J. Differential Equations 252 (2012) 4786–4841

Contents lists available at SciVerse ScienceDirect

Journal of Differential Equations

www.elsevier.com/locate/jde

Hunting French ducks in a noisy environment

Nils Berglund [a,1], Barbara Gentz [b,2], Christian Kuehn [c,*]

[a] MAPMO, CNRS – UMR 6628, Université d'Orléans, Fédération Denis Poisson, FR 2964, B.P. 6759, 45067 Orléans Cedex 2, France
[b] Faculty for Mathematics, University of Bielefeld, P.O. Box 10 01 31, 33501 Bielefeld, Germany
[c] Max Planck Institute for Physics of Complex Systems, Noethnitzer Str. 38, 01187 Dresden, Germany

ARTICLE INFO

Article history:
Received 15 November 2010
Revised 4 January 2012
Available online 7 February 2012

MSC:
primary 37H20, 34E17
secondary 60H10

Keywords:
Singular perturbation
Fast–slow system
Invariant manifold
Dynamic bifurcation
Folded node
Canard
Mixed-mode oscillation
Random dynamical system
First-exit time
Concentration of sample paths

ABSTRACT

We consider the effect of Gaussian white noise on fast–slow dynamical systems with one fast and two slow variables, containing a folded node singularity. In the absence of noise, these systems are known to display mixed-mode oscillations, consisting of alternating large- and small-amplitude oscillations. We quantify the effect of noise and obtain critical noise intensities beyond which the small-amplitude oscillations become hidden by fluctuations. Furthermore we prove that the noise can cause sample paths to jump away from so-called canard solutions with high probability before deterministic orbits do. This early-jump mechanism can drastically influence the local and global dynamics of the system by changing the mixed-mode patterns.

1. Introduction

Our main focus of study are stochastic dynamical systems with multiple time scales. In particular, we are going to study a special bifurcation ("a folded node") in a three-dimensional fast–slow stochastic differential equation (SDE) with one fast and two slow variables. The detailed technical discussion including all relevant definitions and precise statements and proofs of our results starts in Section 2.

* Corresponding author. Current address: Vienna University of Technology, Institute for Analysis and Scientific Computing, 1040 Vienna, Austria.

[1] Supported by ANR project MANDy, Mathematical Analysis of Neuronal Dynamics, ANR-09-BLAN-0008-01.
[2] Supported by CRC 701, Spectral Structures and Topological Methods in Mathematics at the University of Bielefeld.

0022-0396/$ – see front matter
doi:10.1016/j.jde.2012.01.015

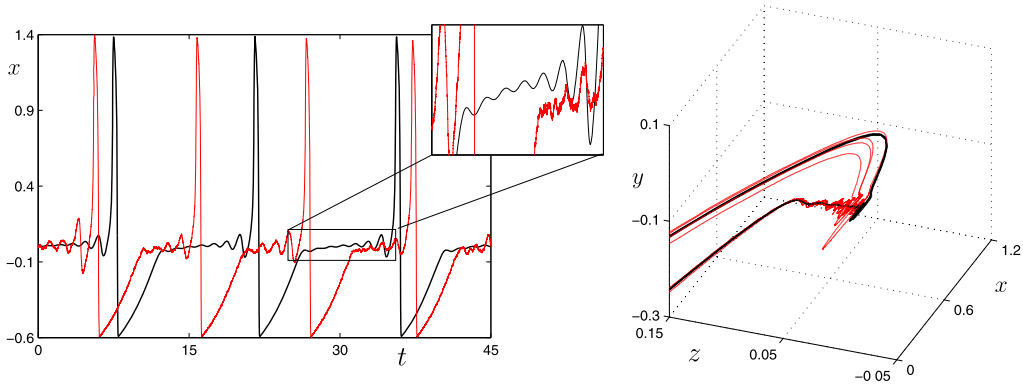

Fig. 1. Comparison of a deterministic (black) and stochastic (red) time trajectory. The left-half of the figure shows the time series exhibiting a 1^7 MMO; see also the zoom near the SAOs. The right part shows part of the trajectory in three-dimensional phase space and illustrates the early jumps of the stochastic sample path. (For interpretation of colors in this figure, the reader is referred to the web version of this article.)

In this section we want to outline our motivation and state our main results in a non-technical way. There are two main motivations for our work:

(M1) We want to develop an analogue to the intricate deterministic bifurcation theory for random dynamical systems by linking stochastic sample path techniques and the well understood deterministic theory.
(M2) A detailed analysis of noise effects in multi-scale stochastic systems is often crucial in applications; in particular, many biological systems have widely separated time scales and are influenced by various random effects.

We are going to describe our two main motivations in more detail, starting with (M2). Complex oscillatory patterns have been discovered in many different applications. Chemical systems [96,32,73] and neuronal dynamics [97,55,29] provide ample examples. Recent work has shown [33,65] that fast–slow systems can be used to model a wide variety of oscillatory patterns. A classification of local and global fast–slow "mechanisms" can be used to analyze each pattern. Mixed-mode oscillations (MMOs) alternate between small-amplitude oscillations (SAOs) and large-amplitude oscillations (LAOs). Fig. 1 shows a typical MMO time series where the deterministic time series has been generated by an MMO model proposed in [30]. The time series shows alternations between $L = 1$ LAOs and $s = 7$ SAOs which is denoted as the MMO pattern $L^s = 1^7$.

Although a deterministic model is able to explain a few experiments (see e.g. experimental results by Hudson et al. [63]) it fails to accurately model realistic MMOs due to the presence of noise (see e.g. the experiments by Dickson et al. [38]). Fig. 1 also shows a sample path which is a stochastic version of the deterministic orbit perturbed by Gaussian white noise. Two important observations that can be made about the stochastic MMO pattern are

- Part of the SAOs become indistinguishable from the random fluctuations so that counting SAOs below a certain amplitude is impossible.
- The stochastic sample path typically jumps before the deterministic solution makes an LAO.

In this work, we are going to provide rigorous formulations and the proofs of both observations.

Our motivation (M1) is to relate deterministic and stochastic methods for fast–slow systems. Within the last two decades substantial progress on deterministic fast–slow systems has been made. The analysis of hyperbolic fast dynamics has been completed in a series of works by Fenichel [44] (see also [68]) at the end of the 1970s; the theory focuses upon perturbations of normally hyperbolic critical manifolds to nearby slow manifolds. Near bifurcation points of the fast dynamics major

developments in the 1980s used nonstandard analysis [31,14] and asymptotic methods [43,8,9,89]. In the 1990s, two main geometric methods were introduced into multiple-time-scale systems. Switching between hyperbolic fast and slow dynamics was analyzed using the exchange lemma [67,107]. The geometry of non-hyperbolic or singular dynamics has been uncovered using the blow-up technique [42,41]. Since then the blow-up method has been applied successfully for many singular perturbation problems [76,77,74,75]. In particular, folded nodes [105,111,78,72] are a very interesting case since they occur already in generic fast–slow systems with one fast and two slow variables and have a highly nontrivial unfolding. Near a folded node the canard phenomenon occurs, i.e., orbits stay near a repelling manifold for a long time. This effect can generate MMOs [30,33]; see also the discussion for (M2) above and Fig. 1.

Noise acting on a system with multiple time scales can induce new phenomena such as early transitions [108,104,102,66,83], stochastic resonance [16,94,88,45,49,50] and bursting oscillations [103,84, 62]. The mathematical theory of fast–slow stochastic differential equations has been quickly developing in the last decade. Classical work on random perturbations of dynamical systems [48], which mainly focused on large-deviation aspects, allows to describe situations with a time scale separation exponentially large in the noise intensity [46,47,64,58,59]. A different approach, based on a detailed description of sample-path properties, applies to situations with time scale separation and noise intensity of comparable magnitude [22,23,20,19]. This method led to a general theory for the behavior of sample paths near normally hyperbolic invariant manifolds [24,25]. Other approaches include [69], which is based on moment estimates, and [99], which adopts the viewpoint of random dynamical systems in the sense of [6]. The associated methods and results have important applications in climate dynamics [16,94,106,21], the theory of critical transitions [98,80], classical and quantum atomic physics [3,4,2] and neuroscience [109,86,87,85,26,17]. In particular, canards in the stochastic FitzHugh–Nagumo system describing the action potential of neurons have been considered from the points of view of large deviations [37,40], and of convergence of sample paths [101]. Stochastic MMOs have also been considered in certain planar systems [91,92] and in coupled oscillators [113].

The theory of stochastic differential equations with higher-dimensional singularities and multiple slow variables is not yet as advanced. Here we make a first step towards bridging this gap between the generic higher-dimensional deterministic theory and stochastic sample paths analysis. The nontechnical statements of our two main results are:

(R1) Sample paths near a folded node stay inside a tubular neighborhood of an attracting deterministic solution. The neighborhood is explicitly defined by the covariance matrix of a linearized process. The relation between the noise level, the time scale separation and a system parameter determines when small oscillations near a folded node become indistinguishable from noisy fluctuations. This relation can be calculated explicitly to lowest asymptotic order.

(R2) Sample paths near a folded node typically escape from a repelling deterministic solution earlier than their deterministic counterparts. The escape time can be estimated with high probability and depends on the same parameters as the relation in (R1).

Both results have important implications from theoretical and applied perspectives. In particular, we can show how to control stochastic sample paths near a multi-dimensional bifurcation point. Therefore it is expected that the methods we develop have a much wider applicability beyond folded nodes, e.g. to singular Hopf bifurcations [53] or other stochastic bifurcation problems [25,5]. The precise quantitative estimates on the relation between noise level and a parameter controlling the number of SAOs are immediately useful in applications. Furthermore, the effect of early jumps could potentially regularize the complicated flow maps near a folded node [52,56] and simplify the local–global decomposition of return maps [81].

The paper is organized as follows. In Section 2 we review the necessary theory for deterministic fast–slow systems and fix the notation. In Section 3 we state the known results about folded nodes and explain why they produce small-amplitude oscillations. In Section 4 we consider a variational equation around a special canard solution, called the weak canard. The solution of the variational equation can be transformed into a “canonical form” which allows us to prove a result on the spacing of canard solutions up cross-sections near or including the folded node. The proof is postponed to

Appendix A. In Section 5 we develop the main setup for stochastic fast–slow SDEs and state a result on attracting normally hyperbolic slow manifolds away from bifurcation points. Several notations that we use throughout are introduced as well. In Section 6 we state our main results (R1)–(R2) for stochastic folded nodes. The result (R1) on covariance tubes for the linearized process is proved in Appendix B. The influence of nonlinear contributions is dealt with in Appendix C. The result (R2) on early jumps is proven in Appendix D. Section 7 develops numerical simulations to visualize the analytical results. We conclude by giving a summary of parameter regimes and discussing the influence of early jumps on the global return mechanism and LAOs in Section 8.

