From random Poincaré maps to stochastic mixed-mode-oscillation patterns

Nils Berglund* Barbara Gentz†‡ Christian Kuehn§

December 22, 2013

Abstract

We quantify the effect of Gaussian white noise on fast–slow dynamical systems with one fast and two slow variables, which display mixed-mode oscillations owing to the presence of a folded-node singularity. The stochastic system can be described by a continuous-space, discrete-time Markov chain, recording the returns of sample paths to a Poincaré section. We provide estimates on the kernel of this Markov chain, depending on the system parameters and the noise intensity. These results yield predictions on the observed random mixed-mode oscillation patterns. An unexpected result of the analysis is that in certain cases, noise may increase the number of small-amplitude oscillations between consecutive large-amplitude oscillations.

Mathematical Subject Classification. 37H20, 34E17 (primary), 60H10 (secondary)
Keywords and phrases. Singular perturbation, fast–slow system, dynamic bifurcation, folded node, canard, mixed-mode oscillation, return map, random dynamical system, first-exit time, concentration of sample paths, Markov chain.

1 Introduction

Oscillation patterns with large variations in amplitude occur frequently in dynamical systems, differential equations and their applications. A class of particular interest are mixed-mode oscillations (MMOs) which are patterns consisting of alternating structures of small-amplitude and large-amplitude oscillations. Typical applications arise from chemical systems such as the Belousov–Zhabotinskii reaction [47], the peroxidase–oxidase reaction [25] and autocatalytic reactions [71] as well as from neuroscience, e.g. stellate cells [28], Hodgkin–Huxley-type neurons [73] and pituitary cells [80]. A remarkable number of models for these phenomena lead to differential equations with multiple timescales see e.g. [66, 27, 84, 40]. Frequently, it suffices to consider two timescales and study fast–slow ordinary differential equations (ODEs) which already provide many generic mechanisms leading to MMOs. For a detailed review of the topic we refer to the survey [26], the special issue [23], and references therein. The basic idea is that a local mechanism

*MAPMO, CNRS – UMR 7349, Université d'Orléans, Fédération Denis Poisson – FR 2964, B.P. 6759, 45067 Orléans Cedex 2, France.
†Faculty of Mathematics, University of Bielefeld, Universitätsstr. 25, 33615 Bielefeld, Germany.
‡Research supported by CRC 701 "Spectral Structures and Topological Methods in Mathematics".
§Vienna University of Technology, Institute for Analysis and Scientific Computing, 1040 Vienna, Austria.

induces the small-amplitude oscillations (SAOs) while a global return mechanism leads to large-amplitude oscillations (LAOs). In this introduction, we shall just outline the main ideas; the precise development of our set-up and results starts in Section 2.

For a deterministic trajectory, we can symbolically write an MMO as a sequence

$$\cdots L_{j-1}^{s_{j-1}} L_j^{s_j} L_{j+1}^{s_{j+1}} \cdots \tag{1.1}$$

where $L_j^{s_j}$ denotes L_j LAOs followed by s_j SAOs, etc. For example, a periodic solution alternating between 2 SAOs and 1 LAO would be $\cdots 1^2 1^2 1^2 \cdots$ or simply 1^2 with the periodicity understood. A prototypical mechanism to generate SAOs are folded-node singularities [24] which are generic in three-dimensional ODEs with one fast and two slow variables [77, 7]. For the global return mechanism, one frequently encounters a relaxation-type structure induced by a cubic (or S-shaped) fast-variable nullcline, also called the critical manifold, which was studied extensively already in the context of van der Pol-type oscillators; see e.g. [51, 35, 20, 24] and the references therein. Non-degenerate folds, folded-node singularities and S-shaped critical manifolds form the basic deterministic building blocks for the work in this paper. However, let us mention already here that the stochastic techniques we develop in this paper could potentially be adapted to other cases such as singular Hopf bifurcation and folded saddle-nodes [38, 58], bursting oscillations [48, 30], tourbillon structures [81, 26], and other global return mechanisms [40, 61]. Although it is certainly of high interest to study all these cases, it seems to us that the combination of folded singularities and relaxation oscillations is a natural first step as both components are basic elements which occur in a large variety of different models [26].

While in some experiments, remarkably clear MMO patterns have been observed [42, 47], in many other cases the SAOs in the patterns appear noisy [28, 65]. Weak noise acting on a dynamical system is known to induce a variety of phenomena, ranging from small fluctuations around deterministic solutions to large excursions in phase space, as shown, e.g., in stochastic resonance [12, 34, 64] and transitions near tipping points [74, 60, 4]. In the context of oscillatory patterns, the effect of noise on MMO patterns in low-dimensional prototypical models has been studied, for instance, in [52, 76, 69, 87, 43, 62], using numerical simulations, bifurcation theory and asymptotic descriptions of the Fokker–Planck equation.

This work concerns the effect of noise on fast–slow differential equations with one fast and two slow variables, containing a folded-node singularity and an S-shaped critical manifold responsible for the global return mechanism. The resulting stochastic differential equations (SDEs) show a subtle interplay between noise, local and global dynamics, which requires a careful analysis of the behaviour of stochastic sample paths. Our approach builds upon our earlier work [17] which in turn was based upon a pathwise approach to fast–slow SDEs [14, 15].

Our main focus is the derivation of estimates for the Poincaré (or return) map of the stochastic system, for a conveniently chosen two-dimensional section Σ. Deterministic return maps in the presence of folded-node singularities have been analyzed, e.g., in [37, 54]. Although the two-dimensional Poincaré map is invertible, the strong contraction near attracting critical manifolds implies that it is close to a one-dimensional, usually non-invertible map. Figure 1 (a) shows an example of such a one-dimensional deterministic return map $z_n \mapsto z_{n+1}$. The apparent discontinuities are in fact points where the map's graph displays very narrow dips, due to the presence of so-called canard orbits. Canards are particular solutions of the system staying close to both the attracting and repelling

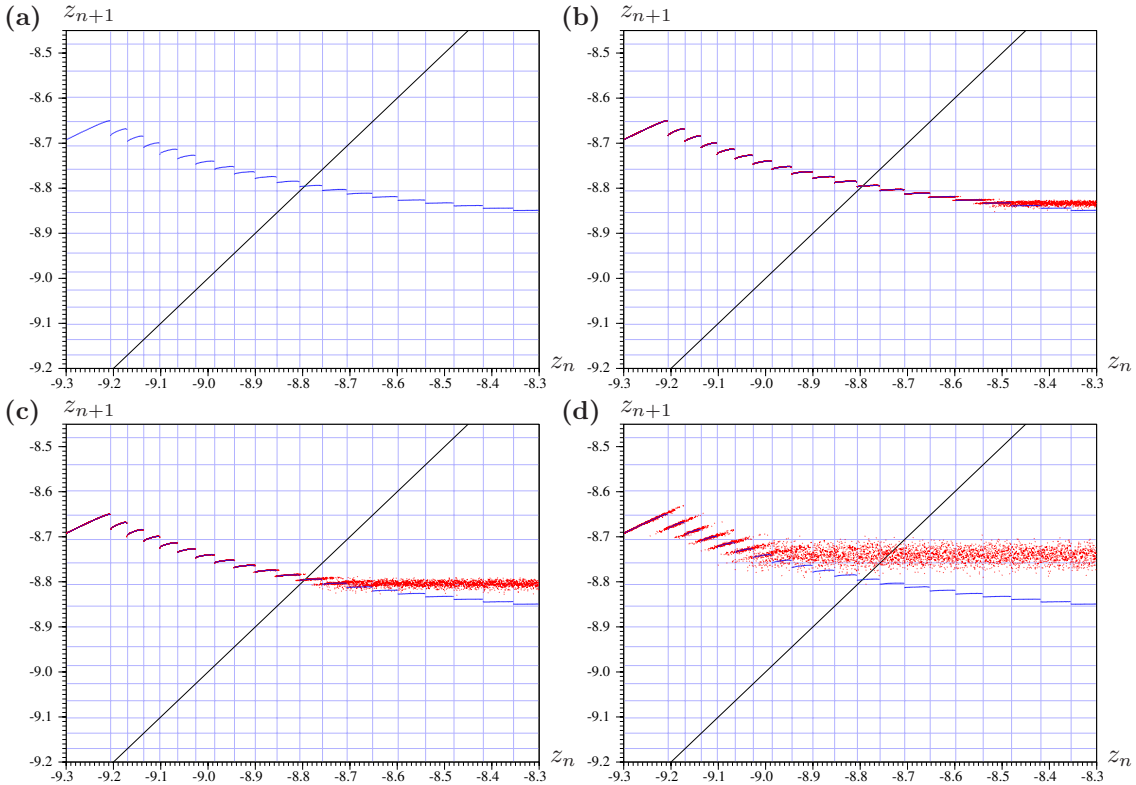

FIGURE 1. z-coordinate of the first-return map on the section Σ_1 for the Koper model (7.1). Parameter values are $k = -10$, $\lambda = -7.6$, $\epsilon_2 = 0.7$, $\epsilon_1 = 0.01$, with noise intensities **(a)** $\sigma = \sigma' = 0$, **(b)** $\sigma = \sigma' = 2 \cdot 10^{-7}$, **(c)** $\sigma = \sigma' = 2 \cdot 10^{-5}$, and **(d)** $\sigma = \sigma' = 2 \cdot 10^{-3}$. The horizontal and vertical lines mark the location of canards.

parts of the critical manifold [10, 7, 11], which separate the phase space into sectors of rotation characterized by different numbers of SAOs [24].

The concept of return maps has been extended to stochastic systems, see for instance [85, 18, 44]. This requires some care, because the rapid fluctuations of stochastic sample paths prevent one from using their very first return to Σ to define the map. Instead, one has to consider the first return after a suitably defined, sufficiently large excursion in phase space has taken place. With these precautions, successive intersections $X_0, X_1, X_2, \dots$ of sample paths with Σ define a continuous-space, discrete-time Markov chain. The distribution of X_{n+1} is obtained from the distribution of X_n via integration with respect to a transition kernel K. Under suitable regularity assumptions [6], the theory of harmonic measures ensures that the kernel K admits a smooth density k, so that the evolution of X_n is specified by an integral equation, namely

$$\mathbb{P}\big\{X_{n+1} \in A \mid X_n = x\big\} = \int_A k(x,y)\,\mathrm{d}y =: K(x,A) \tag{1.2}$$

holds for all Borel sets $A \subset \Sigma$; see for instance [16, Sections 5.2 and 5.3]. The main aim of the present work is to provide estimates on the kernel K. Part of the mathematical challenge is due to the fact that the deterministic flow is not a gradient flow, and thus the stochastic system is irreversible.

Figure 1 (b)–(d) shows simulated stochastic return maps for increasing noise intensity. For each value of z_n, the red points indicate the value of z_{n+1} for 10 different realizations

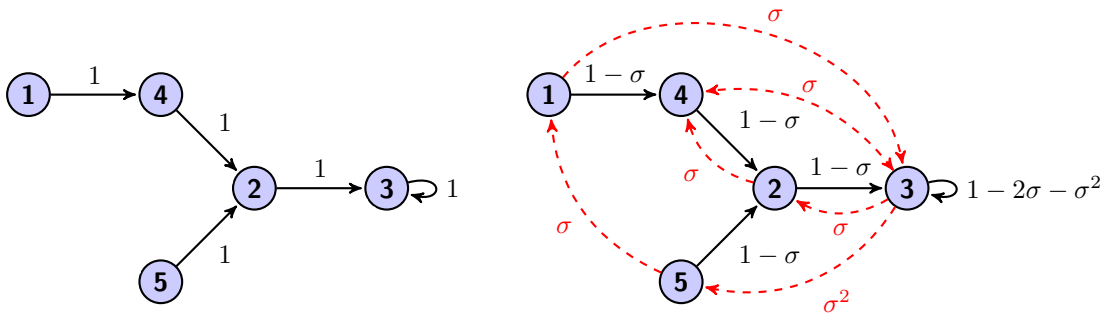

FIGURE 2. A singularly perturbed Markov chain.

of the noise. The deterministic return map is plotted in blue for comparison. Several interesting phenomena can be observed:

1. The size of fluctuations increases with the noise intensity;
2. Orbits in sectors with a small number of SAOs (inner sectors) are less affected by noise than those in sectors with a large number of SAOs (outer sectors);
3. There is a saturation effect, in the sense that for large enough SAO numbers, the typical value of the stochastic return map and its spreading become independent of the sector;
4. The saturation effect sets in earlier for larger noise intensities.

While the first phenomenon is not surprising, the other observed features are remarkable, and can lead to non-intuitive effects. In the example shown in Figure 1, the deterministic map has a stable fixed point in the 11th sector, so that the system will display a stable MMO pattern 1^{11}. For sufficiently strong noise, it will asymptotically operate in the 12th sector, displaying the pattern 1^{12}, with occasional transitions to 1^{11} and 1^{13}. Thus adding noise may increase the number of observed SAOs.

The estimates on the kernel we provide in this work yield quantitative information on the above phenomena. In particular, we obtain estimates on the typical size of fluctuations as a function of noise intensity and sector number, and on the onset of the saturation phenomenon.

The structure of this article is the following: After introducing the deterministic set-up in Section 2, we provide first estimates on noise-induced fluctuations in Section 3. Sections 4 and 5 extend the analysis to a neighbourhood of the regular fold and of the folded node, respectively. Section 6 combines all the local estimates to provide quantitative results on the kernel. The main results are:

- **Theorem 6.1** *(global return map)* quantifies the effect of noise during the global return phase;
- **Theorem 6.2** *(local map for inner sectors)* provides estimates on noise-induced fluctuations for orbits starting near a folded node in sectors with small SAO number; together with Theorem 6.1 it yields bounds on the size of fluctuations of the Poincaré map in all inner sectors;
- **Theorem 6.4** *(local map for outer sectors)* gives similar estimates for orbits starting in sectors with a large SAO number; in particular, it proves the saturation effect.

Finally, Section 7 illustrates these results with numerical simulations for the Koper model.

The results obtained here are a first step towards the understanding of stochastic MMOs, that calls for further work. In particular, it would be desirable to obtain a more precise description of the possible MMO patterns. Let us mention two possible ways to achieve this:

1. *Singularly perturbed Markov chains:* Consider the ideal case where the sectors of rotation form a Markov partition, meaning that the image of each sector is entirely contained in a sector. Then the dynamics can be described by a topological Markov chain between sectors, see Figure 2. Such a chain will in general not be irreducible, for instance, for the chain shown in Figure 2, state 3 is an absorbing state. In the presence of noise, however, new transitions between sectors appear, typically yielding an irreducible chain. In this sense, the chain for the stochastic system is a singular perturbation of its deterministic limit. For weak, non-vanishing noise, transitions between all states may become possible, but transition times diverge as the noise intensity goes to zero. Methods allowing to determine transition rates in singularly perturbed Markov chains for small positive noise have been developed, for instance, in [75, 41, 5, 86].
2. *Metastable transitions between periodic orbits:* Consider a situation where the deterministic system admits several stable periodic orbits, each corresponding to an MMO pattern. Weak noise will induce rare transitions between these orbits. The theory of large deviations [33] provides a way to estimate the exponential asymptotics of transition rates (Arrhenius' law [3]), via a variational problem. In the reversible case, Kramers' law [31, 53] provides a more precise expression for transition rates, which are related to exponentially small eigenvalues of the diffusion's infinitesimal generator, see for instance [21, 22], and [13] for a recent review. For irreversible systems, such precise expressions for transition rates are not available. However, the spectral-theoretic approach may still yield useful information, as in similar irreversible problems involving random Poincaré maps [18, 16].

Notations: We write $|\cdot|$ to denote the absolute value and $\|\cdot\|$ for the Euclidean norm. For $x \in \mathbb{R}$ we write $\lceil x \rceil$ for the smallest integer not less than x. Furthermore, for $a, b \in \mathbb{R}$ we use $a \wedge b := \min\{a, b\}$ and $a \vee b := \max\{a, b\}$. Regarding asymptotics, we use $\mathcal{O}(\cdot)$ in the usual way, i.e., we write $f(x) = \mathcal{O}(g(x))$ as $x \to x^*$ if and only if $\limsup_{x \to x^*} \left|f(x)/g(x)\right| < \infty$. The shorthand $f(x) \asymp g(x)$ is used whenever $f(x) = \mathcal{O}(g(x))$ and $g(x) = \mathcal{O}(f(x))$ hold simultaneously. Furthermore, by $f(x) \ll g(x)$ we indicate that $\lim_{x \to x^*} \left|f(x)/g(x)\right| = 0$. Vectors are assumed to be column vectors and v^{T} denotes the transpose of a vector v.

Acknowledgements: C.K. would like to thank the Austrian Academy of Sciences (ÖAW) for support via an APART fellowship as well as the European Commission (EC/REA) for support by a Marie-Curie International Re-integration Grant. B.G. and C.K. thank the MAPMO at the Université d'Orléans, N.B. and C.K. thank the CRC 701 at the University of Bielefeld for kind hospitality and financial support.

2 Mixed-mode oscillations – The setup

In this section, we shall outline a typical setup for deterministic mixed-mode oscillations based upon three-dimensional fast–slow systems of the form

$$\begin{aligned} \epsilon \tfrac{\mathrm{d}x}{\mathrm{d}s} &= \epsilon \dot{x} = f(x,y,z;\epsilon) \,, \\ \tfrac{\mathrm{d}y}{\mathrm{d}s} &= \dot{y} = g_1(x,y,z;\epsilon) \,, \\ \tfrac{\mathrm{d}z}{\mathrm{d}s} &= \dot{z} = g_2(x,y,z;\epsilon) \,, \end{aligned} \tag{2.1}$$

where $(x,y,z) \in \mathbb{R}^3$ and $0 < \epsilon \ll 1$ is a small parameter. Throughout, we shall make the following assumption:

(A0) The functions $f, g_1, g_2 : \mathbb{R}^4 \to \mathbb{R}$ are of class C^3.

In particular, (A0) implies that on a fixed compact set there exist uniform bounds on f, g_1, g_2. We remark that the system (2.1) is allowed to depend smoothly upon further system parameters $\mu \in \mathbb{R}^p$ although we do not indicate this dependence in the notation. The critical set of (2.1) is

$$C_0 = \{(x,y,z) \in \mathbb{R}^3 : f(x,y,z;0) = 0\} \,. \tag{2.2}$$

Motivated by several applications, such as the Hodgkin–Huxley model [45, 73], the Koper model [51, 59], the forced van der Pol equation [79, 36] and the Rössler model [72], we will assume that the geometric structure of the critical set is an S-shaped smooth manifold; see also Figure 3. More precisely, this assumption can be stated as follows:

(A1) Suppose C_0 is a smooth manifold composed of five smooth submanifolds,

$$C_0 = C_0^{a-} \cup L^- \cup C_0^r \cup L^+ \cup C_0^{a+} \,, \tag{2.3}$$

where the two-dimensional submanifolds $C_0^{a\pm}$ are normally hyperbolic attracting, while the two-dimensional submanifold C_0^r is normally hyperbolic repelling, i.e.,

$$\frac{\partial f}{\partial x}(p;0) < 0 \quad \forall p \in C_0^{a\pm} \quad \text{and} \quad \frac{\partial f}{\partial x}(p;0) > 0 \quad \forall p \in C_0^r \,, \tag{2.4}$$

and $L^\pm$ are one-dimensional smooth fold curves consisting of generic fold points

$$f(p;0) = 0 \,, \quad \frac{\partial f}{\partial x}(p;0) = 0 \,, \quad \frac{\partial^2 f}{\partial x^2}(p;0) \neq 0 \,, \quad \begin{pmatrix} \partial_y f(p;0) \\ \partial_z f(p;0) \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad \forall p \in L^\pm \,. \tag{2.5}$$

Without loss of generality we assume from now on that $\partial_y f(p;0) \neq 0$ for all $p \in L^\pm$.

Fenichel theory [32] shows that for $\epsilon > 0$, the critical submanifolds $C_0^{a\pm}$ and C_0^r perturb to invariant manifolds $C_\epsilon^{a\pm}$ and C_ϵ^r, which are ϵ-close to $C_0^{a\pm}$ and C_0^r in points bounded away from the fold curves $L^\pm$.

Setting $\epsilon = 0$ in (2.1) leads to the slow subsystem

$$\begin{aligned} 0 &= f(x,y,z;0) \,, \\ \dot{y} &= g_1(x,y,z;0) \,, \\ \dot{z} &= g_2(x,y,z;0) \,, \end{aligned} \tag{2.6}$$

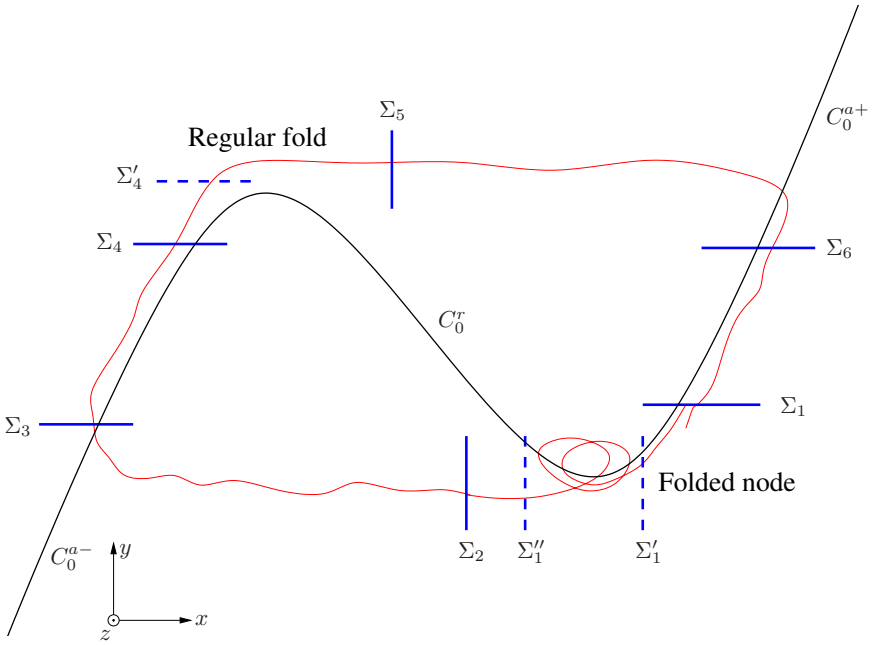

FIGURE 3. Sketch illustrating the definition of the different sections. The horizontal coordinate is x, the vertical one is y, and z points out of the plane.

which is solved by the so-called slow flow. Differentiating f implicitly with respect to s yields

$$\frac{\partial f}{\partial x}\dot{x} = -\frac{\partial f}{\partial y}\dot{y} - \frac{\partial f}{\partial z}\dot{z} = -\frac{\partial f}{\partial y}g_1 - \frac{\partial f}{\partial z}g_2 \tag{2.7}$$

for the slow flow, so that the slow subsystem (2.6) can be written as

$$\begin{aligned} \frac{\partial f}{\partial x}\,\dot{x} &= -\frac{\partial f}{\partial y}g_1 - \frac{\partial f}{\partial z}g_2\ , \\ \dot{z} &= g_2\ , \end{aligned} \tag{2.8}$$

where it is understood that all functions are evaluated at points $(x, y, z; \epsilon) = (p; 0)$ with $p \in C_0$. One may use that (2.8) can locally be written as a closed system by applying the implicit-function theorem to express C_0 as a graph, e.g. $y = h(x, z)$, near the fold as $\frac{\partial f}{\partial y} \neq 0$.

Observe that (2.8) is singular on the fold curves as $\frac{\partial f}{\partial x} = 0$ on $L^{\pm}$. The desingularized slow subsystem is obtained by multiplying the right-hand side of (2.8) by $\frac{\partial f}{\partial x}$ and applying a rescaling of time. It reads

$$\begin{aligned} \dot{x} &= -\frac{\partial f}{\partial y}g_1 - \frac{\partial f}{\partial z}g_2\ , \\ \dot{z} &= \frac{\partial f}{\partial x}g_2\ . \end{aligned} \tag{2.9}$$

We make the following further assumptions:

(A2) Suppose all fold points on L^- satisfy the normal switching condition [67, 78]

$$\begin{pmatrix} \frac{\partial f}{\partial y}(p;0) \\ \frac{\partial f}{\partial z}(p;0) \end{pmatrix} \cdot \begin{pmatrix} g_1(p;0) \\ g_2(p;0) \end{pmatrix} \neq 0 \qquad \forall p \in L^-. \tag{2.10}$$

Furthermore, assume that the projections of $L^{\pm}$ along the x-coordinate onto $C_0^{a\mp}$, which are also called the drop curves, are transverse to the slow flow.

(A3) Assume that the normal switching condition fails only at a unique singularity $p^* \in L^+$ and p^* is a node equilibrium point of (2.9); in this case, p^* is called a folded node (or folded-node singularity) [9, 77].

Let us stress that the above geometric assumptions (A1)–(A3), as well as several further assumptions to follow, provide a convenient framework but that the deterministic and stochastic techniques we present here apply to a much wider range of multiscale systems displaying oscillatory patterns.

On the fast timescale $t = s/\epsilon$ the limit $\epsilon \to 0$ of (2.1) leads to the fast subsystem

$$\begin{aligned} \tfrac{\mathrm{d}x}{\mathrm{d}t} &= x' = f(x, y, z; 0) \,, \\ \tfrac{\mathrm{d}y}{\mathrm{d}t} &= y' = 0 \,, \\ \tfrac{\mathrm{d}z}{\mathrm{d}t} &= z' = 0 \,, \end{aligned} \tag{2.11}$$

which is solved by the fast flow. It is helpful to decompose the singular limit flows and their perturbations into several parts; see Figure 3 for an illustration. In particular, we consider the sections of the form

$$\begin{aligned} \Sigma_i &:= \{(x, y, z) \in \mathbb{R}^3 : x = x_i, y \in [y_{i,a}, y_{i,b}], z \in [z_{i,a}, z_{i,b}]\} \,, \qquad i \in \{2, 5\} \,, \\ \Sigma_i &:= \{(x, y, z) \in \mathbb{R}^3 : y = y_i, x \in [x_{i,a}, x_{i,b}], z \in [z_{i,a}, z_{i,b}]\} \,, \qquad i \in \{1, 3, 4, 6\} \,, \end{aligned} \tag{2.12}$$

for $x_{i,a} < x_{i,b}$, $y_{i,a} < y_{i,b}$, $z_{i,a} < z_{i,b}$ suitably chosen to capture the return map. For an appropriate choice of the constants x_i and y_i (see below or consider the approach in [59]), there are well-defined maps from Σ_i to Σ_j.

(A4) The geometry of the flow-induced maps and sections is as shown in Figure 3.

In particular, Assumption (A4) implies that there is an $\mathcal{O}(1)$ transition time on the slow timescale from Σ_3 to Σ_4 as well as from Σ_6 to Σ_1. (A4) incorporates that there is an $\mathcal{O}(1)$ spatial separation between each pair of fold/drop curves and it guarantees there is an $\mathcal{O}(1)$ transition time on the fast timescale from Σ_2 to Σ_3 as well as from Σ_5 to Σ_6. Furthermore, we exclude the case of a singular Hopf bifurcation [38, 39], where an equilibrium of the full system (2.1) may occur in the neighbourhood of a folded node.

There are four distinct important parts of the flow to analyze:

(I) the flow near the folded node $\Sigma_1 \to \Sigma_2$,
(II) the fast segment $\Sigma_2 \to \Sigma_3$,
(III) the slow-flow region $\Sigma_3 \to \Sigma_4$ near C_ϵ^{a-}, and
(IV) the non-degenerate fold via $\Sigma_4 \to \Sigma_5$.

The map $\Sigma_5 \to \Sigma_6$ can be covered by the same techniques as $\Sigma_2 \to \Sigma_3$, and $\Sigma_6 \to \Sigma_1$ is similar to $\Sigma_3 \to \Sigma_4$.

The geometry of flow maps and the possible generation mechanisms for mixed-mode oscillations under the assumptions (A0)–(A4) are well-known; see for example [24, 26]. A main idea is that twisting of the slow manifolds C_ϵ^{a+} and C_ϵ^r near a folded node generates SAOs and the global return mechanism via the S-shaped critical manifold induces the LAOs. Fixed points of a full return map, say $\Sigma_1 \to \Sigma_1$, correspond to MMOs with a certain pattern

$$\cdots L_k^{s_k} L_{k+1}^{s_{k+1}} \cdots L_{k+l}^{s_{k+l}} L_{k+1}^{s_{k+1}} \cdots \tag{2.13}$$

The main question we address in this paper is how noise influences the patterns (2.13). We are going to split the analysis into two main parts. In Section 3 we provide basic estimates and consider the *global* part of the return map. Sections 4–5 address *local* dynamics in the regions near the regular fold and the folded node.

