,f} = (x_{0,f}, y_{0,f})$. Similarly, $\beta \in H^1([t_0, t_f]; \mathbb{R}^{n-1})$, so that $\beta(t_0) = y_0$ and $\beta(t_f) = y_f$.

6. We denote the Jacobian of a function $F(x, y)$ as $\nabla F(x, y)$. For the Jacobian comprised of only the partial derivatives with respect to the components of $y$, we write $\nabla_y F(x, y)$.

7. We indicate that a variable is being treated as fixed or pre-determined by separating it with a semicolon. E.g., if the value of $y$ has been set, then $F(x; y)$ is only a function of $x$.

We also make the following assumptions for each smooth vector field $F^\pm$. We will use these assumptions to prove two lemmas for the mollified vector field $F^\varepsilon$ in Section 3 which are necessary to establish existence of minimizers to the FW functional. Note, these assumptions also apply in the non-autonomous case but the explicit dependence on time in the argument of $F$ is suppressed.

**Assumptions.** Let $F^\pm : \mathbb{R}^n \setminus \{x = 0\} \to \mathbb{R}^n$ be smooth vector fields. We assume the following properties about $F^\pm$.

1. *(Growth conditions)* There exist $R_1, c_1, c_3, c_5 > 0$ and $c_2, c_4, c_6 \in \mathbb{R}$ such that $c_5 < c_1$ and for some $1 < p < \infty$, $|x| > R_1$ implies

$$
\begin{align*}
|F^-(x)| &\geq c_1|x|^p + c_2, \quad \text{if } x < 0, \\
|F^+(x)| &\geq c_1|x|^p + c_2, \quad \text{if } x > 0,
\end{align*}
$$

and

$$
\begin{align*}
\left|\frac{\partial F^-}{\partial x}(x)\right| &\leq c_3|x|^p + c_4, \quad \text{if } x < 0, \\
\left|\frac{\partial F^+}{\partial x}(x)\right| &\leq c_3|x|^p + c_4, \quad \text{if } x > 0,
\end{align*}
$$

and

$$
|F^+(0, y) - F^-(0, y)| \leq c_5|y|^p + c_6.
$$

2. *(Asymptotically inward-flowing)* There exist $R_2, c_7 > 0$ such that $|x| > R_2$ implies $F^\pm(x) \neq 0$ and

$$
\begin{align*}
\langle F^-(x), r(x) \rangle &< -c_7|F^-(x)|, \quad \text{if } x \leq 0, \\
\langle F^+(x), r(x) \rangle &< -c_7|F^+(x)|, \quad \text{if } x \geq 0,
\end{align*}
$$

where $r(x) = x/|x|$ is the normalized outward-pointing radial vector at $x$.

3 **Mollified Freidlin-Wentzell rate functional**

As discussed in Section 1, the standard form of the Freidlin-Wentzell rate functional cannot be directly used to calculate most probable paths in piecewise-smooth systems. In this section, we seek to ameliorate this deficiency by studying the convergence of the minimizer of the Freidlin-Wentzell functional, where we approximate Equation (3) as a smooth vector field obtained by mollification in the $x-$direction by a smooth, compactly supported, symmetric kernel. Specifically, by smoothing $F$ via mollification in $x$ over a region of width $2\varepsilon$, we obtain a sequence of functionals $I_\varepsilon^{(t_0, t_f)}$ and a corresponding sequence of minimizers $\alpha_\varepsilon \in A^{(t_0, t_f)}$, which we prove converge — up to a subsequence — to a minimizer of a limiting functional $I^{(t_0, t_f)}$ in the limit as $\varepsilon \to 0$. 


3.1 Mollified functional and existence of a minimum

We begin this section by recalling the definition of a (Friedrichs) mollifier, following the presentation in [65]. Let \( \zeta: \mathbb{R} \to \mathbb{R} \) be a smooth even bump function, defined by

\[
\zeta(x) = \begin{cases} 
Ch(x), & |x| < 1, \\
0, & |x| \geq 1,
\end{cases}
\]

(14)

where \( h: \mathbb{R} \to \mathbb{R} \) is a smooth function satisfying \( h(x) \geq 0, h(-1) = h(1) = 0 \) and all of its derivatives vanish at \( x = -1 \) and \( x = 1 \), e.g., \( h(x) = \exp \left( 1/(|x|^2 - 1) \right) \). Additionally, \( C > 0 \) is determined so that \( \int_{-\infty}^{\infty} \zeta(x) \, dx = 1 \); i.e., \( C = 1/\int_{-\infty}^{\infty} h(x) \, dx \). For each \( \varepsilon > 0 \), we define a sequence of functions \( \zeta_\varepsilon(x) = \zeta(x/\varepsilon)/\varepsilon \) which satisfy \( \int_{-\varepsilon}^{\varepsilon} \zeta_\varepsilon(x) \, dx = 1 \) and are compactly supported on \([-\varepsilon, \varepsilon]\). The mollification \( \mathbf{F}^\varepsilon \) of \( \mathbf{F} \) in \( x \) is then defined by convolution with \( \zeta_\varepsilon \) in the \( x \)-direction:

\[
\mathbf{F}^\varepsilon(x, y) = \zeta_\varepsilon(x) \ast \mathbf{F}(x, y) = \int_{-\infty}^{0} \zeta_\varepsilon(x-s) \mathbf{F}^-(s, y) \, ds + \int_{0}^{\infty} \zeta_\varepsilon(x-s) \mathbf{F}^+(s, y) \, ds,
\]

\[
= \begin{cases} 
-\int_{-\varepsilon}^{0} \zeta_\varepsilon(u) \mathbf{F}^-(x-u, y) \, du, & x \leq -\varepsilon, \\
\int_{-\varepsilon}^{\varepsilon} \zeta_\varepsilon(u) \mathbf{F}^+(x-u, y) \, du + \int_{x}^{\varepsilon} \zeta_\varepsilon(u) \mathbf{F}^-(x-u, y) \, du, & |x| < \varepsilon, \\
\int_{-\varepsilon}^{\varepsilon} \zeta_\varepsilon(u) \mathbf{F}^+(x-u, y) \, du, & x \geq \varepsilon,
\end{cases}
\]

(15)

where \( \mathbb{1}_A \) is the indicator function on the set \( A \). The mollified Freidlin-Wentzell rate functional \( I_{\varepsilon}^{(t_0,t_f)}: \mathcal{A}^{(t_0,t_f)} \mapsto \mathbb{R} \) is then defined by:

\[
I_{\varepsilon}^{(t_0,t_f)}[\mathbf{\alpha}] = \int_{t_0}^{t_f} |\dot{\mathbf{\alpha}} - \mathbf{F}^\varepsilon(\mathbf{\alpha})|^2 \, dt.
\]

(16)

Note, by construction, \( \mathbf{F}^\varepsilon(x) \) is a smooth function that converges pointwise to \( \mathbf{F}(x) \) as \( \varepsilon \to 0 \) except at \( x = 0 \) [65]. However, since \( \mathbf{F}(x) \) is not continuous this convergence is necessarily not uniform and thus it is not a priori clear which properties of \( \mathbf{F} \) are preserved under mollification. The following lemmas ensure that \( \mathbf{F}^\varepsilon \) also has \( p \)-growth and is asymptotically inward-flowing. These properties are crucial to establishing the existence of a minimum for \( I_{\varepsilon}^{(t_0,t_f)} \), but to streamline our main results, the proofs of these lemmas are in Appendix A.

**Lemma 3.1** (\( p \)-growth). There exists \( \varepsilon^* \) such that for all \( \varepsilon > 0 \) satisfying \( \varepsilon < \varepsilon^* \), there exist \( R_1^\varepsilon, c_1^\varepsilon > 0 \) and \( c_2^\varepsilon \in \mathbb{R} \) such that \( |x| > R_1^\varepsilon \) implies

\[
|\mathbf{F}^\varepsilon(x)| \geq c_1^\varepsilon |x|^p + c_2^\varepsilon
\]

for some \( p > 1 \).

**Lemma 3.2** (Asymptotically inward-flowing). For all \( \varepsilon > 0 \), there exist \( R_2^\varepsilon, c_2^\varepsilon > 0 \) such that \( |x| > R_2^\varepsilon \) implies \( \mathbf{F}^\varepsilon(x) \neq 0 \) and

\[
\langle \mathbf{F}^\varepsilon(x), \mathbf{r}(x) \rangle < -c_2^\varepsilon |\mathbf{F}^\varepsilon(x)|,
\]

where \( \mathbf{r}(x) = x/|x| \) is the normalized outward-pointing radial vector at \( x \).
We now prove the existence of a minimizer of \( I_\varepsilon^{(t_0,t_f)} \) using Tonelli’s direct method of the calculus of variations; see, e.g., [57, 66]. This procedure consists of first proving that a minimizing sequence of \( I_\varepsilon^{(t_0,t_f)} \) is bounded with respect to the \( H^1 \) norm and thus has a weakly convergent subsequence in the \( H^1 \) topology. Note, it is necessary to work in the \( H^1 \) topology since closed and bounded sets are not necessarily compact in the strong topology. Nevertheless, upon passing to a subsequence, we can establish a limit point for the minimizing sequence. Finally, by proving that \( I_\varepsilon^{(t_0,t_f)} \) is lower semi-continuous with respect to weak convergence in the \( H^1 \) topology, it follows that limit points of the minimizing sequence are indeed minimizers.

To begin the procedure for the direct method of the calculus of variations outlined in the previous paragraph, we first recall the definition of weak convergence in the \( H^1 \) topology in the context of our problem.

**Definition 3.3 (Weak convergence in \( H^1 \) [67]).** A sequence \( \{ \alpha_m \}_{m=1}^{\infty} \subset H^1([t_0,t_f]; \mathbb{R}^n) \) converges weakly to \( \alpha^* \in H^1([t_0,t_f]; \mathbb{R}^n), \) written

\[
\alpha_m \rightharpoonup \alpha^* \quad \text{in} \quad H^1([t_0,t_f]; \mathbb{R}^n)
\]

provided \( \alpha_m \rightharpoonup \alpha^* \) in \( L^2([t_0,t_f]; \mathbb{R}^n) \) and \( \alpha_m \rightharpoonup \alpha^* \) in \( L^2([t_0,t_f]; \mathbb{R}^n). \) That is, for all \( \omega \in L^2([t_0,t_f]; \mathbb{R}^n), \)

\[
\lim_{m \to \infty} \int_{t_0}^{t_f} \langle \alpha_m, \omega \rangle \, dt = \int_{t_0}^{t_f} \langle \alpha^*, \omega \rangle \, dt \quad \text{and} \quad \lim_{m \to \infty} \int_{t_0}^{t_f} \langle \dot{\alpha}_m, \omega \rangle \, dt = \int_{t_0}^{t_f} \langle \dot{\alpha}^*, \omega \rangle \, dt.
\]

The following lemma establishes that boundedness of \( I_\varepsilon^{(t_0,t_f)}[\alpha_m] \) implies boundedness of \( \alpha_m \) with respect to the \( H^1 \) norm.

**Lemma 3.4.** There exists \( \varepsilon^* > 0 \) such that if \( \varepsilon < \varepsilon^* \) and if \( \alpha_m \in A \) satisfies \( I_\varepsilon^{(t_0,t_f)}[\alpha_m] < M_1 \) for some \( M_1 > 0, \) then there exists \( M_2 > 0 \) such that \( \| \alpha_m \|_{H^1} < M_2. \)

**Proof.** For contradiction, suppose \( \alpha_m \) is unbounded with respect to the \( H^1 \) norm. Therefore, there exists a subsequence \( \alpha_{m_k} \) such that \( \| \alpha_{m_k} \|_{H^1} \to \infty. \)

1. Suppose there exists \( M > 0 \) such that \( \| \alpha_{m_k} \|_{\infty} < M. \) Consequently, since \( F^\varepsilon \) is smooth there exists \( C_1 > 0 \) such that \( |F^\varepsilon(\alpha_{m_k})| < C_1. \) By the Cauchy-Schwarz inequality,

\[
I_\varepsilon^{(t_0,t_f)}[\alpha_{m_k}] = \int_{t_0}^{t_f} \left[ |\dot{\alpha}_{m_k}|^2 - 2\langle \alpha_{m_k}, F^\varepsilon(\alpha_{m_k}) \rangle + |F^\varepsilon(\alpha_{m_k})|^2 \right] \, dt
\]

\[
\geq \int_{t_0}^{t_f} \left[ |\dot{\alpha}_{m_k}|^2 - 2\langle \alpha_{m_k}, F^\varepsilon(\alpha_{m_k}) \rangle \right] \, dt
\]

\[
\geq \| \alpha_{m_k} \|^2 - 2C_1(t_f - t_0)^{\frac{1}{2}} \| \alpha_{m_k} \|.
\]

Therefore, since \( \| \alpha_{m_k} \|_{H^1} \to \infty, \) it follows from Poincaré’s inequality that \( \| \alpha_{m_k} \| \to \infty. \) Thus, \( I_\varepsilon^{(t_0,t_f)} \to \infty, \) which is a contradiction.

2. Suppose \( \| \alpha_{m_k} \|_{\infty} \) is unbounded. Upon passing to another subsequence which we also label \( \alpha_{m_k}, \) it follows that \( \| \alpha_{m_k} \|_{\infty} \to \infty. \) Applying the elementary inequality \( a^2 + b^2 \geq 2ab, \) it follows that

\[
I_\varepsilon^{(t_0,t_f)}[\alpha_{m_k}] \geq 2 \int_{t_0}^{t_f} \left[ \| \dot{\alpha}_{m_k} \| |F^\varepsilon(\alpha_{m_k})| - \langle \alpha_{m_k}, F^\varepsilon(\alpha_{m_k}) \rangle \right] \, dt
\]

\[
= 2 \int_{t_0}^{t_f} \left[ |\dot{\alpha}_{m_k}||F^\varepsilon(\alpha_{m_k})| \sin^2 \left( \frac{\theta}{2} \right) \right] \, dt,
\]

where \( \theta \) is the angle between \( \alpha_{m_k} \) and \( F^\varepsilon(\alpha_{m_k}) \).
where $\theta(t)$ is the angle between $\alpha_{m_k}(t)$ and $\mathbf{F}^\varepsilon(\alpha_{m_k}(t))$.

We now choose $\varepsilon^\ast, R_1^\varepsilon$ as in Lemma 3.1 and $R_2^\varepsilon$ as in Lemma 3.2, set $R^\varepsilon = \max\{R_1^\varepsilon, R_2^\varepsilon\}$, and let $A_k = \{ t : |\alpha_{m_k}(t)| > R^\varepsilon \text{ and } \frac{d}{dt}|\alpha_{m_k}(t)| \geq 0 \}$. It follows from Lemmas 3.1 and 3.2 that there exists a constant $C_2 > 0$ such that

$$I_{\varepsilon}^{(t_0,t_f)}[\alpha_{m_k}] \geq C_2 \int_{A_k} |\dot{\alpha}_{m_k}||\mathbf{F}^\varepsilon(\alpha_{m_k})|dt$$

$$\geq C_2 \int_{A_k} |\dot{\alpha}_{m_k}||\alpha_{m_k}|^{p-1}dt$$

$$= C_2 \left( \max_{t \in [t_0,t_f]} |\alpha_{m_k}| - (R^\varepsilon)^{-1} \right).$$

Applying the Cauchy-Schwarz inequality and the fundamental theorem of calculus, we obtain

$$I_{\varepsilon}^{(t_0,t_f)}[\alpha_{m_k}] \geq C_2 \int_{A_k} \langle \dot{\alpha}_{m_k}, \alpha_{m_k} \rangle |\alpha|^{p-1}dt$$

$$= C_2 \int_{A_k} \frac{d}{dt}|\alpha|^{p+1}dt$$

$$= C_2 \max_{t \in [t_0,t_f]} |\alpha_{m_k}|^{p+1} - (R^\varepsilon)^{p+1}.$$

Since $\|\alpha_{m_k}\|_\infty \to \infty$ implies $\max_{t \in [t_0,t_f]} |\alpha_{m_k}(t)| \to \infty$, it follows that $I_{\varepsilon}^{(t_0,t_f)}[\alpha_{m_k}] \to \infty$, which is a contradiction. It follows from items 1 and 2 that $\alpha_m$ is bounded with respect to the $H^1$ norm.