2. Fast–slow systems

We are only going to give a brief introduction to multiple-time-scale dynamics. A detailed reference covering many more topics is currently being written [82]; other excellent references are [90,51] for asymptotic methods and [44,7,68] for geometric methods. Many important discoveries were first made using nonstandard analysis [14,39]; in particular, many results we review in Section 3 were discovered by Benoît [15,11,12]. However, we are not going to use any nonstandard methods and focus on the geometric viewpoint.

A `fast-slow system` of ordinary differential equations (ODEs) is given by

$$\begin{aligned} \epsilon \frac{\mathrm{d}x}{\mathrm{d}s} &= \epsilon \dot{x} = f(x, y, \mu, \epsilon), \\ \frac{\mathrm{d}y}{\mathrm{d}s} &= \dot{y} = g(x, y, \mu, \epsilon), \end{aligned} \tag{2.1}$$

where $(x, y) \in \mathbb{R}^m \times \mathbb{R}^n$ are phase space coordinates, $\mu \in \mathbb{R}^p$ are parameters and $0 < \epsilon \ll 1$ represents the ratio of time scales. We shall assume that f, g are sufficiently smooth. By a rescaling we can change from the `slow time` s to the `fast time` $t = s/\epsilon$; this transforms (2.1) to

$$\begin{aligned} \frac{\mathrm{d}x}{\mathrm{d}t} &= x' = f(x, y, \mu, \epsilon), \\ \frac{\mathrm{d}y}{\mathrm{d}t} &= y' = \epsilon g(x, y, \mu, \epsilon). \end{aligned} \tag{2.2}$$

Remark. The more common notation for the slow time would be τ but we shall reserve τ for stopping times of stochastic processes; see Section 5.

The first step to analyze fast–slow systems is to consider the `singular limit` $\epsilon \to 0$. From (2.2) we obtain

$$\begin{aligned} x' &= f(x, y, \mu, 0), \\ y' &= 0, \end{aligned} \tag{2.3}$$

which is an ODE for the fast variables x where the slow variables y act as parameters. We call (2.3) the `fast subsystem` or `layer equations`; the associated flow is called the `fast flow`. Considering $\epsilon \to 0$ in (2.1) we find a differential-algebraic equation for the slow y-variables

$$\begin{aligned} 0 &= f(x, y, \mu, 0), \\ \dot{y} &= g(x, y, \mu, 0), \end{aligned} \tag{2.4}$$

called the `slow subsystem` or `reduced system`; the flow induced by (2.4) is called the `slow flow`. The slow subsystem is defined on the `critical manifold`

$$C_0 := \{(x, y) \in \mathbb{R}^{m+n} : f(x, y, \mu, 0) = 0\}.$$

Observe that C_0 can also be interpreted as a manifold of equilibria for the fast subsystem. Note also that C_0 does not have to be a manifold [75] but we only consider the manifold case in this paper. If the Jacobian matrix $(D_x f)(p)$ has maximal rank at $p \in C_0$ then the implicit function theorem describes C_0 locally as a graph

$$h_0 : \mathbb{R}^n \to \mathbb{R}^m, \qquad f\big(h_0(y), y, \mu, 0\big) = 0$$

near p. This allows us to write the slow subsystem more concisely as

$$\dot{y} = g\big(h_0(y), y, \mu, 0\big). \tag{2.5}$$

We can strengthen the assumption on $(D_x f)(p)$ and require that it is a hyperbolic matrix, i.e. $(D_x f)(p)$ has no eigenvalues with zero real part. In this case we say that C_0 is `normally hyperbolic` at p. If all the eigenvalues of $(D_x f)(p)$ have negative (positive) real parts we say that C_0 is `attracting (repelling)` with respect to the fast variables. The following theorem shows that normal hyperbolicity is the key regularity assumption for fast–slow systems.

Theorem 2.1 *(Fenichel's Theorem). (See [44].) Suppose M_0 is a compact normally hyperbolic submanifold (possibly with boundary) of the critical manifold C_0 and that $f, g \in C^r$, $1 \leqslant r < \infty$. Then for $\epsilon > 0$ sufficiently small the following holds*:

(F1) *There exists a locally invariant manifold M_ϵ diffeomorphic to M_0. Local invariance means that M_ϵ can have boundaries through which trajectories enter or leave.*
(F2) *M_ϵ has a Hausdorff distance of $\mathcal{O}(\epsilon)$ from M_0.*
(F3) *The flow on M_ϵ converges to the slow flow as $\epsilon \to 0$.*
(F4) *M_ϵ is C^r-smooth and can locally be given as a graph $h_\epsilon : \mathbb{R}^n \to \mathbb{R}^m$.*
(F5) *M_ϵ is normally hyperbolic with the same stability properties with respect to the fast variables as M_0.*
(F6) *M_ϵ is usually not unique. In regions that remain at a fixed distance from the boundary of M_ϵ, all manifolds satisfying* (F1)–(F5) *lie at a Hausdorff distance $\mathcal{O}(\mathrm{e}^{-K/\epsilon})$ from each other for some $K > 0$ with $K = \mathcal{O}(1)$.*

We call a perturbed manifold M_ϵ a `slow manifold`. Sometimes we refer to "the slow manifold" despite the non-uniqueness (F6) as it will be often irrelevant which of the $\mathcal{O}(\mathrm{e}^{-K/\epsilon})$-close manifolds we pick. A simple example where normal hyperbolicity fails is given by

$$\begin{aligned} \epsilon \dot{x} &= y - x^2, \\ \dot{y} &= \mu - x. \end{aligned} \tag{2.6}$$

The critical manifold $C_0 = \{(x, y) \in \mathbb{R}^2 : y = x^2\}$ splits into three parts $C = C_0^a \cup \{(0, 0)\} \cup C_0^r$ where

$$C_0^a = C_0 \cap \{x > 0\} \quad \text{and} \quad C_0^r = C_0 \cap \{x < 0\}.$$

C_0^a is attracting and C_0^r is repelling. At $(x, y) = (0, 0)$ the critical manifold is not normally hyperbolic and has a generic `fold singularity` [76]; observe that $(x, y) = (0, 0)$ is a fold (or saddle-node) bifurcation of the fast subsystem. Fig. 2 illustrates the dynamics near the fold point of (2.6).

To calculate the slow subsystem on C_0 we could consider the two graphs $x = h_0(y) = \pm\sqrt{y}$ as suggested by (2.5). For (2.6) it is more convenient to differentiate $y = x^2$ implicitly with respect to s. This gives

$$\dot{y} = 2x\dot{x} \quad \Rightarrow \quad \dot{x} = \frac{\mu - x}{2x},$$

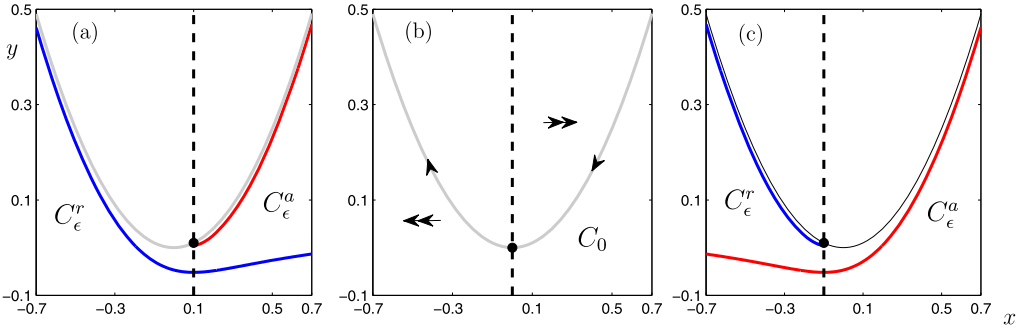

Fig. 2. Planar fold near a singular Hopf bifurcation of (2.6); for (a) and (c) we have fixed $\epsilon = 0.05$. (a) $\mu = 0.1$: The equilibrium point $(x_0, y_0) = (\mu, \mu^2)$ is determined as the intersection of C_0 (grey) and the nullcline $\{x = \mu\}$ (dashed black). The equilibrium is stable and the slow manifolds C_ϵ^a (red) and C_ϵ^r (blue) do not interact. (b) $\mu = 0$: Only the slow flow on C_0 (single arrow) and the fast flow (double arrows) are shown. C_0 coincides with a maximal singular canard. (c) $\mu = -0.1$: After the Hopf bifurcation the slow manifolds have "exchanged sides" suggesting an intersection for some μ near 0. (For interpretation of colors in this figure, the reader is referred to the web version of this article.)

which shows that the slow flow is undefined at $(0, 0)$ if $\mu \neq 0$. Fenichel's Theorem provides slow manifolds C_ϵ^a and C_ϵ^r. A major step in theory of fast–slow systems was to analyze the dynamics of (2.6) depending on the value of μ [42,8,77]. Note that for $\mu = 0$ the slow flow is well-defined and there is a special trajectory that passes from C_0^a to C_0^r and that a singular Hopf bifurcation [77,28] occurs for $\mu = 0$ and $0 < \epsilon \ll 1$. The slow manifolds can be extended under the flow into the fold point region. Comparing Fig. 2(a) to Fig. 2(c) we expect that there is a parameter value μ for which the slow manifolds intersect/coincide. This intersection marks what has become known as a canard explosion [77,79].

More generally, suppose that γ_ϵ is a trajectory of a fast–slow system (2.1)–(2.2) then we call γ_ϵ a `maximal canard` if it lies in the intersection of an attracting and a repelling slow manifold; for $\epsilon = 0$ we also refer to γ_0 as a `maximal singular canard`. Canards in planar fast–slow systems are of codimension one whereas for higher-dimensional systems we do not need an additional parameter. In the next section we are going to focus on canards in three dimensions.

3. Folded nodes

A general three-dimensional fast–slow system with one fast variable and two slow variables can be written as

$$\begin{aligned} \epsilon \dot{x} &= f(x, y, z, \mu, \epsilon), \\ \dot{y} &= g_1(x, y, z, \mu, \epsilon), \\ \dot{z} &= g_2(x, y, z, \mu, \epsilon). \end{aligned} \tag{3.1}$$

We assume that the critical manifold $C_0 = \{(x, y, z) \in \mathbb{R}^3 : \ f(x, y, z, \mu, 0) = 0\}$ of (3.1) is a folded surface near the origin; suitable non-degeneracy conditions [33,105] are

$$\begin{aligned} &f(0, 0, 0, \mu, 0) = 0, \qquad f_x(0, 0, 0, \mu, 0) = 0, \\ &f_y(0, 0, 0, \mu, 0) \neq 0, \qquad f_{xx}(0, 0, 0, \mu, 0) \neq 0, \end{aligned} \tag{3.2}$$

where subscripts denote partial derivatives.