3 The stochastic system

3.1 Estimating stochastic deviations

As a stochastic extension to (2.1) we consider the fast–slow SDE

$$\begin{aligned} \mathrm{d}x_s &= \frac{1}{\epsilon} f(x_s, y_s, z_s)\,\mathrm{d}s + \frac{\sigma}{\sqrt{\epsilon}} F(x_s, y_s, z_s)\,\mathrm{d}W_s\ , \\ \mathrm{d}y_s &= g_1(x_s, y_s, z_s)\,\mathrm{d}s + \sigma' G_1(x_s, y_s, z_s)\,\mathrm{d}W_s\ , \\ \mathrm{d}z_s &= g_2(x_s, y_s, z_s)\,\mathrm{d}s + \sigma' G_2(x_s, y_s, z_s)\,\mathrm{d}W_s\ , \end{aligned} \tag{3.1}$$

where $(W_s)_{s\geqslant 0}$ is a k-dimensional standard Brownian motion on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The maps

$$F(x,y,z) \in \mathbb{R}^{1\times k}, \qquad G(x,y,z) = \begin{pmatrix} G_1(x,y,z) \\ G_2(x,y,z) \end{pmatrix} \in \mathbb{R}^{2\times k} \tag{3.2}$$

may depend on ϵ, and are assumed to be C^1 and to satisfy the usual bounded-growth condition guaranteing existence of a unique strong solution of (3.1). We shall adopt the shorthand notation to write just (x, y, z) instead of $(x, y, z; \epsilon)$.

We will assume that the diffusion coefficients satisfy the following uniform ellipticity assumption:

(A5) Let

$$D(x,y,z) = \begin{pmatrix} FF^{\mathrm{T}}(x,y,z) & FG^{\mathrm{T}}(x,y,z) \\ GF^{\mathrm{T}}(x,y,z) & GG^{\mathrm{T}}(x,y,z) \end{pmatrix} \in \mathbb{R}^{3\times 3} \tag{3.3}$$

be the diffusion matrix. There exist constants $c_+ \geqslant c_- > 0$ such that

$$c_- \left\|\xi\right\|^2 \leqslant \langle \xi, D(x,y,z)\xi\rangle \leqslant c_+ \left\|\xi\right\|^2 \qquad \forall \xi \in \mathbb{R}^3 \quad \forall (x,y,z)^{\mathrm{T}} \in \mathbb{R}^3\ . \tag{3.4}$$

Remark 3.1. In fact, most of our results remain valid under a weaker hypoellipticity assumption (cf. [6, p. 175] – this weaker condition is needed for the random Poincaré map to have a smooth density). The only result that requires the lower bound in (3.4) is Theorem 6.4, which relies on the early-escape result [17, Theorem 6.4]. See [46] for recent work under weaker assumptions.

Finally we make the following assumption on the noise intensities:

(A6) Assume $0 < \sigma = \sigma(\epsilon) \ll 1$ and $0 < \sigma' = \sigma'(\epsilon) \ll 1$.

In fact, in the course of the analysis, we will encounter more restrictive conditions of the form $\sigma = \mathcal{O}(\epsilon^\alpha)$, $\sigma' = \mathcal{O}(\epsilon^\beta)$ with $\alpha, \beta > 0$. The most stringent of these conditions will be needed for the analysis near the folded node, and requires $\sigma, \sigma' = \mathcal{O}(\epsilon^{3/4})$.

The main goal is to establish bounds on the noise-induced deviation from a deterministic solution. In [15, Theorem 5.1.18], rather precise bounds for the deviation near

normally hyperbolic critical manifolds are derived. We want to adapt these to the other phases of motion. As it turns out, the leading-order effect of noise occurs near the folded-node singularity. Therefore, it will be sufficient to determine the order of magnitude of noise-induced deviations during other phases of the dynamics, as a function of σ, σ' and ϵ.

We fix a deterministic reference solution $(x_s^{\mathrm{det}}, y_s^{\mathrm{det}}, z_s^{\mathrm{det}})$ and set

$$\xi_s = x_s - x_s^{\mathrm{det}}, \qquad \eta_s = \begin{pmatrix} y_s \\ z_s \end{pmatrix} - \begin{pmatrix} y_s^{\mathrm{det}} \\ z_s^{\mathrm{det}} \end{pmatrix}, \qquad \zeta_s = \begin{pmatrix} \xi_s \\ \eta_s \end{pmatrix}. \tag{3.5}$$

As initial condition we choose $(\xi_0, \eta_0) = (0, 0)$ as it corresponds to $(x_0^{\mathrm{det}}, y_0^{\mathrm{det}}, z_0^{\mathrm{det}})$. Substituting in (3.1) and Taylor-expanding, we obtain a system of the form

$$\mathrm{d}\zeta_s = \frac{1}{\epsilon}\mathcal{A}(s)\zeta_s \, \mathrm{d}s + \begin{pmatrix} \frac{\sigma}{\sqrt{\epsilon}}\mathcal{F}(\zeta_s, s) \\ \sigma'\mathcal{G}(\zeta_s, s) \end{pmatrix} \mathrm{d}W_s + \begin{pmatrix} \frac{1}{\epsilon} b_\xi(\zeta_s, s) \\ b_\eta(\zeta_s, s) \end{pmatrix} \mathrm{d}s \,, \tag{3.6}$$

where

$$\mathcal{A}(s) \in \mathbb{R}^{3\times 3} \,, \qquad \mathcal{F}(\zeta_s, s),\ b_\xi(\zeta_s, s) \in \mathbb{R} \,, \qquad \mathcal{G}(\zeta_s, s),\ b_\eta(\zeta_s, s) \in \mathbb{R}^2 \,. \tag{3.7}$$

The nonlinear terms b_ξ and b_η satisfy $b_\cdot(\zeta_s, s) = \mathcal{O}(\|\zeta\|^2)$ as $\|\zeta\| \to 0$. The matrix $\mathcal{A}(s)$ of the system linearized around the deterministic solution has the structure

$$\mathcal{A}(s) = \begin{pmatrix} a(s) & c_1(s) \\ \epsilon c_2(s) & \epsilon B(s) \end{pmatrix} \,, \tag{3.8}$$

where $a(s) = \frac{\partial f}{\partial x}(x_s^{\mathrm{det}}, y_s^{\mathrm{det}}, z_s^{\mathrm{det}})$ and so on, so that in particular $c_1(s) \in \mathbb{R}^{1\times 2}$, $c_2(s) \in \mathbb{R}^{2\times 1}$ and $B(s) \in \mathbb{R}^{2\times 2}$. Let

$$U(s, r) = \begin{pmatrix} U_{\xi\xi}(s, r) & U_{\xi\eta}(s, r) \\ U_{\eta\xi}(s, r) & U_{\eta\eta}(s, r) \end{pmatrix} \tag{3.9}$$

denote the principal solution of the linear system $\epsilon\dot{\zeta} = \mathcal{A}(s)\zeta$. Then the solution of (3.6) can be written in the form

$$\begin{aligned} \xi_s = {} & \frac{\sigma}{\sqrt{\epsilon}} \int_0^s U_{\xi\xi}(s, r)\mathcal{F}(\zeta_r, r) \, \mathrm{d}W_r + \sigma' \int_0^s U_{\xi\eta}(s, r)\mathcal{G}(\zeta_r, r) \, \mathrm{d}W_r \\ & + \frac{1}{\epsilon} \int_0^s U_{\xi\xi}(s, r) b_\xi(\zeta_r, r) \, \mathrm{d}r + \int_0^s U_{\xi\eta}(s, r) b_\eta(\zeta_r, r) \, \mathrm{d}r \,, \end{aligned} \tag{3.10}$$

and

$$\begin{aligned} \eta_s = {} & \frac{\sigma}{\sqrt{\epsilon}} \int_0^s U_{\eta\xi}(s, r)\mathcal{F}(\zeta_r, r) \, \mathrm{d}W_r + \sigma' \int_0^s U_{\eta\eta}(s, r)\mathcal{G}(\zeta_r, r) \, \mathrm{d}W_r \\ & + \frac{1}{\epsilon} \int_0^s U_{\eta\xi}(s, r) b_\xi(\zeta_r, r) \, \mathrm{d}r + \int_0^s U_{\eta\eta}(s, r) b_\eta(\zeta_r, r) \, \mathrm{d}r \,. \end{aligned} \tag{3.11}$$

In both equations, we expect the stochastic integrals to give the leading contribution to the fluctuations. They can be estimated by the Bernstein-type inequality Lemma A.2. The magnitude of the other integrals can then be shown to be smaller, using a direct estimate which is valid as long as the system does not exit from the region where the nonlinear terms are negligible; see e.g. [17, p. 4826] or [14, Theorem 2.4].

In order to carry out this program, we need estimates on the elements of the principal solution U. Note that the ξ-components are in principle larger than the η-components,

but this is compensated by the fact that $x_s^{\det}$ spends most of the time in the vicinity of stable critical manifolds. The following ODEs will play an important rôle:

$$\begin{aligned}\epsilon\dot{p}_1 &= c_1(s) + a(s)p_1 - \epsilon p_1 B(s) - \epsilon p_1 c_2(s) p_1 \;,\\ \epsilon\dot{p}_2 &= c_2(s) - a(s)p_2 + \epsilon B(s) p_2 - \epsilon p_2 c_1(s) p_2 \;.\end{aligned} \tag{3.12}$$

Here $p_1(s) \in \mathbb{R}^{1\times 2}$ and $p_2(s) \in \mathbb{R}^{2\times 1}$. If $a(s)$ is bounded away from 0, standard singular perturbation theory implies that these ODEs admit solutions $p_1(s)$ and $p_2(s)$ of order 1 (and in fact $p_1(s)$ close to $-a(s)^{-1}c_1(s)$). If $a(s)$ approaches 0 or changes sign, this need no longer be the case, but there may still be solutions such that $\epsilon|p_1(s)p_2(s)|$ remains small.

Lemma 3.2. *Assume $s - r \leqslant \mathcal{O}(1)$ and that the ODEs (3.12) admit solutions such that $\epsilon|p_1(u)p_2(u)|$ is bounded for $u \in [r,s]$ by a function $\rho(\epsilon)$ satisfying $\lim_{\epsilon\to 0}\rho(\epsilon) = 0$. Let $\alpha(s,r) = \int_r^s a(u)\,\mathrm{d}u$. Then for sufficiently small ϵ,*

$$\begin{aligned}U_{\xi\xi}(s,r) &= \big[\mathrm{e}^{(\alpha(s,r)+\mathcal{O}(\epsilon))/\epsilon} - \epsilon\; p_1(s) V p_2(r)\big](1+\mathcal{O}(\rho))\;,\\ U_{\xi\eta}(s,r) &= \big[-\mathrm{e}^{(\alpha(s,r)+\mathcal{O}(\epsilon))/\epsilon}\, p_1(r) + p_1(s) V\big](1+\mathcal{O}(\rho))\;,\\ U_{\eta\xi}(s,r) &= \epsilon\big[\mathrm{e}^{(\alpha(s,r)+\mathcal{O}(\epsilon))/\epsilon}\, p_2(s) - V p_2(r)\big](1+\mathcal{O}(\rho))\;,\\ U_{\eta\eta}(s,r) &= \big[V - \epsilon\,\mathrm{e}^{(\alpha(s,r)+\mathcal{O}(\epsilon))/\epsilon}\, p_2(s)p_1(r)\big](1+\mathcal{O}(\rho))\;,\end{aligned} \tag{3.13}$$

where $V = V(s,r)$ is the principal solution of the system

$$\dot{\eta} = [B(s) + c_2(s)p_1(s)]\,\eta\;. \tag{3.14}$$

Proof. Consider the matrix

$$S(s) = \begin{pmatrix} 1 & p_1(s) \\ \epsilon p_2(s) & \mathbb{1}\end{pmatrix}\;. \tag{3.15}$$

Then the equations (3.12) imply

$$\epsilon\dot{S} = \mathcal{A}S - SD \quad \text{with} \quad D(s) = \begin{pmatrix} d_1(s) & 0 \\ 0 & \epsilon D_2(s)\end{pmatrix}\;, \tag{3.16}$$

where the blocks $d_1(s) \in \mathbb{R}$ and $\epsilon D_2(s) \in \mathbb{R}^{2\times 2}$ are given by

$$\begin{aligned}d_1(s) &= a(s) + \epsilon c_1(s) p_2(s)\;,\\ \epsilon D_2(s) &= \epsilon B(s) + \epsilon c_2(s) p_1(s)\;.\end{aligned} \tag{3.17}$$

Consider now the variable $\zeta_1 = S(s)^{-1}\zeta$. If $\epsilon\dot{\zeta} = \mathcal{A}(s)\zeta$, then (3.16) implies

$$\epsilon\dot{\zeta}_1 = D(s)\zeta_1\;. \tag{3.18}$$

The principal solution of this equation is block-diagonal, with blocks $\mathrm{e}^{\frac{1}{\epsilon}\int_r^s d_1(u)\,\mathrm{d}u}$ and $V(s,r)$, where V is the principal solution of $\dot{\eta} = D_2(s)\eta$. The principal solution of the original equation is then given by

$$U(s,r) = S(s)\begin{pmatrix} \mathrm{e}^{\frac{1}{\epsilon}\int_r^s d_1(u)\,\mathrm{d}u} & 0 \\ 0 & V(s,r)\end{pmatrix} S(r)^{-1}\;. \tag{3.19}$$

Furthermore, we have

$$S(s)^{-1} = \begin{pmatrix} 1 & -p_1(s) \\ -\epsilon p_2(s) & \mathbb{1} \end{pmatrix} \begin{pmatrix} [1-\epsilon p_1(s)p_2(s)]^{-1} & 0 \\ 0 & [\mathbb{1}-\epsilon p_2(s)p_1(s)]^{-1} \end{pmatrix} . \tag{3.20}$$

Computing the matrix product in (3.19) yields the result. Note that more precise expressions for the matrix elements can be obtained if needed. □

To describe the size of fluctuations, for given $h, h_1 > 0$ we introduce stopping times

$$\begin{aligned} \tau_\xi &= \inf\bigl\{ s > 0 \colon |\xi_s| > h \bigr\} , \\ \tau_\eta &= \inf\bigl\{ s > 0 \colon \|\eta_s\| > h_1 \bigr\} . \end{aligned} \tag{3.21}$$

Proposition 3.3. *Suppose the assumptions of Lemma 3.2 are satisfied with p_1, p_2 bounded uniformly in ϵ. Given a finite time horizon T of order 1 on the slow timescale, there exist constants $\kappa, h_0 > 0$ such that whenever $h, h_1 \leqslant h_0$, $h_1^2 \leqslant h_0 h$ and $h^2 \leqslant h_0 h_1$,*

$$\mathbb{P}\bigl\{ \tau_\xi \wedge \tau_\eta < s \bigr\} \leqslant \left\lceil \frac{s}{\epsilon} \right\rceil \Bigl(\mathrm{e}^{-\kappa h^2/\sigma^2} + \mathrm{e}^{-\kappa h^2/(\sigma')^2} + \mathrm{e}^{-\kappa h_1^2/(\sigma')^2} + \mathrm{e}^{-\kappa h_1^2/(\epsilon\sigma^2)} \Bigr) \tag{3.22}$$

holds for all $s \leqslant T$.

Proof. Denote by ξ_s^i, $i = 0, 1, 2, 3$, the four terms on the right-hand side of (3.10). We will start by estimating ξ_s^0 and ξ_s^1. Since p_1, p_2 are assumed to be bounded, we may choose $\rho(\epsilon)$ of order ϵ in (3.13), and $U_{\eta\xi}$ is of order ϵ, while the other elements of U are of order 1 at most. By Lemma A.2 and the bounds on $U_{\xi\xi}$ and $U_{\xi\eta}$, there exists a constant $M > 0$ such that

$$\mathbb{P}\biggl\{ \sup_{0\leqslant r\leqslant s} |\xi_r^0| > h \biggr\} \leqslant \left\lceil \frac{s}{\epsilon} \right\rceil \mathrm{e}^{-h^2/(M\sigma^2)} \quad \text{and} \quad \mathbb{P}\biggl\{ \sup_{0\leqslant r\leqslant s} |\xi_r^1| > h \biggr\} \leqslant \left\lceil \frac{s}{\epsilon} \right\rceil \mathrm{e}^{-h^2/(M(\sigma')^2)} . \tag{3.23}$$

Indeed, to estimate ξ_s^0 we first use that on any short time interval $s \in [s_1, s_2]$ with $|s_2 - s_1| \leqslant \epsilon$, the stochastic process $\xi_s^0 = U_{\xi\xi}(s, s_2)\mathcal{M}_s$ is close to the martingale $\mathcal{M}_s$, defined by

$$\mathcal{M}_s = \frac{\sigma}{\sqrt{\epsilon}} \int_0^s U_{\xi\xi}(s_2, r) \mathcal{F}(\zeta_r, r) \, \mathrm{d}r , \tag{3.24}$$

since $|U_{\xi\xi}(s, s_2)|$ remains of order 1 on these time intervals. First using (3.13) and (A0) and then our choice of ρ and Lemma A.1, we see that the martingale's variance is bounded by

$$\frac{\sigma^2}{\epsilon} \int_0^s \bigl| U_{\xi\xi}(s_2, r) (\mathcal{F}\mathcal{F}^{\mathrm{T}})(\zeta_r, r) U_{\xi\xi}^{\mathrm{T}}(s_2, r) \bigr| \, \mathrm{d}r \leqslant \widetilde{M} \frac{\sigma^2}{\epsilon} \int_0^s \Bigl[\mathrm{e}^{(2\alpha(s,r)+\mathcal{O}(\epsilon))/\epsilon} + \rho^2 \, \|V\|^2 \Bigr] \, \mathrm{d}r \tag{3.25}$$

for some positive constant $\widetilde{M}$. Thus the variance is at most of order σ^2 for all $s \in [s_1, s_2]$. Now the first inequality in (3.23) follows immediately from the Bernstein-type estimate Lemma A.2. The prefactor in (3.23) simply counts the number of intervals $[s_1, s_2]$ needed to cover $[0, s]$, see e.g. [15, Proposition 3.15] for a detailed proof in a simpler, one-dimensional setting. The second inequality in (3.23) is shown similarly.

Furthermore, we have $|\xi_s^2| + |\xi_s^3| \leqslant M'(h^2 + h_1^2)$ for a constant $M' > 0$ and $s \leqslant \tau_\xi \wedge \tau_\eta$. From this, together with Gronwall's lemma, we deduce that there exists a constant $M > 0$ such that

$$\mathbb{P}\big\{\tau_\xi < s \wedge \tau_\eta\big\} \leqslant \left\lceil \frac{s}{\epsilon} \right\rceil \exp\left\{ -\frac{[h - M'(h^2 + h_1^2)]^2}{M\sigma^2} \right\} + \left\lceil \frac{s}{\epsilon} \right\rceil \mathrm{e}^{-h^2/(M(\sigma')^2)} \ . \tag{3.26}$$

In a similar way, we find

$$\mathbb{P}\big\{\tau_\eta < s \wedge \tau_\xi\big\} \leqslant \left\lceil \frac{s}{\epsilon} \right\rceil \exp\left\{ -\frac{[h_1 - M'(h^2 + h_1^2)]^2}{M(\sigma')^2} \right\} + \left\lceil \frac{s}{\epsilon} \right\rceil \mathrm{e}^{-h_1^2/(M\epsilon\sigma^2)} \ . \tag{3.27}$$

Choosing h_0 small enough, we can ensure that the terms $M'(h^2 + h_1^2)$ are negligible, and the result follows by taking the sum of the last two estimates. □

The size of typical fluctuations is given by the values of h, h_1 for which the probability (3.22) starts getting small, namely $h \gg \sigma \vee \sigma'$ and $h_1 \gg \sigma' \vee \sigma\sqrt{\epsilon}$. We conclude that fluctuations have size $\sigma + \sigma'$ in the fast direction, and $\sigma' + \sigma\sqrt{\epsilon}$ in the slow direction. Note that for simplicity we ignore the logarithmic contribution arising from the prefactor $\lceil s/\epsilon \rceil$.

We now want to estimate the noise-induced spreading for the Poincaré map, starting on the section Σ_2 after the folded node, and arriving on the section Σ_1 before the folded node. As described in Section 2 we decompose the map into several maps, see Figure 3, and estimate the spreading for each map separately. This means that we fix an initial condition on each section, and estimate the deviation of the stochastic sample paths from the deterministic solution when it first hits the next section.

3.2 The fast segments

The fast segments are given by $\Sigma_2 \to \Sigma_3$ and $\Sigma_5 \to \Sigma_6$. By Assumption (A4) there exists a slow time T_0 of order ϵ in which the deterministic solution starting on Σ_2 reaches a neighbourhood of order 1 of the stable critical manifold. In this neighbourhood, the linearization $a(s) = \frac{\partial f}{\partial x}(x_s^{\mathrm{det}}, y_s^{\mathrm{det}}, z_s^{\mathrm{det}})$ is negative and of order 1. To reach an ϵ-neighbourhood of the critical manifold, an additional slow time T_1 of order $\epsilon|\log\epsilon|$ is required. By the drop-curve transversality assumption (A2) and using (A4), it takes another slow time T_2 of at most order 1 to reach the section Σ_3. For $T := T_0 + T_1 + T_2$ we thus have

$$a(s) \leqslant \begin{cases} a_1 & \text{for all } s \\ -a_2 & \text{for } c_1\epsilon \leqslant s \leqslant T \end{cases} \tag{3.28}$$

for some positive constants a_1, a_2, c_1. This implies that whenever $T \geqslant s > r \geqslant 0$,

$$\alpha(s, r) \leqslant c_2\epsilon \tag{3.29}$$

for a constant c_2, and furthermore $\alpha(s, r)$ is negative as soon as s is larger than a constant times ϵ.

Consider now the equations (3.12) for p_1 and p_2. We will show that p_1 remains bounded on $[0, T]$ and that there exists a particular solution p_2 which also remains bounded on $[0, T]$. For p_1, we proceed in two steps:

- For $0 \leqslant s \leqslant c_1\epsilon$, $p_1(s)$ can grow at most by an amount of order 1.

- For $c_1\epsilon < s \leqslant T$, since $a(s)$ is negative, we can use standard singular perturbation theory to show that $p_1(s)$ remains of order 1, and in fact approaches $c_1(s)/|a(s)|$.

For $p_2(s)$, we change the direction of time and consider the equation

$$\dot{p}_2 = -c_2(T-s) + a(T-s)p_2 - \epsilon B(T-s)p_2 + \epsilon p_2 c_1(T-s)p_2 \; . \tag{3.30}$$

We know that $a(T-s)$ is negative, bounded away from 0, except for a time interval of length $c_1\epsilon$ near T. Thus we conclude that there exists a particular solution which remains bounded, of order 1, on the whole time interval. Therefore Lemma 3.2 shows that $U_{\xi\xi}$, $U_{\xi\eta}$ and $U_{\eta\eta}$ remain bounded (in norm), of order 1, and that $U_{\eta\xi}$ remains of order ϵ for $0 \leqslant r < s \leqslant T$. As a consequence, we can apply Proposition 3.3 as is, with the result that on the section Σ_3,

- the spreading in the fast direction is of order $\sigma + \sigma'$,
- the spreading in the slow z-direction is of order $\sigma' + \sigma\sqrt{\epsilon}$.

3.3 The slow segments

The slow segments are given by $\Sigma_3 \to \Sigma_4$ and $\Sigma_6 \to \Sigma_1$. The analysis of the previous subsection can actually be extended to these segments, because $a(t)$ is always negative, bounded away from 0. The conclusions on typical spreading are the same:

- the spreading in the fast direction is of order $\sigma + \sigma'$,
- the spreading in the slow z-direction is of order $\sigma' + \sigma\sqrt{\epsilon}$.

Note that [15, Theorem 5.1.18] provides a more precise description of the dynamics, by constructing more precise covariance tubes. The qualitative conclusion on typical spreading is the same as above.

4 The regular fold

4.1 Approach

The regular fold corresponds to the transition $\Sigma_4 \to \Sigma_5$. We again fix a deterministic solution, now starting on Σ_4. We choose the origin of the coordinate system on the regular fold L^- and the origin of time in such a way that at time $s = 0$, $(y_0^{\rm det}, z_0^{\rm det}) = (0, 0)$.

Recall from the deterministic analysis (see e.g. [55, 68] for the two-dimensional and [78, 67] for the three-dimensional case) that, given $s_0 < 0$ of order 1,

- for $s_0 \leqslant s \leqslant -\epsilon^{2/3}$, the distance of $x_s^{\rm det}$ to the critical manifold grows like $\epsilon/|s|^{1/2}$;
- there exists a $c_1 > 0$ such that $x_s^{\rm det} \asymp \epsilon^{1/3}$ for $-\epsilon^{2/3} \leqslant s \leqslant c_1\epsilon^{2/3}$;
- there exists a $c_2 > 0$ such that $x_s^{\rm det}$ reaches order 1 before time $c_2\epsilon^{2/3}$.