Finally, we use the direct method of the calculus of variations to prove the existence of a minimum.

**Theorem 3.5 (Existence of minimizer of the mollified functional).** There exists an $\varepsilon^\ast > 0$ such that if $\varepsilon < \varepsilon^\ast$, then there exists an $\alpha^\ast \in A^{(t_0,t_f)}$ such that for all $\alpha \in A^{(t_0,t_f)}$,

$$I_{\varepsilon}^{(t_0,t_f)}[\alpha^\ast] \leq I_{\varepsilon}^{(t_0,t_f)}[\alpha].$$

**Proof.** Let $\varepsilon > 0$, and let $\alpha_m$ be a minimizing sequence of $I_{\varepsilon}^{(t_0,t_f)}[\alpha]$; i.e.,

$$\lim_{m \to \infty} I_{\varepsilon}^{(t_0,t_f)}[\alpha_m] = \inf_{\alpha \in A^{(t_0,t_f)}} I_{\varepsilon}^{(t_0,t_f)}[\alpha].$$

Consequently, there exists $M_1 > 0$ such that $I_{\varepsilon}^{(t_0,t_f)}[\alpha_m] < M_1$ and thus by Lemma 3.4 there exists $M_2 > 0$ such that $\|\alpha_m\|_{H^1} < M_2$. Thus, by the Banach–Alaoglu theorem and the Rellich-Kondrachov theorem there exists $\alpha^\ast \in A^{(t_0,t_f)}$ and a subsequence $\alpha_{m_k}$ such that $\alpha_{m_k} \overset{L^2}{\to} \alpha^\ast$ and $\alpha_{m_k} \overset{L^\infty}{\to} \alpha^\ast$ as $k \to \infty$ [68]. Since the integrand appearing in $I_{\varepsilon}^{(t_0,t_f)}$ is convex with respect to $\alpha$, it follows that $I_{\varepsilon}^{(t_0,t_f)}$ is lower semicontinuous with respect to weak convergence [67], and therefore

$$I_{\varepsilon}^{(t_0,t_f)}[\alpha^\ast] \leq \lim_{k \to \infty} \inf I_{\varepsilon}^{(t_0,t_f)}[\alpha_{m_k}] = \lim_{m \to \infty} I_{\varepsilon}^{(t_0,t_f)}[\alpha_m] = \inf_{\alpha \in A^{(t_0,t_f)}} I_{\varepsilon}^{(t_0,t_f)}[\alpha].$$

\[\square\]
3.2 Euler-Lagrange equations

Now that the existence of minimizers is established for the Freidlin-Wentzell rate \( I^{(t_0,t_f)}_\varepsilon[\alpha] \) corresponding to the mollified system \( \dot{x} = F^\varepsilon \), we can establish a necessary condition for minimizers, which is that \( \alpha \in A^{(t_0,t_f)} \) is a \( C^2 \) integral curve of the flow generated by the following system of Euler-Lagrange (EL) equations:

\[
\dot{x} = F_t^\varepsilon + (\nabla F^\varepsilon - \nabla F^\varepsilon^\top) \dot{x} + \nabla F^\varepsilon^\top F^\varepsilon,
\]

where the subscript \( t \) denotes the partial derivative with respect to time. It is convenient to introduce the scaling \( z = x/\varepsilon \), \( z \in [-1,1] \), so the EL Equations (17) become the fast-slow system:

\[
\begin{align*}
\varepsilon \dot{z} &= F_1^\varepsilon + (\nabla_y F_1^\varepsilon - G_z^\varepsilon) \dot{y} + G_z^\varepsilon^\top G^\varepsilon, \\
\dot{y} &= G_1^\varepsilon + (G_z^\varepsilon - \nabla_y F_1^\varepsilon) \varepsilon \dot{z} + (\nabla_y G^\varepsilon - \nabla_y G^\varepsilon^\top) \dot{y} + \nabla_y G^\varepsilon^\top G^\varepsilon.
\end{align*}
\]

Introducing the conjugate momenta \( \varphi = \varepsilon \dot{z} - F_1^\varepsilon \) and \( \psi = \dot{y} - G^\varepsilon \) leads to the Hamiltonian form,

\[
\begin{align*}
\varepsilon \dot{z} &= F_1^\varepsilon + \varphi, \\
\dot{y} &= G^\varepsilon + \psi, \\
\varphi &= -\varepsilon \dot{F}_{1,z}^\varepsilon - \langle \psi, G_z^\varepsilon \rangle, \\
\dot{\psi} &= -\varepsilon \nabla_y F_1^\varepsilon - \langle \psi, \nabla_y G^\varepsilon \rangle.
\end{align*}
\]

The most probable path of System (2),(3) corresponds to solutions of the EL Equations (18) subject to boundary conditions for appropriate values of \( t \). In the case studies of Sections 4 and 5, we use the gradient flow to numerically solve System (18) when an explicit solution is not available.

3.3 Limiting functional

In this section, we consider the problem of computing an appropriate limiting functional \( \bar{T}^{(t_0,t_f)} : A^{(t_0,t_f)} \mapsto \mathbb{R} \). Theorem 3.5 guarantees the existence of a minimizer of \( I^{(t_0,t_f)}_\varepsilon[\alpha] \) for fixed \( \varepsilon \) and is given by \( \alpha^\varepsilon = \arg \min_{\alpha \in A^{(t_0,t_f)}} I^{(t_0,t_f)}_\varepsilon[\alpha(t)] \). If we now consider the sequence of minimizers \( \alpha^\varepsilon \) and suppose there exists \( \alpha^* \in A^{(t_0,t_f)} \) such that \( \alpha^\varepsilon \rightharpoonup H^1 \alpha^* \), an appropriate limiting functional \( \bar{T}^{(t_0,t_f)} \) should have the property that \( \alpha^* \) minimizes \( \bar{T}^{(t_0,t_f)} \). Introduced by de Giorgi, the \( \Gamma \)-limit is the precise notion of a limiting functional that ensures this property [56, 57]. The following definition of a \( \Gamma \)-limit is adapted from [57] for our specific problem.

**Definition 3.6** (\( \Gamma \)-Convergence [57]). We say that \( I^{(t_0,t_f)}_\varepsilon \) \( \Gamma \)-converges to \( \bar{T}^{(t_0,t_f)} : A^{(t_0,t_f)} \mapsto \mathbb{R} \) with respect to \( H^1 \) weak convergence if for all \( \alpha \in A^{(t_0,t_f)} \), we have

1. (liminf inequality) for every sequence \( \alpha^\varepsilon \in A^{(t_0,t_f)} \) satisfying \( \alpha^\varepsilon \rightharpoonup H^1 \alpha \),

\[
\bar{T}^{(t_0,t_f)}[\alpha] \leq \liminf_{\varepsilon \to 0} I^{(t_0,t_f)}_\varepsilon[\alpha^\varepsilon];
\]

2. (recovery sequence) there exists a sequence \( \alpha^\varepsilon \rightharpoonup H^1 \alpha \) such that

\[
\lim_{\varepsilon \to 0} I^{(t_0,t_f)}_\varepsilon[\alpha^\varepsilon] = \bar{T}^{(t_0,t_f)}[\alpha].
\]
If \( I_{\epsilon}^{(t_0,t_f)} \) \( \Gamma \)-converges to \( \overline{T}^{(t_0,t_f)} \) we write \( \Gamma - \lim_{\epsilon \to 0} I_{\epsilon}^{(t_0,t_f)} = \overline{T}^{(t_0,t_f)} \).

Showing both the liminf inequality and the existence of a recovery sequence establish that both the limsup and the liminf of \( I_{\epsilon}^{(t_0,t_f)} \) over all weakly convergent sequences exists and is equal to \( \overline{T}^{(t_0,t_f)} \). Indeed, these two conditions ensure that \( I_{\epsilon}^{(t_0,t_f)} \) has a common lower bound, the lower bound is optimal, and the limiting functional is precisely this lower bound. Moreover, the definition of \( \Gamma \)-convergence is constructed in such a manner that ensures the convergence of minimizers of \( I_{\epsilon}^{(t_0,t_f)} \) to the minimizer of \( \overline{T}^{(t_0,t_f)} \) if an additional equicoercivity condition is satisfied. The following theorem adapted from [57] ensures this convergence for our specific problem.

**Theorem 3.7.** Suppose \( I_{\epsilon}^{(t_0,t_f)} \) \( \Gamma \)-converges to \( \overline{T}^{(t_0,t_f)} : A^{(t_0,t_f)} \to \mathbb{R} \) as \( \epsilon \to 0 \) with respect to \( H^1 \) weak convergence and suppose for all \( \alpha^{\epsilon}_* \) that minimize \( I_{\epsilon}^{(t_0,t_f)} \) there exists a constant \( M > 0 \) such that \( \|\alpha^{\epsilon}_*\|_{H^1} < M \). Then there exists an \( \alpha^* \in A^{(t_0,t_f)} \) such that

\[
\inf_{\alpha \in A^{(t_0,t_f)}} \overline{T}^{(t_0,t_f)}[\alpha] = \overline{T}^{(t_0,t_f)}[\alpha^*].
\]

Moreover, every limit of a subsequence of \( \alpha^{\epsilon}_* \) as \( \epsilon \to 0 \) is a minimizer of \( \overline{T}^{(t_0,t_f)} \).

One of the challenges with proving a \( \Gamma \)-limit is that the target limiting functional must be formed from an educated guess based on analysis of minimizers. In the system of interest, it is natural to expect that the appropriate \( \Gamma \)-limit will correspond to the standard FW functional except for intervals of time on which the curve tracks \( \Sigma \). However, for \( \alpha \in A^{(t_0,t_f)} \) satisfying \( \{ t \in \mathbb{R} : \alpha(t) = 0 \} \) has nonzero measure, the recovery sequence must be constructed to minimize the functional for times in which \( \alpha(t) \in \Sigma^\epsilon = \{ x \in \mathbb{R}^n : |x| < \epsilon \} \). To compute the appropriate \( \Gamma \)-limit for \( I_{\epsilon}^{(t_0,t_f)} \), we introduce the following intervals of time:

\[
I[\alpha] = \{ t \in (t_0,t_f) : \alpha(t) \notin \Sigma \} = \{ t \in (t_0,t_f) : \alpha_1(t) \neq 0 \},
\]

\[
I^\epsilon[\alpha] = \{ t \in (t_0,t_f) : \alpha(t) \notin \Sigma^\epsilon \} = \{ t \in (t_0,t_f) : |\alpha_1(t)| \geq \epsilon \},
\]

\[
I_{\Sigma}^\epsilon[\alpha] = \{ t \in (t_0,t_f) : \alpha(t) \in \Sigma \} = \{ t \in (t_0,t_f) : \alpha_1(t) = 0 \},
\]

\[
I_{\Sigma^+}[\alpha] = \{ t \in (t_0,t_f) : \alpha(t) \in \Sigma^+ \} = \{ t \in (t_0,t_f) : \alpha_1(t) = 0 \text{ and } F_1^+(0,\beta) > 0 \},
\]

\[
I_{\Sigma^-}[\alpha] = \{ t \in (t_0,t_f) : \alpha(t) \in \Sigma^- \} = \{ t \in (t_0,t_f) : \alpha_1(t) = 0 \text{ and } F_1^-(0,\beta) < 0 \},
\]

\[
I^\pm[\alpha] = \{ t \in (t_0,t_f) : \alpha(t) \in \Sigma^\pm \} = \{ t \in (t_0,t_f) : |\alpha_1(t)| < \epsilon \},
\]

\[
I_{\pm}[\alpha] = \{ t \in (t_0,t_f) : \alpha(t) \in S^\pm \},
\]

\[
I^\pm_{\pm}[\alpha] = \{ t \in (t_0,t_f) : \alpha(t) \in S^\pm_\pm \},
\]

as well as introduce the function \( \Lambda : \mathbb{R} \to \mathbb{R} \), defined by

\[
\Lambda(z) = \frac{1}{2} + \int_0^z \zeta(u) \, du.
\]

Note that \( \Lambda(z) = 1 \) if \( z > 1 \), \( \Lambda(z) = 0 \) if \( z < -1 \), and \( \Lambda(z) \) smoothly transitions from 0 to 1 on the interval \( [-1,1] \). We now state the following \( \Gamma \)-limit result.

**Theorem 3.8.** There exists \( \overline{T}^{(t_0,t_f)} : A^{(t_0,t_f)} \to \mathbb{R} \) such that

\[
\Gamma - \lim_{\epsilon \to 0} I_{\epsilon}^{(t_0,t_f)} = \overline{T}^{(t_0,t_f)},
\]
where for $\alpha = (\alpha, \beta) \in A(t_0, t_f)$,

$$
\mathcal{I}^{(t_0,t_f)}[\alpha] = \int_{\mathcal{I}^+_\alpha} \| \dot{\alpha}(t) - \mathbf{F}(\alpha(t)) \|^2 dt \\
+ \int_{\mathcal{I}^-} \min_{\lambda \in [0,1]} \left\{ \left[ \lambda F_1^+(0, \beta) + (1 - \lambda) F_1^-(0, \beta) \right]^2 \\
+ \left| \dot{\beta} - \lambda \mathbf{G}^+(0, \beta) - (1 - \lambda) \mathbf{G}^-(0, \beta) \right|^2 \right\} dt.
$$

(20)

Proof. For fixed $\alpha = (\alpha, \beta) \in A(t_0, t_f)$, we first show the existence of a recovery sequence, $\alpha^\varepsilon$. We will construct $\alpha^\varepsilon$ over three intervals, $\mathcal{I}^+_\varepsilon[\alpha], \mathcal{I}^-_\varepsilon[\alpha]$, and $\mathcal{I}^0_\varepsilon[\alpha]$, by

$$
\alpha^\varepsilon(t) = \alpha(t) 1_{\{ t \in \mathcal{I}^+_\varepsilon[\alpha] \}} + \alpha(t) 1_{\{ t \in \mathcal{I}^-_\varepsilon[\alpha] \}} + (\varepsilon z(t), \beta(t)) 1_{\{ t \in \mathcal{I}^0_\varepsilon[\alpha] \}} \\
= \alpha(t) 1_{\{ \alpha \in S^\varepsilon_+ \}} + \alpha(t) 1_{\{ \alpha \in S^\varepsilon_- \}} + (\varepsilon z(t), \beta(t)) 1_{\{ \varepsilon z, \beta \in V^\varepsilon \}},
$$

where $z(t)$ is given by

$$
z = \arg \min_{z' \in [-1,1]} \left\{ \left[ \Lambda(z') F_1^+(0, \beta) + (1 - \Lambda(z')) F_1^-(0, \beta) \right]^2 \\
+ \left| \dot{\beta} - \Lambda(z') \mathbf{G}^+(0, \beta) - (1 - \Lambda(z')) \mathbf{G}^-(0, \beta) \right|^2 \right\}.
$$

(21)