The critical manifold again decomposes into three parts

$$C_0 = C_0^r \cup L \cup C_0^a,$$

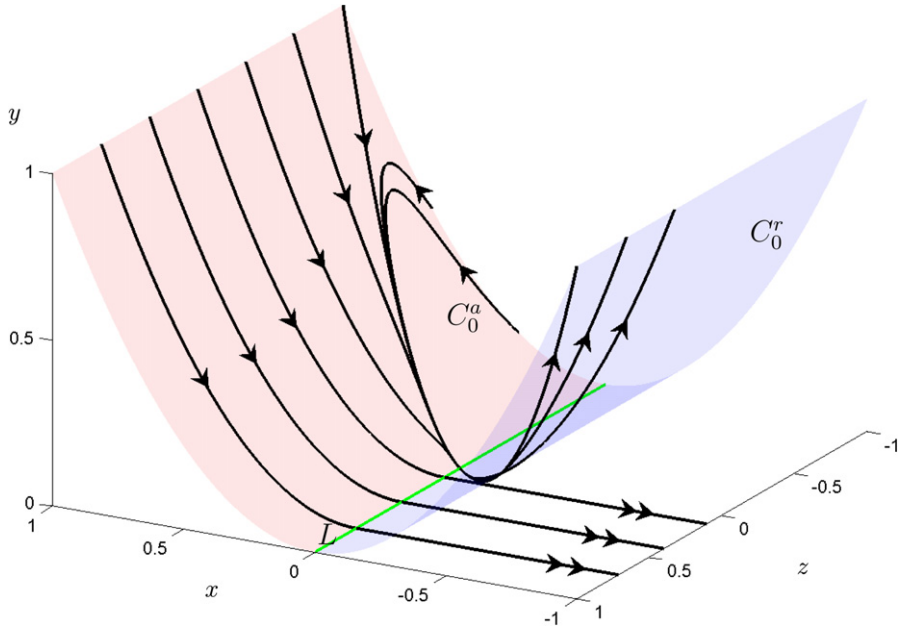

Fig. 3. Singular limit $\epsilon = 0$ for the normal form (3.8) with $\mu = 0.15$. The attracting manifold C_0^a (red), the fold line L (green) and the repelling manifold C_0^r (blue) partition the critical manifold. Trajectories of the slow (fast) subsystem are indicated by single (double) arrows. (For interpretation of colors in this figure, the reader is referred to the web version of this article.)

where $C_0^r = C_0 \cap \{f_x > 0\}$ is repelling, $C_0^a = C_0 \cap \{f_x < 0\}$ is attracting and $L = C_0 \cap \{f_x = 0\}$ is the curve of fold points; see Fig. 3. Note that the assumption $f_y(0, 0, 0, \mu, 0) \neq 0$ in (3.2) implies that the fold curve L can be locally parametrized by z. To obtain the slow subsystem we again differentiate $f(x, y, z, \mu, 0) = 0$ implicitly with respect to s

$$\dot{x} f_x + \dot{y} f_y + \dot{z} f_z = 0.$$

This implies that the slow subsystem is

$$\begin{aligned} f_x \dot{x} &= -f_y g_1 - f_z g_2, \\ \dot{z} &= g_2, \end{aligned} \tag{3.3}$$

where all functions are evaluated for $p = (x, y, z) \in C_0$ and $\epsilon = 0$. On L the ODE (3.3) is singular but we can rescale time $s \mapsto -s/f_x$ to obtain the `desingularized slow subsystem`

$$\begin{pmatrix} \dot{x} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} f_y g_1 + f_z g_2 \\ -f_x g_2 \end{pmatrix} \Big|_{p \in C_0}. \tag{3.4}$$

Note that the time rescaling has reversed the orientation of trajectories of (3.3) on C_0^r and that (3.4) is a well-defined planar ODE. We define

$$l(z) := (f_y g_1 + f_z g_2)|_{p \in L}$$

and make the assumptions that

$$l(0) = 0, \tag{3.5}$$

$$l(z) \neq 0, \quad \text{for } z \neq 0. \tag{3.6}$$

Observe that (3.5) and $f_x(0,0,0,\mu,0) = 0$ imply that $(x,z) = (0,0)$ is an equilibrium point for (3.4) that lies on the fold curve. We say that $(x,y,z) = (0,0,0)$ is a folded singularity. Points $(x,y,z \neq 0) \in L$ are called jump points as trajectories have to make a transition from the slow to the fast flow at these points; the condition (3.6) is also called the `normal switching condition`. Generic folded singularities can be classified according to their equilibrium type into `folded saddles`, `folded foci` and `folded nodes` [105,11]. Folded nodes are the most interesting folded singularities. Without loss of generality we may assume that the folded node is stable for (3.4) with associated eigenvalues λ_1, λ_2 for the linearization of (3.4).

Proposition 3.1. *(See [30,111].) Suppose that* (3.1) *satisfies* (3.2), (3.5), (3.6) *and that we are in the folded node scenario. Then there exist a smooth coordinate change and a smooth change of time which bring* (3.1) *near* $(x,y,z) = (0,0,0)$ *into the form*

$$\begin{aligned} \epsilon \dot{x} &= y - x^2 + \mathcal{O}\big(yx^2, x^3, xyz\big) + \epsilon \mathcal{O}(x,y,z,\epsilon), \\ \dot{y} &= -(\mu+1)x - z + \mathcal{O}\big(y, \epsilon, (x+z)^2\big), \\ \dot{z} &= \frac{\mu}{2}, \end{aligned} \tag{3.7}$$

with $\lambda_1 = -\mu$ *and* $\lambda_2 = -1$.

We shall show in Section 4 that the terms in (3.7) denoted by $\mathcal{O}(\cdot)$ are indeed higher order for the analysis near the folded node. Hence we can work with the normal form

$$\begin{aligned} \epsilon \dot{x} &= y - x^2, \\ \dot{y} &= -(\mu+1)x - z, \\ \dot{z} &= \frac{\mu}{2}. \end{aligned} \tag{3.8}$$

The critical manifold of (3.8) is

$$C_0 = \big\{(x,y,z) \in \mathbb{R}^3 \colon\ y = x^2\big\}.$$

We also write $C_0 = \{x = \pm\sqrt{y} =: h_0(y,z)\}$ if we parametrize C_0 over the slow variables. The critical manifold splits into three components

$$C_0 = C_0^a \cup L \cup C_0^r$$

where $C_0^a = C_0 \cap \{x > 0\}$ is attracting, $C_0^r = C_0 \cap \{x < 0\}$ is repelling and L is now a line of fold points.

The slow subsystem is

$$\begin{aligned} 2x\dot{x} &= -(\mu+1)x - z, \\ \dot{z} &= \frac{\mu}{2}. \end{aligned} \tag{3.9}$$

The desingularized slow subsystem is (see also Fig. 4(a))

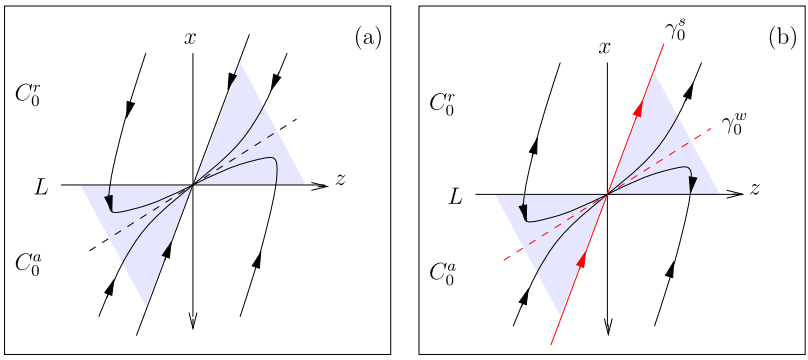

Fig. 4. (a) The desingularized slow flow (3.10) is sketched for some $\mu > 0$ with a stable node at the origin. (b) The slow flow (3.9) is illustrated. The strong eigendirection (solid red) defines the strong singular canard γ_0^s; the weak eigendirection (dashed red) defines the weak singular canard γ_0^w. (For interpretation of colors in this figure, the reader is referred to the web version of this article.)

$$\begin{aligned} \dot{x} &= -(\mu+1)x - z, \\ \dot{z} &= \mu x. \end{aligned} \tag{3.10}$$

The system (3.10) is linear with an equilibrium point at $(x, z) = (0, 0)$. The eigenvalues are $(\lambda_1, \lambda_2) := (-1, -\mu)$. We assume that

$$\mu \in (0, 1)$$

so that $(0, 0)$ is a stable node for the desingularized slow subsystem (3.10). Hence we also denote the eigenvalues as

$$\lambda_1 = -1 =: \lambda_s \quad \text{and} \quad \lambda_2 = -\mu =: \lambda_w$$

to emphasize the strong and weak eigendirections. Note that $\mu = \lambda_w/\lambda_s$ precisely represents the ratio of eigenvalues and attains all resonances $\mu^{-1} \in \mathbb{N}$ for $\mu \in (0, 1)$. The associated (unnormalized) eigenvectors are

$$\gamma_0^s = (-1/\mu, 1)^T \quad \text{and} \quad \gamma_0^w = (-1, 1)^T, \tag{3.11}$$

which also represent directions for two maximal singular canards; see Fig. 4. Observe that the `singular strong canard` γ_0^s and L bound a funnel region on C_0^a of trajectories that all flow into the folded node; see Fig. 4. The funnel region has an opening angle $\cos^{-1}(\mu/\sqrt{1+\mu^2})$ which converges to $\pi/2$ as $\mu \to 0$. The funnel on C_0^a is located in the $\{x > 0, z < 0\}$-quadrant and the `singular weak canard` γ_0^w is given by the anti-diagonal $\{z = -x\}$.

Theorem 3.2. *(See [105,11,15,111].) Suppose* (3.1) *has a generic folded node* (*i.e. Proposition* 3.1 *applies*). *Then for $\epsilon > 0$ sufficiently small the following holds*:

(C1) *The singular strong canard γ_0^s always perturbs to a maximal canard γ_ϵ^s. If $\mu^{-1} \notin \mathbb{N}$ then the singular weak canard γ_0^w also perturbs to a maximal canard γ_ϵ^w. We call γ_ϵ^s and γ_ϵ^w* `primary canards`.

(C2) *Suppose $k > 0$ is an integer such that*

$$2k + 1 < \mu^{-1} < 2k + 3 \quad \text{and} \quad \mu^{-1} \neq 2(k+1).$$

Then, in addition to $\gamma_\epsilon^{s,w}$, there are k other maximal canards, which we call `secondary canards`.

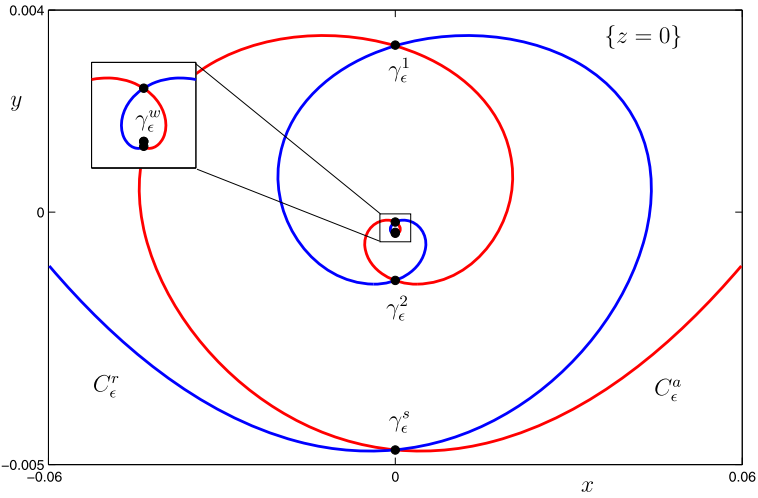

Fig. 5. Canards and slow manifolds near a folded node for (3.8) with $(\mu, \epsilon) = (0.08, 0.01)$ on the cross-section $\{z = 0\}$. The primary canard γ_ϵ^s and the first two secondary canards $\gamma_\epsilon^{1,2}$ are labeled. We also show a zoom near the primary weak canard γ_0^w. All maximal canards are indicated by black dots.