In this section, we consider the transition $\Sigma_4 \to \Sigma_4'$, where Σ_4' is a section on which $y = c_1\epsilon^{2/3}$. In this region, the linearization $a(s) = \frac{\partial f}{\partial x}(x_s^{\rm det}, y_s^{\rm det}, z_s^{\rm det})$ satisfies

$$a(s) \asymp -(|s|^{1/2} + \epsilon^{1/3}) \; . \tag{4.1}$$

Lemma 4.1. *There are solutions of the equations* (3.12) *satisfying*

$$\|p_1(s)\|, \|p_2(s)\| = \mathcal{O}\biggl(\frac{1}{|a(s)|}\biggr) = \mathcal{O}\biggl(\frac{1}{|s|^{1/2}+\epsilon^{1/3}}\biggr) \qquad \text{for } s_0 \leqslant s \leqslant c_1\epsilon^{2/3}. \tag{4.2}$$

Proof. For $p_1(s)$, we first consider the equation $\epsilon\dot{p}_1 = a(s)p_1 + c_1(s)$, whose solution we know behaves as above, see Lemma A.1 (or [15, pp. 87–88]). Regular perturbation theory allows us to extend the estimate to the full equation for p_1. In the case of p_2, we change the direction of time, and thus consider an equation similar to the equation for p_1 on an interval $[-c_1\epsilon^{2/3}, -s_0]$. The above bound can be obtained, e.g., by scaling space by $\epsilon^{1/3}$ and time by $\epsilon^{2/3}$ on $[-c_1\epsilon^{2/3}, \epsilon^{2/3}]$, and using integration by parts on the remaining time interval. □

Corollary 4.2. *For all $s_0 \leqslant r \leqslant s \leqslant c_1\epsilon^{2/3}$, the principal solution $U(s,r)$ satisfies*

$$|U_{\xi\xi}(s,r)| = \mathcal{O}\biggl(\mathrm{e}^{\alpha(s,r)/\epsilon} + \frac{\epsilon}{(|s|^{1/2}+\epsilon^{1/3})(|r|^{1/2}+\epsilon^{1/3})}\biggr)\,, \tag{4.3}$$

$$\|U_{\xi\eta}(s,r)\| = \mathcal{O}\biggl(\frac{\mathrm{e}^{\alpha(s,r)/\epsilon}}{|r|^{1/2}+\epsilon^{1/3}} + \frac{1}{|s|^{1/2}+\epsilon^{1/3}}\biggr)\,, \tag{4.4}$$

$$\|U_{\eta\xi}(s,r)\| = \mathcal{O}\biggl(\epsilon\biggl[\frac{\mathrm{e}^{\alpha(s,r)/\epsilon}}{|s|^{1/2}+\epsilon^{1/3}} + \frac{1}{|r|^{1/2}+\epsilon^{1/3}}\biggr]\biggr)\,, \tag{4.5}$$

$$\|U_{\eta\eta}(s,r)\| = \mathcal{O}(1)\,. \tag{4.6}$$

Proof. We can apply Lemma 3.2, since $\epsilon|p_1(s)p_2(s)| = \mathcal{O}(\epsilon^{1/3})$. Recall that the matrix V occurring in (3.13) is the principal solution of $\dot{\eta} = D_2(s)\eta$, where

$$D_2(s) = B(s) + c_2(s)p_1(s) = \mathcal{O}(|a(s)|^{-1})\,. \tag{4.7}$$

It follows that

$$\frac{\mathrm{d}}{\mathrm{d}s}\|\eta_s\|^2 = 2(\eta_1\dot{\eta}_1 + \eta_2\dot{\eta}_2) \leqslant \frac{M}{|a(s)|}\|\eta_s\|^2 \tag{4.8}$$

for some constant $M > 0$, so that Gronwall's Lemma implies

$$\|\eta_s\|^2 \leqslant \|\eta_{s_0}\|^2 \exp\biggl\{\int_{s_0}^{s} \frac{M}{|a(u)|}\,\mathrm{d}u\biggr\}\,. \tag{4.9}$$

A direct computation using (4.1) shows that for $s \leqslant c_1\epsilon^{2/3}$ the integral has order 1, and thus the principal solution V has order 1 as well. Then the result follows from Lemma 3.2. □

Proposition 4.3. *There exist constants $\kappa, h_0 > 0$ such that whenever $h \leqslant h_0\epsilon^{1/3}$, $h_1 \leqslant h_0$, $h^2 \leqslant h_0 h_1$ and $h_1^2 \leqslant h_0 h\epsilon^{1/3}$,*

$$\begin{aligned}&\mathbb{P}\bigl\{\tau_\xi \wedge \tau_\eta < c_1\epsilon^{2/3}\bigr\}\\ &\qquad\leqslant \biggl\lceil\frac{1}{\epsilon}\biggr\rceil \Bigl(\mathrm{e}^{-\kappa h^2/(\sigma^2\epsilon^{-1/3})} + \mathrm{e}^{-\kappa h^2/((\sigma')^2\epsilon^{-2/3})} + \mathrm{e}^{-\kappa h_1^2/(\sigma')^2} + \mathrm{e}^{-\kappa h_1^2/(\epsilon|\log\epsilon|\sigma^2)}\Bigr)\,.\end{aligned} \tag{4.10}$$

Proof. Estimate (4.3) and Lemma A.1 imply

$$\frac{1}{\epsilon}\int_{s_0}^{s} U_{\xi\xi}(s,r)^2\,\mathrm{d}r = \mathcal{O}\biggl(\frac{1}{|s|^{1/2}+\epsilon^{1/3}}\biggr) \leqslant \mathcal{O}\left(\epsilon^{-1/3}\right) \tag{4.11}$$

for $s_0 + \mathcal{O}(1) \leqslant s \leqslant c_1\epsilon^{2/3}$. Indeed, the term $\mathrm{e}^{\alpha(s,r)/\epsilon}$ yields a contribution of this order, while the second term in (4.3) gives a contribution of order $\epsilon|\log\epsilon|/(|s|^{1/2}+\epsilon^{1/3})$, which is smaller. Next, we estimate

$$\int_{s_0}^{s} \|U_{\xi\eta}(s,r)\|^2\,\mathrm{d}r = \mathcal{O}\biggl(\frac{1}{|s|+\epsilon^{2/3}}\biggr) \leqslant \mathcal{O}\left(\epsilon^{-2/3}\right)\ , \tag{4.12}$$

where the main contribution now comes from the second term in (4.4). We also obtain

$$\frac{1}{\epsilon}\int_{s_0}^{s} \|U_{\eta\xi}(s,r)\|^2\,\mathrm{d}r = \mathcal{O}\bigl(\epsilon|\log\epsilon|\bigr)\ , \tag{4.13}$$

where the main contribution comes from the second term in (4.5). Finally

$$\int_{s_0}^{s} \|U_{\eta\eta}(s,r)\|^2\,\mathrm{d}r = \mathcal{O}(1)\ . \tag{4.14}$$

Similarly, we obtain the estimates

$$\begin{aligned}
\frac{1}{\epsilon}\int_{s_0}^{s} |U_{\xi\xi}(s,r)|\,\mathrm{d}r &= \mathcal{O}\biggl(\frac{1}{|s|^{1/2}+\epsilon^{1/3}}\biggr)\ ,\\
\int_{s_0}^{s} \|U_{\xi\eta}(s,r)\|\,\mathrm{d}r &= \mathcal{O}\biggl(\frac{1}{|s|^{1/2}+\epsilon^{1/3}}\biggr)\ ,\\
\frac{1}{\epsilon}\int_{s_0}^{s} \|U_{\eta\xi}(s,r)\|\,\mathrm{d}r &= \mathcal{O}(1)\ ,\\
\int_{s_0}^{s} \|U_{\eta\eta}(s,r)\|\,\mathrm{d}r &= \mathcal{O}(1)\ .
\end{aligned} \tag{4.15}$$

We can now adapt the proof of Proposition 3.3 to the present situation. Recall the definitions of τ_ξ, τ_η from (3.21). We denote again by ξ_s^i, $i = 0,1,2,3$, the four terms on the right-hand side of (3.10). Let $T = c_1\epsilon^{2/3}$. The Bernstein-type estimate Lemma A.2 and (4.11) yield

$$\mathbb{P}\biggl\{\sup_{s_0\leqslant r\leqslant T}|\xi_r^0| > h\biggr\} \leqslant \left\lceil\frac{1}{\epsilon}\right\rceil \mathrm{e}^{-h^2/(M\sigma^2\epsilon^{-1/3})}\ , \tag{4.16}$$

and similarly

$$\mathbb{P}\biggl\{\sup_{s_0\leqslant r\leqslant T}|\xi_r^1| > h\biggr\} \leqslant \left\lceil\frac{1}{\epsilon}\right\rceil \mathrm{e}^{-h^2/(M(\sigma')^2\epsilon^{-2/3})}\ . \tag{4.17}$$

Furthermore, using (4.15) we obtain $|\xi_s^2| + |\xi_s^3| \leqslant M'\epsilon^{-1/3}(h^2+h_1^2)$ for $s \leqslant \tau_\xi \wedge \tau_\eta$. From this we can deduce

$$\mathbb{P}\bigl\{\tau_\xi < T\wedge\tau_\eta\bigr\} \leqslant \left\lceil\frac{1}{\epsilon}\right\rceil\left(\exp\biggl\{-\frac{[h-M'\epsilon^{-1/3}(h^2+h_1^2)]^2}{M\sigma^2\epsilon^{-1/3}}\biggr\} + \mathrm{e}^{-h^2/(M(\sigma')^2\epsilon^{-2/3})}\right)\ . \tag{4.18}$$

In a similar way, we get

$$\mathbb{P}\bigl\{\tau_\eta < T\wedge\tau_\xi\bigr\} \leqslant \left\lceil\frac{1}{\epsilon}\right\rceil\left(\exp\biggl\{-\frac{[h_1-M'(h^2+h_1^2)]^2}{M\sigma^2\epsilon|\log\epsilon|}\biggr\} + \mathrm{e}^{-h_1^2/(M(\sigma')^2)}\right)\ . \tag{4.19}$$

This concludes the proof. □

The condition $h_1^2 \leqslant h_0 h \epsilon^{1/3}$ together with $h \leqslant h_0 \epsilon^{1/3}$ imposes that we can take h_1 at most of order $\epsilon^{1/3}$. For the typical spreadings, we obtain

- in the fast direction:
$$\frac{\sigma}{\epsilon^{1/6}} + \frac{\sigma'}{\epsilon^{1/3}} \,, \tag{4.20}$$
- in the slow direction:
$$\sigma' + \sigma\sqrt{\epsilon|\log\epsilon|} \,. \tag{4.21}$$

For the bound (4.10) to be useful, we need the spreading in the fast direction to be small compared to $\epsilon^{1/3}$, because of the condition on h. This yields the conditions

$$\sigma \ll \epsilon^{1/2} \,, \qquad \sigma' \ll \epsilon^{2/3} \,. \tag{4.22}$$

The term $\sigma/\epsilon^{1/6}$ of the x-spreading and the condition $\sigma \ll \epsilon^{1/2}$ are expected, because they already occur when there is no noise acting on the slow variables (see [15, Section 3.3]). The term $\sigma'/\epsilon^{1/3}$ and the condition $\sigma' \ll \epsilon^{2/3}$ are due to the coupling with the slow variables.

Remark 4.4. By using sharper estimates on the size of the linear terms ξ_r^0 and ξ_r^1 (cf. Remark A.3), one can in fact show that the typical spreading in the x-direction grows like

$$\frac{\sigma}{|s|^{1/4} + \epsilon^{1/6}} + \frac{\sigma'}{|s|^{1/2} + \epsilon^{1/3}} \,. \tag{4.23}$$

4.2 Normal form

Before analysing the behaviour during the jump, we make a preliminary transformation to normal form near the fold. Recall that $t = s/\epsilon$ denotes the fast timescale.

Proposition 4.5. *Near a regular fold on L^- satisfying the assumptions (A1)–(A2) there exists a smooth change of coordinates such that* (3.1) *is locally given by*

$$\begin{aligned}
\mathrm{d}x_t &= \left[y_t + x_t^2 + \mathcal{O}(z_t, \left\|(x_t, y_t)^{\mathrm{T}}\right\|^3, \epsilon, \sigma^2)\right] \mathrm{d}t + \sigma \widehat{F}_1(x_t, y_t, z_t)\, \mathrm{d}W_t + \sigma'\sqrt{\epsilon}\, \widehat{F}_2(x_t, y_t, z_t)\, \mathrm{d}W_t \,, \\
\mathrm{d}y_t &= \epsilon \hat{g}_1(x_t, y_t, z_t; \epsilon, \sigma')\, \mathrm{d}t \qquad\qquad + \sigma'\sqrt{\epsilon}\, \widehat{G}_1(x_t, y_t, z_t)\, \mathrm{d}W_t \,, \\
\mathrm{d}z_t &= \epsilon \hat{g}_2(x_t, y_t, z_t; \epsilon)\, \mathrm{d}t \qquad\qquad + \sigma'\sqrt{\epsilon}\, \widehat{G}_2(x_t, y_t, z_t)\, \mathrm{d}W_t \,,
\end{aligned} \tag{4.24}$$

where $\hat{g}_1 = g_1 + \mathcal{O}((\sigma')^2)$ and

$$\hat{g}_1(0,0,0;0,0) = 1 \,, \qquad \hat{g}_2(0,0,0;0) = 0 \,. \tag{4.25}$$

Proof. The result is a stochastic analogue of the transformation result for deterministic systems. We extend the proof presented by Szmolyan and Wechselberger in [78, pp. 73–74] and [82, pp. 8–10] to the stochastic case.

First, we may use a translation of coordinates so that the neighbourhood of L^- is chosen with center $(0,0,0) \in L^-$. From the normal switching condition Assumption (A2) we may assume without loss of generality that $g_1(0,0,0;0) \neq 0$; indeed, if $g_1(0,0,0;0) = 0$ then $g_2(0,0,0;0) \neq 0$ and we may exchange the names of the two slow variables. Next, define a coordinate change

$$z =: \bar{z} + \gamma y \quad \text{with} \quad \gamma = \frac{g_2(0,0,0;0)}{g_1(0,0,0;0)} \neq 0 \,. \tag{4.26}$$

This yields

$$\begin{aligned}\mathrm{d}\bar{z}_t = \mathrm{d}z_t - \gamma\,\mathrm{d}y_t = \epsilon\big[g_2(x,y,\bar{z}+\gamma y;\epsilon) - \gamma g_1(x,y,\bar{z}+\gamma y;\epsilon)\big]\,\mathrm{d}t \\ + \sigma'\sqrt{\epsilon}\big[G_1(x,y,\bar{z}+\gamma y;\epsilon) - \gamma G_2(x,y,\bar{z}+\gamma y;\epsilon)\big]\,\mathrm{d}W_t\,. \qquad (4.27)\end{aligned}$$

Introducing new maps $\bar{g}_2 = g_2 - \gamma g_1$ and $\bar{G}_2 = G_1 - \gamma G_2$ and then dropping all the overbars from the notation yields a stochastic fast–slow system of the form (3.1) which now satisfies

$$g_1(0,0,0;0) \neq 0 \qquad \text{and} \qquad g_2(0,0,0;0) = 0. \qquad (4.28)$$

The next step is to rectify the fold curve. By the implicit-function theorem there exists a parametrization of L^- by $(\xi(z),\eta(z),z)$ for $z \in \mathcal{I} \subset \mathbb{R}$ where $\mathcal{I}$ is a suitable interval. The transformation

$$(x,y,z) = (\bar{x}+\xi(z), \bar{y}+\eta(z), \bar{z}) \qquad (4.29)$$

rectifies the fold curve in new coordinates $(\bar{x},\bar{y},\bar{z})$, in the sense that $\bar{f}(0,0,\bar{z}) = 0$. Itô's formula shows

$$\begin{aligned}\mathrm{d}\bar{x}_t &= \big[\bar{f}(\bar{x}_t,\bar{y}_t,\bar{z}_t) + \mathcal{O}(\epsilon(\sigma')^2)\big]\,\mathrm{d}t + \big[\sigma\bar{F}(\bar{x}_t,\bar{y}_t,\bar{z}_t) - \sigma'\sqrt{\epsilon}(\partial_z\xi)\bar{G}_2(\bar{x}_t,\bar{y}_t,\bar{z}_t)\big]\,\mathrm{d}W_t \\ \mathrm{d}\bar{y}_t &= \big[\epsilon\bar{g}_1(\bar{x}_t,\bar{y}_t,\bar{z}_t) + \mathcal{O}(\epsilon(\sigma')^2)\big]\,\mathrm{d}t + \sigma'\sqrt{\epsilon}\big[\bar{G}_1(\bar{x}_t,\bar{y}_t,\bar{z}_t) - (\partial_z\eta)\bar{G}_2(\bar{x}_t,\bar{y}_t,\bar{z}_t)\big]\,\mathrm{d}W_t\,,\end{aligned} \qquad (4.30)$$

where $\bar{f}(\bar{x},\bar{y},\bar{z}) = a\bar{y} + b\bar{x}^2 + c\bar{x}\bar{y} + d\bar{y}^2 + \mathcal{O}(z, \|(\bar{x},\bar{y})^{\mathrm{T}}\|^3)$. By a scaling of $\bar{x},\bar{y}$ and time, we can achieve that $a = b = 1$ and $\bar{g}_1(0,0,0) = 1$.

The final step is a normal-form transformation $\hat{x} = \bar{x} - \frac{1}{2}c\bar{x}^2 - d\bar{x}\bar{y}$, which eliminates the terms of order $\bar{x}\bar{y}$ and $\bar{y}^2$ in the drift term of $\mathrm{d}\bar{x}_t$. Applying again Itô's formula yields the result. □

Remark 4.6. It is possible to further simplify the drift term, in such a way that for $\epsilon = 0$ and $\sigma = 0 = \sigma'$, $g_1(x,y,z) = g(z) + g_{11}(x,y,z)$ where $g_{11}(0,0,z) = 0$ and $g_2(0,0,z) = 0$, see [82, pp. 9–10] and [78, p. 73]. However, this introduces a diffusion term of order σ in $\mathrm{d}y_t$, which we want to avoid.

4.3 Neighbourhood and escape

We determine now the size of fluctuations during the "jump phase" of sample paths starting on Σ_4', until they hit the section Σ_5 which is located at a distance of order 1 in the x-direction from the fold. Before giving a rigorous estimate, we briefly recall some well-known deterministic asymptotics as they are going to motivate several choices in the analysis of the stochastic dynamics.

The lowest-order approximation for the deterministic dynamics near the planar fold is

$$\epsilon\frac{\mathrm{d}x}{\mathrm{d}y} = y + x^2\,, \qquad (4.31)$$

which is just the classical Ricatti equation; see [68, pp. 68–72] or [57, p. 100]. Setting $y = \epsilon^{2/3}\theta$ and $x = \epsilon^{1/3}\tilde{x}$ removes ϵ and yields

$$\frac{\mathrm{d}\tilde{x}}{\mathrm{d}\theta} = \theta + \tilde{x}^2 \qquad (4.32)$$

as the system of first approximation [2, p. 175] which also appears as the key asymptotic problem in the blow-up analysis [55, p. 293] of the non-degenerate fold. It is known [68, pp. 68–72] that there exists an orbit $\theta(\tilde{x})$ of (4.32) with

$$\theta(\tilde{x}) = -\tilde{x}^2 - \frac{1}{2\tilde{x}} + \mathcal{O}\left(\frac{1}{\tilde{x}^4}\right) \qquad \text{as } x \to -\infty \,, \tag{4.33}$$

$$\theta(\tilde{x}) = \theta^* - \frac{1}{\tilde{x}} + \mathcal{O}\left(\frac{1}{\tilde{x}^3}\right) \qquad \text{as } x \to \infty \,, \tag{4.34}$$

which is the extension of the attracting slow manifold through the fold region; the constant θ^* is the horizontal asymptote which can be expressed as the zero of suitable Bessel functions. However, if we look at the variational equation of (4.32) around $\theta(\tilde{x})$ to leading order it follows that

$$\frac{\mathrm{d}\xi}{\mathrm{d}\theta} = 2\,\frac{1}{\theta^* - \theta}\xi, \qquad \text{as } \tilde{x} \to \infty \;\; (\text{or } \theta \to \theta^*). \tag{4.35}$$

The solution is given by

$$\xi(\theta) \cong \frac{1}{(\theta^* - \theta)^2}\xi(\theta_0) \,. \tag{4.36}$$

This growth of the linearization in the fast direction turns out to be too fast to apply directly the same method to control stochastic sample paths as in the previous cases. However, we do not need such a precise control of fluctuations in the fast direction. It is sufficient to show that sample paths are likely to stay in a tube around the deterministic solution, with some specific extension in the slow directions y and z. To do so, we will compare the random process with different deterministic solutions on successive time intervals $[\theta_n, \theta_{n+1}]$ during which fluctuations in the fast direction remain bounded (cf. Figure 4). The expression (4.36) shows that a possible choice are geometrically accumulating θ_n of the form $\theta_n = \theta^* - 2^{-n}$. During the interval $[\theta_n, \theta_{n+1}]$, the deterministic solution $x(\theta)$ moves by a distance of order $\epsilon^{1/3}(2^{n+1} - 2^n) = \epsilon^{1/3}2^n$. For $x(\theta)$ to reach order 1, we need to choose n of order $|\log \epsilon|$.

To make the last idea rigorous, we write the system (4.24) on the timescale $\theta = \epsilon^{1/3}t = \epsilon^{-2/3}s$ as

$$\begin{aligned}
\mathrm{d}x_\theta &= \frac{1}{\epsilon^{1/3}}\hat{f}(x_\theta, y_\theta, z_\theta)\,\mathrm{d}\theta \qquad &&+ \frac{\sigma}{\epsilon^{1/6}}\widehat{F}_1(x_\theta, y_\theta, z_\theta)\,\mathrm{d}W_\theta + \sigma'\epsilon^{1/3}\widehat{F}_2(x_\theta, y_\theta, z_\theta)\,\mathrm{d}W_\theta \,, \\
\mathrm{d}y_\theta &= \epsilon^{2/3}\hat{g}_1(x_\theta, y_\theta, z_\theta; \epsilon, \sigma')\,\mathrm{d}\theta &&+ \sigma'\epsilon^{1/3}\widehat{G}_1(x_\theta, y_\theta, z_\theta)\,\mathrm{d}W_\theta \,, \\
\mathrm{d}z_\theta &= \epsilon^{2/3}\hat{g}_2(x_\theta, y_\theta, z_\theta; \epsilon)\,\mathrm{d}\theta \qquad &&+ \sigma'\epsilon^{1/3}\widehat{G}_2(x_\theta, y_\theta, z_\theta)\,\mathrm{d}W_\theta \,,
\end{aligned} \tag{4.37}$$

where $\hat{f}(x, y, z) = y + x^2 + \mathcal{O}(z, \left\|(x, y)^{\mathrm{T}}\right\|^3, \epsilon, \sigma^2)$. Given $\delta > 0$ of order 1, there exists a δ_0 of order 1 such that by restricting the analysis to a cube of size δ_0, we may assume that $|\hat{g}_1 - 1| < \delta$ and $|\hat{g}_2| < \delta$.

For convenience, we set $\Sigma_0^* = \{(x_0, y_0, z_0)\} \subset \Sigma_4'$, where we recall that the initial condition satisfies $x_0 \asymp -\epsilon^{1/3}$, $y_0 = c_1\epsilon^{2/3}$ for some $c_1 > 0$, and that we may assume $|z_0| \ll \epsilon^{2/3}$. For $n \geqslant 1$ and $\epsilon > 0$ such that $\epsilon^{1/3}2^n < \delta_0$, we introduce sets

$$\Sigma_n^* = \left\{(x, y, z) \colon x = \epsilon^{1/3}2^n, (y, z) \in D_n\right\} \,, \tag{4.38}$$

see Figure 4. The sets D_n are defined inductively as follows:

$$D_1 = (y_1 - c_2\epsilon^{2/3}, y_1 + c_2\epsilon^{2/3}) \times (-c_2\epsilon^{2/3}, c_2\epsilon^{2/3}) \,, \tag{4.39}$$

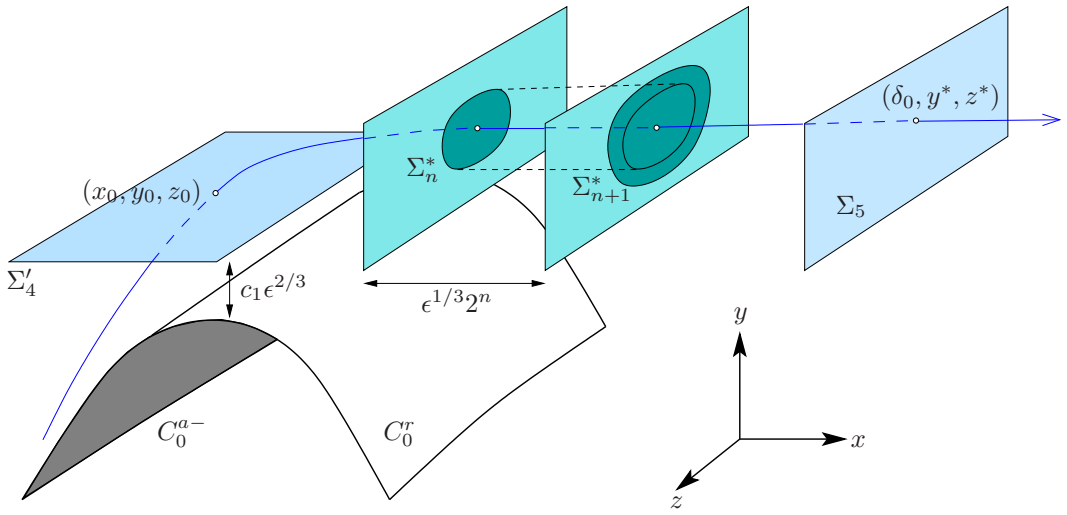

FIGURE 4. Geometry of sections near the regular fold.

where y_1 is such that $(2\epsilon^{1/3}, y_1, z_1)$ belongs to the deterministic orbit starting in (x_0, y_0, z_0), and $c_2 < c_1$ is a sufficiently small constant. Given D_n, the next set D_{n+1} is chosen as the $c_2\epsilon^{2/3}2^{-n/2}$-neighbourhood of the image of D_n under the deterministic Poincaré map from Σ_n^* to Σ_{n+1}^*. It is not difficult to show that for sufficiently small δ and c_2, the time needed for the deterministic flow to go from Σ_n^* to Σ_{n+1}^* is of order $\theta = \mathcal{O}(2^{-n})$. During this time, y and z vary by $\mathcal{O}(\epsilon^{2/3}2^{-n})$ at most, and thus

$$\operatorname{diam}(D_{n+1}) \leqslant \operatorname{diam}(D_n) + \mathcal{O}(\epsilon^{2/3}2^{-n/2}) \,. \tag{4.40}$$

The geometric decay in $2^{-n/2}$ shows that the diameter of the D_n has a uniform bound of order $\epsilon^{2/3}$. In fact, by taking δ small, we can make the extension of D_n in the z-direction small.

We return to the stochastic system (4.37). Fix n. For an initial condition $(x_n, y_n, z_n) \in \Sigma_n^*$, we denote by $(x_\theta^{\det}, y_\theta^{\det}, z_\theta^{\det})$ and $(x_\theta, y_\theta, z_\theta)$ the deterministic and stochastic solutions starting in (x_n, y_n, z_n). We write $\mathbb{P}^{(x_n,y_n,z_n)}$ whenever we wish to stress the initial condition. Consider the stopping times

$$\begin{aligned}
\tau_{n+1} &= \inf\{\theta\colon (x_\theta, y_\theta, z_\theta) \in \Sigma_{n+1}^*\} \,, \\
\tau_{n+1}^{\det} = \tau^{\det} &= \inf\{\theta\colon (x_\theta^{\det}, y_\theta^{\det}, z_\theta^{\det}) \in \Sigma_{n+1}^*\} \,, \\
\tau_\xi = \tau_\xi^{(n)}(h) &= \inf\{\theta\colon |x_\theta - x_\theta^{\det}| > h2^{-n/2}\} \,, \\
\tau_\eta = \tau_\eta^{(n)}(h_1) &= \inf\{\theta\colon \|(y_\theta, z_\theta) - (y_\theta^{\det}, z_\theta^{\det})\| > h_1 2^{-n/2}\} \,.
\end{aligned} \tag{4.41}$$

We first establish that sample paths are likely to go from Σ_n^* to Σ_{n+1}^* in a time of order $\theta = \mathcal{O}(2^{-n})$, as in the deterministic case.

Lemma 4.7. *There exist $h_0, c, c_2, \kappa > 0$, not depending on n, such that for all initial conditions $(x_n, y_n, z_n) \in \Sigma_n^*$ and $h \leqslant h_0\epsilon^{1/3}$, $h_1 \leqslant c_2\epsilon^{2/3}$,*

$$\mathbb{P}^{(x_n,y_n,z_n)}\bigl\{\tau_{n+1} \wedge \tau_\xi^{(n)}(h) \wedge \tau_\eta^{(n)}(h_1) > c2^{-n}\bigr\} \leqslant \exp\biggl\{-\kappa\frac{2^{3n}\epsilon}{\sigma^2 + (\sigma')^2\epsilon}\biggr\} \,. \tag{4.42}$$

Proof. First note that $\tau_{n+1} > c2^{-n}$ implies that either x does not reach the level $\epsilon^{1/3}2^{n+1}$ before time $c2^{-n}$ or that x does reach $\epsilon^{1/3}2^{n+1}$ at a stopping time $\tau^x_{n+1} \leqslant c2^{-n}$ while $(y_{\tau^x_{n+1}}, z_{\tau^x_{n+1}}) \notin D_{n+1}$.

Let us estimate the probability that $\tau^x_{n+1} > c2^{-n}$. Note that h_0 and c_2 can be chosen sufficiently small to guarantee that $\hat{f} \geqslant \epsilon^{2/3}2^{2n-2}$ for all times $\theta \leqslant \tau_\xi \wedge \tau_\eta$. From the representation

$$x_{c2^{-n}} = \epsilon^{1/3}2^n + \frac{1}{\epsilon^{1/3}} \int_0^{c2^{-n}} \hat{f}\,\mathrm{d}\theta + \frac{\sigma}{\epsilon^{1/6}} \int_0^{c2^{-n}} \widehat{F}_1\,\mathrm{d}W_\theta + \sigma'\epsilon^{1/3} \int_0^{c2^{-n}} \widehat{F}_2\,\mathrm{d}W_\theta \tag{4.43}$$

we find that

$$\begin{aligned}
&\mathbb{P}\{x_{c2^{-n}} < \epsilon^{1/3}2^{n+1}, \tau_\xi \wedge \tau_\eta > c2^{-n}\} \\
&\quad\leqslant \mathbb{P}\biggl\{ \frac{\sigma}{\epsilon^{1/6}} \int_0^{c2^{-n}} \widehat{F}_1\,\mathrm{d}W_\theta + \sigma'\epsilon^{1/3} \int_0^{c2^{-n}} \widehat{F}_2\,\mathrm{d}W_\theta < -\left(\frac{1}{4}c - 1\right)\epsilon^{1/3}2^n \biggr\} \\
&\quad\leqslant \exp\biggl\{ -\frac{(\frac{1}{4}c-1)^2\epsilon^{2/3}2^{2n}}{Mc2^{-n}[\sigma^2\epsilon^{-1/3} + (\sigma')^2\epsilon^{2/3}]} \biggr\}
\end{aligned} \tag{4.44}$$

for some constant $M > 0$, provided $c > 4$. In the last line, we used the fact that if $M_t = \int_0^t F(s,\cdot)\,\mathrm{d}W_t$ with integrand $F(s,\omega)$ bounded in absolute value by a constant K, then Novikov's condition [49, pp. 198–199] is satisfied and thus

$$Z_t = \exp\biggl\{ \gamma M_t - \frac{\gamma^2}{2} \int_0^t F(s,\omega)^2\,\mathrm{d}s \biggr\} \tag{4.45}$$

is a martingale for any $\gamma > 0$. It follows that for $h > 0$,

$$\mathbb{P}\bigl\{M_t > h\bigr\} \leqslant \mathbb{P}\bigl\{Z_t > \mathrm{e}^{\gamma h - \gamma^2K^2t/2}\bigr\} \leqslant \mathrm{e}^{-\gamma h + \gamma^2K^2t/2}\,\mathbb{E}\{Z_t\} = \mathrm{e}^{-\gamma h + \gamma^2K^2t/2}\,, \tag{4.46}$$

where we used Markov's inequality and the fact that a martingale has constant expectation.

Thus we obtained a bound on the probability of x not reaching $\epsilon^{1/3}2^{n+1}$ despite of ξ and η remaining small. It remains to consider the case $(y_{\tau^x_{n+1}}, z_{\tau^x_{n+1}}) \notin D_{n+1}$ for $\tau^x_{n+1} \leqslant c2^{-n}$.