To prove convergence of $\mathcal{I}^{(t_0,t_f)}[\alpha^\varepsilon]$ to $\mathcal{I}^{(t_0,t_f)}[\alpha]$ and $\alpha^\varepsilon \overset{H^1}{\rightarrow} \alpha$, we first consider the interval of time $\mathcal{I}^\varepsilon_+[\alpha]$. Since $\alpha^\varepsilon(t) 1_{\{ \alpha \in S^\varepsilon_+ \}} = \alpha(t) 1_{\{ \alpha \in S^\varepsilon_+ \}}$, it follows that $\alpha^\varepsilon(t) 1_{\{ \alpha \in S^\varepsilon_+ \}} \overset{H^1}{\rightarrow} \alpha(t) 1_{\{ \alpha \in S^\varepsilon_+ \}}$. Moreover, we have

$$
\int_{\mathcal{I}^\varepsilon_+[\alpha]} \| \dot{\alpha} - \mathbf{F}^\varepsilon(\alpha^\varepsilon) \|^2 dt - \int_{\mathcal{I}^+_\alpha} \| \dot{\alpha} - \mathbf{F}^+(\alpha) \|^2 dt \\
= \int_{\mathcal{I}^\varepsilon_+[\alpha] \cap \mathcal{I}^+_\alpha} \left(2\langle \dot{\alpha}, \mathbf{F}^\varepsilon(\alpha) - \mathbf{F}^\varepsilon(\alpha^\varepsilon) \rangle + \| \mathbf{F}^\varepsilon(\alpha) - \mathbf{F}^\varepsilon(\alpha^\varepsilon) \|^2 - \| \mathbf{F}^+(\alpha) \|^2 \right) dt \\
+ \int_{\mathcal{I}^\varepsilon_+[\alpha] \cap \mathcal{I}^-_\alpha} \| \dot{\alpha} - \mathbf{F}^\varepsilon(\alpha) \|^2 dt - \int_{\mathcal{I}^+_\alpha \setminus \mathcal{I}^\varepsilon_+[\alpha]} \| \dot{\alpha} - \mathbf{F}^+(\alpha) \|^2 dt.
$$

Since $\mathbf{F}^\varepsilon(x,y) 1_{\{ x \geq \varepsilon \}} = \int_{-1}^1 \zeta(v) \mathbf{F}^+(x - v, y) dv 1_{\{ x \geq \varepsilon \}}$ and $\int_{-1}^1 \zeta(v) \mathbf{F}^+(x - v, y) dv$ converges uniformly to $\mathbf{F}^+(x,y)$ as $\varepsilon \to 0$, the integral $\int_{\mathcal{I}^\varepsilon_+[\alpha] \cap \mathcal{I}^-_\alpha} \langle \dot{\alpha}, \mathbf{F}^\varepsilon(\alpha) - \mathbf{F}^\varepsilon(\alpha^\varepsilon) \rangle dt$ vanishes as $\varepsilon \to 0$. By Minkowski’s inequality and the triangle inequality,

$$
\left| \int_{\mathcal{I}^\varepsilon_+[\alpha] \cap \mathcal{I}^+_\alpha} \| \mathbf{F}^\varepsilon(\alpha) \|^2 dt - \| \mathbf{F}^+(\alpha) \|^2 dt \right| \\
\leq \int_{\mathcal{I}^\varepsilon_+[\alpha] \cap \mathcal{I}^-_\alpha} \| \mathbf{F}^\epsilon(\alpha) \|^2 - \| \mathbf{F}^+(\alpha) \|^2 dt \\
\leq \int_{\mathcal{I}^\varepsilon_+[\alpha] \cap \mathcal{I}^-_\alpha} \| \mathbf{F}^\epsilon(\alpha) - \mathbf{F}^+(\alpha) \|^2 dt,
$$

and thus by the same uniform convergence argument as above vanishes as $\varepsilon \to 0$. Since the measures of $\mathcal{I}^\varepsilon_+[\alpha] \setminus \mathcal{I}^+_\alpha$ and $\mathcal{I}^+_\alpha \setminus \mathcal{I}^\varepsilon_+[\alpha]$ vanish as $\varepsilon \to 0$ it follows that the remaining integrals also
vanish. The analogous case for $I_{-} [\alpha]$ is identical. Hence,

$$
\lim_{\varepsilon \to 0} I_{-}^{(\varepsilon, t_f)} \{ \alpha \varepsilon \mathbb{1}_{\{ \alpha \varepsilon \in \mathcal{S}_g \}} \} = \lim_{\varepsilon \to 0} \int_{I_{-}[\alpha]} \| \dot{\alpha} - \mathbf{F}^\varepsilon (\alpha) \|^2 \, dt
$$

$$
= \int_{I_{-}[\alpha]} \| \dot{\alpha} - \mathbf{F}(\alpha) \|^2 \, dt
$$

$$
= \mathcal{I}^{(l_0, t_f)} \{ \alpha \mathbb{1}_{\{ \alpha \in \Sigma \}} \}.
$$

We now consider the interval of time $I_{+}^{\varepsilon} [\alpha]$. Note that using the substitution $v = u/\varepsilon$, we have

$$
\lim_{\varepsilon \to 0} \int_{I_{+}^{\varepsilon} \{ \{\varepsilon z < \varepsilon \}} \frac{1}{\varepsilon} \int_{-1}^{1} \zeta(v) \mathbf{F}^\varepsilon (\varepsilon z - \varepsilon v, y) \, dv + \int_{-1}^{1} \zeta(v) \mathbf{F}^\varepsilon (\varepsilon z - \varepsilon v, y) \, dv
$$

$$
= \Lambda (z) \mathbf{F}^+(0, y) + (1 - \Lambda (z)) \mathbf{F}^-(0, y).
$$

Let

$$
\mathbf{H}(y; z) = \Lambda (z) \mathbf{F}^+(0, y) + (1 - \Lambda (z)) \mathbf{F}^-(0, y) = \lambda \mathbf{F}^+(0, y) + (1 - \lambda) \mathbf{F}^-(0, y),
$$

where we set $z(t)$ as in Equation (21), so that $\lambda = \Lambda (z) \in [0, 1]$ is a constant. Thus, in general

$$
\lim_{\varepsilon \to 0} \mathbf{F}^\varepsilon (\varepsilon z, y) = \mathbf{H}(z, y),
$$

where $z$ varies in $[-1, 1]$, but we evaluate $\mathbf{H}$ at a particular $z$–value to correspond with the $\Gamma$–limit. Since $\mathbf{F}^\varepsilon (t) \mathbb{1}_{\{ \alpha \varepsilon \in \Sigma \}} = (\varepsilon z, \beta) \in \mathcal{A}^{(l_0, t_f)}$ it follows that

$$
\int_{I_{+}^{\varepsilon}[\alpha]} \| \dot{\alpha} - \mathbf{F}^\varepsilon (\alpha) \|^2 \, dt - \int_{I_{-}[\alpha]} \| \dot{\alpha} - \mathbf{H}(\beta; z) \|^2 \, dt
$$

$$
= \int_{I_{+}^{\varepsilon} \cap I_{-}} \| \dot{\alpha} - \mathbf{F}^\varepsilon (\alpha) \|^2 - \| \dot{\alpha} - \mathbf{H}(\beta; z) \|^2 \, dt
$$

$$
+ \int_{I_{+}^{\varepsilon} \setminus I_{-}} \| \dot{\alpha} - \mathbf{F}^\varepsilon \|^2 \, dt - \int_{I_{-}[\alpha] \setminus I_{+}^{\varepsilon}} \| \dot{\alpha} - \mathbf{H} \|^2 \, dt.
$$

As above, the last two integrals vanish as $\varepsilon \to 0$, since the measure of $I_{+}^{\varepsilon} \setminus I_{-}$ converges to zero and $I_{-} \setminus I_{+}^{\varepsilon} = \emptyset$, and the integrands are bounded. Since $I_{+}^{\varepsilon} \cap I_{-} = I_{-}$,

$$
\int_{I_{+}^{\varepsilon} \cap I_{-}} \| \dot{\alpha} - \mathbf{F}^\varepsilon (\alpha) \|^2 - \| \dot{\alpha} - \mathbf{H}(\beta; z) \|^2 \, dt
$$

$$
= \int_{I_{-}} \| \dot{\alpha} \|^2 - 2(\dot{\alpha}, \mathbf{F}^\varepsilon (\alpha)) + |\mathbf{F}^\varepsilon (\alpha) |^2 - |\dot{\alpha} |^2 + 2(\dot{\alpha}, \mathbf{H}(\beta; z)) - |\mathbf{H}(\beta; z) |^2 \, dt.
$$

Since $\dot{\alpha} \mathbb{1}_{\{ \alpha \in \Sigma \}} \overset{H^1}{\to} \dot{\alpha} \mathbb{1}_{\{ \alpha \in \Sigma \}}$ and $\mathbf{F}^\varepsilon (\alpha) \mathbb{1}_{\{ \alpha \in \Sigma \}} \overset{H^1}{\to} \mathbf{H}(\beta; z) \mathbb{1}_{\{ \alpha \in \Sigma \}},$

$$
\int_{I_{-}} | \dot{\alpha} |^2 - |\dot{\alpha} |^2 \, dt = \int_{l_0}^{t_f} | \dot{\alpha} \mathbb{1}_{\{ \alpha \in \Sigma \}} |^2 - | \dot{\alpha} \mathbb{1}_{\{ \alpha \in \Sigma \}} |^2 \, dt \leq \int_{l_0}^{t_f} \| (\dot{\alpha} - \dot{\alpha}) \mathbb{1}_{\{ \alpha \in \Sigma \}} \|^2 \, dt \to 0
$$

as $\varepsilon \to 0$, and

$$
\lim_{\varepsilon \to 0} \int_{I_{+}^{\varepsilon}} | \mathbf{F}^\varepsilon (\alpha) |^2 - | \mathbf{H}(\beta; z) |^2 \, dt = \lim_{\varepsilon \to 0} \int_{l_0}^{t_f} | \mathbf{F}^\varepsilon \mathbb{1}_{\{ \alpha \in \Sigma \}} |^2 - | \mathbf{H} \mathbb{1}_{\{ \alpha \in \Sigma \}} |^2 \, dt
$$

$$
\leq \lim_{\varepsilon \to 0} \int_{l_0}^{t_f} \| (\mathbf{F}^\varepsilon - \mathbf{H}) \mathbb{1}_{\{ \alpha \in \Sigma \}} \|^2 \, dt
$$

$$
= 0,
$$

14
where we have used Minkowski’s inequality and the triangle inequality, as above. Also, as a consequence of the uniform boundedness of sequences in $H^1$, the product of a weakly convergent sequence and a strongly convergent sequence is weakly convergent [67], and thus

$$
\int_{I_\Sigma} \langle \dot{\alpha}, H(\beta; z) \rangle - \langle \dot{\alpha}_\varepsilon, F^\varepsilon(\alpha_\varepsilon) \rangle \, dt = \int_{t_0}^{t_f} \langle \dot{\alpha}, H(\beta; z) \rangle 1_{\{\alpha \in \Sigma\}} - \langle \dot{\alpha}_\varepsilon, F^\varepsilon(\alpha_\varepsilon) \rangle 1_{\{\alpha \in \Sigma^c\}} \, dt \to 0
$$
as $\varepsilon \to 0$. Hence,

$$
\lim_{\varepsilon \to 0} I^\varepsilon_{[t_0, t_f]} [\alpha_\varepsilon 1_{\{\alpha_\varepsilon \in \Sigma^c\}}] = \lim_{\varepsilon \to 0} \int_{I_\Sigma} \|\dot{\alpha}_\varepsilon - F^\varepsilon(\alpha_\varepsilon)\|^2 \, dt
$$

$$
= \int_{I_\Sigma} \|\dot{\alpha} - H(\beta; z)\|^2 \, dt
$$

$$
= \int_{I_\Sigma} \min_{\lambda \in [0,1]} \left\{ |\lambda F^+_1(0, \beta) + (1 - \lambda) F^-_1(0, \beta)|^2
\right. \left. + |\dot{\beta} - \lambda G^+(0, \beta) - (1 - \lambda) G^-(0, \beta)|^2 \right\} \, dt
$$

$$
= T^\varepsilon_{[t_0, t_f]} [\alpha 1_{\{\alpha \in \Sigma\}}].
$$

We now prove the liminf inequality. For the portion of $\alpha$ in $S_\pm$, the liminf inequality follows directly from the recovery sequence calculation, since we chose $\alpha_\varepsilon = \alpha$ for any $\alpha \in A^{(t_0, t_f)}$. For the portion of $\alpha$ in $\Sigma$, we have that for every $\alpha_\varepsilon, \alpha \in A^{(t_0, t_f)}$ satisfying $\alpha_\varepsilon 1_{\{\alpha_\varepsilon \in \Sigma^c\}} \xrightarrow{H^1} \alpha 1_{\{\alpha \in \Sigma\}}$,

$$
\lim_{\varepsilon \to 0} I^\varepsilon_{[t_0, t_f]} [\alpha_\varepsilon 1_{\{\alpha_\varepsilon \in \Sigma^c\}}] = \lim_{\varepsilon \to 0} \int_{t_0}^{t_f} |\dot{\alpha}_\varepsilon 1_{\{\alpha_\varepsilon \in \Sigma^c\}} - F^\varepsilon(\alpha_\varepsilon 1_{\{\alpha_\varepsilon \in \Sigma^c\}})|^2 \, dt
$$

$$
= \int_{I_\Sigma[\alpha]} |\dot{\alpha} 1_{\{\alpha \in \Sigma\}} - H(\alpha 1_{\{\alpha \in \Sigma\}})|^2 \, dt
$$

$$
\geq \int_{I_\Sigma[\alpha]} \min_{\lambda \in [0,1]} \left\{ |\lambda F^+_1(0, \beta) + (1 - \lambda) F^-_1(0, \beta)|^2
\right. \left. + |\dot{\beta} - \lambda G^+(0, \beta) - (1 - \lambda) G^-(0, \beta)|^2 \right\} \, dt.
$$

Then by the lower semicontinuity of the rate functional with respect to weak convergence,

$$
\lim_{\varepsilon \to 0} I^\varepsilon_{[t_0, t_f]} [\alpha 1_{\{\alpha \in \Sigma\}}] = \int_{I_\Sigma} \min_{\lambda \in [0,1]} \left\{ |\lambda F^+_1(0, \beta) + (1 - \lambda) F^-_1(0, \beta)|^2
\right. \left. + |\dot{\beta} - \lambda G^+(0, \beta) - (1 - \lambda) G^-(0, \beta)|^2 \right\} \, dt
$$

$$
\leq \liminf_{\varepsilon \to 0} I^\varepsilon_{[t_0, t_f]} [\alpha_\varepsilon 1_{\{\alpha_\varepsilon \in \Sigma^c\}}].
$$

In other words, this result shows that in the smooth regions $S_\pm$, the minimum of the rate functional $I^{(t_0, t_f)}$ by $\alpha(t)$, where $\alpha$ is the solution to the EL Equations (18). Then in the switching manifold $\Sigma$, the rate functional is minimized by some $\alpha(t; z) = (0, \beta(t; z)) \in A^{(t_0, t_f)}$, where $z$ is defined using Equation (21). Note that the critical manifold of the fast-slow System (18) is defined implicitly as solutions in $\mathbb{R}^n$ to following system of equations

$$
M_0 \subseteq \left\{ \begin{array}{l}
0 = \Lambda(z) F^+_1(0, y) + (1 - \Lambda(z)) F^-_1(0, y) + \varphi \\
0 = \varphi (F^+_1(0, y) + F^-_1(0, y)) + \langle \psi, G^+(0, y) + G^-(0, y) \rangle
\end{array} \right\},
$$

15
so when the conjugate momentum has $\varphi = 0$, $\Lambda(z)F^+_1(0, y) + (1 - \Lambda(z))F^-_1(0, y) = 0$. Thus, the path that minimizes the rate functional in $\Sigma$ is given by $\alpha = (0, \beta)$, where

$$\dot{\beta} = \Lambda(z)G^+(0, \beta) + (1 - \Lambda(z))G^-(0, \beta). \tag{22}$$

Practically, to obtain the most probable path in the switching manifold, we first determine the value of $z$ so that $\Lambda(z)F^+_1(0, y) + (1 - \Lambda(z))F^-_1(0, y) = 0$ for each $y \in \mathbb{R}^{n-1}$. This fixes $z$, so we may calculate the path using Equation (22).