(C3) *The secondary canards converge to the strong primary canard as* $\epsilon \to 0$.
(C4) *The primary weak canard of a folded node undergoes a transcritical bifurcation for odd* $\mu^{-1} \in \mathbb{N}$ *and a pitchfork bifurcation for even* $\mu^{-1} \in \mathbb{N}$.

We emphasize the results (C1)–(C3) which will be of major importance for our stochastic analysis; (C4) describes the behavior near resonances and will not be considered here. The next theorem provides a geometric viewpoint for the generation of maximal canards near a folded node. We say that a `twist` corresponds to a half rotation (i.e. a rotation by angle π).

Theorem 3.3. *(See [105,111].) If* $2k + 1 < \mu^{-1} < 2k + 3$, *for some* $k \in \mathbb{N}$, *and* $\mu^{-1} \neq 2(k + 1)$ *then the following holds*:

(C5) *The primary strong canard* γ_0^s *twists once around the primary weak canard* γ_ϵ^w.
(C6) *The* j*-th secondary canard* γ_ϵ^j, $1 \leqslant j \leqslant k$, *twists* $2j + 1$ *times around the primary weak canard* γ_ϵ^w.
(C7) *The twisting/rotation occurs in an* $\mathcal{O}(\sqrt{\epsilon})$ *neighborhood of the folded node for* (3.1).

In particular, the slow manifolds C_ϵ^a and C_ϵ^r start to spiral near the folded node creating transversal intersections away from resonances. For visualizations of these manifolds in several different contexts see [34,33,35].

In Fig. 5 we show the slow manifolds for (3.8) near the folded node on the cross-section $\{z = 0\}$. The manifolds have been computed by forward integration and using the symmetry

$$(x, y, z, s) \mapsto (-x, y, -z, -s). \tag{3.12}$$

The center of rotation is the weak canard γ_0^w. Since $\mu = 0.08$ we know by using (C2) that there are five secondary canards in Fig. 5. Five intersections are indeed detected numerically but $\gamma_0^{4,5}$ are very close to γ_0^w on $\{z = 0\}$. All secondary canards γ_0^j approach γ_0^w when $z < 0$ near the folded node region on C_ϵ^a. The canards move away from each other for $z > 0$ (see Fig. 3). The next theorem shows that the maximal canards organize the rotational properties of trajectories passing through a folded node region.

Theorem 3.4. *(See [30].) Fix two sections*

$$\Sigma^1 := \{(x, y, z) \in \mathbb{R}^3 \colon y = K_1\} \quad \text{for some } 0 < K_1 = \mathcal{O}(1),$$
$$\Sigma^2 := \{(x, y, z) \in \mathbb{R}^3 \colon y = K_2\epsilon\} \quad \text{for some } 0 < K_2 = \mathcal{O}(1)$$

for the system (3.8)*. Consider the intersection points of maximal canards in* $\Sigma^1 \cap C_\epsilon^a$*. Let k be the number of secondary canards. Then for* $\epsilon > 0$ *sufficiently small the following holds*:

(C8) *The secondary canards are* $\mathcal{O}(\epsilon^{(1-\mu)/2})$ *close to the primary strong canard.*
(C9) *There exist* $(k+1)$ *so-called* `sectors of rotation` I_j, $1 \leqslant j \leqslant k+1$, *between the two primary canards labeled in increasing order starting from the strong primary canard. The size of the sectors* I_j *for* $1 \leqslant j \leqslant k$ *is* $\mathcal{O}(\epsilon^{(1-\mu)/2})$ *while the size of the sector* I_{k+1} *is* $\mathcal{O}(1)$.
(C10) *The Poincaré map from* Σ^1 *to* Σ^2 *is a contraction with rate* $\mathcal{O}(\epsilon^{(1-\mu)/(2\mu)})$.
(C11) *All maximal canards are separated by* $\mathcal{O}(\sqrt{\epsilon})$ *in their z-coordinate on* Σ^2.

Remark. We note that the results from Theorems 3.2, 3.3 and 3.4 also extend to higher-dimensional fast–slow systems with at least two slow variables and at least one fast variable [112,30] but that they do not provide a detailed analysis of canards beyond the section Σ^2.

Theorem 3.4 provides sectors of rotation that organize the twisting of trajectories near the folded node. Once we know which sector an orbit enters we can predict the number of oscillations. Note that the oscillations can be classified as "small oscillations" due to (C7). Global returns can induce so-called `mixed-mode oscillations (MMOs)` which are found in a wide variety of applications; see [33] for a review of MMO mechanisms in multiple-time-scale systems.

4. Canard spacing

Theorem 3.4 describes the spacing of maximal canards away from the folded node region. Since we are also interested in their spacing on the cross-section $\{z = 0\}$ depending on μ we need a refined analysis near the folded node. The key component in the proofs of Theorems 3.2, 3.3 and 3.4 is a rescaling of (3.8) near the folded node

$$(x, y, z, s) = (\sqrt{\epsilon}\bar{x}, \epsilon\bar{y}, \sqrt{\epsilon}\bar{z}, \sqrt{\epsilon}\bar{s}) \tag{4.1}$$

which can also be interpreted as a zoom-in (or "blow-up transformation") [105,111,30]. We shall not introduce the blow-up method here but restrict ourselves to the analysis of the rescaled system

$$\begin{aligned} \dot{\bar{x}} &= \bar{y} - \bar{x}^2 + \mathcal{O}(\sqrt{\epsilon}), \\ \dot{\bar{y}} &= -(\mu+1)\bar{x} - \bar{z} + \mathcal{O}(\sqrt{\epsilon}), \\ \dot{\bar{z}} &= \frac{\mu}{2}. \end{aligned} \tag{4.2}$$

Therefore the $\mathcal{O}(\cdot)$-terms in (3.7) are indeed of higher order for the analysis near the folded node. Neglecting the small ϵ-dependent terms and dropping the overbars in (4.2) for notational convenience yields

$$\begin{aligned} \dot{x} &= y - x^2, \\ \dot{y} &= -(\mu+1)x - z, \\ \dot{z} &= \frac{\mu}{2}. \end{aligned} \tag{4.3}$$

The ODEs (4.3) are our main focus of study in this section. Note that (C11) in Theorem 3.4 implies that the maximal canards are all $\mathcal{O}(1)$ finitely separated for (4.3) when they are an $\mathcal{O}(1)$ distance away from $(x, y, z) = (0, 0, 0)$. Observe that we can always solve the last equation

$$z(s) = \frac{\mu}{2}(s - s_0) + z_0$$

where $z_0 = z(s_0)$ denotes the initial z-coordinate of the trajectory at the initial time $s = s_0$. Hence we can view z as a time variable and re-write (4.3) if necessary as a planar non-autonomous ODE

$$\begin{aligned} \mu \frac{\mathrm{d}x}{\mathrm{d}z} &= 2y - 2x^2, \\ \mu \frac{\mathrm{d}y}{\mathrm{d}z} &= -2(\mu + 1)x - 2z. \end{aligned} \tag{4.4}$$

Our first goal is to quantify the intersections of canard solutions with the section $\{z = 0\}$. Therefore we are going to focus on the analysis of orbits arising as perturbations of slow subsystem trajectories inside the funnel on C_0^a and assume

$$z_0 \leqslant z \leqslant 0.$$

Let $(x^*(z), y^*(z))$ be any solution of (4.4) and set

$$u = (u_1, u_2) := \left(x - x^*,\ y - y^*\right)$$

to derive the `variational equation`

$$\begin{aligned} \mu \frac{\mathrm{d}u_1}{\mathrm{d}z} &= -4u_1 x^* + 2u_2 - 2u_1^2, \\ \mu \frac{\mathrm{d}u_2}{\mathrm{d}z} &= -2(\mu + 1)u_1. \end{aligned} \tag{4.5}$$

A key observation by Benoît was that there are some special solutions to (4.3).

Lemma 4.1. *(See [13].) The ODE* (4.3) *admits two polynomial solutions*

$$\left(x(s), y(s), z(s)\right) = \left(\frac{\lambda}{2}s, \frac{\lambda^2}{4}s^2 + \frac{\lambda}{2}, \frac{\mu}{2}s\right) \tag{4.6}$$

with $\lambda \in \{-\mu, -1\}$ *corresponding to the two primary singular canards.*

Lemma 4.1 can be checked by direct differentiation of (4.6). We know that the weak canard is the center of rotation. From (4.6) with $\lambda = -\mu$ we find the variational equation around the weak canard is given by

$$\mu \frac{\mathrm{d}u}{\mathrm{d}z} = \begin{pmatrix} 4z & 2 \\ -2(\mu + 1) & 0 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} + \begin{pmatrix} -2u_1^2 \\ 0 \end{pmatrix}. \tag{4.7}$$

We are interested in the detailed interaction of other maximal canards with the weak canard.

Proposition 4.2. *If* $\mu > 0$ *is sufficiently small and* $z < 0$ *is bounded away from* 0 *then solutions* $u = u(z)$ *of* (4.7) *are attracted exponentially fast to* $\{u_1 = 0 = u_2\}$.

Proof. Augmenting the variational equation (4.7) by $\dot{z} = 0$ gives an autonomous fast–slow system with two fast variables $u = (u_1, u_2)$ and one slow variable z (since μ is sufficiently small). The critical manifold is

$$\big\{(u_1, u_2, z) \in \mathbb{R}^3 \colon\ u_1 = 0,\ u_2 = 0\big\}.$$

Computing the linearization with respect to the fast variables gives a matrix

$$\begin{pmatrix} 4z & 2 \\ -2 & 0 \end{pmatrix}$$

with eigenvalues $2(z \pm \sqrt{z^2 - 1})$. Hence if $z < 0$ both eigenvalues have negative real parts and the result follows from Fenichel's Theorem. □

For $0 < \mu \ll 1$, Proposition 4.2 allows us to reduce the study of the nonlinear variational equation (4.7) to a linear one by dropping the higher-order term $-2u_1^2$. This yields the linear non-autonomous ODE

$$\mu \frac{\mathrm{d}u}{\mathrm{d}z} = \underbrace{\begin{pmatrix} 4z & 2 \\ -2(\mu+1) & 0 \end{pmatrix}}_{=:A(z)} u = A(z)u. \tag{4.8}$$

In particular, we must show what happens to solutions $u = u(u_1, u_2)$ near the weak canard and when z is not bounded away from 0.