By (4.37), the lower bound on $\hat{f}$ and the fact that $x^{\mathrm{det}}_{\tau^{\mathrm{det}}} = \epsilon^{1/3}2^{n+1} = x_{\tau^x_{n+1}}$, we see that on the set $\Omega' = \{\omega \in \Omega\colon \tau^{\mathrm{det}} \vee \tau^x_{n+1}(\omega) \leqslant \tau_\xi(\omega) \wedge \tau_\eta(\omega)\}$,

$$\begin{aligned}
|y^{\mathrm{det}}_{\tau^x_{n+1}} - y^{\mathrm{det}}_{\tau^{\mathrm{det}}}| &\leqslant \mathit{const}\ \epsilon^{2/3}|\tau^x_{n+1} - \tau^{\mathrm{det}}| \leqslant \frac{\mathit{const}\ \epsilon^{2/3}}{\epsilon^{1/3}2^{2n-2}}|x^{\mathrm{det}}_{\tau^x_{n+1}} - x^{\mathrm{det}}_{\tau^{\mathrm{det}}}| \\
&\leqslant \mathcal{O}(\epsilon^{1/3}h2^{-5n/2}) \leqslant \mathcal{O}(h_0\epsilon^{2/3}2^{-5n/2})
\end{aligned} \tag{4.47}$$

and $|y_{\tau^x_{n+1}} - y^{\mathrm{det}}_{\tau^x_{n+1}}| \leqslant h_1 2^{-n/2} = c_2\epsilon^{2/3}2^{-n/2}$. Similar estimates hold for the z-coordinate. Since $(y^{\mathrm{det}}_{\tau^{\mathrm{det}}}, z^{\mathrm{det}}_{\tau^{\mathrm{det}}})$ belongs to the image of D_n under the deterministic Poincaré map, we conclude that $(y_{\tau^x_{n+1}}, z_{\tau^x_{n+1}})$ belongs to an $\epsilon^{2/3}2^{-n/2}$-neighbourhood of this image. Thus $\tau^x_{n+1} = \tau_{n+1}$ on Ω'. Choosing c large enough to guarantee $\tau^{\mathrm{det}} \leqslant c2^{-n}$ concludes the proof. □

The next result gives a bound on fluctuations of sample paths, up to time $c2^{-n}$.

Lemma 4.8. *There exist $M, h_0 > 0$ such that for all initial conditions $(x_n, y_n, z_n) \in \Sigma_n^*$ and all $h, h_1 > 0$ satisfying $h \leqslant h_0 \epsilon^{1/3} 2^{5n/2}$, $h_1 \leqslant h_0 \epsilon^{-1/3} 2^{7n/2}$, $h^2 \leqslant h_0 \epsilon^{-1/3} 2^{7n/2} h_1$ and $h_1^2 \leqslant h_0 \epsilon^{1/3} 2^{5n/2} h$,*

$$\begin{aligned}\mathbb{P}^{(x_n,y_n,z_n)}\big\{\tau_\xi^{(n)}(h) \wedge \tau_\eta^{(n)}(h_1) < c2^{-n}\big\} \\ \leqslant \exp\left\{-\frac{h^2}{M(\sigma^2 + (\sigma')^2\epsilon)\epsilon^{-1/3}}\right\} + \exp\left\{-\frac{h^2}{M(\sigma')^2 2^{-2n}}\right\} \\ + \exp\left\{-\frac{h_1^2}{M(\sigma^2 + (\sigma')^2\epsilon)\epsilon 2^{-2n}}\right\} + \exp\left\{-\frac{h_1^2}{M(\sigma')^2\epsilon^{2/3}}\right\}. \end{aligned} \tag{4.48}$$

Proof. The proof is similar to the proof of Proposition 4.3. First note that the linearization $a(\theta) = \partial_x \hat{f}(x_\theta^{\det}, y_\theta^{\det}, z_\theta^{\det})$ has order $x_\theta^{\det}$, satisfying $x_\theta^{\det} \leqslant const\ \epsilon^{1/3} 2^{k+1}$ for $\theta \leqslant \tau_{k+1}^{\det}$ for any k. Since $\tau_{k+1}^{\det} \asymp 2^{-k}$ for all k, $x_\theta^{\det}$ remains of order $\epsilon^{1/3} 2^n$ for all $\theta \leqslant c2^{-n}$. Thus $a(\theta) = \mathcal{O}(\epsilon^{1/3} 2^n)$ for all $\theta \leqslant c2^{-n}$, which implies

$$\alpha(\theta, \phi) = \int_\phi^\theta a(u)\,\mathrm{d}u \leqslant \mathcal{O}(\tau^{\det} \epsilon^{1/3} 2^n) \leqslant \mathcal{O}(2^{-n} \epsilon^{1/3} 2^n) = \mathcal{O}(\epsilon^{1/3}) \tag{4.49}$$

for all $0 \leqslant \phi \leqslant \theta \leqslant c2^{-n}$. Using this, one shows that the analogue of (3.12) admits solutions $p_1(\theta), p_2(\theta) = \mathcal{O}(\epsilon^{-1/3} 2^{-n})$, so that by Lemma 3.2,

$$\begin{aligned} U_{\xi\xi}(\theta, \phi) &= \mathcal{O}(1)\,, & U_{\xi\eta}(\theta, \phi) &= \mathcal{O}(\epsilon^{-1/3} 2^{-n})\,, \\ U_{\eta\xi}(\theta, \phi) &= \mathcal{O}(\epsilon^{2/3} 2^{-n})\,, & U_{\eta\eta}(\theta, \phi) &= \mathcal{O}(1)\,, \end{aligned} \tag{4.50}$$

for $0 \leqslant \phi \leqslant \theta \leqslant c2^{-n}$. It follows from computations similar to those yielding (4.18) that

$$\mathbb{P}\big\{\tau_\xi < c2^{-n} \wedge \tau_\eta\big\} \leqslant \exp\left\{-\frac{[h - M\epsilon^{-1/3} 2^{-5n/2}(h^2 + h_1^2)]^2}{M(\sigma^2 + (\sigma')^2\epsilon)\epsilon^{-1/3}}\right\} + \mathrm{e}^{-h^2/(M(\sigma')^2 2^{-2n})}\,. \tag{4.51}$$

Since we are working on rather short time intervals, we can approximate the stochastic integral by the same Gaussian martingale on the whole time interval. Thus there is no subexponential prefactor of the type $\lceil\cdot\rceil$.

In a similar way, using the fact that $\epsilon^{1/3} \leqslant \delta_0 2^{-n}$, we get

$$\mathbb{P}\big\{\tau_\eta < c2^{-n} \wedge \tau_\xi\big\} \leqslant \exp\left\{-\frac{[h_1 - M\epsilon^{1/3} 2^{-7n/2}(h^2 + h_1^2)]^2}{M(\sigma^2 + (\sigma')^2\epsilon)\epsilon 2^{-2n}}\right\} + \mathrm{e}^{-h_1^2/(M(\sigma')^2\epsilon^{2/3})}\,. \tag{4.52}$$

The conditions on h, h_1 guarantee that the terms in $(h^2 + h_1^2)$ are negligible. □

The conditions on h and h_1 are illustrated in Figure 5.

Putting the preceeding two results together, we obtain the following estimate on the spreading of sample paths when they hit Σ_5.

Proposition 4.9. *Denote by (y^*, z^*) the point where the deterministic solution starting in $(x_0, y_0, z_0) \in \Sigma_4'$ first hits $\Sigma_5 = \{x = \delta_0\}$. Then there exist $C, \kappa, h_0 > 0$ such that for any h_1 satisfying $h_1 \leqslant h_0 \epsilon^{2/3}$, the stochastic sample path starting in (x_0, y_0, z_0) first hits Σ_5 at time $\tau = \tau_{\Sigma_5}$ at a point $(\delta_0, y_\tau, z_\tau)$ such that*

$$\mathbb{P}\left\{\|(y_\tau, z_\tau) - (y^*, z^*)\| > h_1\right\} \leqslant C|\log\epsilon|\left[\exp\left\{-\frac{\kappa h_1^2}{\sigma^2\epsilon + (\sigma')^2\epsilon^{1/3}}\right\} + \exp\left\{-\frac{\kappa\epsilon}{\sigma^2 + (\sigma')^2\epsilon}\right\}\right]. \tag{4.53}$$

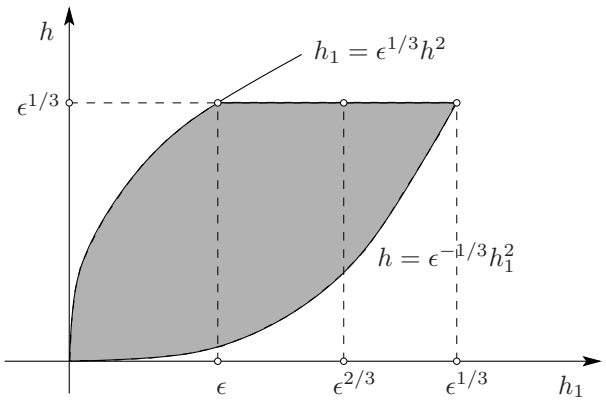

FIGURE 5. The shaded area is the set of (h_1, h) satisfying the conditions given in Lemma 4.8 (if $h_0 = 1$ and $n = 0$). Lemma 4.7 requires in addition that $h_1 \leqslant \epsilon^{2/3}$.

Proof. Let N be the largest integer such that $\epsilon^{1/3} 2^N \leqslant \delta_0$, and $\tau^{(n)} = \tau_\xi^{(n)}(h) \wedge \tau_\eta^{(n)}(h_1/2)$ for $n = 1, \dots, N$, where $h = h_0(\epsilon^{1/3} \wedge \epsilon^{-1/6} h_1^2)$ is taken as large as possible, cf. Figure 5.

If τ_{h_1} denotes the first time the stochastic sample path leaves a tube of size h_1 around the deterministic solution, the left-hand side of (4.53) can be bounded above by $\mathbb{P}\{\tau_{h_1} < \tau_{\Sigma_5}\}$. Since

$$\bigcap_{n=1}^{N} \{\tau_{n+1} \leqslant c2^{-n} \wedge \tau^{(n)}\} \subset \left\{ \tau_{\Sigma_5} \leqslant \sum_{n=1}^{N} c2^{-n} \wedge \tau_{h_1} \right\}, \tag{4.54}$$

we have the bound

$$\mathbb{P}\{\tau_{h_1} < \tau_{\Sigma_5}\} \leqslant \mathbb{P}\left\{ \tau_{h_1} \wedge \sum_{n=1}^{N} c2^{-n} < \tau_{\Sigma_5} \right\} \leqslant \sum_{n=1}^{N} \mathbb{P}\{\tau_{n+1} > c2^{-n} \wedge \tau^{(n)}\}. \tag{4.55}$$

Each term of the sum is bounded by $\mathbb{P}\{\tau_{n+1} \wedge \tau^{(n)} > c2^{-n}\} + \mathbb{P}\{\tau^{(n)} < c2^{-n} \wedge \tau_{n+1}\}$, so that the result follows from the last two lemmas. By distinguishing the cases $h_1 \geqslant \epsilon$ and $h_1 \leqslant \epsilon$, one checks that our choice of h implies that the terms in h^2 are negligible, compared to at least one of the two summands on the right-hand side of (4.53). □

This result implies that the spreading in the y- and z-directions on Σ_5, for a given initial condition on Σ_4', is of order

$$\sigma\sqrt{\epsilon} + \sigma'\epsilon^{1/6}. \tag{4.56}$$

5 The folded node

In this section we analyze the transition $\Sigma_1 \to \Sigma_2$ of sample paths in a neighbourhood of the folded-node point p^*. For convenience, we translate the origin of the coordinate system to p^*. We will decompose the transition into three parts, by introducing further sections $\Sigma_1' = \{x = \delta\sqrt{\epsilon}\}$ and $\Sigma_1'' = \{x = -\delta\sqrt{\epsilon}\}$, where δ is a small constant of order 1. The transitions $\Sigma_1 \to \Sigma_1'$, $\Sigma_1' \to \Sigma_1''$, and $\Sigma_1'' \to \Sigma_2$ are analyzed, respectively, in Subsection 5.2, in Subsections 5.3 and 5.4, and in Subsection 5.5.

5.1 Normal form

We start by making a preliminary transformation to normal form near the folded node point p^*. Recall once again that $t = s/\epsilon$ denotes the fast timescale.

Proposition 5.1. *Near a folded-node point $p^* \in L^+$ satisfying the assumptions (A1) and (A3), there exist a smooth change of coordinates and a random time change such that* (3.1) *is locally given by*

$$\begin{aligned} \mathrm{d}x_t &= \hat{f}(x_t, y_t, z_t; \epsilon, \sigma, \sigma')\,\mathrm{d}t + \left[\sigma \widehat{F}_1(x_t, y_t, z_t) + \sigma'\sqrt{\epsilon}\,\widehat{F}_2(x_t, y_t, z_t)\right] \mathrm{d}W_t\,, \\ \mathrm{d}y_t &= \epsilon \hat{g}_1(x_t, y_t, z_t; \epsilon, \sigma')\,\mathrm{d}t \;+ \sigma'\sqrt{\epsilon}\,\widehat{G}_1(x_t, y_t, z_t)\,\mathrm{d}W_t\,, \\ \mathrm{d}z_t &= \frac{1}{2}\epsilon\mu\,\mathrm{d}t \qquad\qquad\qquad\qquad + \sigma'\sqrt{\epsilon}\,\widehat{G}_2(x_t, y_t, z_t)\,\mathrm{d}W_t\,, \end{aligned} \tag{5.1}$$

where $\mu \in (0,1)$ is the ratio of weak and strong eigenvalues at the folded node (see also [17, p. 4793] or [24, p. 48]), and

$$\begin{aligned} \hat{f}(x, y, z; \epsilon, \sigma, \sigma') &= y - x^2 + \mathcal{O}(\left\|(x,y,z)^{\mathrm{T}}\right\|^3, \epsilon\left\|(x,y,z)^{\mathrm{T}}\right\|, \sigma^2, (\sigma')^2\epsilon)\,, \\ \hat{g}_1(x, y, z; \epsilon, \sigma') &= -(\mu+1)x - z + \mathcal{O}(y, (x+z)^2, \epsilon, (\sigma')^2)\,, \end{aligned} \tag{5.2}$$

while the diffusion matrices $\widehat{F}_1, \widehat{F}_2, \widehat{G}_1, \widehat{G}_2$ all remain of order 1.

Proof. The result is again a stochastic analogue of the transformation result for deterministic systems, see [82, pp. 8–10], as well as [83].

We start by translating the origin of the coordinate system to the folded-node point p^*. Note that the failure of the normal-switching condition (A2) implies that the vectors $(\frac{\partial f}{\partial y}, \frac{\partial f}{\partial z})(0)$ and $g(0)$ are orthogonal. We may thus rotate coordinates in such a way that $g_1(0) = 0$ and $\frac{\partial f}{\partial z}(0) = 0$. This rotation does not change the order of magnitude of the diffusion coefficients $\sigma'\sqrt{\epsilon}\,G_1$ and $\sigma'\sqrt{\epsilon}\,G_2$.

Calculating the linearization of the desingularized slow flow (2.9) and using Assumption (A1), we see that $g_2(0) \neq 0$, since otherwise $p^* = 0$ would not be a node for (2.9) as required by Assumption (A3). We can thus carry out locally a random time change given by

$$\mathrm{d}\tilde{t} = \frac{g_2(x_t, y_t, z_t)}{g_2(0,0,0)}\,\mathrm{d}t\,. \tag{5.3}$$

Lemma A.4 in Appendix A shows that this time change yields a system in which all drift coefficients have been multiplied by $g_2(0,0,0)/g_2(x,y,z)$, and all diffusion coefficients have been multiplied by $[g_2(0,0,0)/g_2(x,y,z)]^{1/2}$. We may thus assume that $g_2(x,y,z)$ is constant and equal to $g_2(0,0,0)$ in (3.1).

The remainder of the proof is similar to the proof of Proposition 4.5. A transformation $x = \bar{x} + \xi(z)$, $y = \bar{y} + \eta(z)$ rectifies the fold curve, i.e. $f(\xi(z), \eta(z), z) = 0$ and $\frac{\partial f}{\partial x}(\xi(z), \eta(z), z) = 0$ in a neighbourhood of $z = 0$, and thus

$$\bar{f}(\bar{x}, \bar{y}, z) = a\bar{y} + b\bar{x}^2 + c\bar{x}\bar{y} + d\bar{y}^2 + e\bar{y}z + k\epsilon + \mathcal{O}(\|(\bar{x}, \bar{y}, z)\|^3, \epsilon z, (\sigma')^2\epsilon)\,. \tag{5.4}$$

The standard form of f and g_1 can then be achieved by combining a translation of x by $\mathcal{O}(\epsilon)$, a scaling of space and a near-identity transformation $\hat{x} = \bar{x} - \frac{1}{2}c\bar{x}^2 - d\bar{x}\bar{y} - e\bar{x}z$ (cf. [82, pp. 9–10]). These transformations do not change the order of the diffusion coefficients for y and z. □

5.2 Approach

In this section, we consider solutions of the normal form (5.1), starting at a fast time $s_0 \asymp -1$ on Σ_1, as long as $x \geqslant \mathcal{O}(\sqrt{\epsilon})$. We fix a deterministic solution $(x_s^{\det}, y_s^{\det}, z_s^{\det})$ which is sufficiently close to the strong canard to display SAOs when approaching the folded-node point p^*. From the deterministic analysis we know that

$$a(s) = \partial_x f(x_s^{\det}, y_s^{\det}, z_s^{\det}) = -2x_s^{\det} = cs + \mathcal{O}(s^2) \qquad \text{for } s_0 \leqslant s \leqslant -\sqrt{\epsilon}\,, \tag{5.5}$$

where c is a constant of order 1. Scaling time if necessary, we may assume that $c = 1$. The linearization of the deterministic system at $(x_s^{\det}, y_s^{\det}, z_s^{\det})$ has the form $\epsilon\dot{\zeta} = \mathcal{A}(s)\zeta$, where

$$\mathcal{A}(s) = \begin{pmatrix} A(s) & c_1(s) \\ 0 & 0 \end{pmatrix}, \qquad A(s) = \begin{pmatrix} 2a(s) & 1 + \mathcal{O}(s^2) \\ -\epsilon(1+\mu) + \mathcal{O}(\epsilon s) & \mathcal{O}(\epsilon) \end{pmatrix}, \tag{5.6}$$

and $c_1(s) = (1 + \mathcal{O}(s^2), \mathcal{O}(\epsilon))^{\mathrm{T}}$. We have used the fact that $s^2 \geqslant \epsilon$ to simplify the expression of the error terms.

For $s \leqslant -\sqrt{\epsilon}$, the eigenvalues of $A(s)$ behave like s and $\epsilon/|s|$. This implies that while for $s \asymp -1$, the variable x is faster than both y and z, $\dot{x}$ and $\dot{y}$ become of comparable order $1/\sqrt{\epsilon}$ as s approaches $-\sqrt{\epsilon}$. This is the well-known effect that one may extend the normally hyperbolic theory slightly near fold points from $x \asymp 1$ up to a neighbourhood with $x \asymp \sqrt{\epsilon}$, see [24, pp. 48–49]. Instead of blocking y and z as in (3.5), we write

$$\xi_s = \begin{pmatrix} x_s \\ y_s \end{pmatrix} - \begin{pmatrix} x_s^{\det} \\ y_s^{\det} \end{pmatrix}, \qquad \eta_s = z_s - z_s^{\det}, \qquad \zeta_s = \begin{pmatrix} \xi_s \\ \eta_s \end{pmatrix}, \tag{5.7}$$

since $\dot{x}$ and $\dot{y}$ eventually become comparable. Then ζ_s obeys a system of the form

$$\mathrm{d}\zeta_s = \frac{1}{\epsilon}\mathcal{A}(s)\zeta_s\,\mathrm{d}s + \begin{pmatrix} \frac{\sigma}{\sqrt{\epsilon}}\mathcal{F}_1(\zeta_s, s) + \sigma'\mathcal{F}_2(\zeta_s, s) \\ \sigma'\mathcal{G}_1(\zeta_s, s) \\ \sigma'\mathcal{G}_2(\zeta_s, s) \end{pmatrix} \mathrm{d}W_s + \begin{pmatrix} \frac{1}{\epsilon}b_x(\zeta_s, s) \\ b_y(\zeta_s, s) \\ 0 \end{pmatrix} \mathrm{d}s\,. \tag{5.8}$$

The principal solution of $\epsilon\dot{\zeta} = \mathcal{A}(s)\zeta$ has the block structure

$$U(s,r) = \begin{pmatrix} V(s,r) & \dfrac{1}{\epsilon}\displaystyle\int_s^r V(s,u)c_1(u)\,\mathrm{d}u \\ 0 & 1 \end{pmatrix}, \tag{5.9}$$

where $V(s,r)$ denotes the principal solution of $\epsilon\dot{\xi} = A(s)\xi$.

Lemma 5.2. *For $s_0 \leqslant r \leqslant s \leqslant -\sqrt{\epsilon}$, the matrix elements of $V(s,r)$ satisfy*

$$\begin{aligned}
V_{xx}(s,r) &= \mathcal{O}\biggl(\frac{|a(r)|^{1+\mu}}{|a(s)|^{1+\mu}}\,\mathrm{e}^{\alpha(s,r)/\epsilon}\biggr), \\
V_{xy}(s,r) &= \mathcal{O}\biggl(\frac{|a(r)|^{\mu}}{|a(s)|^{1+\mu}}\,\mathrm{e}^{\alpha(s,r)/\epsilon} + \frac{|a(s)|^{\mu}}{|a(r)|^{1+\mu}}\biggr), \\
V_{yx}(s,r) &= \mathcal{O}\biggl(\epsilon\,\frac{|a(r)|^{1+\mu}}{|a(s)|^{2+\mu}}\,\mathrm{e}^{\alpha(s,r)/\epsilon} + \epsilon\,\frac{|a(s)|^{1+\mu}}{|a(r)|^{2+\mu}}\biggr), \\
V_{yy}(s,r) &= \mathcal{O}\biggl(\frac{|a(s)|^{1+\mu}}{|a(r)|^{1+\mu}} + \epsilon\,\frac{|a(r)|^{\mu}}{|a(s)|^{2+\mu}}\,\mathrm{e}^{\alpha(s,r)/\epsilon}\biggr),
\end{aligned} \tag{5.10}$$

where

$$\alpha(s,r) = \int_r^s \operatorname{Tr} A(u)\,\mathrm{d}u = (s^2 - r^2) + \mathcal{O}((s-r)(s^2+r^2))\,. \tag{5.11}$$

In the particular case where $A(s) = \left(\begin{smallmatrix} 2s & 1 \\ -\epsilon(1+\mu) & 0 \end{smallmatrix}\right)$, the equation $\epsilon\dot{\xi} = A(s)\xi$ for $\xi_s = (\xi_{1,s}, \xi_{2,s})^{\mathrm{T}}$ is equivalent to a Weber equation

$$\epsilon \frac{\mathrm{d}^2\xi_1}{\mathrm{d}s^2} - 2s\frac{\mathrm{d}\xi_1}{\mathrm{d}s} + (\mu - 1)\xi_1 = 0\,, \tag{5.12}$$

and the estimates (5.10) follow directly from the asymptotics of parabolic cylinder functions [1, p. 689]; see also [77, p. 449]. In Appendix B, we provide a proof of Lemma 5.2 valid in the general case, which does not rely on these asymptotics. With the above estimates, we obtain the following result on the size of fluctuations during the approach phase.

Proposition 5.3. *Define the stopping times*

$$\begin{aligned} \tau_{\xi,1} &= \inf\{s > s_0 \colon |\xi_{1,s}| > h\}\,, \\ \tau_{\xi,2} &= \inf\{s > s_0 \colon |\xi_{2,s}| > h_1\}\,, \\ \tau_{\eta} &= \inf\{s > s_0 \colon |\eta_{s}| > h_2\}\,. \end{aligned} \tag{5.13}$$

There exist constants $\kappa, h_0 > 0$ such that for all $s_0 \leqslant s \leqslant -\sqrt{\epsilon}$, and all $h, h_1, h_2 > 0$ satisfying $h^2 + h_1^2 + h_2^2 \leqslant h_0|s|h$ and $h^2 + h_1^2 + h_2^2 \leqslant h_0 h_1$,

$$\begin{aligned} \mathbb{P}\big\{\tau_{\xi,1} \wedge \tau_{\xi,2} \wedge \tau_\eta < s\big\} \leqslant \left\lceil \frac{s - s_0}{\epsilon} \right\rceil \Bigg[&\exp\left\{ -\frac{\kappa h^2}{(\sigma^2 + (\sigma')^2)\,|s|^{-1}} \right\} \\ &+ \exp\left\{ -\frac{\kappa h_1^2}{\sigma^2\epsilon\,|s|^{-1} + (\sigma')^2\,|s|} \right\} + \exp\left\{ -\frac{\kappa h_2^2}{(\sigma')^2} \right\} \Bigg]\,. \end{aligned} \tag{5.14}$$

Proof. The proof is similar to the proof of Proposition 4.3, so we omit the details. Let us just remark that when evaluating the elements in (5.9), one encounters integrals of the form

$$\int_{s_0}^{s} \frac{|s|^\mu}{|u|^{1+\mu}}\,\mathrm{d}u = \frac{1 - y^\mu}{\mu} = |\log y|\,\frac{1 - \mathrm{e}^{-\mu|\log y|}}{\mu\,|\log y|} \tag{5.15}$$

where $y = |s|\,/\,|s_0|$. The fraction on the right-hand side being bounded, the integral is bounded by a constant times $|\log(|s|\,/\,|s_0|)|$. □

Under the condition $\sigma, \sigma' = \mathcal{O}(\epsilon^{3/4})$, we obtain the typical spreadings

- $(\sigma + \sigma')\,|s|^{-1/2}$ in the x-direction, which reaches order $(\sigma + \sigma')\epsilon^{-1/4}$ for $s \asymp -\sqrt{\epsilon}$,
- $\sigma\epsilon^{1/2}\,|s|^{-1/2} + \sigma'\,|s|^{1/2}$ in the y-direction, which reaches order $(\sigma + \sigma')\epsilon^{1/4}$ for $s \asymp -\sqrt{\epsilon}$,
- and σ' in the z-direction.

Note carefully that the integrals in (5.15) become unbounded when $s \to 0$ as $\mu \in (0,1)$ so we cannot use the same methods to control sample paths closer to the folded node.

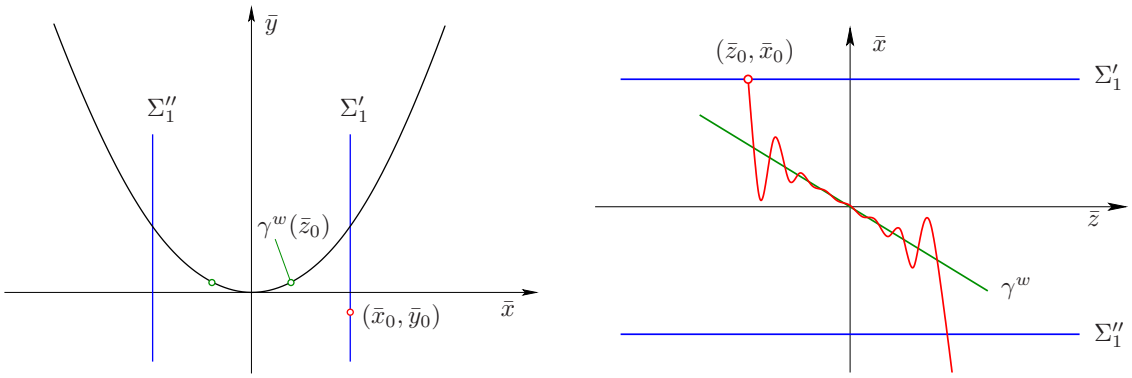

FIGURE 6. Sketch of the geometry of the orbits near the folded-node singularity.