To make explicit the connection of this result to the dynamics imposed on the switching manifold by the Filippov convex combination, recall the definition of sliding flow given by Equations (5),(6). Here, we may fix $\lambda = \Lambda(z)$ so that the first component of the sliding flow in Equation (5) becomes zero. The remaining components describing the sliding flow in Equation (5) then either coincide with Equation (22) or its time-reversed flow, with $t \rightarrow -t$. That is, Theorem 3.8 indicates that the intersection of the most probable path with the switching manifold, $\alpha(t)1_{\{\alpha \in \Sigma\}}$, may correspond to the solution of the flow given by $\dot{y} = G^s(0, y; z)$, where $G^s$ is the convex combination

$$G^s(0, y; z) = \lambda G^+(0, y; z) + (1 - \lambda)G^-(0, y; z),$$

and $\lambda \in [0, 1]$. Thus, the most probable path in the switching manifold precisely recovers the Filippov dynamics with this particular form of $\lambda$ (or its time-reversed dynamics). This result is independent of the mollifier imposed, as long as it fits the form defined in Equation (14), though the path does depend on the mollifier via $\Lambda(z)$.

Furthermore, this indicates that the functional given by Equation (20) has no contribution during times when the path slides in a sliding region, but it does when the path slides in a crossing region.

### 4 Case Study 1: Linear system in 2D

In this section, we explore the application of the most probable path derived in Section 3 to a simple piecewise-linear system in $\mathbb{R}^2$. Consider System (2),(3), where $F^\pm : \mathbb{R}^2 \mapsto \mathbb{R}^2$ are defined as

$$F^+(x, y) = \begin{pmatrix} f^+ \\ g^+ \end{pmatrix} = \begin{pmatrix} a(x - 1) + b(y - \eta) \\ c(x - 1) + y - \eta \end{pmatrix},$$

$$F^-(x, y) = \begin{pmatrix} f^- \\ g^- \end{pmatrix} = \begin{pmatrix} p(x + 1) + q(y + \eta) \\ r(x + 1) + y + \eta \end{pmatrix}. \tag{23}$$

Here, $a, b, c, p, q, r, \eta \in \mathbb{R}$ are parameters. This splits up the plane into two smooth regions separated by the switching manifold $\Sigma = \{(x, y) \mid x = 0\}$. This system has been nondimensionalized to scale out a parameter.

This case study is intended to represent a prototypical example of a planar model with a one-dimensional switching manifold and linear or linearizable dynamics near the switch; see, e.g., [1, 6, 39, 40, 43, 44, 49]. For example, the stochastic Stommel model for thermohaline convection switches between temperature- and salinity-driven ocean circulation [1, 49]. Here the vector field is nonlinear, but it is reasonable to consider linearizing the vector field near the switch.

We begin by investigating the dynamics of the deterministic skeleton, System (4),(23), as the parameter $\eta$ varies. For all values of $\eta$, there are two stable fixed points separated by the switching manifold $\Sigma$. For most values of $\eta$, dynamics fall under two broad categories: (1) the two fixed points are the only asymptotic states possible for $x_0 \in \mathbb{R}^2 \setminus \Sigma$, and (2) there is a sliding-cycle surrounding each fixed point, with a crossing-cycle surrounding all sliding-cycles (see Figure 2(f)).
For intermediate values of $\eta$ we observe both standard and nonstandard discontinuity-induced bifurcations.

Taking the two most prevalent cases (1) and (2), we perform Monte Carlo simulations of System (2),(23). In the case of attracting sliding, Monte Carlo simulations match the most probable path derived by minimizing the functional Equation (20), which slides along $\Sigma$. However, in the case of repelling sliding, Monte Carlo simulations match a local minimizer that does not slide, instead crossing $\Sigma$ at the onset of sliding.

4.1 Deterministic dynamics

In System (4),(23), there is one equilibrium in each smooth region, at $x_\pm = (\pm 1, \pm \eta)$. Figure 2(a) summarizes the linear stability analysis of $x_+$; due to the symmetry of the system, the linear stability of $x_-$ is identical if we consider the vertical axis as $qr$ and the horizontal axis as $p$. The solid line separates saddle ($a < bc$) from non-saddle equilibria. For non-saddle equilibria, the dashed line further separates stable equilibria ($a < -1$) from unstable equilibria. Stable spirals occur in the yellow region and stable nodes occur in the blue region, which are separated by the curve $bc = -(a-1)^2/4$.

Figure 2(b) shows the bifurcations in the system as $\eta$ varies, in the case where $x_\pm$ are stable spirals ($\eta$ does not affect the linear stability of $x_\pm$). The remainder of this section discusses dynamics shown in this bifurcation diagram. We identify bifurcations using existing terminology where possible, though some bifurcations do not appear to have existing classification. Note, although we discuss the time-reversed dynamics in this section, in general most probable paths are not time-reversible unless there is a gradient structure [69].

For large $\eta$, in addition to the two stable foci $x_\pm$, the system has three limit cycles of the piecewise-smooth flow: one stable crossing limit cycle with large amplitude and two unstable sliding limit cycles around each focus. See Figure 2(f) for a visualization of the phase portrait with all three limit cycles. In (b) the points at which the crossing limit cycle crosses $\Sigma$ are shown as brown curves, and the maximum and minimum of the intersection of each sliding limit cycle with the (repelling) sliding region are dashed orange and purple curves, respectively. Although dynamics on $\Sigma$ are not defined $a$ priori, if we impose a sliding flow using the Filippov convex method, with

$$\lambda(y) = \frac{-a + b(y - \eta)}{-a + b(y - \eta) - p - q(y + \eta)},$$

then there are at most two pseudoequilibria $x_s = (x_s, y_s)$, with $x_s = 0$ and $y_s$ given by the root(s) of the quadratic $[\lambda(y_s)(r + y_s + \eta) + (1 - \lambda(y_s))(-c + y_s - \eta)]$. For $a = p$, $b = q$, and $c = r$ there is one pseudoequilibrium, at the origin.

As we decrease $\eta$ from 0.8, the two sliding cycles collide (though not at a point of tangency). Due to the repelling sliding region and nonuniqueness of solution curves leaving the sliding region, this collision leads to higher period behavior for both limit cycles while preserving the period-1 behavior, in that on the limit cycle a solution may either cycle around one focus, or both, or switch between the two. Although similar to a period-doubling bifurcation, in this case the limit cycle may have any number of periods, and the stability is unchanged by the bifurcation. This bifurcation is labeled as (IV) in Figure 2(b); for brevity, we refer to it as a ‘sliding cycle-nonunique period’ bifurcation. We indicate $\eta$–values with potential higher-period limit cycles using gray shading. In reverse time, in which the unstable sliding cycles become stable and the repelling sliding region becomes attracting, this behaves like a period-doubling bifurcation of two sliding limit cycles. Neither the repelling nor attracting sliding cases appear to fit within recent classifications of piecewise-smooth bifurcations; see [70–72].
Figure 2: (a) Stability plane for $x_+$. Shaded regions have stable nodes (blue) or stable spirals (yellow). (b) A bifurcation diagram for System (4),(23). Red lines are stable foci and the black line is a pseudosaddle. Brown curves are the crossing points of the large crossing limit cycle. The orange and purple dashed curves are the upper and lower limits of the sliding portion of the two (unstable) sliding limit cycles. The grey shaded region indicates $\eta$–values where the two sliding limit cycles may have both period-1 and period-2. Vertical dashed lines indicate bifurcations identified: a visible tangency bifurcation ($VV_2$), a crossing cycle-tangency bifurcation ($I$), two sliding cycle-tangency bifurcations ($II$) and ($III$), and a sliding cycle-nonunique period bifurcation ($IV$). (c)-(f) Phase portraits for illustrative values of $\eta$. Colors correspond to (b), with the period-2 unstable sliding limit cycle indicated as a dashed gray curve. Parameters used in all plots are $a = p = -2$, $b = q = -7$, and $c = r = 1$.

At ($III$) in Figure 2(b), the point of tangency for the right sliding limit cycle collides with the maximum of the left sliding limit cycle on $\Sigma$, annihilating the left sliding cycle and thereby the higher-period limit cycles as well. However, the trajectory that formed the left sliding limit cycle persists, though it can no longer slide. Then at ($II$) a similar bifurcation occurs when the minimum of the right sliding cycle collides with the tangent point of the left formerly-sliding cycle trajectory, breaking the right limit cycle.

At ($I$) in Figure 2(b), the crossing limit cycle collides with both the left and right formerly-sliding cycle trajectories, breaking the cycle. Although the collision does not occur at the tangency points of the formerly-sliding trajectories, for brevity we refer to this as a ‘crossing cycle-tangency’ bifurcation. As with ($IV$), bifurcations ($I$) – ($III$) do not appear to have been previously classified.
Finally, at $\eta = 0$ the two points of tangency at the switching manifold collide and slide past each other, causing the repelling sliding region to become an attracting sliding region. In the context of Kuznetsov, Rinaldi, and Gragnani’s classification of bifurcations in planar piecewise-smooth systems, this appears to be a $VV_2$–type bifurcation [70].

4.2 Most probable paths

We determine the most probable path through the switching manifold by smoothing out the boundary, then taking the limit of the resulting most probable path as our smoothing parameter goes to zero to return to the original piecewise-smooth system. We define the mollification of the vector field as in Equation (15), so that $F^\varepsilon(x, y) = (f^\varepsilon(x, y), g^\varepsilon(x, y))$, where

$$
f^\varepsilon(x) = \begin{cases} 
f^-(x), & x \leq -\varepsilon, \\
f^+(x)\Lambda\left(\frac{x}{\varepsilon}\right) + f^-(x)\left(1 - \Lambda\left(\frac{x}{\varepsilon}\right)\right) - a\Phi(-\varepsilon, x) - p\Phi(x, \varepsilon), & |x| < \varepsilon, \\
f^+(x), & x \geq \varepsilon, 
\end{cases}
$$

$$
g^\varepsilon(x) = \begin{cases} 
g^-(x), & x \leq -\varepsilon, \\
g^+(x)\Lambda\left(\frac{x}{\varepsilon}\right) + g^-(x)\left(1 - \Lambda\left(\frac{x}{\varepsilon}\right)\right) - c\Phi(-\varepsilon, x) - r\Phi(x, \varepsilon), & |x| < \varepsilon, \\
g^+(x), & x \geq \varepsilon. 
\end{cases}
$$

Here, $\Lambda(x/\varepsilon)$ is defined as in Equation (19), and we define $\Phi(x_1, x_2)$ as

$$
\Phi(x_1, x_2) = \int_{x_1}^{x_2} u\zeta_\varepsilon(u) \, du.
$$

Note that for $|x| < \varepsilon$, if $a = p$ the second two terms of $f^\varepsilon$ vanish, and if $c = r$ the second two terms of $g^\varepsilon$ vanish. There are now two switching manifolds in the system, $\Sigma^\varepsilon_\pm = \{x \in \mathbb{R}^2 : x = \pm\varepsilon\}$, but the flow is smooth across them. Note that due to the linearity of the system, mollification does not affect the fixed points $x_\pm$ as long as $\varepsilon < 1$.

The persistence of the pseudoequilibrium $x_\ast$ through mollification is not as straightforward as that of $x_\pm$. Although a regular equilibrium may limit to a pseudoequilibrium as a smooth system becomes piecewise-smooth, nonlinear dynamics on the switching manifold may lead to multiple possible phenomena in smoothing out a piecewise-smooth field [53]. However, for the particular choice of $F$ in this case study, after mollification there is a saddle equilibrium at the origin, which was the pre-mollification location of the pseudosaddle.

We introduce the scaling $z = x/\varepsilon$, $z \in [-1, 1]$, to simplify the limit as $\varepsilon \to 0$. With this change of variables, System (4),(24) becomes a fast-slow system,

$$
\varepsilon\dot{z} = f^\varepsilon(\varepsilon z, y), \\
\dot{y} = g^\varepsilon(\varepsilon z, y).
$$

Minimizers of the rate functional $I^{(t_0,t_f)}$ (Equation (1)) in this rescaled flow are solutions of the EL equations

$$
\varepsilon\ddot{z} = \dot{y} \left( f_y^\varepsilon - g_y^\varepsilon \right) + f^\varepsilon f_z^\varepsilon + g^\varepsilon g_z^\varepsilon, \\
\ddot{y} = \varepsilon\dot{z} \left( g_z^\varepsilon - f_y^\varepsilon \right) + f^\varepsilon f_y^\varepsilon + g^\varepsilon g_y^\varepsilon,
$$

19
Most probable paths in piecewise-smooth SDEs

This ensures that the stable fixed points \( x_{\pm} \) persist through the rescaling, extended by two additional components due to \( \varphi \) and \( \psi \). Fixed points in this extended system are given by \( z_{\pm} = (z_{\pm}, y_{\pm}, 0, 0) = (\pm 1/\varepsilon, \pm \eta, 0, 0) \).

The projection onto the \( xy \)-plane of the solution to System (25), which is the most probable path of System (4),(23), is shown in Figure 3. Note that we have used \( x \) instead of \( z \) to plot the most probable path for several values of \( \varepsilon \) together. Figure 3(a) shows the most probable path for the mollified rescaled system, System (25) with \( \varepsilon = 0.5 \), including the deterministic flow from \( x_s \) to \( x_- \). Figure 3(b) shows the most probable path for decreasing values of \( \varepsilon \), and Figure 3(c) shows a zoomed-in region of (b). As illustrated in (b) and (c), this is consistent with the minimizer of the rate functional for the piecewise-smooth system as \( \varepsilon \to 0 \). Here and in the remaining figures for this case study, we have used the function \( h(x) = \exp \left( \frac{1}{x^2 - 1} \right) \) to construct the mollifier. See Appendix B for the derivation of the most probable path as the solution of System (25).

We can numerically validate the most probable path using Monte Carlo simulations of System (2),(23) with additive noise using the Euler-Maruyama scheme. Note that this system has a discontinuous and non-monotonic drift coefficient. Euler-Maruyama has been shown to converge in probability for such systems [73, 74], but results for stronger convergence remain elusive; see [75] and references therein. Since we are concerned with convergence of the mean of the distribution of paths, Euler-Maruyama is sufficient for this investigation.

We performed Monte Carlo simulations for two values of \( \eta \): \( \eta = 1 \), where there is repelling sliding along \( \Sigma \), and \( \eta = -2 \), where there is attracting sliding. See Figure 4 for histograms of the solutions to System (2),(23) that ‘tip’ from \( x_+ \) to a neighborhood of \( x_- \). For (a) \( \eta = -2 \), transition paths tip over the basin boundary \( \Sigma \) by tracking the attracting sliding region until it ends. Recall that in this parameter regime, there are only two limit points in the deterministic system. For (b) \( \eta = 1 \), transition paths tip to a neighborhood of \( x_- \) in a sequence of basin-crossings, first over the \( x > 0 \) unstable sliding cycle, then primarily following the deterministic flow toward the crossing.
Figure 4: Distribution of tipping events from Monte Carlo simulations and most probable paths (red and black curves) from $x_+$ to $x_-$, with (a) attracting sliding ($\eta = -2$; 6,487 tips with $N = 1.056 \times 10^7$ simulations) and (b) repelling sliding ($\eta = 1$; 1,960 tips with $N = 1.676 \times 10^7$ simulations). Parameters used are $a = p = -2$, $b = q = -7$, $c = r = 1$, and $\sigma = 0.3$. All other curves are as in Figure 2.

limit cycle, but ultimately crossing over the $x < 0$ unstable sliding cycle and approaching $x_-$. For all figures, we have overlaid the most probable path calculated as the path $\alpha$ that minimizes the rate functional, Equation (20).