Remark. The variational equation (4.8) has been analyzed [13,105] by re-writing it as a second-order equation

$$\frac{\mathrm{d}^2 u_1}{\mathrm{d}s^2} - \mu s \frac{\mathrm{d}u_1}{\mathrm{d}s} + u_1 = 0. \tag{4.9}$$

Benoît [13] observed that using the time rescaling $s = \tilde{s}/\sqrt{\mu}$ of (4.9) one gets

$$\frac{\mathrm{d}^2 u_1}{\mathrm{d}\tilde{s}^2} - \tilde{s} \frac{\mathrm{d}u_1}{\mathrm{d}\tilde{s}} + \frac{1}{\mu} u_1 = 0 \tag{4.10}$$

which is referred to as `Weber equation` or `Ricatti-Hermite equation`. The ODE (4.10) has explicit solutions in terms of Hermite polynomials (see [1, p. 781]). Then the asymptotic properties of Hermite polynomials can be used to draw conclusions about the existence of maximal canards.

We develop two alternative ways to describe the variational equation (4.8) directly. This provides new quantitative information about canard solutions and also gives estimates for the spacing of canards on cross-sections near the folded node. Our first approach in Section 4.1 uses averaging and provides a formal result. This result is then stated in Section 4.2 and proven in Appendix A using coordinate changes that are motivated by the formal calculation.

4.1. Averaging

As shown in (4.9) we can consider the variational equation (4.8) as the second-order equation

$$\ddot{u}_1 + u_1 = \mu s \dot{u}_1. \tag{4.11}$$

The form (4.11) suggests to view the problem as a damped oscillator; we assume that

$$-\frac{2}{\mu} < s_0 \leqslant s \leqslant 0. \tag{4.12}$$

Note that $\mu s = -2$ is precisely critical damping and for $\mu s \in (-2, 0)$ the oscillator is underdamped. Therefore we expect that (4.11) describes harmonic oscillations with damping/contraction that is non-uniform in time. We make a change to polar coordinates

$$(u_1, \dot{u}_1) = \big(r(s)\cos\big(s + \psi(s)\big), r(s)\sin\big(s + \psi(s)\big)\big)$$

which yields the ODEs

$$\begin{aligned} \dot{r} &= \mu s r \sin^2(s + \psi), \\ \dot{\psi} &= \mu s \cos(s + \psi)\sin(s + \psi). \end{aligned} \tag{4.13}$$

To simplify the analysis, consider the time rescaling $s = -\sqrt{-\tilde{s}}$. This converts (4.13) to

$$\begin{aligned} \frac{\mathrm{d}r}{\mathrm{d}\tilde{s}} &= -\frac{\mu r}{2}\sin^2(-\sqrt{-\tilde{s}} + \psi), \\ \frac{\mathrm{d}\psi}{\mathrm{d}\tilde{s}} &= -\frac{\mu}{2}\cos(-\sqrt{-\tilde{s}} + \psi)\sin(-\sqrt{-\tilde{s}} + \psi). \end{aligned} \tag{4.14}$$

We consider (4.14) on each time subinterval

$$I_j := \big[-\big((j+2)\pi\big)^2, -(j\pi)^2\big], \quad \text{for } j \in \{0, 2, 4, 6, \ldots\}$$

by viewing (4.14) as a vector field on $\mathbb{R}^+ \times (I_j/\sim)$ where the equivalence relation $\sim$ identifies the endpoints of I_j. Then the vector field is in the form for averaging [110]. The averaged equations are

$$\frac{\mathrm{d}r_j}{\mathrm{d}\tilde{s}} = -\frac{\mu r_j}{2|I_j|}\int_{I_j} \sin^2(-\sqrt{-\tilde{s}} + \psi_j)\,\mathrm{d}\tilde{s} = -\frac{(2j+2)\pi + \sin(2\psi_j)}{2(4j+4)\pi}\mu r_j, \tag{4.15}$$

$$\frac{\mathrm{d}\psi_j}{\mathrm{d}\tilde{s}} = -\frac{\mu}{2|I_j|}\int_{I_j} \cos(-\sqrt{-\tilde{s}} + \psi_j)\sin(-\sqrt{-\tilde{s}} + \psi_j)\,\mathrm{d}\tilde{s} = -\frac{\mu\pi\cos(2\psi_j)}{2(4j+4)\pi^2}. \tag{4.16}$$

In particular, we find that if we take a formal limit $j \to \infty$ the equation for the radius is

$$\frac{\mathrm{d}r_\infty}{\mathrm{d}\tilde{s}} = -\frac{\mu}{4}r_\infty. \tag{4.17}$$

Remark. Observe that one could also view (4.15)–(4.16) as an autonomous vector field and formally average over the angle ψ_j to get

$$\frac{\mathrm{d}r_j}{\mathrm{d}\tilde{s}} = -\frac{(j+1)}{(4j+4)}\mu r_j, \qquad \frac{\mathrm{d}\psi_j}{\mathrm{d}\tilde{s}} = 0$$

and then take the limit $j \to \infty$.

The solution of (4.17) is given by

$$r_\infty(\tilde{s}) = r_\infty(\tilde{s}_0)\mathrm{e}^{-\frac{1}{4}\mu(\tilde{s}-\tilde{s}_0)} = r_\infty(s_0)\mathrm{e}^{\frac{1}{4}\mu(s^2-s_0^2)} = r_\infty(z_0)\mathrm{e}^{(z^2-z_0^2)/\mu} \tag{4.18}$$

which shows that the leading-order behavior of the solutions to the variational equation (4.8) for $z < 0$ consists of a contraction towards the weak canard given by (4.18) combined with a rotation. If we assume that $\psi_0(0) = 0$, i.e. the rotation ends at angle 0 on section $\{z = 0\}$, then solving (4.16) yields

$$\psi_0(\tilde{s}) = -\tan^{-1}\left[\tanh\left[\frac{\tilde{s}\mu}{8(1+j)\pi}\right]\right].$$

In principle we can now calculate $\psi_0(-(2\pi)^2)$, use this result as an initial condition for ψ_1 and then repeat the process to get a very detailed description of the rotational properties of trajectories near a folded node.

4.2. Diagonalization

For the proof of our formal calculation we consider the variational equation in first-order form (4.8). The matrix $A(z)$ has eigenvalues

$$2z \pm 2\mathrm{i}\omega(z), \quad \text{where } \omega(z) = \sqrt{1 - z^2 + \mu}.$$

We assume that $|z| < 1$ so that $\omega(z)$ is real and bounded away from 0. Furthermore $A(z)$ has trace $4z$. We expect that the solution $u(z)$ for (4.8) consists of a contraction and a rotation for

$$1 < z_0 \leqslant z < 0. \tag{4.19}$$

Observe that (4.19) corresponds to the condition (4.12).

Theorem 4.3 *(Canonical form). There exists a matrix*

$$S(z) = \frac{1}{\sqrt{\omega(z)}}\begin{pmatrix} -z+\omega(z) & -z-\omega(z) \\ 1 & 1 \end{pmatrix} + \mathcal{O}(\mu) \tag{4.20}$$

such that the coordinate change $u(z) = S(z)\tilde{u}(z)$ transforms the variational equation (4.8) *into canonical form*

$$\mu\frac{\mathrm{d}\tilde{u}}{\mathrm{d}z} = \begin{pmatrix} a(z) & \varpi(z) \\ -\varpi(z) & a(z) \end{pmatrix}\tilde{u}, \tag{4.21}$$

where

$$\begin{aligned} a(z) &= 2z + \mathcal{O}\big(\mu^2\big), \\ \varpi(z) &= 2\omega(z) + \mathcal{O}(\mu). \end{aligned} \tag{4.22}$$

As a consequence, we can write the solution of (4.8) *in the form*

$$u(z) = \mathrm{e}^{\alpha(z,z_0)/\mu}S(z)U(z,z_0)S(z_0)^{-1}u(z_0) \tag{4.23}$$

for $|z| < 1$ and $|z_0| < 1$, where

$$\alpha(z, z_0) = \int_{z_0}^{z} a(s)\,\mathrm{d}s = z^2 - z_0^2 + \mathcal{O}\big(\mu^2\big) \tag{4.24}$$

and $U(z, z_0)$ is the orthogonal matrix

$$U(z, z_0) = \begin{pmatrix} \cos(\varphi(z,s)/\mu) & \sin(\varphi(z,s)/\mu) \\ -\sin(\varphi(z,s)/\mu) & \cos(\varphi(z,s)/\mu) \end{pmatrix}, \qquad \varphi(z, z_0) = \int_{z_0}^{z} \varpi(s)\,\mathrm{d}s + \mathcal{O}(\mu). \tag{4.25}$$

The proof is given in Appendix A. We remark that Theorem 4.3 easily extends to linearization around solutions other than the weak canard.

4.3. The distance estimate

From the solution of the variational equation in Theorem 4.3 we can give asymptotic estimates on the distance of the secondary canards to the weak canard on the section

$$\Sigma^0 := \big\{(x, y, z) \in \mathbb{R}^3 \colon\ z = 0\big\}.$$

Theorem 4.4 *(Canard spacing). The distance of the k-th secondary canard γ^k to γ^w on Σ^0 is given by $\mathcal{O}(\mathrm{e}^{-c(2k+1)^2\mu})$ as $\mu \to 0$ where $0 < c = \mathcal{O}(1)$ is a constant.*

Proof. By Theorem 3.3, part (C5), the k-th secondary canard makes $(2k+1)/2$ twists around the weak canard between $z_0 \leqslant z \leqslant 0$. Using this fact together with the rotational part (4.25) of the solution (4.23) in Theorem 4.3 we find the condition

$$\frac{\varphi(0, z_0)}{\mu} = \frac{1}{\mu}\int_{z_0}^{0} \varpi(s)\,\mathrm{d}s \stackrel{!}{=} \pi\,\frac{2k+1}{2} \tag{4.26}$$

to leading order. Using concavity of $z \mapsto \omega(z)$ for $z \in (-1, 0)$ a direct approximation of the integral in (4.26) is given by

$$-\frac{\pi}{4} z_0(1+\mu) + \mathcal{O}\big(\mu^2\big) \leqslant \int_{z_0}^{0} \omega(s)\,\mathrm{d}s \leqslant -z_0\left(1 + \frac{\mu}{2}\right) + \mathcal{O}\big(\mu^2\big). \tag{4.27}$$

Hence as $\mu \to 0$ we get that

$$z_0 \in \left[-(2k+1)\mu, -\frac{\pi}{4}(2k+1)\mu\right] =: \big[z_0^{(1)}, z_0^{(2)}\big] \tag{4.28}$$

where $z_0^{(1)}$ is the start point estimate for maximal canards that start to spiral around γ^w near $z \gtrapprox -1$ and $z_0^{(2)}$ the estimate for maximal canards that start to spiral around γ^w near $z \lessapprox 0$. In particular, we find that $z_0 = \mathcal{O}((2k+1)\mu)$. Note that the contraction term towards γ^w in (4.23) for $-1 < z_0 \leqslant z < 0$ has order

$$\mathcal{O}\big(\mathrm{e}^{(z^2 - z_0^2)/\mu}\big). \tag{4.29}$$

The result follows upon evaluating (4.29) on Σ^0 and by substituting (4.28). □

5. Stochastic fast–slow systems

As for deterministic fast–slow systems we only give a brief introduction to stochastic fast–slow systems. For a detailed introduction to stochastic multiple-time-scale dynamics consider [25,82]; we are also going to assume standard result on stochastic differential equations (SDEs) [95,70]. All SDEs and stochastic integrals in this paper are considered in the Itô interpretation.