5.3 Neighbourhood – Deterministic dynamics

In this section we briefly describe the behaviour of solutions of the normal form (5.1) in the deterministic case $\sigma = \sigma' = 0$. Recall that standard results (see [77, Section 4], [26, Theorem 2.3] and the foundational work [7, 8]) imply the existence of two primary canards and k_μ secondary canards [83] where

$$2k_\mu + 1 < \mu^{-1} < 2k_\mu + 3 \,, \tag{5.16}$$

and each canard lies in $C_\epsilon^r \cap C_\epsilon^{a+}$. One strategy to prove the existence of canards, as first suggested for the planar case in [29], is to look for transversal intersections of C_ϵ^r and C_ϵ^{a+} by extending the manifolds via the blow-up method [77] into a region near the folded node where the blow-up reduces to the scaling (or zoom-in) transformation

$$x = \sqrt{\epsilon}\,\bar{x} \,, \quad y = \epsilon\,\bar{y} \,, \quad z = \sqrt{\epsilon}\,\bar{z} \,. \tag{5.17}$$

The scaling (5.17) transforms the deterministic version of the normal form (5.1) to

$$\begin{aligned} \mu \frac{\mathrm{d}\bar{x}}{\mathrm{d}\bar{z}} &= 2\bar{y} - 2\bar{x}^2 + \mathcal{O}(\sqrt{\epsilon}) \,, \\ \mu \frac{\mathrm{d}\bar{y}}{\mathrm{d}\bar{z}} &= -2(1+\mu)\bar{x} - 2\bar{z} + \mathcal{O}(\sqrt{\epsilon}) \,. \end{aligned} \tag{5.18}$$

We consider henceforth the dynamics for $\epsilon = 0$, as results can be extended to small positive ϵ by regular perturbation theory. Note that the system is symmetric under the transformation

$$(\bar{x}, \bar{y}, \bar{z}) \mapsto (-\bar{x}, \bar{y}, -\bar{z}) \,. \tag{5.19}$$

The normal form admits a particular solution γ^w given by

$$\bar{x} = -\bar{z} \,, \qquad \bar{y} = \bar{z}^2 - \frac{\mu}{2} \,, \tag{5.20}$$

which is called the singular weak canard (there is also a singular strong canard, given by $\bar{x} = -\bar{z}/\mu$, $\bar{y} = (\bar{z}/\mu)^2 - 1/2$). Generic solutions twist a certain number of times around the weak canard, see Figure 6 for an illustration. One possibility to prove the persistence of the weak and strong canards as well as secondary canards is to analyse the zeros of the variational Weber equation as shown in [77]. To also obtain estimates on

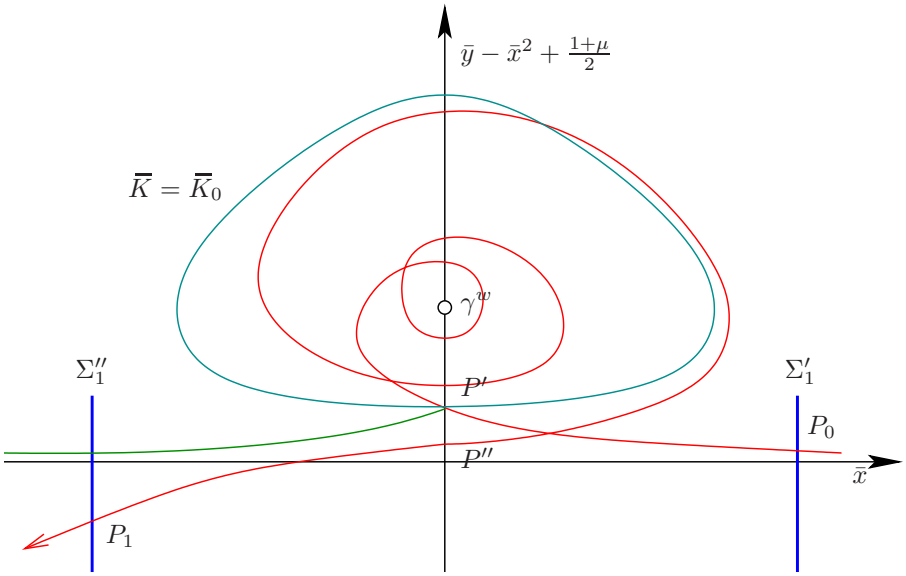

FIGURE 7. The map from Σ_1' to Σ_1'' is decomposed into three phases.

individual non-canard orbits, our aim is to determine the map from an initial condition $P_0 = (\delta, \bar{y}_0, \bar{z}_0) \in \Sigma_1'$, close to the attracting slow manifold, to the first-hitting point $P_1 = (-\delta, \bar{y}_1, \bar{z}_1) \in \Sigma_1''$. The key tool will be suitable coordinate transformations; we note that although the method only provides a small refinement of previous results, it has the advantage of being quite explicit so we choose to record the results here. We will proceed in three steps, see Figure 7:

(S1) Estimate the coordinates of $P' = (0, \bar{y}', \bar{z}')$, the first-hitting point of $\{\bar{x} = 0\}$.

(S2) Use an averaging-type transformation to describe the rotations of this orbit around the weak canard, until the last-hitting point $P'' = (0, \bar{y}'', \bar{z}'')$ of $\{\bar{x} = 0\}$.

(S3) Determine the map from P'' to P_1.

For Steps (S1) and (S3), it is useful to introduce the rectified coordinate $\eta = \bar{y} - \bar{x}^2 + (1+\mu)/2$. The normal form (with $\epsilon = 0$) in rectified coordinates then reads

$$\begin{aligned} \mu \frac{\mathrm{d}\bar{x}}{\mathrm{d}\bar{z}} &= 2\eta - (1+\mu) \,, \\ \mu \frac{\mathrm{d}\eta}{\mathrm{d}\bar{z}} &= -4\bar{x}\eta - 2\bar{z} \,. \end{aligned} \tag{5.21}$$

Lemma 5.4. *Fix constants $0 < \beta < \alpha \leqslant 1$ and $0 < L \leqslant \sqrt{(\alpha-\beta)|\log \mu|/2}$. Then for $\mu > 0$ small enough the orbit of (5.21) passing through a point $(\bar{x}, \eta, \bar{z}) = (0, \eta_*, \bar{z}_*)$ with $|\eta_*| \leqslant \mu^\alpha$ and $|\bar{z}_*| \leqslant \mu^\beta$ is given by*

$$\begin{aligned} \eta(\bar{x}) &= \mathrm{e}^{2\bar{x}^2} \biggl[\eta_* + \bar{z}_* \int_0^{2\bar{x}} \mathrm{e}^{-u^2/2} \, \mathrm{d}u + \mathcal{O}(\mu) \biggr] \bigl[1 + \mathcal{O}(\mu^\beta L^2) \bigr] \,, \\ \bar{z}(\bar{x}) &= \bar{z}_* + \mu \bar{x} \bigl[1 + \mathcal{O}(\mu^\beta) \bigr] \,, \end{aligned} \tag{5.22}$$

for all $\bar{x} \in [-L, L]$.

Proof. For $\eta = \mathcal{O}(\mu^\beta)$ the equations (5.21) can be rewritten in the form

$$\begin{aligned} \frac{\mathrm{d}\bar{z}}{\mathrm{d}\bar{x}} &= \mu[1 + \mathcal{O}(\mu^\beta)] \,, \\ \frac{\mathrm{d}\eta}{\mathrm{d}\bar{x}} &= a(\bar{x})\eta + 2\bar{z}[1 + \mathcal{O}(\mu^\beta)] \,, \end{aligned} \tag{5.23}$$

where $a(\bar{x}) = 4\bar{x}[1 + \mathcal{O}(\mu^\beta)]$. Integrating the first equation yields the expression for $\bar{z}(\bar{x})$. To obtain the expression for $\eta(\bar{x})$, observe that $\exp(\int_0^{\bar{x}} a(y)\,\mathrm{d}y) = \mathrm{e}^{2\bar{x}^2}[1 + \mathcal{O}(\mu^\beta L^2)]$ and solve the equation for η by variation of the constant. □

From this result we immediately see that the map from P_0 to P' is given by

$$\begin{aligned} \eta' &= \mathrm{e}^{-2\delta^2}\,\eta_0[1 + \mathcal{O}(\delta^2\mu^\beta)] - \bar{z}_0 \int_0^{2\delta} \mathrm{e}^{-u^2/2}\,\mathrm{d}u + \mathcal{O}(\mu) \,, \\ \bar{z}' &= \bar{z}_0 + \mathcal{O}(\delta\mu) \,, \end{aligned} \tag{5.24}$$

provided $\eta_0, \bar{z}_0 = \mathcal{O}(\mu^\beta)$ for some $\beta \in (0,1)$. Similarly, the map from P'' to P_1 is given by

$$\begin{aligned} \eta_1 &= \mathrm{e}^{2\delta^2}\left[\eta'' - \bar{z}'' \int_{-2\delta}^{0} \mathrm{e}^{-u^2/2}\,\mathrm{d}u + \mathcal{O}(\mu)\right][1 + \mathcal{O}(\delta^2\mu^\beta)] \,, \\ \bar{z}_1 &= \bar{z}'' + \mathcal{O}(\delta\mu) \,, \end{aligned} \tag{5.25}$$

provided $\eta'' = \mathcal{O}(\mu^\alpha)$ and $\bar{z}'' = \mathcal{O}(\mu^\beta)$ for some choice of $0 < \beta < \alpha \leqslant 1$. In addition, Lemma 5.4 shows that for sufficiently small $\bar{z}$, the distance at $\bar{x} = 0$ between the invariant manifolds C_ϵ^{a+} and C_ϵ^r has order $\bar{z}$. This follows from the fact that orbits in C_ϵ^{a+} should be such that $\eta(\bar{x})$ is close to 0 for large positive $\bar{x}$, while orbits in C_ϵ^r should be such that $\eta(\bar{x})$ is close to 0 for large negative $\bar{x}$.

We now turn to Step (S2), estimating the map from P' to P''. The difference u between a general solution of (5.18) and the weak canard γ^w satisfies the variational equation

$$\begin{aligned} \mu\frac{\mathrm{d}u_1}{\mathrm{d}\bar{z}} &= 4\bar{z}u_1 + 2u_2 - 2u_1^2 \,, \\ \mu\frac{\mathrm{d}u_2}{\mathrm{d}\bar{z}} &= -2(1+\mu)u_1 \,. \end{aligned} \tag{5.26}$$

Consider the variable

$$K := \left[1 + \frac{2}{1+\mu}(u_2 - u_1^2)\right] \mathrm{e}^{-2u_2/(1+\mu)} \,, \tag{5.27}$$

which is a first integral of the system when $\bar{z} = 0$. In fact, K is just a version of the classical first integral near planar degenerate folded singularities (cf. [56, Lemma 3.3; Figure 3.2], [54, Figure 5], [69, Figure 2]). Although K is not a first integral for arbitrary $\bar{z}$ it turns out that the variable K is still very useful for obtaining explicit control over the oscillations near the folded node. A short computation yields

$$\mu\frac{\mathrm{d}K}{\mathrm{d}\bar{z}} = -\frac{16\bar{z}}{1+\mu}u_1^2\,\mathrm{e}^{-2u_2/(1+\mu)} \,. \tag{5.28}$$

In [17, Section D.2] we provided an averaging result valid in a small neighbourhood of the weak canard (for the stochastic case). The following result extends this to the larger domain $\{K > 0\}$.

Proposition 5.5 (Averaged system)**.** *Set* $\rho(K) = (1-K)(1+|\log K|^{3/2})$*. For* $\bar z$ *in a neighbourhood of* 0 *and* $K>0$*, there exist a variable* $\overline K = K + \mathcal{O}(|\bar z|\,\rho(K))$*, an angular variable* φ*, a function* $\bar g$ *and constants* $c_\pm > 0$ *such that*

$$\frac{c_-}{1+|\log K|^{1/2}} \leqslant \mu \frac{\mathrm{d}\varphi}{\mathrm{d}\bar z} \leqslant c_+(1+|\log K|^{1/2}) \tag{5.29}$$

for $K \geqslant \mathcal{O}(|\bar z|)$ *and*

$$\frac{\mathrm{d}\overline K}{\mathrm{d}\varphi} = \bar z \bar g(\overline K, \bar z) + \mathcal{O}\Big((\mu+\bar z^2)\rho(\overline K)\Big)\,, \tag{5.30}$$

where $c_-(1-\overline K) \leqslant -\bar g(\overline K, \bar z) \leqslant c_+(1-\overline K)$*.*

We give the proof in Appendix C.

The averaged equation (5.30) is similar to the equation describing dynamic pitchfork or Hopf bifurcations, which display a bifurcation delay. Initially, i.e. when $\bar z = \bar z' < 0$, $\overline K$ has a value $\overline K_0 > 0$ of order $\bar z_0$. As long as $\bar z < 0$, $\overline K$ will keep increasing, and thus get so close to 1 that a time of order $|\bar z'|$ is needed, once $\bar z$ becomes positive, for $\overline K$ to decrease to the value $\overline K_0$ again. We set

$$\tau = \inf\{\bar z > 0 \colon \overline K = \overline K_0\}\,. \tag{5.31}$$

Note that the error term in (5.30) is no longer negligible when $|\bar z|$ is of order μ, but this only results in a shift of the delay by a quantity of order μ, which will be negligible.

Corollary 5.6. *Let* $\rho_1 = (\mu + \bar z_0^2)|\log \bar z_0|^{3/2}$*. We have*

$$\tau = -\bar z' + \mathcal{O}(\rho_1)\,. \tag{5.32}$$

Furthermore, the change in angle is given by $\varphi(\tau) - \varphi(\bar z') = \phi(\bar z')/\mu$*, where* ϕ *is monotonically decreasing for* $\bar z' < 0$ *and satisfies*

$$\frac{2c_-}{1+|\log \bar z_0|^{1/2}}\big[|\bar z'| + \mathcal{O}(\rho_1)\big] \leqslant \phi(\bar z') \leqslant 2c_+(1+|\log \bar z_0|^{1/2})\big[|\bar z'| + \mathcal{O}(\rho_1)\big]\,. \tag{5.33}$$

Proof. Set $Q = 1 - \overline K$. As long as $\bar z \leqslant \tau$, we can bound $|\log \overline K|$ by $|\log \bar z_0|$ and write

$$\frac{\mathrm{d}Q}{\mathrm{d}\varphi} \leqslant Q\Big[c_+\bar z + \mathcal{O}\big(Q(\mu+\bar z^2)|\log \bar z_0|^{3/2}\big)\Big]\,. \tag{5.34}$$

Using (5.29) we obtain

$$\frac{\mathrm{d}Q}{\mathrm{d}\bar z} \leqslant \frac{1}{\mu} Q\Big[c_+^2\bar z + \mathcal{O}\big(Q(\mu+\bar z^2)|\log \bar z_0|^{3/2}\big)\Big]\big[1+\mathcal{O}(|\log \bar z_0|^{1/2})\big]\,. \tag{5.35}$$

Integrating, we arrive at

$$Q(\bar z) \leqslant Q(\bar z')\exp\left\{\frac{c_+^2}{2\mu}(\bar z - \bar z')\big[\bar z + \bar z' + \mathcal{O}(\rho_1)\big]\big[1+\mathcal{O}(|\log \bar z_0|^{1/2})\big]\right\}\,. \tag{5.36}$$

This shows that $\tau \geqslant -\bar z' + \mathcal{O}(\rho_1)$. Using the corresponding lower bounds, we also get $\tau \leqslant -\bar z' - \mathcal{O}(\rho_1)$. This proves (5.32), and (5.33) follows by using (5.29) again. □

We can now draw consequences on the Poincaré map from the last results. If

$$\phi(\bar{z}') = 2\pi n\mu \;, \qquad n \in \mathbb{N} \;, \tag{5.37}$$

then the orbit will hit the plane $\{\bar{x} = 0\}$ at $P'' = P'$, which is on (or very near) the repelling slow manifold C_ϵ^r, cf. (5.24). Therefore (5.37) gives a condition on $\bar{z}'$ (and thus on $\bar{z}_0$) for the orbit being a canard. If, on the other hand,

$$\phi(\bar{z}') = (2\pi n - \theta)\mu \;, \qquad 0 < \theta < 2\pi \;, \tag{5.38}$$

the orbit will leave the set $\{\overline{K} > \overline{K}_0\}$ far from the repelling slow manifold C_ϵ^r, see Figure 7. One can then use Proposition 5.5 to estimate $\bar{z}''$, which is of the form $\bar{z}'' = \tau + \mathcal{O}(\mu\theta)$. As θ increases from 0 to 2π, P'' moves downwards until it approaches the continuation of the attracting slow manifold C_ϵ^{a+}.

Once orbits have hit $\bar{x} = 0$ at some point P'' below P', one can use (5.25) to follow their future evolution. Note in particular that the domain $\{\bar{x} < 0, \eta < 0\}$ is positively invariant, so that once orbits have reached this domain they will stay bounded away from the repelling slow manifold.

5.4 Neighbourhood – Stochastic dynamics

We now consider the stochastic dynamics of sample paths starting on Σ_1' up to the first time they hit the section $\Sigma_1'' = \{x = -\delta\sqrt{\epsilon}\,\}$. The first step is again to apply the scaling (or zoom-in)

$$x = \sqrt{\epsilon}\,\bar{x} \;, \quad y = \epsilon\,\bar{y} \;, \quad z = \sqrt{\epsilon}\,\bar{z} \;, \quad \frac{1}{2}\mu\sqrt{\epsilon}\,t = \theta \;, \tag{5.39}$$

which transforms the normal form (5.1) into

$$\begin{aligned}
\mathrm{d}\bar{x}_\theta &= \frac{2}{\mu}\big[\bar{y}_\theta - \bar{x}_\theta^2 + \mathcal{O}(\sqrt{\epsilon}\,)\big]\,\mathrm{d}\theta &&+ \sqrt{\frac{2}{\mu}}\,\big[\bar{\sigma}\widehat{F}_1(\bar{x}_\theta, \bar{y}_\theta, \bar{z}_\theta) + \bar{\sigma}'\sqrt{\epsilon}\,\widehat{F}_2(\bar{x}_\theta, \bar{y}_\theta, \bar{z}_\theta)\big]\,\mathrm{d}W_\theta \;, \\
\mathrm{d}\bar{y}_\theta &= \frac{2}{\mu}\big[-(1+\mu)\bar{x}_\theta - \bar{z}_\theta + \mathcal{O}(\sqrt{\epsilon}\,)\big]\,\mathrm{d}\theta &&+ \sqrt{\frac{2}{\mu}}\,\bar{\sigma}'\widehat{G}_1(\bar{x}_\theta, \bar{y}_\theta, \bar{z}_\theta)\,\mathrm{d}W_\theta \;, \\
\mathrm{d}\bar{z}_\theta &= \mathrm{d}\theta &&+ \sqrt{\frac{2}{\mu}}\,\bar{\sigma}'\sqrt{\epsilon}\,\widehat{G}_2(\bar{x}_\theta, \bar{y}_\theta, \bar{z}_\theta)\,\mathrm{d}W_\theta \;,
\end{aligned} \tag{5.40}$$

where

$$\bar{\sigma} = \epsilon^{-3/4}\sigma \qquad \text{and} \qquad \bar{\sigma}' = \epsilon^{-3/4}\sigma' \;. \tag{5.41}$$

The deviation ζ_θ from the deterministic solution $(\bar{x}_\theta^{\mathrm{det}}, \bar{y}_\theta^{\mathrm{det}}, \bar{z}_\theta^{\mathrm{det}})$, defined as in (5.7), satisfies a SDE of the form

$$\mathrm{d}\zeta_\theta = \frac{1}{\mu}\mathcal{A}(\theta)\zeta_\theta\,\mathrm{d}\theta + \frac{1}{\sqrt{\mu}}\begin{pmatrix} \bar{\sigma}\mathcal{F}_1(\zeta_\theta,\theta) + \bar{\sigma}'\sqrt{\epsilon}\,\mathcal{F}_2(\zeta_\theta,\theta) \\ \bar{\sigma}'\mathcal{G}_1(\zeta_\theta,\theta) \\ \bar{\sigma}'\sqrt{\epsilon}\,\mathcal{G}_2(\zeta_\theta,\theta) \end{pmatrix}\mathrm{d}W_\theta + \frac{1}{\mu}\begin{pmatrix} b_x(\zeta_\theta,\theta) \\ \sqrt{\epsilon}\,b_y(\zeta_\theta,\theta) \\ 0 \end{pmatrix}\mathrm{d}\theta \;. \tag{5.42}$$

The principal solution of $\mu\dot{\zeta} = \mathcal{A}(\theta)\zeta$ has a block structure similar to (5.9). Provided we take δ sufficiently small, the upper left block $A(\theta)$ has complex conjugated eigenvalues $a(\theta) \pm 2\,\mathrm{i}\,\omega(\theta)$, where

$$a(\theta) = -2\bar{x}_\theta^{\mathrm{det}} + \mathcal{O}(\sqrt{\epsilon}\,) \;, \qquad \omega(\theta) = \sqrt{1 - (\bar{x}_\theta^{\mathrm{det}})^2 + \mu} + \mathcal{O}(\sqrt{\epsilon}\,) \;. \tag{5.43}$$

By [17, Theorem 4.3], the principal solution $V(\theta,\phi)$ of $\mu\dot{\xi} = A(\theta)\xi$ can be written in the form

$$V(\theta,\theta_0) = \mathrm{e}^{\alpha(\theta,\theta_0)/\mu}\, S(\theta) \begin{pmatrix} \cos(\phi(\theta,\theta_0)/\mu) & \sin(\phi(\theta,\theta_0)/\mu) \\ -\sin(\phi(\theta,\theta_0)/\mu) & \cos(\phi(\theta,\theta_0)/\mu) \end{pmatrix} S(\theta_0)^{-1}\,, \tag{5.44}$$

where

$$S(\theta) = \frac{1}{\sqrt{\omega(\theta)}} \begin{pmatrix} -\theta+\omega(\theta) & -\theta-\omega(\theta) \\ 1 & 1 \end{pmatrix} + \mathcal{O}(\mu)\,, \tag{5.45}$$

and

$$\alpha(\theta,\theta_0) = \int_{\theta_0}^{\theta} a(\psi)\,\mathrm{d}\psi\,, \qquad \phi(\theta,\theta_0) = \int_{\theta_0}^{\theta} 2\omega(\psi)\,\mathrm{d}\psi + \mathcal{O}(\mu)\,. \tag{5.46}$$

The off-diagonal term in the principal solution of $\mu\dot{\zeta} = \mathcal{A}(\theta)\zeta$ has the form

$$\begin{aligned} U_{\xi\eta}(\theta,\theta_0) &= \frac{1}{\mu}\int_{\theta_0}^{\theta} V(\theta,\psi)c_1(\psi)\,\mathrm{d}\psi \\ &= \frac{1}{\mu}\int_{\theta_0}^{\theta} \mathrm{e}^{\alpha(\theta,\psi)/\mu}\left[\cos\left(\frac{\phi(\theta,\psi)}{\mu}\right)v_1 + \sin\left(\frac{\phi(\theta,\psi)}{\mu}\right)v_2\right]\mathrm{d}\psi \end{aligned} \tag{5.47}$$

for some vectors v_1, v_2. Using integration by parts and the fact that the eigenvalues $a \pm \mathrm{i}\,\omega$ are bounded away from 0, one shows that the elements of $U_{\xi\eta}$ are of order 1 at most. The next proposition then follows in the same way as before.

Proposition 5.7. *Define stopping times*

$$\begin{aligned} \tau_\xi &= \inf\{\theta > \theta_0 \colon \|\xi_\theta\| > h\}\,, \\ \tau_\eta &= \inf\{\theta > \theta_0 \colon |\eta_\theta| > h_1\}\,. \end{aligned} \tag{5.48}$$

There exist constants $\kappa, h_0 > 0$ such that for all $\theta_0 \leqslant \theta \leqslant \sqrt{\mu}$, and all $0 < h, h_1 \leqslant \sqrt{\mu}$,

$$\begin{aligned} \mathbb{P}\big\{\tau_\xi \wedge \tau_\eta > \theta\big\} \leqslant \left\lceil \frac{\theta-\theta_0}{\mu} \right\rceil \Bigg(&\exp\left\{ -\frac{\kappa[h - \mu^{-1/2}(h^2+h_1^2)]^2}{(\bar\sigma^2 + (\bar\sigma')^2)\mu^{-1/2}} \right\} \\ &+ \exp\left\{ -\frac{\kappa h^2}{(\bar\sigma^2 + (\bar\sigma')^2)\epsilon\mu^{-1}(\theta-\theta_0)} \right\} + \exp\left\{ -\frac{\kappa h_1^2}{(\bar\sigma')^2\epsilon\mu^{-1}(\theta-\theta_0)} \right\} \Bigg)\,. \end{aligned} \tag{5.49}$$

These estimates show that if the deterministic solution hits Σ_1'' at a point such that $\bar z \leqslant \sqrt{\mu}$, and provided $\epsilon(\theta-\theta_0) \leqslant \sqrt{\mu}$ and $\bar\sigma + \bar\sigma' \ll \mu^{3/4}$, the typical spreadings are

$$\Delta\bar y \asymp (\bar\sigma + \bar\sigma')\left(\frac{1}{\mu^{1/4}} + \frac{\epsilon^{1/2}(\theta-\theta_0)^{1/2}}{\mu^{1/2}}\right) \qquad \text{and} \qquad \Delta\bar z \asymp \frac{\bar\sigma'\epsilon^{1/2}(\theta-\theta_0)^{1/2}}{\mu^{1/2}}\,. \tag{5.50}$$

In particular, if $\theta - \theta_0 \asymp \mu^{1/2}$, going back to original variables we find that provided $\sigma + \sigma' \ll (\epsilon\mu)^{3/4}$, the typical spreadings on Σ_1'' are of order

- $(\sigma+\sigma')(\epsilon/\mu)^{1/4}$ in the y-direction,
- and $\sigma'(\epsilon/\mu)^{1/4}$ in the z-direction.

Remark 5.8. Theorem 6.2 in [17] provides a more precise description of the dynamics, in a slightly simpler setting (in particular without noise on the z-variable): it shows that sample paths concentrate in a "covariance tube" centred in the deterministic solution. The size of the tube is compatible with the above estimates on noise-induced spreading. Such a refined analysis is possible in the present setting as well, but it would require some more work, mainly in order to control the effect of the position-dependence of the noise term.

Remark 5.9. It is possible to extend Estimate (5.49) to slightly larger θ, at the cost of replacing $\mu^{-1/2}$ in the denominator by $\mu^{-1/2}\,\mathrm{e}^{2c\theta^2/\mu}$ for some $c > 0$. This is due to the exponential growth of the variance for $\theta > \sqrt{\mu}$.

5.5 Escape

In this subsection, we fix an initial condition $(-\delta\sqrt{\epsilon}, y_0, z_0) \in \Sigma_1''$, sufficiently close to the folded-node point p^*, and estimate the fluctuations of sample paths up to their first hitting of $\Sigma_2 = \{x = -\delta_0\}$.

Proposition 5.10. *Denote by (y^*, z^*) the point where the deterministic solution starting in $(x_0 = -\delta\sqrt{\epsilon}, y_0, z_0) \in \Sigma_1''$ first hits $\Sigma_2 = \{x = -\delta_0\}$. Assume $y_0 \leqslant x_0^2 - \epsilon(\frac{1+\mu}{2} + c_0)$ for a constant $c_0 > 0$, and $0 \leqslant z_0 \leqslant \mathcal{O}(\sqrt{\epsilon})$. For sufficiently small $\delta, \delta_0 > 0$, there exist $C, \kappa, h_0 > 0$ such that for all $h_1, h_2 > 0$ satisfying $h_1 \leqslant h_0\epsilon$ and $h_2 \leqslant h_0\sqrt{h_1}$, the stochastic sample path starting in (x_0, y_0, z_0) first hits Σ_2 at time $\tau = \tau_{\Sigma_2}$ in a point $(-\delta_0, y_\tau, z_\tau)$ such that*

$$\begin{aligned}
&\mathbb{P}^{(x_0,y_0,z_0)}\left\{|y_\tau - y^*| > h_1 \text{ or } |z_\tau - z^*| > h_2\right\} \\
&\quad\leqslant C\,|\log\epsilon|\left(\exp\left\{-\frac{\kappa h_1^2}{(\sigma^2 + (\sigma')^2)\sqrt{\epsilon}}\right\} + \exp\left\{-\frac{\kappa h_2^2}{(\sigma')^2\sqrt{\epsilon}}\right\} + \exp\left\{-\frac{\kappa\epsilon^{3/2}}{\sigma^2 + (\sigma')^2\epsilon}\right\}\right).
\end{aligned} \tag{5.51}$$

The result remains true uniformly in initial conditions (x_0, y, z) such that $|y - y_0| \leqslant h_1$ and $|z - z_0| \leqslant h_2$.