Note that in Figure 4(b), we indicate both the most probable path followed by the Monte Carlo simulations (black curve) and the predicted family of nonunique most probable paths that slide (red curves). These paths coincide for $x > 0$ but are distinct for $x \leq 0$. The observed most probable path in the left-half plane corresponds to the solutions to the EL equations with boundary conditions $x(t_1) = (0, \max_{y \in \Sigma_R} y)$ and $x(t_f) = x_-$. We calculate such solutions using the gradient flow; that
Most probable paths in piecewise-smooth SDEs

Figure 5: Values of the derived rate functional, Equation (20) for \( \eta = 1 \) and \( F = F^\varepsilon \), as \( \varepsilon \) varies. The black line corresponds to the rate functional value for the crossing most probable path and the red lines correspond to the functional values for the non-unique family of sliding most probable paths.

is, we consider solutions \( \alpha : [t_1, t_f] \mapsto \mathbb{R}^2 \) to

\[
\begin{align*}
\frac{\partial \alpha}{\partial s} &= -\frac{\delta T^{(t_0, t_f)}}{\delta \alpha} = \dot{\alpha} - \nabla F^\top F + (\nabla F - \nabla F^\top) \dot{\alpha}, \\
\alpha(s, t_1) &= (0, \max_{y \in \Sigma_R} y), \\
\alpha(s, t_f) &= x_\varepsilon, \\
\alpha(0, t) &= x_g(t),
\end{align*}
\]

where \( x_g(t) \) is a smooth function satisfying \( x_g(t_1) = (0, \max_{y \in \Sigma_R} y) \) and \( x_g(t_f) = x_- \), and \( s > 0 \) is an artificial time. Along solution curves \( \alpha(s, t) \),

\[
\frac{d}{ds} T^{(t_0, t_f)}[\alpha(s, t)] \leq 0.
\]

Therefore \( T^{(t_0, t_f)} \) is a Lyapunov functional and thus \( \lim_{s \to \infty} \alpha(s, t) = \alpha^*(t) \), where \( \alpha^*(t) \) is a solution to the EL equations, System (25). That is, steady state solutions of Equation (26) solve the Euler-Lagrange equations. The converged solution to the IBVP in Equation (26), approximated using the Forward Time Centered Space finite difference scheme with \( s \) as the time variable and \( t \) as the space variable, is shown as the black curve in Figure 4(b).

From Equation (20) we might expect the predicted most probable paths in Figure 4(b) to minimize \( T^{(t_0, t_f)} \) because \( T^{(t_1, t_f)} = 0 \). However, if we consider the predicted most probable path of the piecewise-smooth system in the context of the mollified system and calculate \( T^{(t_0, t_f)}_\varepsilon \) for several values of \( \varepsilon > 0 \), the contribution of the rate functional as \( \varepsilon \to 0 \) is substantial: the minimum rate functional values for the predicted family of most probable paths are two magnitudes greater than those of the observed most probable path; see Figure 5. This shows anecdotally that although the minimizers of the Freidlin-Wentzell rate functional converge weakly to minimizers of Equation (20), they may not converge strongly. Furthermore, since the repelling sliding region has Lebesgue measure zero, the probability that the Euler-Maruyama simulations lie on \( \Sigma \) at time \( t \) is zero for all \( t \).
5 Case Study 2: Non-autonomous system in 1D

In this section, we extend the most probable path framework to a linear one-dimensional non-autonomous system with periodic forcing. Consider Equation (2), where $F^\pm : \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ are defined by

\[
F^+ = f^+(x, t) = -r_+(x - 1) + A_+ \cos(2\pi t),
\]
\[
F^- = f^-(x, t) = -r_-(x - a) + A_- \cos(2\pi(t - p)).
\] (27)

Here, $r_\pm, A_\pm, p \geq 0$ and $a \leq 0$ are parameters. Since $F$ is a scalar here, we will use the notation $F = f$ throughout this case study. This case study represents a prototypical example of a one-dimensional periodically forced model with a switching manifold in the state variable, which is a typical feature of some low-dimensional conceptual climate models; see, e.g., [2, 4, 45, 46, 48, 63].

Since we are considering periodic forcing, we view the flow generated by System (4),(27) as lying on the cylindrical extended phase space $\mathbb{P} = \mathbb{R} \times S^1$, where $S^1$ denotes a circle. More precisely, we make the change of variable $\tau = t$, augmenting Equation (4) by the additional equation $d\tau/dt = 1$, and apply periodic boundary conditions at $\tau = 0$ and $\tau = 1$. However, for ease of presentation, we will continue to use Equation (4) to represent the dynamics on $\mathbb{P}$. With this convention, $\Sigma$ and $S^\pm$ are again defined as in Section 1 and the underlying SDE is given by

\[
dx = f(x)dt + \sigma dW,\] (28)

where, following Filippov’s convex combination method, the deterministic skeleton is given by

\[
\dot{x} = f(x) = \begin{cases} 
  f^+(x, t), & (x, t) \in S_+ \cup \Sigma_+ , \\
  f^-(x, t), & (x, t) \in S_- \cup \Sigma_- , \\
  0, & (x, t) \in \Sigma_A \cup \Sigma_R .
\end{cases}
\] (29)

5.1 Deterministic dynamics

Since $f(x, t)$ is linear in $x$ for $(x, t) \notin \Sigma$, the general solutions to Equation (7) on $S^\pm$ are given by

\[
x^\pm(t) = C^\pm e^{-r^\pm t} + h^\pm(t),
\]

respectively. Here $C^\pm \in \mathbb{R}$ are integration constants and

\[
h_+(t) = 1 + \frac{r_+ A_+}{r_+^2 + 4\pi^2} \cos(2\pi t) + \frac{2A_+ \pi}{r_+^2 + 4\pi^2} \sin(2\pi t),
\]
\[
h_-(t) = a + \frac{r_- A_-}{r_-^2 + 4\pi^2} \cos(2\pi(t - p)) + \frac{2A_- \pi}{r_-^2 + 4\pi^2} \sin(2\pi(t - p)).
\]

A solution with initial conditions in either $S_+$ or $S_-$ is then constructed by selecting $C_\pm$ to satisfy the initial condition as well as the continuity assumption on $\Sigma$. Note, in this construction solutions to Equation (7) may be non-unique, since solutions can intersect in forward time when tracking $\Sigma_A$ and in backwards time on $\Sigma_R$.

We now consider the various parameter regimes of interest. First, if the following inequalities are satisfied:

\[
A_+ < \sqrt{r_+^2 + 4\pi^2} \quad \text{and} \quad A_- < -a \sqrt{r_-^2 + 4\pi^2},
\] (30)

then $h_\pm(t)$ do not intersect $\Sigma$ and thus are stable limit cycle solutions to Equation (7). Since we are ultimately interested in noise-induced transitions from one stable limit cycle to another, we will
assume these inequalities hold throughout the rest of the case study. Second, the geometry of the basins of attraction $B_\pm$ for $h_\pm(t)$, respectively, and the existence of $\Sigma_A$, $\Sigma_R$, and $\Sigma_\pm$ depend on whether the nullclines of $f$ intersect $\Sigma$. The nullclines are given by curves along which $f^\pm(x,t) = 0$; for this system, this occurs along $n_\pm(t)$, where

$$n_+(t) = 1 + \frac{A_+}{r_+} \cos(2\pi t) \quad \text{and} \quad n_-(t) = a + \frac{A_-}{r_-} \cos(2\pi(t - p)).$$

There are thus four cases to consider which are illustrated in Figure 6:

1. If $A_+ < r_+$ and $A_- < r_-|a|$, then neither nullcline intersects $\Sigma$ and thus $\Sigma_R = \Sigma$ and $\Sigma_\pm = \Sigma_A = \emptyset$; see Figure 6(a). In this case $B_\pm$ correspond to $S_+$ and $S_-.$

2. If $A_+ > r_+$ and $A_- < r_-|a|$, then $n_+$ intersects $\Sigma$ and thus $\Sigma_\pm \neq \emptyset$, $\Sigma_R = \mathbb{R} - \Sigma_-$, and $\Sigma_+ = \Sigma_A = \emptyset$; see Figure 6(b). In this case $B_-$ has a nontrivial intersection with $S_+$ and $\Sigma$.

3. If $A_+ < r_+$ and $A_- > r_-|a|$, then $n_-$ intersects $\Sigma$ and thus $\Sigma_\pm \neq \emptyset$, $\Sigma_R = \Sigma - \Sigma_+$, and $\Sigma_- = \Sigma_A = \emptyset$; see Figure 6(c). In this case $B_+$ has a nontrivial intersection with $S_-$ and $\Sigma$.

4. If $A_+ > r_+$ and $A_- > r_-|x_0|$, then both $n_\pm$ intersect $\Sigma$. In this case $\Sigma_\pm \neq \emptyset$ but, depending on the phase shift $p$, $\Sigma_A$ can be either empty or nonempty; see Figure 6(d)-(f). Moreover, $B_\pm$ can both have a nontrivial intersection with $S_+$ and $S_-.$

### 5.2 Most probable paths

In contrast to Case Study 1, this system does not contain stable fixed points but instead stable limit cycles. Consequently, for $t_0 < t_f$, we consider the problem of determining the most probable path from $h_-(t_0)$ to $h_+(t_f)$. That is, we consider the family of optimization problems parameterized by $t_0$ and $t_f$ and define the most probable transition path from $h_-$ to $h_+$ as the minimum over this family. To do so, we redefine the admissible set of transition paths $\mathcal{A}^{(t_0,t_f)}$ by

$$\mathcal{A}^{(t_0,t_f)} = \{\alpha \in H^1([t_0,t_f]; \mathbb{R}) : \alpha(t_0) = h_-(t_0) \text{ and } \alpha(t_f) = h_+(t_f)\},$$

and consider the optimization problem

$$\inf_{t_0 < t_f} \inf_{\alpha \in \mathcal{A}^{(t_0,t_f)}} \mathcal{T}^{(t_0,t_f)}[\alpha],$$

where it follows from Theorem 3.8 that

$$\mathcal{T}^{(t_0,t_f)}[\alpha] = \int_{I[\alpha]} |\dot{\alpha}(t) - f(\alpha(t),t)|^2 \, dt + \int_{\Sigma[\alpha]} \min_{\lambda \in [0,1]} \left\{ [\lambda f^+(0,t) + (1 - \lambda) f^-(0,t)]^2 \right\} dt. \quad (31)$$

In this case study, the minimum over $\lambda \in [0,1]$ in the second integral of $\mathcal{T}^{(t_0,t_f)}$ can be explicitly calculated. First, for times in which $\alpha$ lies in $\Sigma_A \cup \Sigma_R$, the minimum is obtained by setting $\lambda = f^-(0,t)/\left(f^-(0,t) - f^+(0,t)\right)$. Second, for times in which $\alpha$ lies in $\Sigma_\pm$, $f^+(0,t)$ and $f^-(0,t)$ have the same sign and thus the minimum is obtained when $\lambda = 0$ or $\lambda = 1$. Therefore, it follows that

$$\mathcal{T}^{(t_0,t_f)}[\alpha] = \int_{I[\alpha]} |\dot{\alpha}(t) - f(\alpha(t),t)|^2 dt + \int_{I[\Sigma_\pm]} \min\{|f^+(0,t)|, |f^-(0,t)|\}^2 dt. \quad (32)$$

To study the optimization problem defined by Equation (32) we first present the following lemma which simplifies the analysis.
Figure 6: Example phase portraits for Equations (7)-(8). The red curves show the stable limit cycles $h_{\pm}$, the black curves show the nullclines $n_{\pm}$, the green line is $\Sigma_R$, the blue line is $\Sigma_A$, and the dashed black lines are $\Sigma_{\pm}$. The dark and light gray regions correspond to the basins of attraction $B_{\pm}$ for $h_{\pm}$ respectively.

**Lemma 5.1.** For the vector field defined by Equation (27) with parameters satisfying (30):

$$\inf_{t_0 < t_f} \inf_{\alpha \in A(t_0, t_f)} T^{(t_0, t_f)}[\alpha] = \inf_{\alpha \in A(-\infty, \infty)} T^{(-\infty, \infty)}[\alpha]. \quad (33)$$

**Proof.** Since $f^{\pm}(x, t)$ have linear growth in $x$ and are asymptotically inward-flowing, it follows from Theorem 3.7 that for $t_0 < t_f$ satisfying $t_0 > -\infty$ and $t_f < \infty$ there exists $\alpha^* \in A(t_0, t_f)$ such that

$$T^{(t_0, t_f)}[\alpha^*] = \inf_{\alpha \in A(t_0, t_f)} T^{(t_0, t_f)}[\alpha].$$

Define $\tilde{\alpha} \in A(-\infty, \infty)$ by

$$\tilde{\alpha}(t) = \begin{cases} h^-(t), & t \leq t_0, \\
\alpha^*(t), & t_0 < t < t_f, \\
h^+(t), & t \geq t_f. \end{cases}$$

By construction, $\dot{\tilde{\alpha}}(t) = f(\tilde{\alpha}(t), t)$ for $t \leq t_0$ and $t \geq t_f$, and thus

$$\inf_{\alpha \in A(-\infty, \infty)} T^{(-\infty, \infty)}[\alpha] \leq T^{(-\infty, \infty)}[\tilde{\alpha}] = T^{(t_0, t_f)}[\alpha^*] = \inf_{\alpha \in A(t_0, t_f)} T^{(t_0, t_f)}[\alpha].$$
Figure 7: (a) Non-unique minimizers of $I((−∞,∞))$ (blue curves) overlaid on the deterministic dynamics in the case when $\Sigma = \Sigma_R$. (b) Most probable path overlaid on the probability density generated by Monte-Carlo simulations of Equation (28) in a parameter regime in which $(t_{\text{max}}, 0) \in B_+$. The solid black line corresponds to the most probable path constructed using Theorem 5.2 and the dashed magenta curve is the mean of the probability density. All other curves in both (a) and (b) follow the same convention as in Figure 6. The parameter values used in both (a) and (b) are $r_+ = 2$, $r_- = 3$, $p = 0$, $A_+ = 1$, $A_- = 1$, and $a = -0.5$. We set $\sigma = 0.2$ for the Monte-Carlo simulations.

Since this inequality is true for all values of $t_0 < t_f$ satisfying $t_0 \neq -\infty$ and $t_f \neq \infty$, the result follows.

Interestingly, it follows from Lemma 5.1 that minimizers of Equation (31) are not unique. Specifically, since $f^\pm(x, t) = f^\pm(x, t + T)$ for $T \in \mathbb{Z}$, it follows that if $\alpha^* \in A((−∞,∞))$ is a minimum then $\alpha^*_T \in A((−∞,∞))$ defined by $\alpha^*_T(t) = \alpha^*(t + T)$ is a minimum as well. Nevertheless, when represented as curves in the cylindrical phase space $\Pi$ these curves are identical. However, $\mathcal{T}((−∞,∞))$ can possibly contain an uncountable number of non-unique minimizers even on $\Pi$. This possibility arises when $\Sigma_R$ and the closure of $B_+$, which we denote as $\mathcal{B}_+$, have a non-trivial intersection, and a global minimizer of $\mathcal{T}((−∞,∞))$ intersects this region of $\Sigma_R$. Each member of the uncountable family of global minimizers can then be constructed by joining three curves as follows: the first minimizes the functional to $\Sigma_R$, the second tracks $\Sigma_R$ until reaching an arbitrary point $t = s$ in the intersection of $\Sigma_R$ with $\mathcal{B}_+$, and the third curve leaves $\Sigma_R$ and tracks the drift; see Figure 7(a).

Key to this construction is that for Filippov systems there is a fundamental non-uniqueness in how solution curves exit a repelling sliding region, and minimizers of $\mathcal{T}((−∞,∞))$ can track the drift at no cost.