We associate to the deterministic fast–slow system (2.1) a `stochastic fast-slow system` given by

$$\begin{aligned} \mathrm{d}x_s &= \frac{1}{\epsilon} f(x, y, \mu, \epsilon)\,\mathrm{d}s + \frac{\sigma}{\sqrt{\epsilon}} F(x_s, y_s)\,\mathrm{d}W_s, \\ \mathrm{d}y_s &= g(x, y, \mu, \epsilon)\,\mathrm{d}s + \sigma' G(x_s, y_s)\,\mathrm{d}W_s, \end{aligned} \tag{5.1}$$

where $F : \mathbb{R}^{m+n} \to \mathbb{R}^{m\times k}$, $G : \mathbb{R}^{m+n} \to \mathbb{R}^{n\times k}$, $\{W_s\}_{s\geqslant 0}$ is a standard k-dimensional Brownian motion on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and $\sigma, \sigma' > 0$ are parameters controlling the `noise level`

$$\sqrt{\sigma^2 + \left(\sigma'\right)^2}.$$

We also define $\rho := \sigma'/\sigma$. The initial conditions $(x_0, y_0) := (x_{s_0}, y_{s_0})$ are chosen to be either deterministic or independent of $\{W_s\}_{s\geqslant 0}$. Furthermore we assume that f, g, F, G are sufficiently smooth so that solutions (x_s, y_s) to the SDE (5.1) exist which are continuous and unique. The law of the process (x_s, y_s) is denoted by $\mathbb{P}^{s_0,(x_0,y_0)}$ and the corresponding expectation by $\mathbb{E}^{s_0,(x_0,y_0)}$. Our approach to understand the dynamics of (5.1) is to analyze the time a sample path spends in a given Borel-measurable set $\mathcal{A} \subset \mathbb{R}^{m+n}$. Suppose $(x_0, y_0) \in \mathcal{A}$ and define the `first-exit time` from $\mathcal{A}$ as

$$\tau_{\mathcal{A}} := \inf\bigl\{s \in [s_0, \infty)\colon (x_s, y_s) \notin \mathcal{A}\bigr\}.$$

In this paper we are always going to choose sets $\mathcal{A}$ so that $\tau_{\mathcal{A}}$ is a stopping time with respect to the filtration generated by $\{(x_s, y_s)\}_{s\geqslant s_0}$.

Our first goal is to state an analog of Fenichel's Theorem. This theorem is going to describe the approach of sample paths near an attracting critical manifold bounded away from folded singularities. Suppose the deterministic version of (5.1) with $\sigma = 0 = \sigma'$ has a compact attracting normally hyperbolic critical manifold

$$C_0 = \bigl\{(x, y) \in \mathbb{R}^{m+n}\colon x = h_0(y),\ y \in \mathcal{D}_0\bigr\}.$$

Let C_ϵ be the slow manifold obtained from Fenichel's Theorem 2.1. Our strategy is to construct a neighborhood $\mathcal{B}(r)$ for C_ϵ that contains sample paths with high probability [24,25]. Define

$$\xi_s = x_s - h_\epsilon(y_s). \tag{5.2}$$

Observe that ξ_s measures the deviation of the fast components from C_ϵ. Applying Itô's formula to (5.2) gives:

$$\begin{aligned} \mathrm{d}\xi_s &= \mathrm{d}x_s - D_y h_\epsilon(y_s)\,\mathrm{d}y_s + O\bigl(\left(\sigma'\right)^2\bigr)\,\mathrm{d}s \\ &= \frac{1}{\epsilon}\bigl[f\bigl(h_\epsilon(y_s) + \xi_s, y_s, \mu, \epsilon\bigr) - \epsilon D_y h_\epsilon(y_s) g\bigl(h_\epsilon(y_s) + \xi_s, y_s, \mu, \epsilon\bigr) + \mathcal{O}\bigl(\epsilon\left(\sigma'\right)^2\bigr)\bigr]\,\mathrm{d}s \\ &\quad + \frac{\sigma}{\sqrt{\epsilon}}\bigl[F\bigl(h_\epsilon(y_s) + \xi_s, y_s\bigr) - \rho\sqrt{\epsilon} D_y h_\epsilon(y_s) G\bigl(h_\epsilon(y_s) + \xi_s, y_s\bigr)\bigr]\,\mathrm{d}W_s. \end{aligned} \tag{5.3}$$

From now on, we suppress the arguments μ and ϵ for brevity. Consider the linear approximation of (5.2) in ξ_s, neglect the Itô term $\mathcal{O}(\epsilon(\sigma')^2)$ and replace y_s by its deterministic version y_s^{det} to obtain

$$\begin{aligned} \mathrm{d}\xi_s^0 &= \frac{1}{\epsilon} A_\epsilon\big(y_s^{\mathrm{det}}\big)\xi_s^0\,\mathrm{d}s + \frac{\sigma}{\sqrt{\epsilon}} F_\epsilon^0\big(y_s^{\mathrm{det}}\big)\,\mathrm{d}W_s, \\ \mathrm{d}y_s^{\mathrm{det}} &= g\big(h_\epsilon\big(y_s^{\mathrm{det}}\big), y_s^{\mathrm{det}}\big)\,\mathrm{d}s, \end{aligned} \tag{5.4}$$

where the two matrices A_ϵ and F_ϵ^0 are defined as

$$\begin{aligned} A_\epsilon(y) &= D_x f\big(h_\epsilon(y), y\big) - \epsilon D_y h_\epsilon(y) D_x g\big(h_\epsilon(y), y\big), \\ F_\epsilon^0(y) &= F\big(h_\epsilon(y), y\big) - \rho\sqrt{\epsilon} D_y h_\epsilon(y) G\big(h_\epsilon(y), y\big). \end{aligned}$$

Observe that $A_0(y) = D_x f(h_0(y), y)$ and $F_0^0(y) = F(h_0(y), y)$. To solve (5.4) we pick an initial condition on the slow manifold $(\xi_0^0, y_0^{\mathrm{det}}) = (0, y_0^{\mathrm{det}})$. Then the solution of (5.4) is the Itô integral

$$\xi_s^0 = \frac{\sigma}{\sqrt{\epsilon}} \int_0^s U(s,r) F_\epsilon^0\big(y_r^{\mathrm{det}}\big)\,\mathrm{d}W_r$$

where $U(s,r)$ denotes the principal solution of the homogeneous linear system $\epsilon\dot{\nu} = A_\epsilon(y_s^{\mathrm{det}})\nu$. If we fix a time s then ξ_s^0 is a Gaussian random variable of mean zero and covariance matrix

$$\mathrm{Cov}\big(\xi_s^0\big) = \frac{\sigma^2}{\epsilon} \int_0^s U(s,r) F_\epsilon^0\big(y_s^{\mathrm{det}}\big) F_\epsilon^0\big(y_s^{\mathrm{det}}\big)^T U(s,r)^T\,\mathrm{d}r.$$

Note carefully that $X_s := \sigma^{-2}\,\mathrm{Cov}(\xi_s^0)$ does satisfy a fast–slow ODE given by

$$\begin{aligned} \epsilon\dot{X} &= A_\epsilon(y)X + XA_\epsilon(y)^T + F_\epsilon^0(y)F_\epsilon^0(y)^T, \\ \dot{y} &= g\big(h_\epsilon(y), y\big). \end{aligned} \tag{5.5}$$

The system (5.5) has a critical manifold S_0^ξ given by solving the equation:

$$A_0(y)X + XA_0(y)^T + F_0^0(y)F_0^0(y)^T = 0.$$

By the remarks above we see that this is equivalent to solving:

$$(D_x f)\big(h_0(y), y\big)X + X(D_x f)\big(h_0(y), y\big)^T + F\big(h_0(y), y\big)F\big(h_0(y), y\big)^T = 0. \tag{5.6}$$

Again this manifold can be locally described as a graph

$$S_0^\xi = \big\{(X, y) \in \mathbb{R}^{m+n} \colon X = H_0(y)\big\}.$$

The next lemma shows that S_0^ξ is normally hyperbolic and attracting.

Lemma 5.1. *(See [10].) Let M_1 and M_2 be square matrices of dimensions m with eigenvalues $\lambda_{1,1}, \dots, \lambda_{1,m}$ and $\lambda_{2,1}, \dots, \lambda_{2,m}$ respectively. Then the linear map $L : \mathbb{R}^{m\times m} \to \mathbb{R}^{m\times m}$ defined by*

$$L(X) = M_1 X + X M_2$$

has m^2 eigenvalues given by $\{\lambda_{1,i} + \lambda_{2,j}\}$ for $i, j \in \{1, 2, \dots, m\}$.

Therefore Fenichel's Theorem 2.1 provides us with a slow manifold

$$S_\epsilon^\xi = \big\{(X, y) \in D \subset \mathbb{R}^{m+n} \colon\ X = H_\epsilon(y) = H_0(y) + \mathcal{O}(\epsilon)\big\}.$$

Theorem 5.2. *(See [24,25].) Suppose the norms $\|H_\epsilon(y)\|$ and $\|H_\epsilon^{-1}(y)\|$ are uniformly bounded. Define the neighborhood $\mathcal{B}(r)$ around the deterministic slow manifold as*

$$\mathcal{B}(r) := \big\{(x, y) \in D \colon\ \big|\big[x - h_\epsilon(y)\big] \cdot \big[H_\epsilon^{-1}(y)\big(x - h_\epsilon(y)\big)\big]\big| < r^2\big\}.$$

Then, for $\epsilon > 0$ sufficiently small, sample paths starting on C_ϵ remain in $\mathcal{B}(r)$ with high probability; in particular, we have

$$\mathbb{P}^{s_0,(x_0,y_0)}\{\tau_{\mathcal{B}(r)} < s \wedge \tau_{\mathcal{D}_0}\} < K_1(s,\epsilon)\mathrm{e}^{-K_2 r^2/2\sigma^2} \tag{5.7}$$

with $K_{1,2} > 0$.

Detailed discussions of the factors $K_1(s,\epsilon)$ and K_2 can be found in [24,25]. In particular, $K_1(s,\epsilon)$ grows at most like s^2, while K_2 does not depend on time and can be taken close to 1. The probability in (5.7) thus remains small on long time spans as soon as we choose $r \gg \sigma$. This implies that sample paths stay for exponentially long times near an attracting slow manifold before they jump away or come close to the boundary of C_0, i.e. the y-coordinates leave the set $\mathcal{D}_0$.