Proof. The proof is basically the same as the proof of Proposition 4.9, so we will omit its details. We introduce sections $\Sigma_n^* = \{x = -\delta\sqrt{\epsilon}\,2^n, (y, z) \in D_n\}$, where each D_{n+1} is obtained by enlarging the image of D_n under the deterministic flow by order $h_1 2^{-n/2}$ in the y-direction, and by order $h_2 2^{-n/2}$ in the z-direction. Choosing D_1 as a rectangle of size $2h_1 \times 2h_2$ allows to deal with more general initial conditions. Let $\tau^{(n)}$ denote the first-exit time from a block of dimensions $2h \times 2h_1 \times 2h_2$ centred in a given deterministic solution. Then the analogue of Lemma 4.7 reads

$$\mathbb{P}^{(x_n,y_n,z_n)}\left\{\tau_{n+1} \wedge \tau^{(n)} > c\delta 2^{-n}\mu\right\} \leqslant \exp\left\{-\frac{\kappa\delta^2 2^{3n}\epsilon^{3/2}}{\sigma^2 + (\sigma')^2\epsilon}\right\}, \tag{5.52}$$

while the equivalent of Lemma 4.8 is

$$\begin{aligned}
&\mathbb{P}^{(x_n,y_n,z_n)}\left\{\tau^{(n)} < c\delta 2^{-n}\mu\right\} \\
&\quad\leqslant \exp\left\{-\frac{\kappa h^2\sqrt{\epsilon}}{\delta[\sigma^2 + (\sigma')^2 2^{-2n}]}\right\} + \exp\left\{-\frac{\kappa h_1^2}{\delta\sqrt{\epsilon}[\sigma^2\delta^2 2^{-2n} + (\sigma')^2]}\right\} + \exp\left\{-\frac{\kappa h_2^2}{\delta(\sigma')^2\sqrt{\epsilon}}\right\}.
\end{aligned} \tag{5.53}$$

The remainder of the proof is similar to the proof of Proposition 4.9, after redefining κ. □

This result shows that if $\sigma+\sigma' \ll \epsilon^{3/4}$, for a given initial condition on Σ_1'', the spreading in the y- and z-directions on Σ_2 is of order

$$(\sigma+\sigma')\epsilon^{1/4} \qquad \text{and} \qquad \sigma'\epsilon^{1/4}\,. \tag{5.54}$$

6 From noisy returns to Markov chains

In this section we combine the results from the last three sections to obtain estimates on the kernel K of the random Poincaré map on Σ_1. Table 1 summarizes the results obtained so far. For each part of the dynamics, it shows the typical size of fluctuations when starting in a point on the previous section. Deviations will not necessarily add up because of the contraction during some phases of the motion.

Transition	Analysis in	Δx	Δy	Δz
$\Sigma_2 \to \Sigma_3$	Section 3.2	$\sigma+\sigma'$		$\sigma\sqrt{\epsilon}+\sigma'$
$\Sigma_3 \to \Sigma_4$	Section 3.3	$\sigma+\sigma'$		$\sigma\sqrt{\epsilon}+\sigma'$
$\Sigma_4 \to \Sigma_4'$	Section 4.1	$\dfrac{\sigma}{\epsilon^{1/6}}+\dfrac{\sigma'}{\epsilon^{1/3}}$		$\sigma\sqrt{\epsilon\lvert\log\epsilon\rvert}+\sigma'$
$\Sigma_4' \to \Sigma_5$	Section 4.3		$\sigma\sqrt{\epsilon}+\sigma'\epsilon^{1/6}$	$\sigma\sqrt{\epsilon}+\sigma'\epsilon^{1/6}$
$\Sigma_5 \to \Sigma_6$	Section 3.2	$\sigma+\sigma'$		$\sigma\sqrt{\epsilon}+\sigma'$
$\Sigma_6 \to \Sigma_1$	Section 3.3	$\sigma+\sigma'$		$\sigma\sqrt{\epsilon}+\sigma'$
$\Sigma_1 \to \Sigma_1'$	Section 5.2		$(\sigma+\sigma')\epsilon^{1/4}$	σ'
$\Sigma_1' \to \Sigma_1''$ if $z=\mathcal{O}(\sqrt{\mu\epsilon})$	Section 5.4		$(\sigma+\sigma')(\epsilon/\mu)^{1/4}$	$\sigma'(\epsilon/\mu)^{1/4}$
$\Sigma_1'' \to \Sigma_2$	Section 5.5		$(\sigma+\sigma')\epsilon^{1/4}$	$\sigma'\epsilon^{1/4}$

TABLE 1. Summary of results on the size of fluctuations at the time of first hitting a section Σ_j, when starting from a specific point on Σ_i, under the assumption $\sigma+\sigma' \ll \epsilon^{3/4}$, cf. Figure 3.

6.1 The global return map

The following result describes the global return map $\Sigma_2 \to \Sigma_1$.

Theorem 6.1 (Global return map)**.** *Fix $P_2 = (x_2^*, y_2^*, z_2^*) \in \Sigma_2$. Assume the deterministic orbit starting in P_2 hits Σ_1 for the first time in $P_1 = (x_1^*, y_1^*, z_1^*)$. Then there exist constants $h_0, \kappa, C > 0$ such that for all $h \leqslant h_0$ and $h^2/h_0 \leqslant h_1 \leqslant h$, the stochastic sample path starting in P_2 hits Σ_1 for the first time in a point (x_1, y_1^*, z_1) satisfying*

$$\begin{aligned}&\mathbb{P}^{P_2}\bigl\{|x_1-x_1^*|>h \text{ or } |z_1-z_1^*|>h_1\bigr\}\\ &\leqslant \frac{C}{\epsilon}\biggl(\exp\biggl\{-\frac{\kappa h^2}{\sigma^2+(\sigma')^2}\biggr\}+\exp\biggl\{-\frac{\kappa h_1^2}{\sigma^2\epsilon|\log\epsilon|+(\sigma')^2}\biggr\}+\exp\biggl\{-\frac{\kappa\epsilon}{\sigma^2+(\sigma')^2\epsilon^{-1/3}}\biggr\}\biggr)\,.\end{aligned} \tag{6.1}$$

Proof. Denote by (x_i^*, y_i^*, z_i^*) the deterministic first-hitting point of section Σ_i, and by (x_i, y_i, z_i) the corresponding random first-hitting point. We use similar notations for Σ_4'. We will decompose the dynamics into three main steps, and introduce the events

$$\begin{aligned} \Omega_1(h, h_1) &= \left\{ |x_4' - (x_4')^*| \leqslant h, |z_4' - (z_4')^*| \leqslant h_1 \right\}, \\ \Omega_2(H_1) &= \left\{ \|(y_5, z_5) - (y_5^*, z_5^*)\| \leqslant H_1 \right\}. \end{aligned} \tag{6.2}$$

- *Step 1:* $\Sigma_2 \to \Sigma_3 \to \Sigma_4 \to \Sigma_4'$. Propositions 3.3 and 4.3 can be applied simultaneously, because they are based on the same kind of estimates of the principal solution. This directly yields the bound

$$\mathbb{P}^{P_2}\big(\Omega_1(h, h_1)^c\big) \leqslant \frac{C}{\epsilon}\left(\exp\left\{ -\frac{\kappa h^2}{\sigma^2 \epsilon^{-1/3} + (\sigma')^2 \epsilon^{-2/3}} \right\} + \exp\left\{ -\frac{\kappa h_1^2}{\sigma^2 \epsilon |\log \epsilon| + (\sigma')^2} \right\} \right) \tag{6.3}$$

  for some $C > 0$, which is valid for all h, h_1 satisfying $h \leqslant h_0 \epsilon^{1/3}$, $h_1 \leqslant h_0$, $h^2 \leqslant h_0 h_1$ and $h_1^2 \leqslant h_0 h \epsilon^{1/3}$.

- *Step 2:* $\Sigma_4' \to \Sigma_5$. The difference ζ_θ between two deterministic solutions starting on Σ_4' satisfies a relation of the form

$$\zeta_\theta = U(\theta, \theta_0)\zeta_{\theta_0} + \int_{\theta_0}^{\theta} U(\theta, \phi) b(\zeta_\phi, \phi) \, \mathrm{d}\phi \,, \tag{6.4}$$

  where $b(\zeta, \phi)$ is a nonlinear term. Using the estimates on the principal solution (cf. (4.50) in the proof of Lemma 4.8), one obtains that the deterministic orbit starting in $(\hat{x}_4, (y_4')^*, \hat{z}_4) \in \Sigma_4'$ hits Σ_5 at a point $(x_5^*, \hat{y}_5, \hat{z}_5)$ satisfying

$$\|(\hat{y}_5, \hat{z}_5) - (y_5^*, z_5^*)\| \leqslant M \epsilon^{2/3} |\hat{x}_4 - (x_4')^*| + M |\hat{z}_4 - (z_4')^*| \tag{6.5}$$

  for some constant $M > 0$. Proposition 4.9 yields that for $H_1 \leqslant h_0 \epsilon^{2/3}$,

$$\begin{aligned} &\mathbb{P}^{P_2}\Big(\Omega_1(h, h_1) \cap \Omega_2(H_1)^c\Big) \\ &\qquad \leqslant C|\log \epsilon| \left(\exp\left\{ -\frac{\kappa (H_1 - M[\epsilon^{2/3} h + h_1])^2}{\sigma^2 \epsilon + (\sigma')^2 \epsilon^{1/3}} \right\} + \exp\left\{ -\frac{\kappa \epsilon}{\sigma^2 + (\sigma')^2 \epsilon} \right\} \right). \end{aligned} \tag{6.6}$$

  We now choose $h_1 = H_1/(3M)$ and $h = \epsilon^{1/3} \wedge H_1/(3M\epsilon^{2/3})$. Distinguishing the cases $h = \epsilon^{1/3}$ and $h = H_1/(3M\epsilon^{2/3})$ when using (6.3), this yields

$$\mathbb{P}^{P_2}\big(\Omega_2(H_1)^c\big) \leqslant \frac{C}{\epsilon}\left(\exp\left\{ -\frac{\kappa H_1^2}{\sigma^2 \epsilon |\log \epsilon| + (\sigma')^2} \right\} + \exp\left\{ -\frac{\kappa \epsilon}{\sigma^2 + (\sigma')^2 \epsilon^{-1/3}} \right\} \right), \tag{6.7}$$

  where κ has been redefined.

- *Step 3:* $\Sigma_5 \to \Sigma_6 \to \Sigma_1$. A similar argument as above shows that two deterministic solutions starting at distance H_1 on Σ_5 hit Σ_1 at a distance of order H_1. The result then follows from Proposition 3.3 and (6.7), choosing $H_1 = \epsilon^{2/3} \wedge c h_1$ for a sufficiently small constant c. □

This result is useful if $\sigma \ll \sqrt{\epsilon}$ and $\sigma' \ll \epsilon^{2/3}$. It shows that stochastic sample paths are likely to hit Σ_1 at a distance of order $\sigma + \sigma'$ in the fast directions from the deterministic solution, and at a distance of order $\sigma\sqrt{\epsilon|\log \epsilon|} + \sigma'$ in the slow direction.

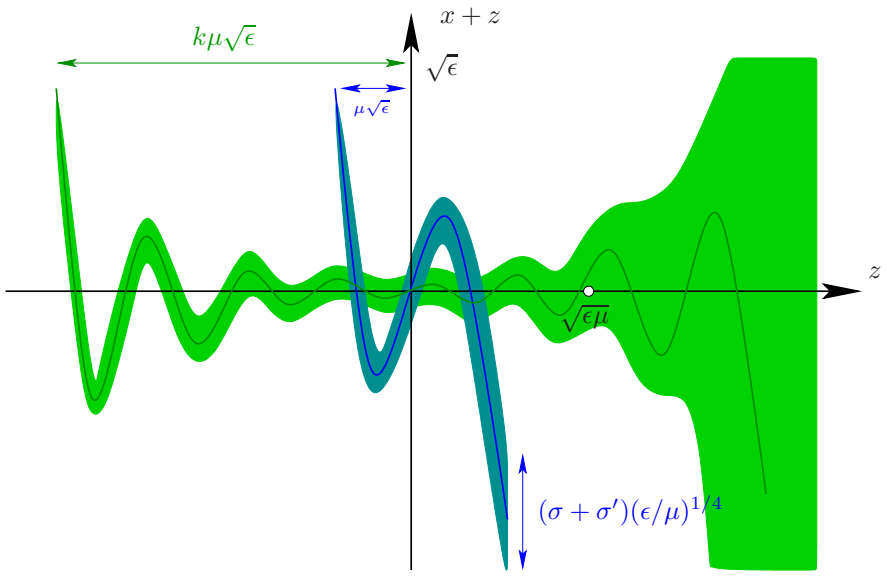

FIGURE 8. Inner (blue) and outer (green) canards. The shaded sets indicate the extension of typical fluctuations.

6.2 The local map

We know from the deterministic analysis (cf. (5.16)) that the section Σ_1 can be subdivided into $k_\mu \simeq 1/(2\mu)$ sectors of rotation. An orbit starting in the kth sector makes $2k+1$ half-turns before hitting Σ_2. The width of the kth sector has order $\epsilon^{(1-\mu)/2}$, cf. [24]. The analysis of Section 5.3 shows that the images of these sectors on Σ_1' have a size of order $\mu\epsilon$.

For the stochastic system, it will be relevant to distinguish between

- **inner sectors**, which are sectors with $k \leqslant \mathcal{O}(1/\sqrt{\mu})$; orbits starting in these sectors hit Σ_1' for $z \leqslant \mathcal{O}(\sqrt{\epsilon\mu})$;
- **outer sectors**, which are sectors with $k > \mathcal{O}(1/\sqrt{\mu})$.

Theorem 6.2 (Local return map for inner sectors)**.** *Fix $P_0 = (x_0^*, y_0^*, z_0^*) \in \Sigma_1$, and suppose that P_0 lies in an inner sector. Assume the deterministic orbit starting in P_0 hits Σ_2 for the first time in $P_2 = (x_2^*, y_2^*, z_2^*)$. Further assume $\epsilon \leqslant \mu$. Then there exist constants $h_0, \kappa, C > 0$ such that for all $h \leqslant h_0\epsilon\sqrt{\mu}$ and $h_2 \leqslant h_0(\sqrt{h_1} \wedge \sqrt{\epsilon\mu})$, the stochastic sample path starting in P_0 hits Σ_2 for the first time in a point (x_2^*, y_2, z_2) satisfying*

$$\begin{aligned}\mathbb{P}^{P_0}&\big\{|y_2 - y_2^*| > h_1 \text{ or } |z_2 - z_2^*| > h_2\big\} \\ &\leqslant \frac{C}{\epsilon}\left(\exp\left\{-\frac{\kappa h_1^2\sqrt{\mu}}{(\sigma^2+(\sigma')^2)\sqrt{\epsilon}}\right\} + \exp\left\{-\frac{\kappa h_2^2}{(\sigma')^2}\right\} + \exp\left\{-\frac{\kappa\epsilon^{3/2}}{\sigma^2+(\sigma')^2}\right\}\right). \quad (6.8)\end{aligned}$$

Proof. We introduce the events

$$\begin{aligned}\Omega_1(h_1,h_2) &= \big\{|y_1' - (y_1')^*| \leqslant h_1, |z_1' - (z_1')^*| \leqslant h_2\big\}, \\ \Omega_2(H_1,H_2) &= \big\{|y_1'' - (y_1'')^*| \leqslant H_1, |z_1'' - (z_1'')^*| \leqslant H_2\big\}, \end{aligned} \quad (6.9)$$

where (y_1', z_1') and (y_1'', z_1'') denote the first-hitting points of the stochastic path with Σ_1' and Σ_1'', and the starred quantities are the corresponding deterministic hitting points.

Proposition 5.3 with h of order $\sqrt{\epsilon}$ implies

$$\mathbb{P}^{P_0}(\Omega_1(h_1,h_2)^c) \leqslant \frac{C}{\epsilon}\left(\exp\left\{-\frac{\kappa h_1^2}{(\sigma^2+(\sigma')^2)\sqrt{\epsilon}}\right\}+\exp\left\{-\frac{\kappa h_2^2}{(\sigma')^2}\right\}+\exp\left\{-\frac{\kappa\epsilon^{3/2}}{\sigma^2+(\sigma')^2}\right\}\right) \tag{6.10}$$

for some $C>0$. Using this bound with h_1 of order H_1 and h_2 of order $H_1 \wedge H_2$, together with Proposition 5.7 to estimate the probability of $\Omega_1 \cap \Omega_2^\varepsilon$ and the assumption $\epsilon \leqslant \mu$ yield

$$\mathbb{P}^{P_0}(\Omega_2(H_1,H_2)^c) \leqslant \frac{C}{\epsilon}\left(\exp\left\{-\frac{\kappa H_1^2\sqrt{\mu}}{(\sigma^2+(\sigma')^2)\sqrt{\epsilon}}\right\}+\exp\left\{-\frac{\kappa H_2^2}{(\sigma')^2}\right\}+\exp\left\{-\frac{\kappa\epsilon^{3/2}}{\sigma^2+(\sigma')^2}\right\}\right). \tag{6.11}$$

The result then follows from Proposition 5.10, taking $h_1 = H_1$ and $h_2 = H_2$. □

This result is useful if $\sigma, \sigma' \ll (\epsilon\mu)^{3/4}$. It shows that stochastic sample paths are likely to hit Σ_2 at a distance of order $(\sigma+\sigma')(\epsilon/\mu)^{1/4}$ in the y-direction from the deterministic solution, and at a distance of order σ' in the z-direction. Combining this with Theorem 6.1 on the global return map, we conclude that for initial conditions $P_0 \in \Sigma_1$, starting in a inner sector, stochastic sample paths will return to Σ_1 in a neighbourhood of the deterministic solution, of width

- $\sigma+\sigma'$ in the fast x-direction,
- $\sigma\sqrt{\epsilon|\log\epsilon|}+\sigma'$ in the z-direction.

Remark 6.3. The limitation $k \leqslant \mathcal{O}(1/\sqrt{\mu})$ is due to the fact that Proposition 5.7 is formulated for $\theta \leqslant \sqrt{\mu}$. Using Remark 5.9, Theorem 6.2 can be extended to sectors $k = \sqrt{a|\log(\sigma+\sigma')|/\mu}$, which results in fluctuations of y of order $(\sigma+\sigma')^{1-ca}(\epsilon/\mu)^{1/4}$. This does not affect the order of fluctuations in the z-direction as long as a is small enough.

Finally, we consider what happens to sample paths starting on Σ_1 in an outer sector.

Theorem 6.4 (Local return map for outer sectors)**.** *Fix $P_0 = (x_0^*, y_0^*, z_0^*) \in \Sigma_1$, and assume P_0 lies in an outer sector k, with $k \geqslant k_0/\sqrt{\mu}$ for some $k_0 > 0$. If k_0 is large enough, there exist constants $\kappa, C_0, C, \gamma > 0$ such that the stochastic sample path starting in P_0 hits Σ_2 for the first time in a point (x_2^*, y_2, z_2) satisfying*

$$\mathbb{P}^{P_0}\big\{z_2 \leqslant \sqrt{\epsilon\mu}\big\} \leqslant \frac{C}{\epsilon}\exp\left\{-\frac{\kappa\mu^{1/2}\epsilon^{3/2}}{\sigma^2+(\sigma')^2}\right\}, \tag{6.12}$$

$$\mathbb{P}^{P_0}\big\{z_2 \geqslant z\big\} \leqslant C_0|\log\sigma|^\gamma \exp\left\{-\frac{\kappa(z^2-\epsilon\mu)}{\epsilon\mu\log|\sigma+\sigma'|}\right\} \qquad \forall z \geqslant \sqrt{\epsilon\mu}\,. \tag{6.13}$$

Furthermore, there exist a constant $h_0 > 0$ and an interval I of size of order ϵ, independent of $k \geqslant k_0/\sqrt{\mu}$, such that

$$\mathbb{P}^{P_0}\big\{\mathrm{dist}(y_2, I) > h_1\big\} \leqslant \frac{C}{\epsilon}\left(\exp\left\{-\frac{\kappa h_1^2}{(\sigma^2+(\sigma')^2)\sqrt{\epsilon}}\right\}+\exp\left\{-\frac{\kappa\epsilon^{3/2}}{\sigma^2+(\sigma')^2\epsilon}\right\}\right) \tag{6.14}$$

holds for all $h_1 \leqslant h_0\epsilon$.

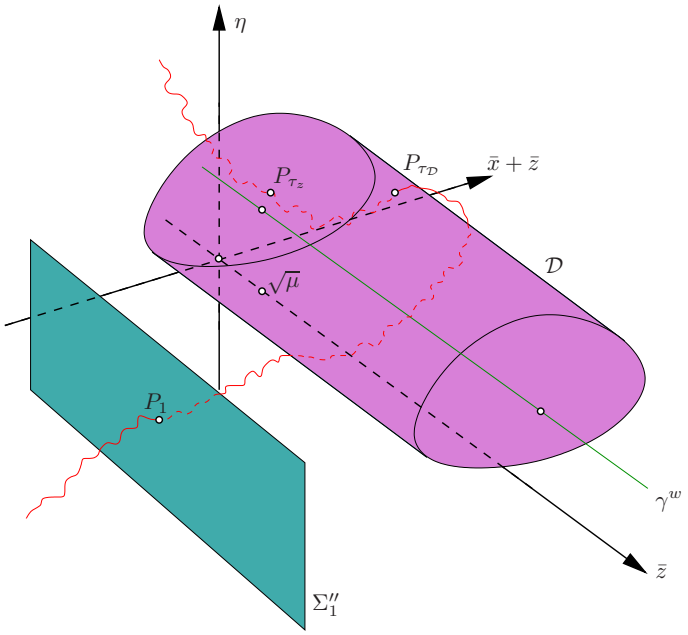

FIGURE 9. Escape of sample paths starting in an outer sector from a neighbourhood of the weak canard γ^w.

Proof. We shall work in the zoomed-in coordinates $(\bar{x}, \bar{y}, \bar{z})$, cf. (5.17). Fix a $K_0 \in (0,1)$ and introduce a neighbourhood $\mathcal{D}$ of the weak canard γ^w given by

$$\mathcal{D} = \big\{(\bar{x}, \bar{y}, \bar{z})\colon K \geqslant K_0, \bar{z} \geqslant \sqrt{\mu}\big\} = \mathcal{D}_0 \times [\sqrt{\mu}, \infty) \,, \tag{6.15}$$

where K is the first integral introduced in (5.27) (recall that u_1 and u_2 measure the deviation of $(\bar{x}, \bar{y})$ from the weak canard) – see Figure 9. The proof proceeds in four main steps:

- *Step 1: Entering $\mathcal{D}$.* From Theorem 4.4 in [17] describing the spacing of canards, we know that the deterministic solution starting in P_0 hits $\mathcal{D}$ in a point P_0^* at a distance of order $\mathrm{e}^{-c(2k+1)^2\mu} < \mathrm{e}^{-4ck_0^2}$ from the weak canard. Taking k_0 sufficiently large, we may assume that P_0^* is bounded away from the boundary $\partial\mathcal{D}_0$. Combining, as in the previous theorem, Proposition 5.3 and Proposition 5.7, we obtain that the stochastic sample path first hits $\mathcal{D}$ at a point P_{τ_z} such that

  $$\mathbb{P}^{P_0}\big\{(\bar{x}_{\tau_z}, \bar{y}_{\tau_z}) \notin \mathcal{D}_0\big\} \leqslant \frac{C}{\epsilon} \exp\left\{-\frac{\kappa\mu^{1/2}\epsilon^{3/2}}{\sigma^2 + (\sigma')^2}\right\} . \tag{6.16}$$

  This proves in particular (6.12).

- *Step 2: Leaving $\mathcal{D}$.* Theorem 6.4 in [17] estimates the probability of sample paths not leaving a neighbourhood of the weak canard. Because we worked in polar coordinates, the result only applied to a small neighbourhood of size proportional to $\sqrt{\bar{z}}$. However by using the coordinate K instead of the distance to γ^w, the same proof applies to the exit from $\mathcal{D}$. It suffices to realize that the nonlinear drift term β in Equation (D.33) of [17] is replaced by a term of order $\bar{z}$ as a consequence of

Proposition 5.5. We thus conclude that the sample path leaves $\mathcal{D}$ at a point $P_{\tau_\mathcal{D}}$ whose $\bar{z}$-coordinate satisfies

$$\mathbb{P}^{P_0}\big\{\bar{z}_{\tau_\mathcal{D}} \geqslant \bar{z}\big\} \leqslant C_0|\log\bar{\sigma}|^\gamma \exp\biggl\{-\frac{\kappa(\bar{z}^2-\mu)}{\mu\log|\bar{\sigma}+\bar{\sigma}'|}\biggr\} \qquad \forall \bar{z} \geqslant \sqrt{\mu} \tag{6.17}$$

for some constants $C_0, \gamma > 0$.

- *Step 3: Transition from $\mathcal{D}$ to Σ_1''.* Since $P_{\tau_\mathcal{D}}$ is at distance of order 1 from the weak canard, we know that the deterministic solutions starting in $P_{\tau_\mathcal{D}}$ will take a time of order μ to reach Σ_1'', in a point that we will denote $P_1^* = (\bar{x}_1^*, \bar{y}_1^*, \bar{z}_1^*)$. Let $P_1 = (\bar{x}_1^*, y_1^*, z_1^*)$ denote the point where the stochastic sample path first hits Σ_1''. Starting from System (5.40) and applying the usual procedure, we obtain the estimate

$$\begin{aligned}\mathbb{P}^{P_{\tau_\mathcal{D}}}\big\{|\bar{y}_1 - \bar{y}_1^*| > \bar{h}_1 \text{ or } |\bar{z}_1 - \bar{z}_1^*| > \bar{h}_2\big\} \\ \leqslant C\biggl(\exp\biggl\{-\frac{\kappa[\bar{h}_1 - M(\bar{h}_1^2+\bar{h}_2^2)]^2}{\bar{\sigma}^2+(\bar{\sigma}')^2}\biggr\} + \exp\biggl\{-\frac{\kappa\bar{h}_2^2}{(\bar{\sigma}')^2\epsilon}\biggr\}\biggr)\,.\end{aligned} \tag{6.18}$$

  Note that this implies fluctuations of size $\bar{\sigma}+\bar{\sigma}'$ in the $\bar{y}$-direction, and of size $\bar{\sigma}'\sqrt{\epsilon}$ in the $\bar{z}$-direction. Going back to original variables, this entails fluctuations of size $(\sigma+\sigma')\epsilon^{1/4}$ in the y-direction, and $\sigma'\epsilon^{1/4}$ in the z-direction.

  We also have to take into account the fact that we do not know the $(\bar{x},\bar{y})$-coordinates of $P_{\tau_\mathcal{D}}$. In fact all exit points on $\partial\mathcal{D}_0$ might have a comparable probability. Hence the coordinate $\bar{y}_1^*$ can vary in an interval I_1, which is the image of $\partial\mathcal{D}_0$ under the deterministic flow. It follows from (5.25) that I_1 has a size of order $\bar{z}_1^*$ in $\bar{y}$-coordinates.

- *Step 4: Transition from Σ_1'' to Σ_2.* If P_1 satisfied $y_1 \leqslant x_1^2 - \epsilon(\frac{1+\mu}{2}+c_0)$ for some $c_0 > 0$, or equivalently $\eta_1 \leqslant -c_0$, we could directly apply Proposition 5.10 to estimate the fluctuations during the last transition step, which would remain of the same order as in Step 3.

  The estimate (5.25) shows that P_1 is too close to the repelling slow manifold C_ϵ^r to apply Proposition 5.10 directly. However, using (5.40), we obtain that the variable η measuring the distance to C_ϵ^r satisfies an equation of the form

$$\mathrm{d}\eta_\theta = \frac{2}{\mu}\Big[-2\bar{x}_\theta\eta_\theta - \bar{z}_\theta + \mathcal{O}\big(\sqrt{\epsilon}(1+|\bar{x}_\theta|)\big)\Big]\,\mathrm{d}\theta + \sqrt{\frac{2}{\mu}}\big[\bar{\sigma}\widetilde{G}_1 + \bar{\sigma}'\widetilde{G}_2\big]\,\mathrm{d}W_\theta\,. \tag{6.19}$$

  We have used that we may assume $\sigma, \sigma' \ll \sqrt{\epsilon}$ to simplify the error term (for larger noise intensities, the main results of the theorem become meaningless). Using the same approach as in [15, Section 3.2] or [17, Section D], one can show that η_θ is likely to leave $[-c_0, c_0]$ in a time θ of order $\mu\sqrt{|\log(\bar{\sigma}+\bar{\sigma}')|}$. During this time interval, $\bar{x}_\theta$ decreases by an amount of order $\sqrt{|\log(\bar{\sigma}+\bar{\sigma}')|}$. Either this exit takes place in the direction of negative η, and we can apply again Proposition 5.10. Or it takes place in the direction of positive η, and the sample path makes one more excursion towards C_ϵ^{a+} (backward canard). In this case we have to use one more time the analysis of Step 3 before applying again Proposition 5.10. Finally, one can check that the deterministic flow maps the set of points where paths escape $\{-c_0 \leqslant \eta \leqslant c_0\}$ to points in Σ_2 with a y-coordinate in an interval I of size $\mathcal{O}(\epsilon)$. □

The important point of Theorem 6.4 is that the bounds on the distribution of (y_2, z_2) are *independent* of the starting sector number k, as soon as $k > k_0/\sqrt{\mu}$. Thus we observe a saturation effect, in the sense that the stochastic Poincaré map becomes independent of the initial condition – see Figure 1. Combining the local result with Theorem 6.1, we see in particular that the size of fluctuations in the z-direction is at most of order

$$\sqrt{\epsilon\mu|\log(\sigma+\sigma')|} + \sigma\sqrt{\epsilon|\log\epsilon|} + \sigma' \,. \tag{6.20}$$

Disregarding logarithms, we observe that unless $\mu < \sigma^2 \wedge ((\sigma')^2/\epsilon)$, the first term will be the dominating one. We conclude that in this regime, the noise-induced fluctuations in the z-direction are at most of order $\sqrt{\epsilon\mu|\log(\sigma+\sigma')|}$. However this bound is certainly not sharp, since it uses $z = \sqrt{\epsilon\mu}$ as lower bound of typical exits from a neighbourhood of the weak canard, which may underestimate typical exit times if the noise is weak.