The following theorem summarizes when explicit minimizers of $\mathcal{T}((−∞,∞))$ can be calculated for this case study. The key to the proof is the fact that since the system is piecewise linear, a change of variables converts the problem of determining the most probable path of the non-autonomous system to $\Sigma$ to that of an autonomous system. Moreover, the autonomous system is a gradient system and thus by Kramers’ law, the most probable transition through $\Sigma$ occurs when the potential difference is a minimum [22, 76]. In this case, this corresponds precisely to when the separation
between $h_-$ and $\Sigma$ is minimal,
\[ t_{\text{max}} = p + \frac{1}{2\pi} \tan^{-1}\left(\frac{2\pi}{r_-}\right) + T, \]
where $T \in \mathbb{Z}$; i.e., times when $h_-$ has a maximum.

**Theorem 5.2.** If $(t_{\text{max}}, 0) \in \overline{B}_+$, where $t_{\text{max}} = p + \frac{1}{2\pi} \tan^{-1}\left(\frac{2\pi}{r_-}\right) + T$, and $\alpha^* \in \mathcal{A}(-\infty, \infty)$ is defined piecewise by
\[
\begin{cases}
\alpha^*(t) = -h_-(t_{\text{max}})e^{r_-(t-t_{\text{max}})} + h_-(t), & t \leq t_{\text{max}}, \\
\hat{\alpha}^*(t) \text{ satisfies (29)}, & t > t_{\text{max}},
\end{cases}
\]
then
\[ 2h_-(t_{\text{max}})^2 = T(-\infty, \infty)[\alpha^*] = \inf_{\alpha \in \mathcal{A}(-\infty, \infty)} T(-\infty, \infty)[\alpha]. \]

**Proof.** We first compute the most probable transition paths to $\Sigma$ by determining the most probable transition paths for the SDE
\[ dx = f^-(x, t)dt + \sigma dW \]
from $h_-(t)$ to $\Sigma$. To do so, it is convenient to introduce the change of variables $y(t) = x(t) - h_-(t)$, yielding
\[ dy = -r_- y dt + \sigma dW, \]
which corresponds to a gradient system with potential $V(y) = r_- y^2/2$. With this change of variables, the relevant admissible set $\mathcal{A}^{(t_0, t_m)}_\Sigma$ and functional $T^{(t_0, t_m)}_\Sigma : \mathcal{A}^{(t_0, t_m)}_\Sigma \mapsto \mathbb{R}$ are defined by
\[
\mathcal{A}^{(t_0, t_m)}_\Sigma = \{ \gamma \in H^1([t_0, t_m]; \mathbb{R}) : \gamma(t_0) = 0, \gamma(t_m) = -h_-(t_m), \text{ and } \gamma(t) < -h_-(t) \},
\]
\[ T^{(t_0, t_m)}_\Sigma[\gamma] = \int_{t_0}^{t_m} |\dot{\gamma}(t) + r_- \gamma(t)|^2 dt = \int_{t_0}^{t_m} |\dot{\gamma}(t) + V'(\gamma(t))|^2 dt. \]
Expanding and integrating, it follows that for all $\gamma \in \mathcal{A}^{(t_0, t_m)}_\Sigma$:
\[
T^{(t_0, t_m)}_\Sigma[\gamma] = \int_{t_0}^{t_m} \left[ \dot{\gamma}(t)^2 + 2\dot{\gamma}(t) V'(\gamma(t)) + V'(\gamma(t))^2 \right] dt
= \int_{t_0}^{t_m} \left[ (\dot{\gamma} - V'(\gamma(t)))^2 + 4 \frac{d}{dt} V(\gamma(t)) \right] dt
= \int_{t_0}^{t_m} \left[ \dot{\gamma} - V'(\gamma(t)) \right]^2 dt + 4 \{ V(\gamma(t_m)) - V(\gamma(t_0)) \}
\geq 4V(\gamma(t_m))
= 2h(t_m)^2.
\]
This lower bound is an equality if $t_0 = -\infty$ and $\dot{\gamma} = V'(y)$, i.e. $\gamma$ tracks the time-reversed dynamics. Consequently, setting $t_m = t_{\text{max}}$ it implies that $\gamma^*(t) = -h_-(t_{\text{max}})e^{r_-(t-t_{\text{max}})}$ minimizes $T^{(-\infty, t_{\text{max}})}_\Sigma$ over all $t_m \in \mathbb{R}$.

Finally, for all $\alpha \in \mathcal{A}(-\infty, \infty)$, if we let $t_c = \min\{t : \alpha(t) = 0\}$ denote the first time $\alpha$ crosses $\Sigma$, then by the above inequality,
\[ T^{(-\infty, \infty)}[\alpha] \geq T^{(-\infty, t_c)}_\Sigma[\alpha - h_-] = 2h_-(t_c)^2 \geq 2h_-(t_{\text{max}})^2. \]
Therefore, since the integrand of $T^{(-\infty, \infty)}[\alpha]$ vanishes for times in which $\dot{\alpha}$ satisfies Equation (29), it follows that if $(t_{\text{max}}, 0) \in \overline{B}_+$ then $\alpha^*$ as constructed in the hypothesis of the theorem achieves this lower bound. \qed
5.3 Comparison of most probable paths to Monte-Carlo simulations

As in Section 4, we compare the most probable paths with Monte-Carlo simulations of tipping events using the Euler-Maruyama scheme to numerically approximate realizations of Equation (28). Specifically, for a realization χ of Equation (28) we define the following tipping time

\[ \tau_+(\chi) = \min \{ t : \chi(t) > h_+(t) \}, \]

and given \( t_0 < t_f \) we say χ is a tipping event on the interval \([t_0, t_f]\) if \( \tau_+(\chi) < t_f \). The Monte-Carlo simulations are then conducted until the distribution of tipping events on \([t_0, t_f]\) converges.

In the case when \((t_{\text{max}}, 0) \in \overline{B}_+\), Theorem 5.2 provides an explicit construction for the most probable paths which can be directly compared with the distribution of tipping events. In Figure 7(b) we plot the most probable path overlaid on the probability density of the tipping events on the interval \([0, 3]\) for the same parameter values used to generate Figure 7(a). In this case, \(\overline{B}_+ = \{ x : x \geq 0 \}\) and thus \((t_{\text{max}}, 0) \in \overline{B}_+\). Figure 7(b) illustrates that there is excellent agreement between the predicted most probable path and the mean and mode of the probability density generated by the Monte-Carlo simulations. Note, however, that the most probable path plotted is one that crosses and does not slide as in the other potential most probable paths illustrated in Figure 7(a). Clearly, while the minimizers of \(\mathcal{T}_m(-\infty, \infty)\) are not unique, the tipping events concentrate about the most probable path that does not slide.

When \((t_{\text{max}}, 0) \notin \overline{B}_+\) the most probable paths cannot be directly computed and thus we again approximate the most probable paths by numerically computing the stationary curves of the gradient flow. In this case, to make the problem numerically tractable, we use the mollified vector field computed using a Gaussian kernel

\[ \zeta_\varepsilon(x) = \frac{1}{\varepsilon \sqrt{2\pi}} \exp \left( -\frac{x^2}{2\varepsilon^2} \right). \tag{34} \]

While \(\zeta_\varepsilon\) does not have compact support, \(\zeta_1(x) < 10^{-16}\) for \(|x| \geq 9\) and thus these kernels serve as an accurate approximation of a function with compact support. With this kernel the mollified vector field for this problem becomes

\[ f_\varepsilon(x, t) = \varepsilon \frac{(r_- - r_+)}{\sqrt{2\pi}} + \frac{1}{2} \left( f_+(x, t) \left( 1 + \text{erf} \left( \frac{x}{\varepsilon \sqrt{2}} \right) \right) + f_-(x, t) \left( 1 - \text{erf} \left( \frac{x}{\varepsilon \sqrt{2}} \right) \right) \right), \tag{35} \]

where \(\text{erf}(x) = 2\pi^{-\frac{1}{2}} \int_0^x \exp(-s^2) \, ds\) denotes the standard error function. The gradient flow in this case is given by the partial differential equation

\[ \frac{\partial \alpha}{\partial s} = -\frac{1}{2} \frac{\delta I}{\delta \alpha} = \frac{\partial^2 \alpha}{\partial t^2} - \frac{\partial f_\varepsilon^x}{\partial t} \bigg|_{(\alpha, t)} - \frac{\partial f_\varepsilon^x}{\partial \alpha} \bigg|_{(\alpha, t)} f_\varepsilon^x(\alpha, t), \tag{36} \]

where \(x_\varepsilon(t)\) is a smooth function satisfying \(x_\varepsilon(t_0) = h_-(t_0)\) and \(x_\varepsilon(t_f) = h_+(t_f)\) and again \(s > 0\) is an artificial time. Note, since we used a Gaussian kernel for the mollification, the functional forms of \(f_\varepsilon^x\) and its various derivatives are in terms of Gaussian and error functions and thus Equation (36) can be numerically solved using a Forward Time Centered Space finite difference scheme.

In Figure 8 we plot the numerically computed most probable path for the mollified vector field (35) overlaid on the density of tipping events for two sets of parameters in which \((t_{\text{max}}, 0) \notin \overline{B}_+\). In
Most probable paths in piecewise-smooth SDEs

Figure 8: Most probable paths overlaid on the probability density generated by Monte-Carlo simulations of Equation (28) in a parameter regime in which \((t_{\text{max}}, 0) \notin B_+\). The solid black line corresponds to the most probable path numerically computed as a stationary curve of Equation (36) and the dashed magenta curve is the mean of the probability density. All other curves follow the same convention as in Figure 7. In (a) the parameters are \(r_+ = 2, r_- = 3, p = 0.25, A_+ = 3, A_- = 1, a = -0.5, \sigma = 0.22, \text{and } \varepsilon = \sigma^4\). In (b) the parameters are \(r_+ = 2, r_- = 3, p = 0.25, A_+ = 4, A_- = 3, a = -0.5, \text{and } \sigma = 0.07, \) and \(\varepsilon = \sigma^4\).

Figure 8(a) the \(n_+\) nullcline intersects \(\Sigma\) and thus \(B_-\) contains regions in which \(x > 0\). Moreover, \((t_{\text{max}}, 0)\) lies in a crossing region. Nevertheless, in this case, the most probable path intersects \(\Sigma\) slightly to the right of \((t_{\text{max}}, 0) \notin \overline{B}_+\) while the mean of the density appears to pass precisely through \((t_{\text{max}}, 0)\). In Figure 8(b) the parameters are selected so that both nullclines \(n_+\) and \(n_-\) intersect \(\Sigma\) and thus, in addition to \(B_-\) having nontrivial intersection with \(x > 0\), \(B_+\) contains regions in which \(x < 0\). Again, \((t_{\text{max}}, 0)\) lies in a crossing region but in this case the most probable path and the mean of the tipping events are both to the left of \((t_{\text{max}}, 0)\). However, in this case there is remarkable agreement between the most probable path and the mean of the tipping events. Moreover, the most probable path indicates that there is “noise-induced” sliding along \(\Sigma\) until reaching the boundary of \(B_+\).

6 Discussion

We have proposed an extension to the Freidlin-Wentzell theory of large deviations for determining most probable transition paths in stochastic differential equations with a piecewise-smooth drift, where the switch in drift dynamics occurs across an \((n - 1)\)-dimensional hyperplane. Without loss of generality, we assume this hyperplane spans the first component, so \(\Sigma = \{x = 0\}\). Namely, we derived an appropriate functional to determine the most probable path(s) between two metastable states on either side of \(\Sigma\). For smooth regions of the drift, the derived functional matches the Freidlin-Wentzell rate functional. However, across \(\Sigma\) there is an additional possible contribution to the rate functional when the most probable path slides in a crossing region of \(\Sigma\).

Smoothing out the piecewise-smooth drift by mollifying it with a compactly supported, smooth, radially symmetric kernel of characteristic width \(\varepsilon\) allowed us to consider most probable paths of
Most probable paths in piecewise-smooth SDEs

the mollified system as minimizers of the Friedlin-Wentzell rate functional. We then considered
the sequence of minimizers of the Friedlin-Wentzell rate functional in the limit as \( \varepsilon \to 0 \). Using
the direct method of calculus of variations, we showed that this sequence of minimizers converges
weakly, and that their limit minimizes our derived functional, which itself is the \( \Gamma \)-limit of the
sequence of rate functionals for the mollified system as \( \varepsilon \to 0 \).

Notably, as we recovered the piecewise-smooth drift in the limit as \( \varepsilon \to 0 \), we encountered
a natural choice regarding which dynamics to impose on the switching manifold \( \Sigma \) and thereby
determine which path across \( \Sigma \) has minimal energy. Following the convention of piecewise-smooth
dynamical systems, we did not choose either side, but rather introduced a scaling variable \( z = x/\varepsilon \in [-1, 1] \), so that as \( \varepsilon \to 0 \) we could still define the drift using a nonzero variable \( z \). Then the
path which contributes minimal energy is tangent to \( \Sigma \) in the first component and also minimizes
the contribution of \( z \) to the drift as \( \varepsilon \to 0 \).
Interpret the limiting drift on \( \Sigma \) as a convex combination
of the drift on either side of \( \Sigma \) for a certain choice of weights, we recovered that the most probable
path follows the Filippov convex combination or its time-reversed dynamics — although we did not
specify that the flow on \( \Sigma \) be defined using this convention.

The derived rate functional, which has a possible contribution in addition to the Friedlin-
Wentzell rate functional that depends on minimizing over some parameter \( \lambda \in [0, 1] \), presents an
intriguing perspective on two fronts: the need for a minimum function in the rate functional and
the emergence of the Filippov convex combination. The Filippov convex combination, however,
is not the only possible choice for imposing a sliding flow on \( \Sigma \): for example, the possibility of
nonlinear sliding has recently been explored as a generalization of Filippov sliding, which may
reflect observed types of sliding behavior that Filippov flow cannot reproduce; see, e.g., [53, 54].
Possible extensions of the rate functional derived here may include exploring the effect of imposing
a nonlinear sliding flow in \( \Sigma \), or a complete derivation of the limiting functional for the case of a
non-autonomous vector field in \( \mathbb{R}^n \).

We used two illustrative case studies to explore implications of most probable paths in a
low-dimensional setting, one case study with 2D linear piecewise-smooth drift and another with
periodically-forced 1D piecewise-smooth drift. In both of these case studies, most probable paths
calculated using minimizers of the derived rate functional generally matched the mean of Monte
Carlo simulations of tipping events between stable states on either side of a switching manifold \( \Sigma \).
Additionally, both cases reveal an implication of the piecewise-smooth drift and possibility of most
probable paths sliding: the most probable path is naturally non-unique when sliding in a repelling
sliding region.

However, there are still open questions arising from regimes where there is disagreement be-
tween theory and Monte Carlo simulations. Although minimizer(s) of the derived rate functional
theoretically equal the most probable path only in the zero noise limit [22], we expect qualitative
agreement for small noise. However, in the linear case study in Section 4, Monte Carlo simulations
do not match the predicted (non-unique) family of most probable paths when the most probable
paths slide in a repelling sliding region. Instead, simulations in which tipping is observed qualita-
tively match the most probable path conditioned to travel from the original fixed point \( x_+ \) to the
onset of sliding, \((0, \max_{y \in \Sigma_R} y)\), but from the onset of sliding they follow the most probable path
in the left-half plane from \((0, \max_{y \in \Sigma_R} y)\) to \( x_- \).

Two possible explanations for this behavior are the need for a smaller noise strength, \( \sigma \), or the
need for a higher-order functional in the smoothing parameter, \( \varepsilon \). It may be that although the
Monte Carlo simulations that tipped only happened for a small percent of simulations (0.12% for
Figure 4(b)), the noise strength \( \sigma \) needs to be smaller to capture the rare event statistics necessary
to slide.