6. Stochastic folded nodes

6.1. Zoom-in

Proposition 3.1 shows that Eq. (3.8) is a normal form for deterministic fast–slow systems with a folded node. We study the associated SDE

$$\begin{aligned} \mathrm{d}x_s &= \frac{1}{\epsilon}\big(y_s - x_s^2\big)\,\mathrm{d}s + \frac{\sigma}{\sqrt{\epsilon}}\,\mathrm{d}W_s^{(1)},\\ \mathrm{d}y_s &= \big[-(\mu+1)x_s - z_s\big]\,\mathrm{d}s + \sigma'\,\mathrm{d}W_s^{(2)},\\ \mathrm{d}z_s &= \frac{\mu}{2}\mathrm{d}s, \end{aligned} \tag{6.1}$$

where $W_s^{(1)}$, $W_s^{(2)}$ are independent standard Brownian motions; to simplify the notation we also define $W_s := (W_s^{(1)}, W_s^{(2)})^T$. Since z plays the role of a time variable we do not add noise to the z-component. We will always assume that the noise terms are of equal order, i.e., that $\rho = \sigma'/\sigma$ is bounded above and below by positive constants. Note that (6.1) fits into the framework of a general fast–slow SDE (5.1) with

$$F(x, y) = (\,1 \quad 0\,) \quad \text{and} \quad G(x, y) = (\,0 \quad 1\,).$$

We apply the zoom-in (or blow-up transformation) given by (4.1) to (6.1) to get

$$\epsilon^{1/2}\,\mathrm{d}\bar{x}_{\bar{s}} = \epsilon^{1/2}\bigl(\bar{y}_{\bar{s}} - \bar{x}_{\bar{s}}^2\bigr)\,\mathrm{d}\bar{s} + \frac{\sigma}{\sqrt{\epsilon}}\,\mathrm{d}W^{(1)}_{\epsilon^{1/2}\bar{s}},$$
$$\epsilon\,\mathrm{d}\bar{y}_{\bar{s}} = \epsilon\bigl[-(\mu+1)\bar{x}_{\bar{s}} - \bar{z}_{\bar{s}}\bigr]\,\mathrm{d}\bar{s} + \sigma'\,\mathrm{d}W^{(2)}_{\epsilon^{1/2}\bar{s}},$$
$$\mathrm{d}\bar{z}_{\bar{s}} = \frac{\mu}{2}\,\mathrm{d}\bar{s}.$$

We use the scaling law of Brownian motion, divide the first equation by $\sqrt{\epsilon}$ and the second one by ϵ and drop the overbars for notational convenience

$$\mathrm{d}x_s = \bigl(y_s - x_s^2\bigr)\,\mathrm{d}s + \frac{\sigma}{\epsilon^{3/4}}\,\mathrm{d}W_s^{(1)},$$
$$\mathrm{d}y_s = \bigl[-(\mu+1)x_s - z_s\bigr]\,\mathrm{d}s + \frac{\sigma'}{\epsilon^{3/4}}\,\mathrm{d}W_s^{(2)},$$
$$\mathrm{d}z_s = \frac{\mu}{2}\,\mathrm{d}s. \tag{6.2}$$

Therefore rescaling the noise intensities as

$$\bigl(\sigma, \sigma'\bigr) = \bigl(\epsilon^{3/4}\bar{\sigma}, \epsilon^{3/4}\bar{\sigma}'\bigr) \tag{6.3}$$

removes ϵ from the equations and yields

$$\mathrm{d}x_s = \bigl(y_s - x_s^2\bigr)\,\mathrm{d}s + \sigma\,\mathrm{d}W_s^{(1)},$$
$$\mathrm{d}y_s = \bigl[-(\mu+1)x_s - z_s\bigr]\,\mathrm{d}s + \sigma'\,\mathrm{d}W_s^{(2)},$$
$$\mathrm{d}z_s = \frac{\mu}{2}\,\mathrm{d}s, \tag{6.4}$$

where the overbars from (6.3) have again been dropped. To study (6.4) we consider the variational equation around a deterministic solution (which may be the weak primary canard, a secondary canard, or even any other solution with initial condition (x_0, y_0, z_0) sufficiently close to C_0^a). This approach is a generalization of Section 5 where we considered the variation around the slow manifold. Viewing (6.2) as a planar non-autonomous system with time $z = (\mu/2)s$ and setting

$$(x_z, y_z) = \bigl(x_z^{\mathrm{det}} + \xi_z, y_z^{\mathrm{det}} + \eta_z\bigr)$$

we get the SDE

$$\mathrm{d}\xi_z = \frac{2}{\mu}\bigl(\eta_z - \xi_z^2 - 2x_z^{\mathrm{det}}\xi_z\bigr)\,\mathrm{d}z + \sigma\sqrt{\frac{2}{\mu}}\,\mathrm{d}W_z^{(1)},$$
$$\mathrm{d}\eta_z = -\frac{2}{\mu}(\mu+1)\xi_z\,\mathrm{d}z + \sigma'\sqrt{\frac{2}{\mu}}\,\mathrm{d}W_z^{(2)}. \tag{6.5}$$

6.2. Covariance tubes

We linearize (6.5) and denote the solution by $\zeta_z^0 = (\xi_z^0, \eta_z^0)^T$ which satisfies the SDE

$$\mathrm{d}\zeta_z^0 = \frac{1}{\mu} \underbrace{\begin{pmatrix} -4x_z^{\mathrm{det}} & 2 \\ -2(\mu+1) & 0 \end{pmatrix}}_{=:A(x_z^{\mathrm{det}})} \zeta_z^0 \,\mathrm{d}z + \frac{\sigma}{\sqrt{\mu}} \underbrace{\begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2}\rho \end{pmatrix}}_{=:F^0} \mathrm{d}W_z. \tag{6.6}$$

Then (6.6) is solved by the following Gaussian process

$$\zeta_z^0 = U(z, z_0)\zeta_{z_0}^0 + \frac{\sigma}{\sqrt{\mu}} \int_{z_0}^{z} U(z, r) F^0 \,\mathrm{d}W_r$$

where $U(z, r)$ is the principal solution to the deterministic homogeneous non-autonomous linear system $\mu\dot{\zeta} = A(x_z^{\mathrm{det}})\zeta$. The two-by-two covariance matrix $\mathrm{Cov}(\zeta_z^0) =: \mathrm{Cov}(z)$ is given by

$$\mathrm{Cov}(z) = \frac{\sigma^2}{\mu} \int_{z_0}^{z} U(z, r)\big(F^0\big)\big(F^0\big)^T U(z, r)^T \,\mathrm{d}r. \tag{6.7}$$

Differentiating $V(z) := \sigma^{-2}\,\mathrm{Cov}(z)$ we find that it satisfies the ODE

$$\mu \frac{\mathrm{d}V}{\mathrm{d}z} = A\big(x_z^{\mathrm{det}}\big)V + V A\big(x_z^{\mathrm{det}}\big)^T + \big(F^0\big)\big(F^0\big)^T \tag{6.8}$$

with initial condition $V(z_0) = 0$. The following result describes the behavior of the solutions of (6.8), in particular as z_s approaches 0.

Theorem 6.1 *(Behavior of the covariance matrix). Fix an initial time $z_0 < 0$. There exist constants $c_+ > c_- > 0$ such that the solution of* (6.8) *with initial condition $V(z_0) = 0$ satisfies*

$$\frac{c_-}{|z|} \leqslant V_{11}(z), V_{22}(z) \leqslant \frac{c_+}{|z|} \qquad \text{for } z_0 + \mathcal{O}\big(\mu|\log\mu|\big) \leqslant z \leqslant -\sqrt{\mu}, \tag{6.9}$$

$$\frac{c_-}{\sqrt{\mu}} \leqslant V_{11}(z), V_{22}(z) \leqslant \frac{c_-}{\sqrt{\mu}} \quad \text{for } -\sqrt{\mu} \leqslant z \leqslant \sqrt{\mu}, \tag{6.10}$$

and

$$\begin{aligned} \big|V_{12}(z)\big| = \big|V_{21}(z)\big| &\leqslant c_+ \quad \text{for } z_0 \leqslant z \leqslant \sqrt{\mu}, \\ \big|V_{22}(z) - V_{11}(z)\big| &\leqslant c_+ \quad \text{for } z_0 \leqslant z \leqslant \sqrt{\mu}. \end{aligned} \tag{6.11}$$

Furthermore, let $\bar{V}(z)$ be any solution of (6.8) *with positive definite initial condition $\bar{V}(z_0)$. More precisely, we require both $\bar{V}(z_0)$ and $\bar{V}(z_0)^{-1}$ to have all elements uniformly bounded in μ. Then its matrix elements satisfy* (6.10) *and* (6.11)*, and* (6.9) *holds for* all $z \in [z_0, -\sqrt{\mu}\,]$.

The proof is given in Appendix B, where we also give some additional information on how the covariance can be approximated by asymptotic expansions.

Eqs. (6.9) and (6.10) show that the variances of ξ_z^0 and η_z^0 grow like $\sigma^2/|z|$ up to time $-\sqrt{\mu}$, and then stay of order $\sigma^2/\sqrt{\mu}$ up to time $\sqrt{\mu}$. Thus we expect the fluctuations of stochastic sample paths

around the deterministic solution to increase when the fold at $z = 0$ is approached. The restriction $z \geqslant z_0 + \mathcal{O}(\mu|\log\mu|)$ is due to the fact that the variances are initially equal to zero, and need some time to build up.

Eqs. (6.11) show that the covariance of ξ_z^0 and η_z^0 remains bounded, of order σ^2, up to time $\sqrt{\mu}$, and the same holds true for the difference between the variances. This implies that fluctuations become more isotropic as z approaches 0.

We now turn to the analysis of the full nonlinear SDE (6.5) satisfied by the difference $\zeta_z = (\xi_z, \eta_z)$ between stochastic sample paths and deterministic solutions. This SDE can be written in vectorial form as

$$\mathrm{d}\zeta_z = \frac{1}{\mu}\big[A\big(x_z^{\mathrm{det}}\big)\zeta_z + b(\zeta_z)\big]\,\mathrm{d}z + \frac{\sigma}{\sqrt{\mu}}F^0\,\mathrm{d}W_z, \tag{6.12}$$

where A and F^0 have been defined in (6.6), and $b(\zeta)^T = (-\xi^2, 0)$ denotes the nonlinear term. We expect the covariance matrix of ζ_z to be close to $\mathrm{Cov}(\zeta_z^0) = \sigma^2 V(z)$, where $V(z)$ is the solution of (6.8) with initial condition $V(z_0) = 0$. Thus sample paths should be concentrated in a tube surrounding the deterministic solution, with elliptical cross-section determined by $V(z)$. Note that the elliptical cross-section becomes close to circular as z approaches 0, since the variances V_{11}, V_{22} are then of larger order than the covariance V_{12} and the difference $V_{22} - V_{11}$.