7 An example – The Koper model

In order to illustrate some of our results numerically, we consider the example of the Koper model [51]. Its deterministic version [59] is given by

$$\begin{aligned} \epsilon_1\dot{x} &= y - x^3 + 3x \,, \\ \dot{y} &= kx - 2(y+\lambda) + z \,, \\ \dot{z} &= \epsilon_2(\lambda + y - z) \,, \end{aligned} \tag{7.1}$$

with parameters k, λ, ϵ_1, ϵ_2. Note that there is a symmetry

$$(x, y, z, \lambda, k) \mapsto (-x, -y - z, -\lambda, k) \tag{7.2}$$

so that we can restrict the parameter space. We shall assume that $0 < \epsilon_1 =: \epsilon \ll 1$ and $\epsilon_2 = 1$ so that (7.1) has the structure (2.1) and Assumption (A0) obviously holds. For a detailed bifurcation analysis we refer to [26, 59].

Of course, if $0 < \epsilon_2 \ll 1$ one may still simulate the three-scale system numerically, and it is even known via explicit asymptotic analysis which MMO patterns one expects to observe in certain classes of three-scale systems [54]. The first variant of the Koper model was a planar system due to Boissonade and De Kepper [19]. Koper [51] introduced the third variable and studied MMOs via numerical continuation. In fact, the system (7.1) has been suggested independently by various other research groups as a standard model for MMOs [24, 38, 50]. Therefore it certainly provides an excellent test case.

The critical manifold of (7.1) is given by $C_0 = \{(x, y, z) \in \mathbb{R}^3 \colon y = c(x)\}$ with $c(x) := x^3 - 3x$, and the two fold curves are $L^{\pm} = \{(x, y, z) \in \mathbb{R}^3 \colon x = \pm 1, y = \mp 2\}$. This yields a decomposition

$$C_0 = C_0^{a-} \cup L^- \cup C_0^r \cup L^+ \cup C_0^{a+} \,, \tag{7.3}$$

where $C_0^{a-} = C_0 \cap \{x < -1\}$, $C^r = C_0 \cap \{-1 < x < 1\}$ and $C_0^{a+} = C_0 \cap \{1 < x\}$ are normally hyperbolic. It is easy to verify that Assumption (A1) is satisfied.

The desingularized slow subsystem is of the form

$$\begin{aligned} \dot{x} &= kx - 2(c(x) + \lambda) + z \,, \\ \dot{z} &= (3x^2 - 3)(\lambda + c(x) - z) \,. \end{aligned} \tag{7.4}$$

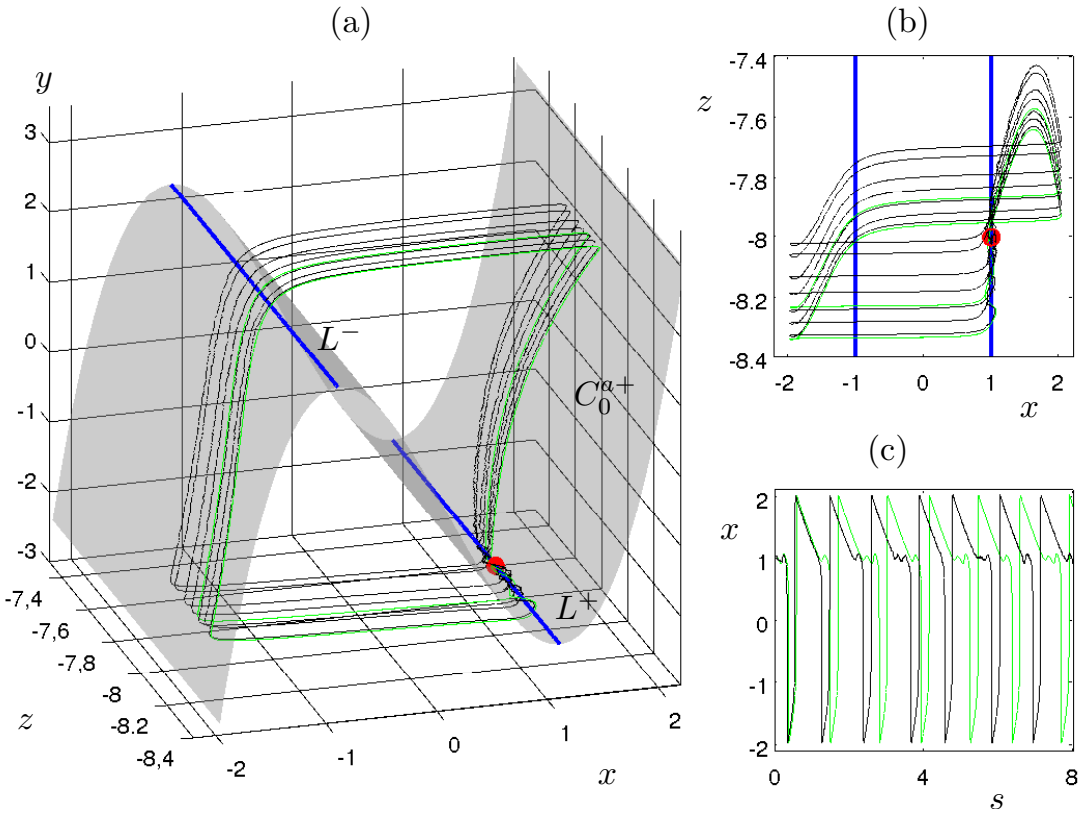

FIGURE 10. Basic structure of the dynamics for the Koper model (7.6). (a) Phase space with deterministic critical manifold C_0 (grey) and the two fold lines $L^{\pm}$ (blue). Two orbits are also shown, one for the deterministic system ($\sigma = 0 = \sigma'$, green) and one for the stochastic system ($\sigma = 0.01 = \sigma'$, black). For both, the parameter values are $\epsilon = 0.01$, $k = -10$, $\lambda = -7$, and M is given by (7.7). (b) Projection of the full system onto the (x, z)-plane. (c) Time series for the two orbits.

Note that in (7.4) the direction of time is reversed on C_0^r. The only folded equilibria are $(x, z) = (1, 2\lambda - 4 - k) \in L^+$ and $(x, z) = (-1, 2\lambda + 4 + k) \in L^-$. From the linearization of the slow subsystem

$$A^{\pm} = \begin{pmatrix} k & 1 \\ 6(2 + k \mp \lambda) & 0 \end{pmatrix} \tag{7.5}$$

at the folded singularities one may determine the parameter values for which we have a folded node on L^+. It turns out that there exist parameter regimes where this is the case, and the only passages of deterministically stable MMO orbits near L^- are via nondegenerate folds [26, 59]. Furthermore, the fast flow is transverse to the relevant drop curves in such a regime [26, 59]. From now on, we shall restrict our attention to this parameter regime so that Assumptions (A2)–(A4) are satisfied in a suitable compact absorbing set in phase space.

Since (7.1) is a phenomenological model, it is not immediate how to derive noise terms

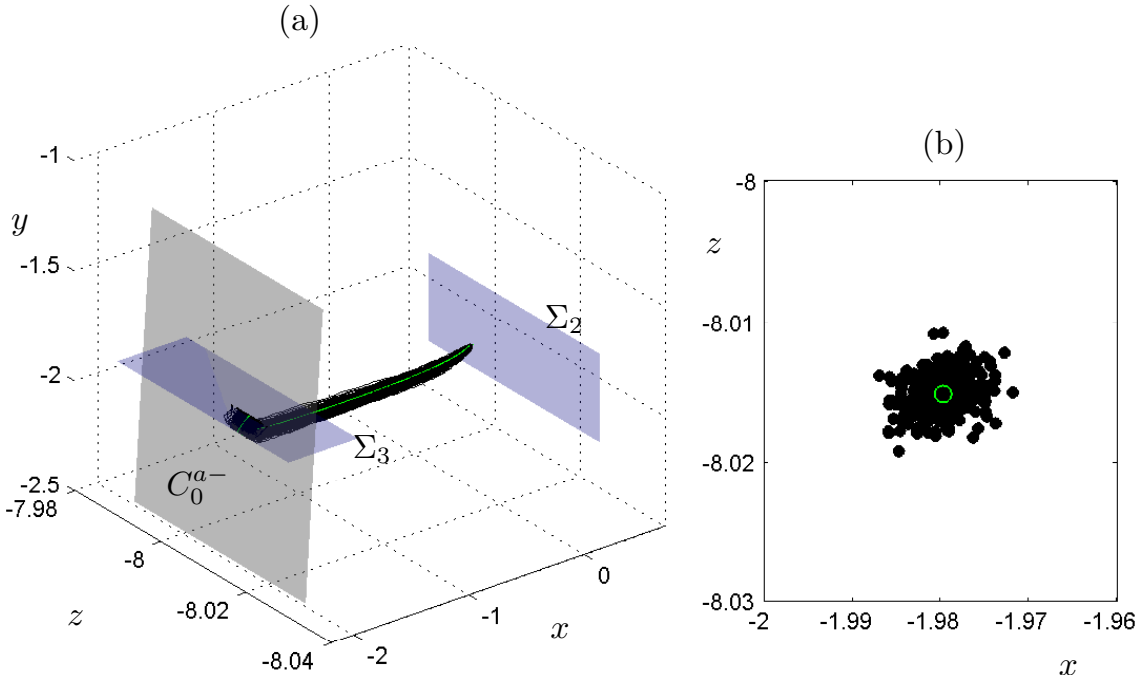

FIGURE 11. Computation of the transition map $\Sigma_2 \to \Sigma_3$ where the sections (blue) have been defined by the conditions $x = 0.5$ and $y = -1.8$ respectively. All paths have been started at $(x, y, z) = (0.5, -2.1, -8)$. There is one deterministic path (green) and 300 different realizations (black) for the stochastic case with M given by (7.7) and $\sigma = 0.1 = \sigma'$. The relevant part of the critical manifold C_0^{a+} (grey) is shown as well. (a) Phase space. (b) View of the landing points on the section Σ_3 where the landing point of the deterministic path is in the center of the green circle.

so we will just choose correlated additive noise as a first benchmark, setting

$$\begin{aligned} \mathrm{d}x_s &= \frac{1}{\epsilon}(y_s - x_s^3 + 3x_s)\,\mathrm{d}s + \frac{\sigma}{\sqrt{\epsilon}} F\,\mathrm{d}W_s\,, \\ \mathrm{d}y_s &= (kx_s - 2(y_s + \lambda) + z_s)\,\mathrm{d}s + \sigma' G_1\,\mathrm{d}W_s\,, \\ \mathrm{d}z_s &= \epsilon_2(\lambda + y_s - z_s)\,\mathrm{d}s + \sigma' G_2\,\mathrm{d}W_s\,, \end{aligned} \tag{7.6}$$

where the Brownian motion $(W_s)_s$ is assumed to be three-dimensional, and $\frac{\sigma}{\sqrt{\epsilon}}F$, $\sigma' G_1$, $\sigma' G_2$ may be viewed as rows of a constant matrix $M \in \mathbb{R}^{3\times 3}$. Figure 10 shows the basic geometry of the Koper model including two orbits computed for $(\epsilon, k, \lambda) = (0.01, -10, -7)$. One of these orbits is deterministic ($\sigma = 0 = \sigma'$) and the other one shows a realization of a stochastic sample path computed for $\sigma = 0.01 = \sigma'$ and

$$M = \begin{pmatrix} F \\ G_1 \\ G_2 \end{pmatrix} = \begin{pmatrix} 1.0 & 0.5 & 0.2 \\ 0.5 & 1.0 & 0.3 \\ 0.2 & 0.3 & 1.0 \end{pmatrix}\,. \tag{7.7}$$

Note that the deterministic orbit exhibits an MMO of type $1^1 1^2$ while the stochastic sample path shows combinations of patterns of the form 1^0, 1^1 and 1^2. Since we proved results about separate phases of the flow we investigate the estimates for each phase as summarized in Table 1.

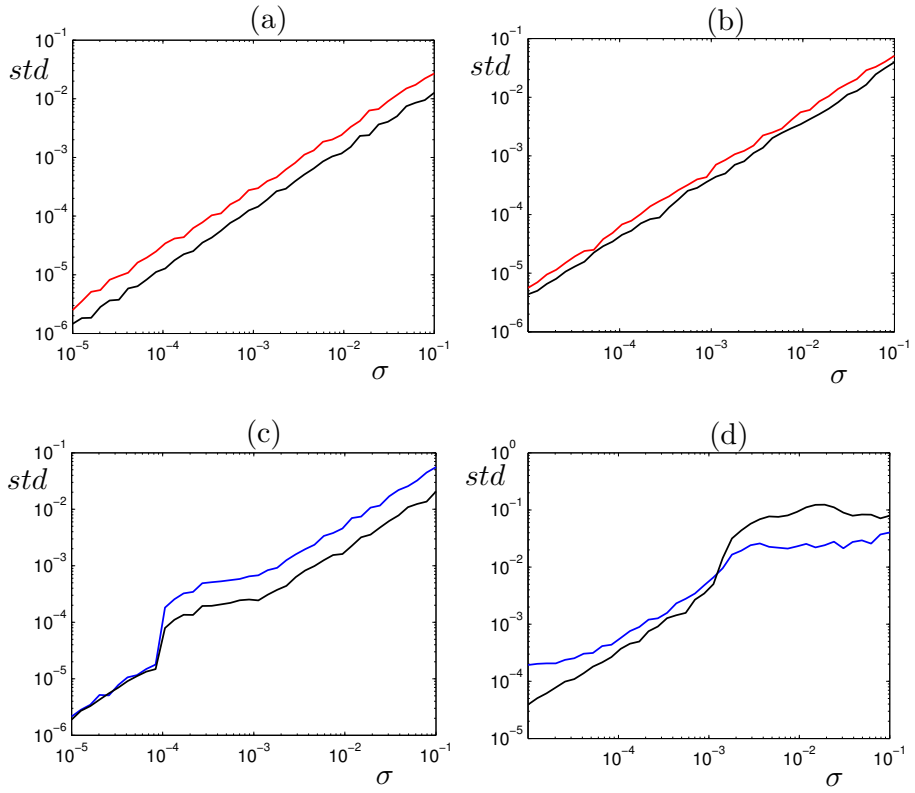

FIGURE 12. Dependence of the standard deviation of the distance between the stochastic and the deterministic transition maps on the noise intensity for an attracting deterministic periodic MMO during various phases of the flow. We have always fixed the matrix M by (7.7), $\epsilon = 0.01$ and $\sigma = \sigma'$. The results are shown on a log–log scale with two directions out of x (red), y (blue) and z (black). The standard deviation *std* has been computed from 100 realizations of sample paths and the domain for the noise level was subdivided into 40 points. (a) Transition map from $\Sigma_2 \to \Sigma_3$ with $\Sigma_2 = \{x = 0.5\}$, $\Sigma_3 = \{y = -1.8\}$ and $(x_0, y_0, z_0) = (0.5, -2.1, -8)$. (b) Transition map from $\Sigma_3 \to \Sigma_4$ with $\Sigma_4 = \{y = 1.8\}$ and $(x_0, y_0, z_0) = (-2, -1.8, -8)$. (c) Transition map from $\Sigma_4 \to \Sigma_5$ with $\Sigma_5 = \{x = -0.5\}$ and $(x_0, y_0, z_0) = (-1.3, 1.8, -7.8)$. (d) Transition map from $\Sigma_1 \to \Sigma_2$ with $\Sigma_1 = \{y = -1.8\}$ and $(x_0, y_0, z_0) = (1.3, -1.8, -7.7)$.

Figure 11 illustrates the map $\Sigma_2 \to \Sigma_3$ which describes the fast flow towards the critical manifold C_0^{a-}. Several stochastic sample paths are compared with the deterministic solution. In Section 3.2 we derived the typical spreading of stochastic sample paths around their deterministic counterpart. It was shown that the typical spreading has an upper bound $\mathcal{O}(\sigma + \sigma')$ in the x-coordinate and $\mathcal{O}(\sigma' + \sigma\sqrt{\epsilon})$ in the z-coordinate. Since the typical spreading can be understood as standard deviation, cf. [15, Prop. 3.1.13], Figure 11(b) confirms that the theoretical results indeed provide upper bounds (note the scaling on the axes and that $\sigma = 0.01 = \sigma'$).

To investigate the scaling results further, we computed sample paths going from Σ_2 to Σ_3 numerically for a much wider range of noise values as shown in Figure 12(a). For the hitting point on Σ_3 we plotted the standard deviation of the hitting point's distance to its deterministic counterpart for the (x, z)-coordinates in a log–log plot for different noise levels with $\sigma = \sigma'$. The slope of 1 for both coordinates in Figure 12(a)–(b) is expected from the upper bounds in Table 1. However, the overall spreading is smaller than expected

since we have started the orbits in the vicinity of an attracting deterministic periodic orbit, cf. Figure 10. Therefore, contraction transverse to the periodic orbit shrinks the stochastic neighbourhood more than the general upper-bound estimates predict. Similarly, we may also study the remaining phases of the flow which are analyzed in Figure 12(b)–(d). We observe not only the correct asymptotic decrease in size of the stochastic neighbourhood as $\sigma \to 0$, but also a larger spreading of sample paths near the folded node, see Figure 12(d). This is related to the mechanism that sample paths may jump only with high probability during certain parts of the SAOs after the folded node; this effect has already been discussed in detail in [17] with associated numerics in [17, Section 7] so we shall not detail it here.

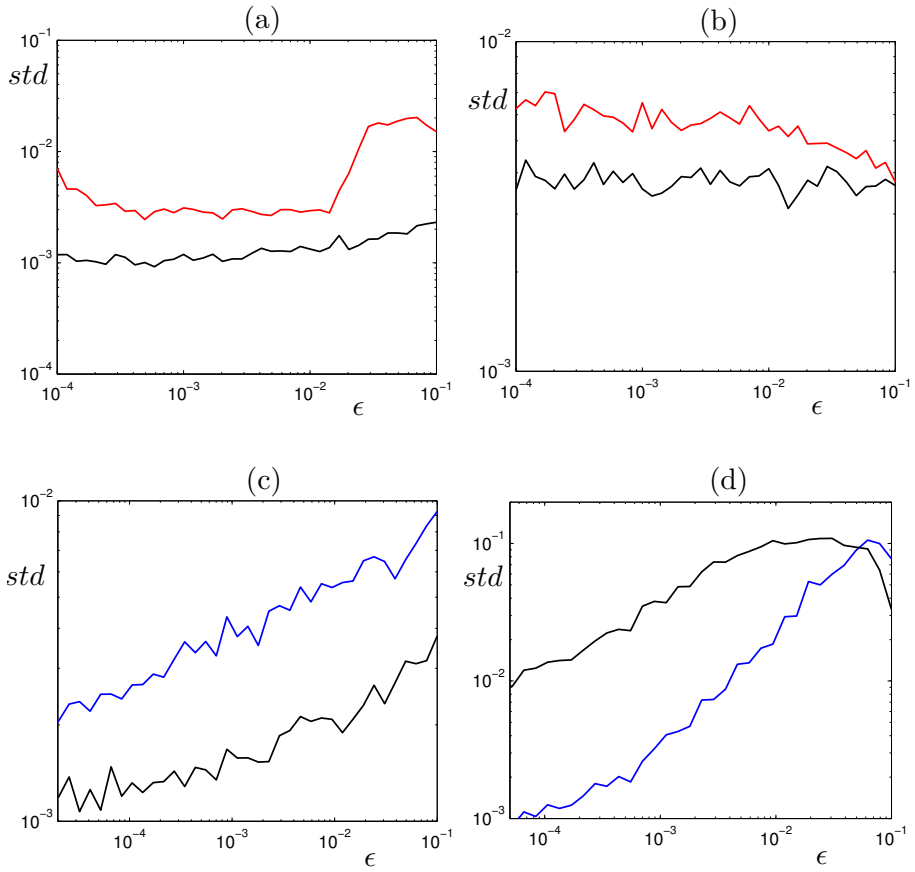

FIGURE 13. Dependence of the standard deviation of the distance between the stochastic and the deterministic transition maps on ϵ for an attracting deterministic periodic MMO during various phases of the flow. We have always fixed $\sigma = 0.01 = \sigma'$ and viewed ϵ as a parameter. Otherwise, the same conventions as in Figure 12 apply.

Via the same strategy as in Figure 12 one may also numerically investigate the dependence upon ϵ. Figure 13 shows the results for this computation. Again, the results are consistent with the derived upper bounds. Figure 13(a)–(b) verifies that for the maps $\Sigma_2 \to \Sigma_3$ and $\Sigma_3 \to \Sigma_4$ the stochastic spreadings of order $\mathcal{O}(\sigma + \sigma')$ and $\mathcal{O}(\sigma\sqrt{\epsilon} + \sigma')$ are dominated by σ, σ' if the noise level is fixed. For the map $\Sigma_4 \to \Sigma_5$ analyzed in Figure 13(c) we expect from Table 1 that the spreading is dominated by a scaling $\mathcal{O}(\epsilon^{1/6})$ since we have fixed $\sigma = 0.01 = \sigma'$ and $\epsilon^{1/2} \ll \epsilon^{1/6}$ as $\epsilon \to 0$. Inspection of Figure 13(c) shows indeed a corresponding slope of approximately 1/6. Figure 13(d) is also consistent

with the expected scaling $\mathcal{O}(\epsilon^{1/4})$ near the folded node for noise level and μ fixed. These results provide very good evidence that our theoretical estimates may also form a practical guideline to analyze the spreading due to noise.

Of course, one may also consider the influence of noise on sample paths for global returns. Figure 1 shows the global return map $\Sigma_1 \to \Sigma_1$. For this computation, in contrast to the previous computations in this section, we have chosen a regime with many secondary canards [17, 83]. Indeed, the parameter values have been fixed to $k = -10$, $\lambda = -7.6$ so that the folded node on L^+ is given by $(x, z) = (1, -9.2)$ with eigenvalues of the matrix A from (7.5) given by $\rho_s < -1 < \rho_w < 0$. The eigenvalue ratio $\mu := \rho_w/\rho_s$ is approximately given by $\mu \approx 0.0252$ so that [26, Thm. 2.3] implies that there are two primary and 19 secondary canards. The deterministic return map has been analyzed numerically in [37, 59], and the structure of the different rotational sectors separated by canard orbits is well understood; see also [40]. However, Figure 1 shows that the attracting deterministic periodic orbit corresponding to a fixed point of the return map, can shift due to noise, even to a higher sector of rotation. This effect can be seen directly from Theorem 6.4 above.

In summary, we may conclude that there is a highly non-trivial interplay between the number of SAOs, the global return map and the noise level (cf. also [17, Cor. 6.3]). As discussed already in Section 1, the natural next step is to consider the analysis of the discrete-time Markov chain on a finite state space of MMO patterns. The results in this paper and in [17] provide the necessary estimates for the kernel of the Markov chain and may form the starting point for future work.

A Technical lemmas

Lemma A.1 (Scaling behaviour)**.** *Assume $a(s) \asymp -(|s|^{1/2} + \epsilon^{1/3})$ for $-1 \leqslant s \leqslant \epsilon^{2/3}$, and define $\alpha(s, r) = \int_r^s a(u)\,\mathrm{d}u$. Then for any $\nu \geqslant -1$,*

$$\int_{s_0}^{s} \mathrm{e}^{\alpha(s,r)/\epsilon} |a(r)|^\nu \,\mathrm{d}r \asymp \epsilon |a(s)|^{2\nu-1} \tag{A.1}$$

holds for $-1 \leqslant s_0 \leqslant s_0 + \mathcal{O}(\epsilon|\log \epsilon|) \leqslant s \leqslant \epsilon^{2/3}$.

Lemma A.2 (Bernstein-type inequality)**.** *Let $W_t^{(1)}, \dots, W_t^{(k)}$ be k independent standard Brownian motions, and let X_t be adapted to the filtration generated by the $W_t^{(j)}$. For measurable functions g_{ij} and deterministic bounds G_i satisfying almost surely*

$$\sum_{j=1}^{k} g_{ij}(X_s, s)^2 \leqslant G_i(s)^2 \,, \qquad V_i(t) = \int_0^t G_i(s)^2 \,\mathrm{d}s < \infty \,, \tag{A.2}$$

consider the n martingales

$$M_t^{(i)} = \sum_{j=1}^{k} \int_0^t g_{ij}(X_s, s) \,\mathrm{d}W_s^{(j)} \,, \qquad i = 1, \dots, n \,. \tag{A.3}$$

Then for any $h > 0$ and any choice of $\gamma_1, \dots, \gamma_n \in [0, 1]$ such that $\gamma_1 + \cdots + \gamma_n = 1$, one has

$$\mathbb{P}\Big\{ \sup_{0 \leqslant s \leqslant t} \|M_s\| \geqslant h \Big\} \leqslant 2 \sum_{i=1}^{n} \exp\Big\{ -\frac{\gamma_i h^2}{2V_i(t)} \Big\} \,. \tag{A.4}$$

Proof. The left-hand side of (A.4) is bounded above by

$$\mathbb{P}\Big\{\sum_{i=1}^{n} \sup_{0\leqslant s\leqslant t} (M_s^{(i)})^2 \geqslant h^2\Big\} \leqslant \sum_{i=1}^{n} \mathbb{P}\Big\{\sup_{0\leqslant s\leqslant t} (M_s^{(i)})^2 \geqslant \gamma_i h^2\Big\} \,. \tag{A.5}$$

Each term in this sum is bounded by $2\,\mathrm{e}^{-\gamma_i h^2/(2V_i(t))}$, cf. [17, Lemma D.8]. □

Remark A.3. It is possible to obtain sharper estimates of the form

$$\mathbb{P}\Big\{\sup_{0\leqslant s\leqslant t} \frac{\|M_s\|}{\|V(s)\|^{1/2}} \geqslant h\Big\} \leqslant C(t,\kappa)\,\mathrm{e}^{-\kappa h^2/2\sigma^2} \tag{A.6}$$

for any $\kappa < 1$, using a decomposition of $[0,t]$ in small intervals (see [15, Section 5.1.2]).

Lemma A.4 (Random time change)**.** *Consider an $\mathbb{R}^n$-valued diffusion $(Y_t)_{t\geqslant 0}$ given by*

$$\mathrm{d}Y_t = f(Y_t)\,\mathrm{d}t + g(Y_t)\,\mathrm{d}W_t \,, \tag{A.7}$$

where $f:\mathbb{R}^n\to\mathbb{R}^n$ is such that the nth component f_n of f satisfies $f_n(y)>0$ for all y. We further assume that $g:\mathbb{R}^n\to\mathbb{R}^{n\times k}$, and that $(W_t)_{t\geqslant 0}$ is a k-dimensional standard Brownian motion. Fix $\gamma>0$ and let

$$\beta(t,\omega) = \gamma\int_0^t f_n(Y_s(\omega))\,\mathrm{d}s \,. \tag{A.8}$$

Then $(Y_{\beta^{-1}}(t))_{t\geqslant 0}$ is equal in distribution to $(X_t)_{t\geqslant 0}$, where X_t satisfies $X_0=Y_0$ and

$$\mathrm{d}X_t = \frac{1}{\gamma f_n(X_t)} f(X_t)\,\mathrm{d}t + \frac{1}{\sqrt{\gamma f_n(X_t)}} g(X_t)\,\mathrm{d}W_t \,. \tag{A.9}$$

If the condition $f_n(y)>0$ is only satisfied in a subset $\mathcal{D}$ of $\mathbb{R}^n$, then the result remains true for $0\leqslant t\leqslant \tau_{\mathcal{D}} = \inf\{s\geqslant 0\colon Y_s\notin\mathcal{D}\}$.