Additionally, when we consider the limiting behavior of the predicted most probable path in the
piecewise-smooth limit as $\varepsilon \to 0$, although the path seen in simulations is associated with nonzero action, for $0 < \varepsilon \ll 1$ the cost of sliding is at least two magnitudes greater than the cost of crossing; see Figure 5 and related discussion in Section 4. This may indicate that in the repelling sliding case, the $\Gamma$–limit of the Friedlin-Wentzell rate functional does not completely characterize the asymptotic behavior of $I_{\varepsilon}^{(t_0,t_f)}$. In such a situation, one may consider an asymptotic development to calculate the appropriate correction terms to the limiting functional; i.e., set

$$I_{\varepsilon}^{(t_0,t_f)} = T^{(t_0,t_f),0} + \varepsilon T^{(t_0,t_f),1} + \varepsilon^2 T^{(t_0,t_f),2} + \cdots + \varepsilon^k T^{(t_0,t_f),k} + o(\varepsilon^k),$$

where $T^{(t_0,t_f),0} = T^{(t_0,t_f)}$ is the $\Gamma$–limit of $I_{\varepsilon}^{(t_0,t_f)}$ and each $T^{(t_0,t_f),\ell}$ for $\ell = 1, 2, \ldots, k$ is a functional recursively defined by the $\Gamma$–limit of $T^{(t_0,t_f),\ell-1}$ and its infimum [77]. One possible extension to the present work is an asymptotic development of a family of functionals $T^{(t_0,t_f),\ell}$, $\ell > 0$, to more completely characterize the asymptotic behavior of the family $I_{\varepsilon}^{(t_0,t_f)}$.

Similarly, in the non-autonomous case study in Section 5, while the predicted most probable path may slide non-uniquely in a repelling sliding region, the mean of the observed paths does not slide. One may expect this behavior due to the sliding region having Lebesgue measure zero; see Figure 7 and the related discussion at the end of Section 4. However, as in the linear case study, the observed most probable path intersects $\Sigma$ at the onset of the sliding region. Additionally, although the derived rate functional in Equation (20) penalizes sliding in a crossing region, in certain parameter regimes we nevertheless observe the mean of the paths exhibiting sliding behavior; see Figure 8(b).

The discrepancies described above for both case studies may also be related to factors not accounted for in the derived rate functional. Such factors arise in the probability density describing the transition paths $\alpha(t)$ from $x(t_0)$ to $x(t_f)$:

$$P(\alpha(t)) = c \exp \left[ -\frac{1}{\sigma^2} \left( \int_{t_0}^{t_f} \| \dot{\alpha}(t) - F(\alpha(t), t) \|^2 dt + \sigma^2 \int_{t_0}^{t_f} \nabla \cdot F(\alpha(t)) dt \right) \right],$$

for some normalization factor $c \in \mathbb{R}$ [26]. Note that the first integral leads to the Freidlin-Wentzell rate functional, but together both integrals are referred to as the Onsager-Machlup rate functional. Neither the normalization factor $c$ nor the Onsager-Machlup rate functional are well-defined for piecewise-smooth systems. In particular, since $F$ is non-differentiable the divergence term requires considering regimes of the $\Gamma$–limit of the mollified $F^\varepsilon$ that incorporate sigma explicitly to ensure convergence. Possible connections between most probable paths across repelling sliding regions and terms in the probability density or nonlinear convex combination will be explored in a subsequent study.

### Appendix A: Inward-pointing flow and $p$–growth of $F^\varepsilon$

In this appendix, we supply the proofs of Lemmas 3.1 and 3.2.

**Lemma 3.1** ($p$–growth). There exists $\varepsilon^*$ such that for all $\varepsilon > 0$ satisfying $\varepsilon < \varepsilon^*$, there exist $R_1^\varepsilon, c_1^\varepsilon > 0$ and $c_2^\varepsilon \in \mathbb{R}$ such that $|x| > R_1^\varepsilon$ implies

$$|F^\varepsilon(x)| \geq c_1^\varepsilon |x|^p + c_2^\varepsilon.$$

**Proof.** For $x = (x, y)$, we split the proof into the cases $|x| > \varepsilon$ and $|x| < \varepsilon$. 

31
First, assume $|x| > \varepsilon$. It follows from the reverse triangle inequality and the fact that $\zeta$ is nonnegative with unit area that
\[
|F^e(x)| - |F(x)| \leq |F^e(x) - F(x)| \\
= \left| \int_{-\varepsilon}^{\varepsilon} \zeta(u) F(x - u, y) \, du - F(x) \right| \\
= \int_{-\varepsilon}^{\varepsilon} \zeta(u) (F(x - u, y) - F(x)) \, du \\
\leq \int_{-\varepsilon}^{\varepsilon} \zeta(u) |F(x - u, y) - F(x)| \, du.
\]
Let $R_1, c_1, c_2, c_3, c_4, c_5 > 0$ be defined as in Assumption 1. If $|x| > R_1 + \varepsilon$, it follows from the mean value theorem and growth condition Equation (11) that
\[
|F^e(x)| - |F(x)| > -\varepsilon (c_3 \max\{(|x + \varepsilon, y|^p, |x - \varepsilon, y|^p\} + c_4) |u| \, du \\
\leq (c_3 \max\{(|x + \varepsilon, y|^p, |x - \varepsilon, y|^p\} + c_4) \int_{-\varepsilon}^{\varepsilon} \zeta(u) |u| \, du \\
= \varepsilon (c_3 \max\{(|x + \varepsilon, y|^p, |x - \varepsilon, y|^p\} + c_4),
\]
and thus
\[
|F^e(x)| - |F(x)| > -\varepsilon (c_3 \max\{(|x + \varepsilon, y|^p, |x - \varepsilon, y|^p\} + c_4).
\]
Consequently, since $|x| > R_1$ and $|x| > \varepsilon$ it follows from Equation (10) that
\[
|F^e(x)| > -\varepsilon (c_3 \max\{(|x + \varepsilon, y|^p, |x - \varepsilon, y|^p\} + c_4) + c_1|x|^p + c_2 \\
> -\varepsilon (c_3(|x + \varepsilon|)^p + c_4) + c_1|x|^p + c_2 \\
> -\varepsilon (c_3(2|x|)^p + c_4) + c_1|x|^p + c_2 \\
= (c_1 - 2p^p c_3 \varepsilon) |x|^p + c_2 - c_4 \varepsilon.
\]
Now, assume $|x| < \varepsilon$ and without loss of generality $x \geq 0$. Again, from the reverse triangle inequality,
\[
|F^e(x)| - |F(x)| \leq \int_{-\varepsilon}^{\varepsilon} \zeta(u) \left( F^+(x - u, y) - F^+(x) \right) \, du + \int_{\varepsilon}^{x} \zeta(u) \left( F^-(x - u, y) - F^+(x) \right) \, du \\
\leq \int_{-\varepsilon}^{\varepsilon} \zeta(u) \left( F^+(x - u, y) - F^+(x) \right) \, du + \int_{\varepsilon}^{x} \zeta(u) \left( F^-(x - u, y) - F^+(x, y) \right) \, du \\
+ \int_{x}^{\varepsilon} \zeta(u) \left( F^-(0, y) - F^+(0, y) \right) \, du + \int_{x}^{\varepsilon} \zeta(u) \left( F^+(0, y) - F^+(y, x) \right) \, du \\
\leq \int_{-\varepsilon}^{\varepsilon} \zeta(u) |F^+(x - u, y) - F^+(x)| \, du + \int_{\varepsilon}^{x} \zeta(u) |F^-(x - u, y) - F^-(x)| \, du \\
+ |F^-(0, y) - F^+(0, y)| + |F^+(0, y) - F^+(x)|.
\]
Therefore, if $|x| > R_1 + \varepsilon$ it follows from the mean value theorem, the growth condition Equation (11), and the jump condition Equation (12) that
\[
|F^e(x)| - |F(x)| \leq \int_{-\varepsilon}^{\varepsilon} \zeta(u) (c_3(|x + \varepsilon, y|^p + c_4) |u| \, du + \int_{x}^{\varepsilon} \zeta(u) (c_3| |x - \varepsilon, y|^p + c_4) |u| \, du \\
+ c_5|y|^p + c_6 + \varepsilon (c_3(|x + \varepsilon, y|^p + c_4) \\
\leq 2\varepsilon c_4 + c_3(|x + \varepsilon, y|^p + c_3| |x - \varepsilon, y|^p + c_5|y|^p + c_6 + \varepsilon (c_3(|x + \varepsilon, y|^p + c_4).
Therefore,
\[ |F^\varepsilon(x)| \geq -c_5 |x|^3 - c_6 - 2\varepsilon c_3 \max\{|2\varepsilon|^p, |y|^p\} - 2\varepsilon c_4 + c_4 + |F(x)| \]
\[ \geq (c_1 - c_5 - 2^{p+1}\varepsilon c_3) |x|^p + (c_2 - c_6 - 2\varepsilon c_4). \]
The same argument can be applied for \( x < 0 \) to obtain an identical lower bound.
Therefore, from these two cases and the assumption that \( c_1 - c_5 > 0 \), if
\[ \varepsilon < \min\{(c_1 - c_5)/(2^{p+1}c_3), c_1/(2^p c_3)\} = (c_1 - c_5)/(2^{p+1}c_3), \]
\[ c_1^\varepsilon = \min\{c_1 - 2^p c_3\varepsilon, c_1 - c_5 - 2^{p+1}\varepsilon c_3\} = c_1 - c_5 - 2^{p+1}\varepsilon c_3, \]
\[ c_2^\varepsilon = \min\{c_2 - c_4\varepsilon, c_2 - c_6 - 2\varepsilon c_4\} = c_2 - 6 - 2\varepsilon c_4, \]
then the result follows.

**Lemma 3.2 (Asymptotically inward-flowing).** For all \( \varepsilon > 0 \) there exist \( R_2^\varepsilon, c_7^\varepsilon > 0 \) such that \( |x| > R_2^\varepsilon \) implies \( F^\varepsilon(x) \neq 0 \) and
\[ \langle F^\varepsilon(x), r(x) \rangle < -c_7^\varepsilon |F^\varepsilon(x)|. \]

**Proof.** By construction,
\[ \langle F^\varepsilon(x), r(x) \rangle = \int_{-\varepsilon}^{\varepsilon} \zeta_\varepsilon(u) (F^-(x - u, y) \mathbb{1}_{\{u \geq x\}}(u) + F^+(x - u, y) \mathbb{1}_{\{u \leq x\}}(u), r(x)) du. \]
For \( x \in \mathbb{R}^n \) and \( u \in (-\varepsilon, \varepsilon) \), let \( \theta(u) \in [0, \pi] \) denote the angle between \( r(x) \) and \( r(x - u, y) \). It follows from the mean value theorem that there exists \( C(x, u) \in (-\varepsilon, \varepsilon) \) such that
\[ |1 - \cos(\theta(u))| = \langle r(x), r(x) \rangle - \langle r(x), r(x - u, y) \rangle \]
\[ = C(x, u) \frac{|y|^2}{|x| \cdot |(C(x, u) - x, y)|^2} u \]
\[ = C(x, u) \frac{|y|^2}{|x|(C(x, u)^2 - 2C(x, u)x + |x|^2)^{\frac{3}{2}}} u \]
\[ \leq \frac{\varepsilon^2}{|x|^2 (1 + (C(x, u)^2 - 2C(x, u)x)/|x|^2)^{\frac{3}{2}}}. \]
Therefore, for all \( u \in (-\varepsilon, \varepsilon) \) and \( \theta^* \in (0, \pi) \) there exists \( R_{\theta^*} > 0 \) such that \( |x| > R_{\theta^*} \) implies \( \theta(x, u) < \theta^* \). Furthermore, since \( F \) is asymptotically inward-flowing there exist \( R_2, c_7 > 0 \) such that if \( |x| > R_2 \) then Inequality (13) is satisfied. Consequently, if \( |x| > \max\{R_2 + \varepsilon, R_{\theta^*}\} \), then for \( u \in (-\varepsilon, \varepsilon) \) it follows that
\[ \frac{(F^-(x - u, y) \mathbb{1}_{\{u \geq x\}}(u) + F^+(x - u, y) \mathbb{1}_{\{u \leq x\}}(u), r(x))}{|(F^-(x - u, y) \mathbb{1}_{\{u \geq x\}}(u) + F^+(x - u, y) \mathbb{1}_{\{u \leq x\}}(u))|} \leq \cos(\cos^{-1}(-c_7) + \theta^*) \]
\[ = -c_7 \cos(\theta^*) - \sqrt{1 - c_7^2} \sin(\theta^*). \]
Hence, if we choose \( \theta^* \) so that \( c_7^\varepsilon = -c_7 \cos(\theta^*) - \sqrt{1 - c_7^2} \sin(\theta^*) > 0 \), there there exists \( R_2^\varepsilon > 0 \) such that \( |x| > R_2^\varepsilon \) implies
\[ \frac{(F^-(x - u, y) \mathbb{1}_{\{u \geq x\}}(u) + F^+(x - u, y) \mathbb{1}_{\{u \leq x\}}(u), r(x))}{|(F^-(x - u, y) \mathbb{1}_{\{u \geq x\}}(u) + F^+(x - u, y) \mathbb{1}_{\{u \leq x\}}(u))|} < -c_7^\varepsilon. \]
The most probable path corresponds to solutions of System (25) subject to appropriate boundary conditions. Since the vector field is now smooth, there must be at least one invariant manifold of $z$ conditions. The most probable path follows the deterministic flow, which does not contribute to $I_{\epsilon}^{(t_0,t_f)}$. See Figure 9 for a diagram illustrating the sequence of mollifying, scaling, and translating transformations used in this derivation. Translating the $z_+$ equilibrium to the origin by defining $\tilde{z} = z - z_+$ and $\tilde{y} = y - y_+$, the boundary conditions for the path from $z_+$ to $\Sigma_+^\epsilon$ (Stage (1)) become

\[
\begin{align*}
\tilde{z}(\infty) &= 0, & \tilde{y}(\infty) &= 0, & \varphi(\infty) &= 0, & \psi(\infty) &= 0, \\
\tilde{z}(0) &= 1 - \frac{1}{\epsilon}, & \tilde{y}(0) &= y_1 - \eta, & \varphi(0) &= \varphi_1, & \psi(0) &= \psi_1.
\end{align*}
\]

The boundary conditions $\varphi_1$ and $\psi_1$ over-determine the system if we consider that the solution must be finite at $t = -\infty$ and $t = 0$, so we allow them to vary. We will discuss choosing $y_1$ for both Stages (1) and (2) below. We then translate the most probable path for Stage (1) back to the original coordinates before continuing to Stage (2), which gives the most probable path from $\Sigma_+^\epsilon$ to $z_s$.

Therefore,

\[
\frac{\langle \mathbf{F}_\epsilon(x), r(x) \rangle}{|\mathbf{F}_\epsilon(x)|} < -c_1^\epsilon \frac{\int_{-\epsilon}^{\epsilon} \zeta(u) \left( \mathbf{F}^-(x - u, y) \mathbb{I}_{\{u \geq x\}}(u) + \mathbf{F}^+(x - u, y) \mathbb{I}_{\{u \leq x\}}(u) \right) du}{\int_{-\epsilon}^{\epsilon} \zeta(u) \left( \mathbf{F}^-(x - u, y) \mathbb{I}_{\{u \geq x\}}(u) + \mathbf{F}^+(x - u, y) \mathbb{I}_{\{u \leq x\}}(u) \right) du} < -c_1^\epsilon \frac{\int_{-\epsilon}^{\epsilon} \zeta(u) \left( \mathbf{F}^-(x - u, y) \mathbb{I}_{\{u \geq x\}}(u) + \mathbf{F}^+(x - u, y) \mathbb{I}_{\{u \leq x\}}(u) \right) du}{\int_{-\epsilon}^{\epsilon} \zeta(u) \left( \mathbf{F}^-(x - u, y) \mathbb{I}_{\{u \geq x\}}(u) + \mathbf{F}^+(x - u, y) \mathbb{I}_{\{u \leq x\}}(u) \right) du} = -c_1^\epsilon.
\]
The boundary conditions for Stage (2) are as follows:

\[
\begin{align*}
  z(0) &= 1, & y(0) &= y_1, & \varphi(0) &= \varphi_1, & \psi(0) &= \psi_1, \\
  z(t_2) &= 0, & y(t_2) &= 0, & \varphi(t_2) &= 0, & \psi(t_2) &= 0,
\end{align*}
\]

where \( t_2 < \infty \) is unknown \textit{a priori}. Solving (25), (B.38) gives us the most probable path for Stage (2). To apply the boundary conditions, we first eliminate coefficients of the solution associated with positive eigenvalues. Applying the boundary condition at \( t = 0 \) then leads to a system with two unknown coefficients and three initial conditions, which is over-determined. If we introduce the condition that along the most probable path the Hamiltonian of the system is zero,

\[
H = f^e \varphi + g^e \psi + \frac{1}{2} \varphi^2 + \frac{1}{2} \psi^2 = 0,
\]

we can eliminate an equation in System (25). It suffices to choose the variable associated with a positive real eigenvalue and an eigenvector with nonzero conjugate momentum. For the parameters used in Figure 3, \( \varphi \) satisfies this condition. There are generally two possible values of \( \varphi \): we choose \( \varphi = -f^e + \sqrt{(f^e)^2 - 2g^e \psi - \psi^2} \), so that \( \varphi = 0 \) when \( \psi = 0 \).

\[\textbf{7 Acknowledgments}\]

The work of JG was supported by the American Institute of Mathematics, and the work of JG and JZ was supported by Wake Forest University. Computations were performed on the Wake Forest University DEAC Cluster, a centrally managed resource with support provided in part by the University.

\[\textbf{References}\]

[1] N Berglund and B Gentz. Metastability in simple climate models: pathwise analysis of slowly driven Langevin equations. \textit{Stoch Dyn}, 2(03):327–356, 2002, doi:doi.org/10.1142/S0219493702000455.

[2] I Eisenman and JS Wettlaufer. Nonlinear threshold behavior during the loss of Arctic sea ice. \textit{Proc Natl Acad Sci USA}, 106(1):28–32, 2009, doi:doi.org/10.1073/pnas.0806887106.

[3] S Wieczorek, P Ashwin, CM Luke, and PM Cox. Excitability in ramped systems: the compost-bomb instability. \textit{Proc R Soc A}, 467(2129):1243–1269, 2010, doi:doi.org/10.1098/rspa.2010.0485.

[4] I Eisenman. Factors controlling the bifurcation structure of sea ice retreat. \textit{J Geophys Res Atmos}, 117(D1), 2012, doi:doi.org/10.1029/2011JD016164.

[5] DH Rothman. Characteristic disruptions of an excitable carbon cycle. \textit{Proc Natl Acad Sci USA}, 116(30):14813–14822, 2019, doi:doi.org/10.1073/pnas.1905164116.

[6] CW Arnscheidt and DH Rothman. Routes to global glaciation. \textit{Proc Roy Soc A}, 476(2239):20200303, 2020, doi:doi.org/10.1098/rspa.2020.0303.

[7] J Lohmann and PD Ditlevsen. Risk of tipping the overturning circulation due to increasing rates of ice melt. \textit{Proc Natl Acad Sci USA}, 118(9), 2021, doi:doi.org/10.1073/pnas.2017989118.

[8] A Vanselow, S Wieczorek, and U Feudel. When very slow is too fast-collapse of a predator-prey system. \textit{J Theor Biol}, 479:64–72, 2019, doi:doi.org/10.1016/j.jtbi.2019.07.008.

[9] PE O’Keeffe and S Wieczorek. Tipping phenomena and points of no return in ecosystems: beyond classical bifurcations. \textit{SIAM J Appl Dyn Syst}, 19(4):2371–2402, 2020, doi:https://doi.org/10.1137/19M1242884.

[10] JM Drake and BD Griffen. Early warning signals of extinction in deteriorating environments. \textit{Nature}, 467(7314):456–459, 2010, doi:https://doi.org/10.1038/nature09389.
Most probable paths in piecewise-smooth SDEs

[11] M Scheffer, EH Van Nes, M Holmgren, and T Hughes. Pulse-driven loss of top-down control: the critical-rate hypothesis. *Ecosystems*, 11(2):226–237, 2008, doi:https://doi.org/10.1007/s10021-007-9118-8.

[12] E Forgoston, S Bianco, LB Shaw, and IB Schwartz. Maximal sensitive dependence and the optimal path to epidemic extinction. *Bull Math Biol*, 73(3):495–514, 2011, doi:https://doi.org/10.1007/s11538-010-9537-0.

[13] IB Schwartz, E Forgoston, S Bianco, and LB Shaw. Converging towards the optimal path to extinction. *J R Soc Interface*, 8(65):1699–1707, 2011, doi:https://doi.org/10.1098/rsif.2011.0159.

[14] L Billings and E Forgoston. Seasonal forcing in stochastic epidemiology models. *Ric di Mat*, 67(1):27–47, 2018, doi:https://doi.org/10.1007/s11587-017-0346-8.

[15] J Hindes and IB Schwartz. Rare events in networks with internal and external noise. *Europhys Lett*, 120(5):56004, 2018, doi:https://doi.org/10.1209/0295-5075/120/56004.

[17] JMT Thompson and J Sieber. Climate tipping as a noisy bifurcation: a predictive technique. *IMA J Appl Math*, 76(1):27–46, 2011, doi:https://doi.org/10.1093/imamat/hxq060.

[18] TM Lenton. Early warning of climate tipping points. *Nat Clim Change*, 1(4):201–209, 2011, doi:https://doi.org/10.1038/nclimate1143.

[21] B Lindner, J García-Ojalvo, A Neiman, and L Schimansky-Geier. Effects of noise in excitable systems. *Phys Rep*, 392(6):321–424, 2004, doi:https://doi.org/10.1016/j.physrep.2003.10.015.

[22] M Chaichian and A Demichev. *Path Integrals in Physics: Stochastic Processes and Quantum Mechanics*. Institute of Physics, 2001.

[23] W Ren and E Vanden-Eijnden. Minimum action method for the study of rare events. *Commun Pure Appl Math*, 57(5):637–656, 2004, doi:10.1002/cpa.20005.

[24] E Forgoston and RO Moore. A primer on noise-induced transitions in applied dynamical systems. *SIAM Rev*, 60(4):969–1009, 2018, doi:10.1137/17M1142028.

[25] C B Muratov and E Vanden-Eijnden. Noise-induced mixed-mode oscillations in a relaxation oscillator near the onset of a limit cycle. *Chaos*, 18(1):015111, 2008, doi:https://doi.org/10.1063/1.2779582.

[26] RS Maier and DL Stein. Limiting exit location distributions in the stochastic exit problem. *SIAM J Appl Math*, 57(3):752–790, 1997, doi:10.1137/S0036139994271753.

[27] M Heymann and E Vanden-Eijnden. The geometric minimum action method: A least action principle on the space of curves. *Commun Pure Appl Math*, 61(8):1052–1117, 2008, doi:10.1002/cpa.20238.

[28] BS Lindley and IB Schwartz. An iterative action minimizing method for computing optimal paths in stochastic dynamical systems. *Physica D*, 255:22–30, 2013, doi:10.1016/j.physd.2013.04.001.
[34] GM Donovan and WL Kath. An iterative stochastic method for simulating large deviations and rare events. *SIAM J Appl Math*, 71(3):903–924, 2011, doi:10.1137/100789373.

[35] DJW Simpson and R Kuske. The influence of localized randomness on regular grazing bifurcations with applications to impacting dynamics. *J Vib Control*, 24(2):407–426, 2018, doi:10.1177/107754631664205.

[36] EJ Staunton and PT Piiroinen. Discontinuity mappings for stochastic nonsmooth systems. *Physica D*, 406:132405, 2020, doi:10.1016/j.physd.2020.132405.

[37] EJ Staunton and PT Piiroinen. Estimating the dynamics of systems with noisy boundaries. *Nonlinear Anal: Hybrid Syst*, 36:100863, 2020, doi:10.1016/j.nahs.2020.100863.

[38] Y Chen, A Baule, H Touchette, and W Just. Weak-noise limit of a piecewise-smooth stochastic differential equation. *Phys Rev E*, 88(5):052103, 2013, doi:10.1103/PhysRevE.88.052103.

[39] Y Chen and W Just. Large-deviation properties of Brownian motion with dry friction. *Phys Rev E*, 90(4):042102, 2014, doi:10.1103/PhysRevE.90.042102.

[40] A Baule, H Touchette, and EGD Cohen. Stick–slip motion of solids with dry friction subject to random vibrations and an external field. *Nonlinearity*, 24(2):351, 2010, doi:10.1088/0951-7715/24/2/001.

[41] A Baule, EGD Cohen, and H Touchette. A path integral approach to random motion with nonlinear friction. *J Phys A Math*, 43(2):025003, 2009, doi:10.1088/1751-8113/43/2/025003.

[42] DJW Simpson and R Kuske. Mixed-mode oscillations in a stochastic, piecewise-linear system. *Physica D*, 240(14-15):1189–1198, 2011, doi:10.1016/j.physd.2011.04.017.

[43] SH Piltz, MA Porter, and PK Maini. Prey switching with a linear preference trade-off. *SIAM J Appl Dyn Syst*, 13(2):658–682, 2014, doi:10.1137/130910920.

[44] L Serdukova, Y Zheng, J Duan, and J Kurths. Metastability for discontinuous dynamical systems under Lévy noise: Case study on Amazonian vegetation. *Sci Rep*, 7(1):1–13, 2017, doi:10.1038/s41598-017-07686-8.

[45] W Moon and JS Wettlaufer. A stochastic dynamical model of Arctic sea ice. *J Clim*, 30(13):5119–5140, 2017, doi:10.1175/JCLI-D-16-0223.1.

[46] F Yang, Y Zheng, J Duan, L Fu, and S Wiggins. The tipping times in an Arctic sea ice system under influence of extreme events. *Chaos*, 30(6):063125, 2020, doi:10.1063/5.0006626.

[47] T Mitsui, M Crucifix, and K Aihara. Bifurcations and strange nonchaotic attractors in a phase oscillator model of glacial–interglacial cycles. *Physica D*, 306:25–33, 2015, doi:10.1016/j.physd.2015.05.007.

[48] KS Morupisi and C Budd. An analysis of the periodically forced PP04 climate model, using the theory of non-smooth dynamical systems. *IMA J Appl Math*, 86(1):76–120, 2021, doi:10.1093/imamat/hxaa039.

[49] AH Monahan. Stabilization of climate regimes by noise in a simple model of the thermohaline circulation. *J Phys Oceanogr*, 32(7):2072–2085, 2002, doi:10.1175/1520-0485(2002)032%3C2072:SOCRBN%3E2.0.CO;2.

[50] M di Bernardo, CJ Budd, AR Champneys, P Kowalczyk, AB Nordmark, GO Tost, and PT Piiroinen. Bifurcations in nonsmooth dynamical systems. *SIAM Rev.*, 50(4):629–701, 2008, doi:10.1137/050625060.

[51] M di Bernardo, CJ Budd, AR Champneys, and P Kowalczyk. *Piecewise-smooth dynamical systems: theory and applications*, volume 163 of *Applied Mathematical Sciences*. Springer Science & Business Media, 2008.

[52] AF Filippov. *Differential equations with discontinuous righthand sides*, volume 18. Springer Science & Business Media, 2008.

[53] MR Jeffrey. Hidden dynamics in models of discontinuity and switching. *Physica D*, 273:34–45, 2014, doi:10.1016/j.physd.2014.02.003.
Most probable paths in piecewise-smooth SDEs

[54] DD Novaes and MR Jeffrey. Regularization of hidden dynamics in piecewise smooth flows. *J Differ Equ*, 259(9):4615–4633, 2015, doi:10.1016/j.jde.2015.06.005.

[55] EN Lorenz. Deterministic nonperiodic flow. *J Atmos Sci*, 20(2):130–141, 1963, doi:10.1175/1520-0469(1963)020%0130:DNF%2.0.CO;2.

[56] G Dal Maso. *An introduction to Gamma-convergence*, volume 8. Springer Science & Business Media, 2012.

[57] A Braides. *Gamma-convergence for Beginners*, volume 22. Clarendon Press, 2002.

[58] FJ Pinski, AM Stuart, and F Theil. $\gamma$-limit for transition paths of maximal probability. *J Stat Phys*, 146(5):955–974, 2012, doi:10.1007/s10955-012-0443-8.

[59] T Li and X Li. Gamma-limit of the Onsager–Machlup functional on the space of curves. *SIAM J Math Anal*, 53(1):1–31, 2021, doi:10.1137/20M1310539.

[60] B Ayanbayev, I Klebanov, HC Lie, and TJ Sullivan. Gamma-convergence of Onsager-Machlup functionals. Part II: With applications to maximum a posteriori estimation in Bayesian inverse problems. *Inverse Probl*, 38(2):025006, 2021, doi:10.1088/1361-6420/ac3f81.

[61] B Ayanbayev, I Klebanov, H C Lie, and TJ Sullivan. Gamma-convergence of Onsager-Machlup functionals. Part II: Infinite product measures on Banach spaces. *Inverse Probl*, 38(2):025006, 2021, doi:10.1088/1361-6420/ac3f82.

[62] K Hill, DS Abbot, and M Silber. Analysis of an Arctic sea ice loss model in the limit of a discontinuous albedo. *SIAM J Appl Dyn Syst*, 15(2):1163–1192, 2016, doi:10.1137/15M1037718.

[63] LC Evans. *Partial differential equations*, volume 19. American Mathematical Society, 2010.

[64] J Jost and X Li-Jost. *Calculus of variations*, volume 64. Cambridge University Press, 1998.

[65] LC Evans. *Weak convergence methods for nonlinear partial differential equations*, volume 74. American Mathematical Society, 1990.

[66] RA Adams and JF Fournier. *Sobolev spaces*. Elsevier, 2003.

[67] T Grafke. String method for generalized gradient flows: Computation of rare events in reversible stochastic processes. *J Stat Mech Theory Exp*, 2019(4):043206, 2019, doi:10.1088/1742-5468/ab11db.

[68] YA Kuznetsov, S Rinaldi, and A Gragnani. One-parameter bifurcations in planar Filippov systems. *Int J Bifurc Chaos Appl Sci Eng*, 13(8):2157–2188, 2003, doi:10.1142/S0218127403007874.

[69] N Berglund. Kramers’ law: Validity, derivations and generalisations. *Markov Process Relat Fields*, 19(3):4915–4935, 2010, doi:10.1088/rsta.2010.0198.

[70] P Glendinning, MR Jeffrey, E Bossolini, J T Lázaro, and JM Olm. *An introduction to piecewise smooth dynamics*. Springer, 2019.

[71] I Gyöngy and N Krylov. Existence of strong solutions for Itô’s stochastic equations via approximations. *Prob Theory Relat Fields*, 105(2):143–158, 1996, doi:10.1007/BF01203833.

[72] I Gyöngy. A note on Euler’s approximations. *Potential Anal*, 8(3):205–216, 1998, doi:10.1023/A:1008605221617.

[73] A Neuenkirch, M Szölgyenyi, and L Szpruch. An adaptive Euler–Maruyama scheme for stochastic differential equations with discontinuous drift and its convergence analysis. *SIAM J Numer Anal*, 57(1):378–403, 2019, doi:10.1137/18M1170017.

[74] N Berglund. Kramers’ law: Validity, derivations and generalisations. *Markov Process Relat Fields*, 19(3):4915–4935, 2010, doi:math-mprf.org/journal/articles/id1304.

[75] G Anzellotti and S Baldo. Asymptotic development by $\gamma$-convergence. *Appl Math Optim*, 27(2):105–123, 1993, doi:10.1007/BF01195977.