Proof. Write $\mathrm{d}X_t = \tilde f(X_t)\,\mathrm{d}t + \tilde g(X_t)\,\mathrm{d}W_t$ for the SDE (A.9). The stochastic process $c(t,\omega) = \gamma f_n(Y_t(\omega))$ is adapted to the filtration $(\mathcal{F}_t)_t$ of the Brownian motion. Since $t\mapsto\beta(t)$ is almost surely invertible, [70, Theorem 8.5.1, p. 154] implies that $(X_t)_t$ is equal in distribution to $(Y_{\beta^{-1}(t)})_t$, with Y_t is given by

$$\mathrm{d}Y_t = u(t,\omega)\,\mathrm{d}t + v(t,\omega)\,\mathrm{d}W_t \,, \tag{A.10}$$

provided

$$\begin{aligned} u(t,\omega) &= c(t,\omega)\tilde f(Y_t) \,, \\ vv^{\mathrm{T}}(t,\omega) &= c(t,\omega)\tilde g\tilde g^{\mathrm{T}}(Y_t) \,. \end{aligned} \tag{A.11}$$

Setting $u(t,\omega)=f(Y_t(\omega))$ and $v(t,\omega)=g(Y_t(\omega))$ this is clearly the case. To prove the last statement, it suffices to consider $Y_{t\wedge\tau_{\mathcal{D}}}$. □

B Proof of Lemma 5.2

We would like to estimate the principal solution of $\epsilon\dot{\xi} = A(s)\xi$ for times $s \leqslant -\sqrt{\epsilon}$, where

$$A(s) = \begin{pmatrix} a_0(s) & d(s) \\ c(s) & \epsilon a_1(s) \end{pmatrix} = \begin{pmatrix} s + \mathcal{O}(s^2) & 1 + \mathcal{O}(s^2) \\ -\epsilon(1+\mu) + \mathcal{O}(\epsilon s) & \epsilon\mathcal{O}(1) \end{pmatrix} . \tag{B.1}$$

By changing s into $-s$ and ξ_2 into $-\xi_2$, we can restrict the analysis to positive s. The proof of the lemma is close in spirit to the proof of [17, Theorem 4.3], and consists in a number of changes of variables bringing the system into diagonal form. A first transformation

$$\xi = \exp\left\{ \frac{1}{2\epsilon} \int_0^s \operatorname{Tr} A(u)\, \mathrm{d}u \right\} \xi_1 =: \mathrm{e}^{\alpha(s)/2\epsilon}\, \xi_1 \tag{B.2}$$

yields the system $\epsilon\dot{\xi}_1 = A_1(s)\xi_1$, where

$$A_1(s) = \begin{pmatrix} \frac{1}{2}a(s) & d(s) \\ c(s) & -\frac{1}{2}a(s) \end{pmatrix} , \qquad a(s) = a_0(s) - \epsilon a_1(s) = s + \mathcal{O}(s^2) . \tag{B.3}$$

Next we set $\xi_1 = S_1(s)\xi_2$, where

$$S_1(s) = \frac{1}{\sqrt{d(s)}} \begin{pmatrix} d(s) & 0 \\ -\frac{1}{2}a(s) + \frac{1}{2}\epsilon\frac{\dot{d}(s)}{d(s)} & 1 \end{pmatrix} , \tag{B.4}$$

which yields $\epsilon\dot{\xi}_2 = A_2(s)\xi_2$, with

$$A_2(s) = \begin{pmatrix} 0 & 1 \\ h(s) & 0 \end{pmatrix} , \tag{B.5}$$

and

$$h = \frac{a^2}{4} + cd + \epsilon\frac{\dot{a}}{2} - \epsilon\frac{a\dot{d}}{2d} + \frac{\epsilon^2}{2}\frac{\ddot{d}}{d} + \frac{3}{4}\epsilon^2\frac{\dot{d}^2}{d^2} = \frac{a^2}{4} - \epsilon\left(\frac{1}{2} + \mu\right) + \mathcal{O}(\epsilon s) . \tag{B.6}$$

Note that this system is equivalent to $\epsilon^2\ddot{\xi}_{2,1} = h(s)\xi_{2,1}$, which reduces to Weber's equation in the particular case $a(s) = s$. The next step is to set $\xi_2 = S_2(s)\xi_3$, where

$$S_2(s) = \frac{1}{\sqrt{2}} \begin{pmatrix} h(s)^{-1/4} & -h(s)^{-1/4} \\ h(s)^{1/4} & h(s)^{1/4} \end{pmatrix} \tag{B.7}$$

is such that $S_2^{-1}A_2S_2$ is diagonal. This yields $\epsilon\dot{\xi}_3 = A_3(s)\xi_3$, with

$$A_3(s) = \begin{pmatrix} h(s)^{1/2} & -\frac{\epsilon}{4}\frac{\dot{h}(s)}{h(s)} \\ -\frac{\epsilon}{4}\frac{\dot{h}(s)}{h(s)} & -h(s)^{1/2} \end{pmatrix} . \tag{B.8}$$

The last transformation $\xi_3 = S_3(s)\xi_4$ makes the system diagonal. This final transformation is given by

$$S_3(s) = \begin{pmatrix} 1 & p_2(s) \\ p_1(s) & 1 \end{pmatrix} , \tag{B.9}$$

where p_1 and p_2 satisfy the ODEs

$$\epsilon \dot{p_1} = -2h(s)^{1/2} p_1 + \frac{\epsilon}{4} \frac{\dot{h}(s)}{h(s)} (p_1^2 - 1) \,, \tag{B.10}$$

$$\epsilon \dot{p_2} = 2h(s)^{1/2} p_2 + \frac{\epsilon}{4} \frac{\dot{h}(s)}{h(s)} (p_2^2 - 1) \,. \tag{B.11}$$

One can show that these ODEs admit solutions of order ϵ/s^2. The resulting system has the form $\epsilon \dot{\xi}_4 = A_4(s)\xi_4$, where

$$A_4(s) = \begin{pmatrix} h(s)^{1/2} - \frac{\epsilon}{4} \frac{\dot{h}(s)}{h(s)} p_1(s) & 0 \\ 0 & -h(s)^{1/2} - \frac{\epsilon}{4} \frac{\dot{h}(s)}{h(s)} p_2(s) \end{pmatrix} \,, \tag{B.12}$$

and the principal solution is thus of the form $V(s,r) = V(s)V(r)^{-1}$, where

$$V(s) = \mathrm{e}^{\alpha(s)/2\epsilon} \, S_1(s) S_2(s) S_3(s) \begin{pmatrix} \mathrm{e}^{\alpha_+(s)/\epsilon} & 0 \\ 0 & \mathrm{e}^{\alpha_-(s)/\epsilon} \end{pmatrix} \tag{B.13}$$

with

$$\alpha_\pm(s) = \pm \int_1^s h(u)^{1/2} \, \mathrm{d}u + \mathcal{O}\left(\frac{\epsilon^2}{s^2}\right) . \tag{B.14}$$

Expanding $h(s)^{1/2}$ and using the fact that $\dot{a}(s) = 1 + \mathcal{O}(s)$, one obtains

$$\mathrm{e}^{\alpha_\pm(s)/\epsilon} \asymp a(s)^{\mp(1/2+\mu)} \, \mathrm{e}^{\pm\alpha(s)/2\epsilon} \,. \tag{B.15}$$

The result follows by evaluating the matrix products in (B.13). □

C Proof of Proposition 5.5

We shall use a parametrization of the level curves of K for $K > 0$ which was introduced in [18]. It is given by

$$\begin{aligned} u_1 &= \sqrt{\frac{1+\mu}{2} |\log K|} \, \sin\varphi \,, \\ u_2 &= u_1^2 + \frac{1+\mu}{2} f\big(X\big) \,, \qquad X := \sqrt{\frac{|\log K|}{2}} \cos\varphi \,, \end{aligned} \tag{C.1}$$

where $f(t)$ is the solution of

$$\log(1 + f(t)) = f(t) - 2t^2 \tag{C.2}$$

satisfying $\operatorname{sign} f(t) = \operatorname{sign} t$. It is easy to check (see [63, Section 5.2.1]) that

- $f(t) = 2t + \mathcal{O}(t^2)$ near $t = 0$;
- $-1 + \mathrm{e}^{-1-2t^2} \leqslant f(t) \leqslant -1 + \mathrm{e}^{-1-2t^2} + \mathcal{O}((\mathrm{e}^{-1-2t^2})^2)$ for $t \leqslant 0$;
- $f(t) \leqslant 2t^2 + \mathcal{O}(t)$ for all $t \geqslant 0$.

Lemma C.1. *The variational equations* (5.26) *are equivalent to*

$$\begin{aligned}\mu\frac{\mathrm{d}K}{\mathrm{d}\bar{z}} &= -8\bar{z}\frac{K|\log K|}{1+f(X)}\sin^2\varphi\;,\\ \mu\frac{\mathrm{d}\varphi}{\mathrm{d}\bar{z}} &= \frac{1}{X}\biggl[\sqrt{1+\mu}f(X) - 4\bar{z}\sqrt{\frac{|\log K|}{2}}\sin\varphi\biggl(\frac{\sin^2\varphi}{1+f(X)}-1\biggr)\biggr]\;.\end{aligned} \tag{C.3}$$

Proof. The first equation is a direct consequence of (5.28), using the fact that $\mathrm{e}^{-f(X)} = K^{\cos^2\varphi}/[1+f(X)]$. The second one is obtained by differentiating the first equation in (C.1) and solving for $\mu\,\mathrm{d}\varphi/\,\mathrm{d}\bar{z}$. □

The first term on the right-hand side of $\mu\,\mathrm{d}\varphi/\,\mathrm{d}\bar{z}$ is of order $f(X)/X$, which varies between $1/|X|$ and $2|X|$. As for the second term, it is easy to see that it is negligible for $X \geqslant 0$. For $X < 0$, setting $u = \sin\varphi$, we have $1 + f(X) = K^{1-u^2}$. The function $u \mapsto u^3K^{u^2-1}$ has a maximal value $1/(K|\log K|^{3/2})$, reached when $u^2 = 3/(2|\log K|)$. This allows to bound the second term for $X < 0$ and proves (5.29).

It follows immediately from (C.3) that we can write

$$\frac{\mathrm{d}K}{\mathrm{d}\varphi} = \bar{z}g(K,\varphi,\bar{z})\;, \tag{C.4}$$

where

$$g(K,\varphi,\bar{z}) = -8\frac{K|\log K|}{1+f(X)}\frac{X\sin^2\varphi}{\sqrt{1+\mu}f(X) - 4\bar{z}R\sin\varphi\Bigl[\frac{\sin^2\varphi}{(1+f(X))}-1\Bigr]}\;, \tag{C.5}$$

and we have set $R = \sqrt{|\log K|/2}$. Note that $\mu\,\mathrm{d}K/\,\mathrm{d}\bar{z}$ has order $(1-K)$ near $K=1$, and order $|\log K|$ near $K=0$, because $1+f(X) \geqslant \mathrm{e}^{-1}K$. This shows that

$$|g(K,\varphi,\bar{z})| \leqslant \mathcal{O}((1-K)(1+|\log K|)^{3/2}) =: \mathcal{O}(\rho(K))\;. \tag{C.6}$$

We now perform the averaging transformation

$$\overline{K} = K + \bar{z}w(K,\varphi,\bar{z})\;, \tag{C.7}$$

where

$$\frac{\partial}{\partial\varphi}w(K,\varphi,\bar{z}) = -g(K,\varphi,\bar{z}) + \bar{g}(K,\bar{z})\;, \qquad \bar{g}(K,\bar{z}) = \frac{1}{2\pi}\int_0^{2\pi} g(K,\varphi,\bar{z})\,\mathrm{d}\varphi\;. \tag{C.8}$$

This yields

$$\frac{\mathrm{d}\overline{K}}{\mathrm{d}\varphi} = \bar{z}\bar{g}(K,\bar{z}) + \bar{z}\frac{\partial w}{\partial K}\frac{\mathrm{d}K}{\mathrm{d}\varphi} + \biggl(\bar{z}\frac{\partial w}{\partial\bar{z}} + w\biggr)\frac{\mathrm{d}\bar{z}}{\mathrm{d}\varphi}\;. \tag{C.9}$$

Recalling that $\mathrm{d}K/\,\mathrm{d}\varphi$ has order $\bar{z}g$ and $\mathrm{d}\bar{z}/\,\mathrm{d}\varphi$ has order $\mu|\log K|^{1/2}$, together with (C.6) allows to bound the last two terms.

Finally, we show that $c_-(1-K) \leqslant -\bar{g} \leqslant c_+(1-K)$ for $K \geqslant c\bar{z}$. The result will then follow by expressing K in terms of $\overline{K}$ (note that $1-K$ and $1-\overline{K}$ are comparable for $K \geqslant c\bar{z}$). It will be sufficient to consider the behaviour of $\bar{g}$ near $K=1$ and near $K=0$. Near $K=1$ we have

$$g(K,\varphi,\bar{z}) = -\frac{4(1-K)}{\sqrt{1+\mu}}\bigl[1+\mathcal{O}(1-K)+\mathcal{O}(\bar{z})\bigr]\sin^2\varphi\;, \tag{C.10}$$

and $\sin^2\varphi$ averages to $1/2$. Near $K = 0$, the integral defining $\bar{g}$ is dominated by φ near π. Performing the change of variables $u = 2R\sin\varphi$, we obtain

$$\begin{aligned}\int_{\pi/2}^{3\pi/2} g(K,\varphi,\bar{z})\,\mathrm{d}\varphi &= -\frac{2\,\mathrm{e}}{\sqrt{1+\mu}}\int_{-2R}^{2R} u^2\,\mathrm{e}^{-u^2/2}\biggl[1+\mathcal{O}(K\,\mathrm{e}^{u^2/2})+\mathcal{O}\biggl(\frac{\bar{z}u^3}{KR^2}\,\mathrm{e}^{-u^2/2}\biggr)\biggr]\\ &= -\frac{2\sqrt{2\pi}\,\mathrm{e}}{\sqrt{1+\mu}}\Bigl[1+\mathcal{O}(K|\log K|)+\mathcal{O}(\bar{z}/(K|\log K|))\Bigr]\,, \qquad \text{(C.11)}\end{aligned}$$

while the integral over $[-\pi/2,\pi/2]$ has order $K|\log K|$. □

References

[1] M. Abramowitz and I. Stegun. *Handbook of Mathematical Functions*. Dover, 10th edition, 1972.

[2] V. I. Arnold, V. S. Afrajmovich, Y. S. Il′yashenko, and L. P. Shil′nikov. *Bifurcation theory and catastrophe theory*. Springer, Berlin, 1999. Translated from the 1986 Russian original, Reprint of the 1994 English edition from the series Encyclopaedia of Mathematical Sciences [ıt Dynamical systems. V, Encyclopaedia Math. Sci., 5, Springer, Berlin, 1994].

[3] S. Arrhenius. On the reaction velocity of the inversion of cane sugar by acids. *J. Phys. Chem.*, 4:226, 1889. In German. Translated and published in: Selected Readings in Chemical Kinetics, M.H. Back and K.J. Laider (eds.), Pergamon, Oxford, 1967.

[4] P. Ashwin, S. Wieczorek, R. Vitolo, and P. Cox. Tipping points in open systems: bifurcation, noise-induced and rate-dependent examples in the climate system. *Phil. Trans. R. Soc. A*, 370:1166–1184, 2012.

[5] K. E. Avrachenkov and J. B. Lasserre. The fundamental matrix of singularly perturbed Markov chains. *Adv. in Appl. Probab.*, 31(3):679–697, 1999.

[6] G. Ben Arous, S. Kusuoka, and D. W. Stroock. The Poisson kernel for certain degenerate elliptic operators. *J. Funct. Anal.*, 56(2):171–209, 1984.

[7] E. Benoît. Systèmes lents-rapides dans $\mathbb{R}^3$ et leurs canards. In *Third Snepfenried geometry conference*, volume 2, pages 159–191. Soc. Math. France, 1982.

[8] E. Benoît. Enlacements de canards. *C.R. Acad. Sc. Paris*, 300(8):225–230, 1985.

[9] E. Benoît. Canards et enlacements. *Publ. Math. IHES*, 72:63–91, 1990.

[10] E. Benoît, J. Callot, F. Diener, and M. Diener. Chasse au canards. *Collect. Math.*, 31:37–119, 1981.

[11] E. Benoît and C. Lobry. Les canards de $\mathbb{R}^3$. *C.R. Acad. Sc. Paris*, 294:483–488, 1982.

[12] R. Benzi, A. Sutera, and A. Vulpiani. The mechanism of stochastic resonance. *J. Phys. A*, 14(11):L453–L457, 1981.

[13] N. Berglund. Kramers' law: Validity, derivations and generalisations. *Markov Process. Related Fields*, 19(3):459–490, 2013.

[14] N. Berglund and B. Gentz. Geometric singular perturbation theory for stochastic differential equations. *J. Differential Equations*, 191:1–54, 2003.

[15] N. Berglund and B. Gentz. *Noise-induced phenomena in slow–fast dynamical systems. A sample-paths approach*. Probability and its Applications. Springer-Verlag, London, 2006.

[16] N. Berglund and B. Gentz. On the noise-induced passage through an unstable periodic orbit II: The general case. Preprint `arXiv/1208.2557`, to appear in SIAM J. Math. Anal., 2013.

[17] N. Berglund, B. Gentz, and C. Kuehn. Hunting French ducks in a noisy environment. *J. Differential Equations*, 252(9):4786–4841, 2012.

[18] N. Berglund and D. Landon. Mixed-mode oscillations and interspike interval statistics in the stochastic FitzHugh–Nagumo model. *Nonlinearity*, 25(8):2303–2335, 2012.

[19] J. Boissonade and P. DeKepper. Transitions from bistability to limit cycle oscillations. Theoretical analysis and experimental evidence in an open chemical system. *J. Phys. Chem.*, 84:501–506, 1980.

[20] K. Bold, C. Edwards, J. Guckenheimer, S. Guharay, K. Hoffman, J. Hubbard, R. Oliva, and W. Weckesser. The forced van der Pol equation 2: Canards in the reduced system. *SIAM Journal of Applied Dynamical Systems*, 2(4):570–608, 2003.

[21] A. Bovier, M. Eckhoff, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes. I. Sharp asymptotics for capacities and exit times. *J. Eur. Math. Soc. (JEMS)*, 6(4):399–424, 2004.

[22] A. Bovier, V. Gayrard, and M. Klein. Metastability in reversible diffusion processes. II. Precise asymptotics for small eigenvalues. *J. Eur. Math. Soc. (JEMS)*, 7(1):69–99, 2005.

[23] M. Brons, T. Kaper, and H. Rotstein. Introduction to focus issue – mixed mode oscillations: experiment, computation, and analysis. *Chaos*, 18:015101, 2008.

[24] M. Brøns, M. Krupa, and M. Wechselberger. Mixed mode oscillations due to the generalized canard phenomenon. *Fields Institute Communications*, 49:39–63, 2006.

[25] H. Degn, L. Olsen, and J. Perram. Bistability, oscillation, and chaos in an enzyme reaction. *Annals of the New York Academy of Sciences*, 316(1):623–637, 1979.

[26] M. Desroches, J. Guckenheimer, C. Kuehn, B. Krauskopf, H. Osinga, and M. Wechselberger. Mixed-mode oscillations with multiple time scales. *SIAM Review*, 54(2):211–288, 2012.

[27] M. Desroches, B. Krauskopf, and H. Osinga. The geometry of mixed-mode oscillations in the Olsen model for the perioxidase-oxidase reaction. *DCDS-S*, 2(4):807–827, 2009.

[28] C. Dickson, J. Magistretti, M. Shalisnky, B. Hamam, and A. Alonso. Oscillatory activity in entorhinal neurons and circuits: Mechanisms and function. *Ann. N.Y. Acad. Sci.*, 911:127–150, 2006.

[29] F. Dumortier and R. Roussarie. Canard cycles and center manifolds. *Memoirs of the American Mathematical Society*, 121(577), 1996.

[30] G. Ermentrout and D. Terman. *Mathematical Foundations of Neuroscience*. Springer, 2010.

[31] H. Eyring. The activated complex in chemical reactions. *Journal of Chemical Physics*, 3:107–115, 1935.

[32] N. Fenichel. Geometric singular perturbation theory for ordinary differential equations. *J. Differential Equations*, 31(1):53–98, 1979.

[33] M. Freidlin and A. Wentzell. *Random Perturbations of Dynamical Systems*. Springer, 1998.

[34] L. Gammaitoni, P. Hänggi, P. Jung, and F. Marchesoni. Stochastic resonance. *Rev. Mod. Phys.*, 70:223–287, 1998.

[35] A. Goryachev, P. Strizhak, and R. Kapral. Slow manifold structure and the emergence of mixed-mode oscillations. *J. Chem. Phys.*, 107(18):2881–2889, 1997.

[36] J. Guckenheimer. Global bifurcations of periodic orbits in the forced van der Pol equation. *in: Global Analysis of Dynamical Systems — Festschrift dedicated to Floris Takens. Eds.: Henk W. Broer, Bernd Krauskopf and Gert Vegter*, pages 1–16, 2003.

[37] J. Guckenheimer. Return maps of folded nodes and folded saddle-nodes. *Chaos*, 18, 2008.

[38] J. Guckenheimer. Singular Hopf bifurcation in systems with two slow variables. *SIAM J. Appl. Dyn. Syst.*, 7(4):1355–1377, 2008.

[39] J. Guckenheimer and P. Meerkamp. Bifurcation analysis of singular hopf bifurcation in $\mathbb{R}^3$. *SIAM J. Appl. Dyn. Syst.*, 11(4):1325–1359, 2012.

[40] J. Guckenheimer and C. Scheper. A geometric model for mixed-mode oscillations in a chemical system. *SIAM J. Appl. Dyn. Sys.*, 10(1):92–128, 2011.

[41] R. Hassin and M. Haviv. Mean passage times and nearly uncoupled Markov chains. *SIAM J. Discrete Math.*, 5(3):386–397, 1992.

[42] T. Hauck and F. Schneider. Mixed-mode and quasiperiodic oscillations in the peroxidase-oxidase reaction. *J. Phys. Chem.*, 97:391–397, 1993.

[43] P. Hitczenko and G. Medvedev. Bursting oscillations induced by small noise. *SIAM J. Appl. Math.*, 69:1359–1392, 2009.

[44] P. Hitczenko and G. Medvedev. The Poincaré map of randomly perturbed periodic motion. *J. Nonlinear Sci.*, 23(5):835–861, 2013.

[45] A. Hodgin and A. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.*, 117:500–505, 1952.

[46] R. Höpfner, E. Löcherbach, and M. Thieullen. Ergodicity for a stochastic Hodgkin–Huxley model driven by Ornstein–Uhlenbeck type input. arXiv:1311.3458v1, 2013.

[47] J. Hudson, M. Hart, and D. Marinko. An experimental study of multiple peak periodic and nonperiodic oscillations in the Belousov–Zhabotinskii reaction. *J. Chem. Phys.*, 71(4):1601–1606, 1979.

[48] E. Izhikevich. Neural excitability, spiking, and bursting. *Int. J. Bif. Chaos*, 10:1171–1266, 2000.

[49] I. Karatzas and S. E. Shreve. *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1991.

[50] A. Kawczynski and P. Strizhak. Period adding and broken Farey tree sequences of bifurcations for mixed-mode oscillations and chaos in the simplest three-variable nonlinear system. *J. of Chem. Phys.*, 112(14):6122–6130, 2000.

[51] M. Koper. Bifurcations of mixed-mode oscillations in a three-variable autonomous Van der Pol–Duffing model with a cross-shaped phase diagram. *Physica D*, 80:72–94, 1995.

[52] E. K. Kosmidis and K. Pakdaman. An analysis of the reliability phenomenon in the FitzHugh–Nagumo model. *J. Comput. Neuroscience*, 14:5–22, 2003.

[53] H. A. Kramers. Brownian motion in a field of force and the diffusion model of chemical reactions. *Physica*, 7:284–304, 1940.

[54] M. Krupa, N. Popovic, and N. Kopell. Mixed-mode oscillations in three time-scale systems: A prototypical example. *SIAM J. Applied Dynamical Systems*, 7(2), 2008.

[55] M. Krupa and P. Szmolyan. Extending geometric singular perturbation theory to nonhyperbolic points - fold and canard points in two dimensions. *SIAM J. Math. Anal.*, 33(2):286–314, 2001.

[56] M. Krupa and P. Szmolyan. Extending slow manifolds near transcritical and pitchfork singularities. *Nonlinearity*, 14:1473–1491, 2001.

[57] M. Krupa and P. Szmolyan. Geometric analysis of the singularly perturbed fold. *In: Multiple-Time-Scale Dynamical Systems*, IMA Vol. 122:89–116, 2001.

[58] M. Krupa and M. Wechselberger. Local analysis near a folded saddle-node singularity. *J. Diff. Eq.*, 248(12):2841–2488, 2010.

[59] C. Kuehn. On decomposing mixed-mode oscillations and their return maps. *Chaos*, 21(3):033107, 2011.

[60] C. Kuehn. A mathematical framework for critical transitions: normal forms, variance and applications. *J. Nonlinear Sci.*, 23(3):457–510, 2013.

[61] C. Kuehn. Loss of normal hyperbolicity of unbounded critical manifolds. arXiv:1204.0947, 2013.

[62] R. Kuske and R. Borowski. Survival of subthreshold oscillations: the interplay of noise, bifurcation structure, and return mechanism. *Discrete and Continous Dynamical Systems S*, 2(4):873–895, 2009.

[63] D. Landon. *Perturbation et excitabilité dans des modèles stochastiques de transmission de l'influx nerveux*. PhD thesis, Université d'Orléans, 2012.
`http://tel.archives-ouvertes.fr/tel-00752088`.

[64] B. McNamara and K. Wiesenfeld. Theory of stochastic resonance. *Phys. Rev. A*, 39:4854–4869, 1989.

[65] M. Mikikian, M. Cavarroc, L. Couedel, Y. Tessier, and L. Boufendi. Mixed-mode oscillations in complex plasma instabilities. *Physical Review Letters*, 100(22), 2008.

[66] A. Milik, P. Szmolyan, H. Loeffelmann, and E. Groeller. Geometry of mixed-mode oscillations in the 3-d autocatalator. *Int. J. of Bif. and Chaos*, 8(3):505–519, 1998.

[67] E. Mishchenko, Y. Kolesov, A. Kolesov, and N. Rozov. *Asymptotic Methods in Singularly Perturbed Systems*. Plenum Press, 1994.

[68] E. Mishchenko and N. Rozov. *Differential Equations with Small Parameters and Relaxation Oscillations*. Plenum Press, 1980. (translated from Russian).

[69] C. Muratov and E. Vanden-Eijnden. Noise-induced mixed-mode oscillations in a relaxation oscillator near the onset of a limit cycle. *Chaos*, 18:015111, 2008.

[70] B. Øksendal. *Stochastic Differential Equations*. Springer, 2003.

[71] V. Petrov, S. Scott, and K. Showalter. Mixed-mode oscillations in chemical systems. *J. Chem. Phys.*, 97(9):6191–6198, 1992.

[72] O. Rössler. Chaos in abstract kinetics: Two prototypes. *Bulletin of Mathematical Biology*, 39:275–289, 1977.

[73] J. Rubin and M. Wechselberger. Giant squid – hidden canard: the 3d geometry of the Hodgin–Huxley model. *Biological Cybernetics*, 97(1), 2007.

[74] M. Scheffer, J. Bascompte, W. Brock, V. Brovkhin, S. Carpenter, V. Dakos, H. Held, E. van Nes, M. Rietkerk, and G. Sugihara. Early-warning signals for critical transitions. *Nature*, 461:53–59, 2009.

[75] P. J. Schweitzer. Perturbation theory and finite Markov chains. *J. Appl. Probability*, 5:401–413, 1968.

[76] J. Su, J. Rubin, and D. Terman. Effects of noise on elliptic bursters. *Nonlinearity*, 17(1):133–157, 2004.

[77] P. Szmolyan and M. Wechselberger. Canards in $\mathbb{R}^3$. *Journal of Differential Equations*, 177:419–453, 2001.

[78] P. Szmolyan and M. Wechselberger. Relaxation oscillations in $\mathbb{R}^3$. *Journal of Differential Equations*, 200:69–104, 2004.

[79] B. Van der Pol. The nonlinear theory of electric oscillations. *Proc. IRE*, 22:1051–1086, 1934.

[80] F. van Goor, D. Zivadinovic, A. Martinez-Fuentes, and S. Stojilkovic. Dependence of pituitary hormone secretion on the pattern of spontaneous voltage-gated calcium influx. *J. Biol. Chem.*, 276:33840–33846, 2001.

[81] G. Wallet. Entrée–sortie dans un tourbillon. *Ann. Inst. Fourier*, 36:157–184, 1986.

[82] M. Wechselberger. *Singularly perturbed folds and canards in* $\mathbb{R}^3$. PhD thesis, Vienna University of Technology, Vienna, Austria, 1998.

[83] M. Wechselberger. Existence and bifurcation of canards in $\mathbb{R}^3$ in the case of a folded node. *SIAM J. Applied Dynamical Systems*, 4(1):101–139, 2005.

[84] M. Wechselberger and W. Weckesser. Bifurcations of mixed-mode oscillations in a stellate cell model. *Physica D*, 238:1598–1614, 2009.

[85] J. Weiss and E. Knobloch. A stochastic return map for stochastic differential equations. *J. Stat. Phys.*, 58(5):863–883, 1990.

[86] G. Yin and Q. Zhang. *Discrete-time Markov Chains: Two-Time-Scale Methods and Applications*. Springer, 2005.

[87] N. Yu, R. Kuske, and Y. Li. Stochastic phase dynamics and noise-induced mixed-mode oscillations in coupled oscillators. *Chaos*, 18:015112, 2008.

Contents