The fact that $V(z_0)$ is not invertible causes some technical complications. Therefore, in the following we let $\bar{V}(z)$ be the solution of (6.8) with an initial condition $\bar{V}(z_0)$ which is positive definite. Observe that the difference between $\bar{V}(z)$ and $V(z)$ decreases exponentially fast. We fix a $z_0 < 0$ and define the covariance tube as

$$\mathcal{B}(r) = \big\{(x, y, z)\colon z_0 \leqslant z \leqslant \sqrt{\mu},\ \big\langle (x, y) - \big(x_z^{\mathrm{det}}, y_z^{\mathrm{det}}\big), \bar{V}(z)^{-1}\big[(x, y) - \big(x_z^{\mathrm{det}}, y_z^{\mathrm{det}}\big)\big]\big\rangle < r^2\big\}. \tag{6.13}$$

The cross-section of $\mathcal{B}(r)$ at any plane $\{z = \mathrm{const}\}$ is an ellipsoid whose axes are determined by $\bar{V}(z)^{-1}$, while the scaling parameter r controls the size of the tube.

Theorem 6.2 *(Concentration of sample paths in the covariance tube). There exist constants $\Delta_0, r_0, \mu_0 > 0$ such that for all $0 < \Delta < \Delta_0$, all $\sigma < r < r_0\mu^{3/4}$ and all $\mu \leqslant \mu_0$,*

$$\mathbb{P}\{\tau_{\mathcal{B}(r)} < z\} \leqslant C_+(z, z_0)\mathrm{e}^{-\kappa_0 r^2/2\sigma^2} \tag{6.14}$$

holds for all $z \leqslant \sqrt{\mu}$, where the exponent κ_0 satisfies

$$\kappa_0 = 1 - \mathcal{O}(\Delta) - \mathcal{O}\big(r\mu^{-3/4}\big) - \mathcal{O}\big(\sigma^2/r^2\big), \tag{6.15}$$

and the prefactor is given by

$$C_+(z, z_0) = \frac{\mathrm{const}}{\Delta\mu}\left(\frac{r}{\sigma}\right)^2 \int_{z_0}^{z} x_s^{\mathrm{det}}\,\mathrm{d}s. \tag{6.16}$$

The proof is given in Appendix C. This result shows that the probability that sample paths leave the covariance tube before time z is small provided we take $r \gg \sigma\log(C_+(z, z_0))$. Indeed, for these values of r, the exponential term dominates the polynomial prefactor. We thus say that sample paths are concentrated in the covariance tube $\mathcal{B}(r)$ for r slightly larger than σ, or that the typical spreading of sample paths is given by $\mathcal{B}(\sigma)$.

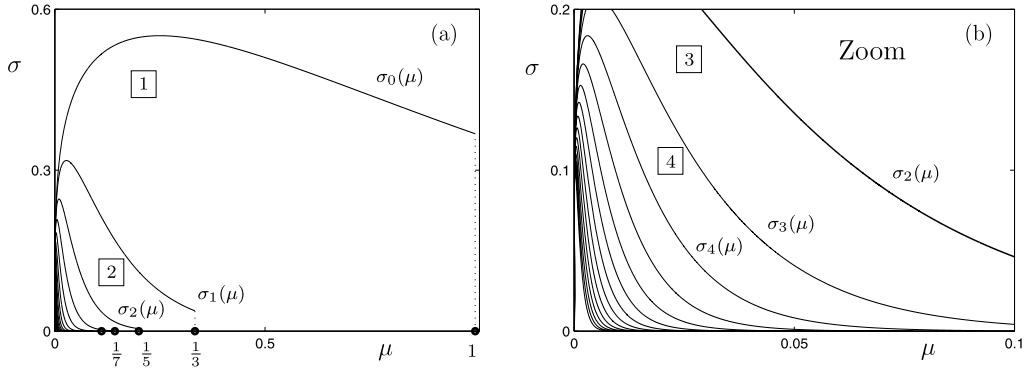

Fig. 6. (μ,σ)-parameter plane. (a) The curves $\sigma_k(\mu)$ are shown for $c_0 = 1$. Above the k-th curve we cannot distinguish the oscillation induced by the k-th canard from noisy fluctuations. The regions of different numbers of SAOs are delimited by the curves $\sigma_k(\mu)$ and the existence requirements for the k-th canard. The number of canard orbits that induce oscillations that can be distinguished from noise is indicated in rectangular boxes. (b) A zoom of (a) is shown that illustrates the structure of thin regions as $\mu \to 0$ and $\sigma \to 0$.

The condition $\sigma < r < r_0\mu^{3/4}$ implies that the theorem only applies to noise intensities smaller than $\mathcal{O}(\mu^{3/4})$. What happens for $\sigma \geqslant \mu^{3/4}$ is that fluctuations become large already some time before the fold line is reached, which completely smears out the small oscillations present in the deterministic case. In fact, it is possible to show that if $\sigma \geqslant \mu^{3/4}$, the bound (6.14) still holds true for $r \leqslant r_0|z|^{3/2}$, and thus sample paths are localized up to times $z \ll -\sigma^{2/3}$.

6.3. Small-amplitude oscillations and noise

Depending on the size of the covariance tubes we can determine when they start to overlap; from this we can deduce consequences for the existence of small-amplitude oscillations (SAOs) in the presence of noise. Recall from Theorem 3.2 that there are two primary canards $\gamma_\epsilon^{s,w}$ and $K-1$ secondary canards γ_ϵ^k for $2K-1 < \mu^{-1} < 2K+1$. We want to consider the rotations around γ_ϵ^w and denote the strong canard by γ_ϵ^0, i.e. $k=0$. By Theorem 4.4 the distance from the weak canard is given by $\mathcal{O}(\exp(-c_0(2k+1)^2\mu))$ for some positive constant $c_0 \in [\pi/4, 1]$. Theorem 6.1 implies that the width of the covariance tubes is given on $z=0$ by $\mathcal{O}(\mu^{-1/4}\sigma)$. Therefore sample paths starting on the k-th canard start to overlap with the weak canard for $\sigma \approx \mu\exp(-c_0(2k+1)^2\mu)$. We define the functions

$$\sigma_k(\mu) := \mu^{1/4}\mathrm{e}^{-c_0(2k+1)^2\mu}.$$

The previous considerations yield the following result.

Corollary 6.3 *(Noisy SAOs). On the section $\{z=0\}$ the covariance tubes of the k-th canard overlap if*

$$\sigma > \sigma_k(\mu), \tag{6.17}$$

i.e. depending on the noise level σ, the canard number and the parameter μ, deterministic SAOs with $2k+1$ twists become indistinguishable from noisy fluctuations if (6.17) *holds.*

In Fig. 6 we show the curves $\sigma_k(\mu)$ and indicate in which regions of the (μ,σ)-parameter plane one can distinguish which number of canards. The curves $\sigma_k(\mu)$ and the existence conditions of canards enclose bounded regions where precisely $k+1$ canards can be distinguished which yields $(2k+1)/2$ twists up to the section $\{z=0\}$; see also Theorem 3.3. We can also study Fig. 6 for fixed σ. In this case decreasing μ first increases the number of visible SAOs and then decreases it again.

6.4. Early jumps

We now turn to the behavior for times $z > \sqrt{\mu}$. For definiteness, we let $(x_z^{\mathrm{det}}, y_z^{\mathrm{det}}) = (-z, z^2 - \mu/2)$ be the weak canard solution, and define the set

$$\mathcal{D}(\eta) = \bigl\{(x, y, z)\colon z \geqslant \sqrt{\mu},\ \bigl(x - x_z^{\mathrm{det}}\bigr)^2 + \bigl(y - y_z^{\mathrm{det}}\bigr)^2 < \eta^2 z\bigr\}. \tag{6.18}$$

$\mathcal{D}(\eta)$ is a tube centered in the weak canard, whose width grows like $\sqrt{z}$. The following result, which is proved in Appendix D, shows that sample paths are unlikely to stay very long in $\mathcal{D}(\eta)$.

Theorem 6.4 *(Escape of sample paths from the primary canard). There exist constants $\kappa = \kappa(\eta) > 0$, $C_0 > 0$ and $\gamma_1, \gamma_2 > 0$ such that, whenever $\sigma |\log \sigma|^{\gamma_1} \leqslant \mu^{3/4}$,*

$$\mathbb{P}\{\tau_{\mathcal{D}(\eta)} > z\} \leqslant C_0 |\log \sigma|^{\gamma_2} \mathrm{e}^{-\kappa(z^2-\mu)/(\mu|\log\sigma|)}. \tag{6.19}$$

The probability that a sample path stays in $\mathcal{D}(\eta)$ thus becomes small as soon as

$$z \gg \sqrt{\mu |\log \sigma| / \kappa}. \tag{6.20}$$

Unless the noise intensity σ is exponentially small in μ, the typical time at which sample paths jump away from the canard is slightly (that is, logarithmically) larger than $\sqrt{\mu}$.

7. Numerics and visualization

In this section we briefly discuss how to compute canard solutions and their associated covariance tubes. Furthermore we visualize the early jumps after passage near a folded node in phase space for a model system with global returns. We also compute the probability density of escaping trajectories on a cross-section for this example. SDEs have been integrated numerically by a standard Euler–Maruyama scheme [60,71]. Deterministic solutions have been computed using a stiff ODE solver [100, 57].

7.1. Covariance tubes

The maximal canards and their associated covariance tubes can be computed. Fig. 7 shows an example for these computations where we used the normal form (6.4) with $\mu = 0.08$.

We compute the deterministic attracting slow manifold C_ϵ^a by forward integration and the deterministic repelling slow manifold C_ϵ^r by backward integration (see [33]) up to the section

$$\Sigma^0 = \bigl\{(x, y, z) \in \mathbb{R}^3\colon z = 0\bigr\}.$$

The primary and secondary maximal canards have been computed as intersections of the slow manifolds $C_\epsilon^a \cap C_\epsilon^r$; see also Section 3 and [54,36]. The resulting maximal canards (thick green curves) are shown in Fig. 7. We also computed a sample path (thin red curves) for each maximal canard starting at the same point as the maximal canard with $z_0 = -1$; the noise values were fixed at $\sigma = 0.008 = \sigma'$. The tubes defined by the covariance are shown in grey and have been computed using integration of the covariance differential equation (6.8); the section Σ^0 is drawn in yellow for better orientation.

Fig. 7(a) shows a side view that illustrates how the different primary canards $\gamma_{s,w}^\epsilon$ and secondary canards γ_ϵ^j are organized with respect to z. We have only started to draw the covariance tubes $\mathcal{B}(r)$ with $r^2 = 0.02$ a bit beyond the initial values at $x_0 = 1$. It is clearly visible how the canards and the tubes are attracted towards the weak canard and then start to rotate around it. Fig. 7(b) shows a front view towards the section Σ^0. This view shows nicely how the tubes grow with increasing z-values and that the ellipses defined by the covariance matrix are indeed close to circular. Furthermore we can

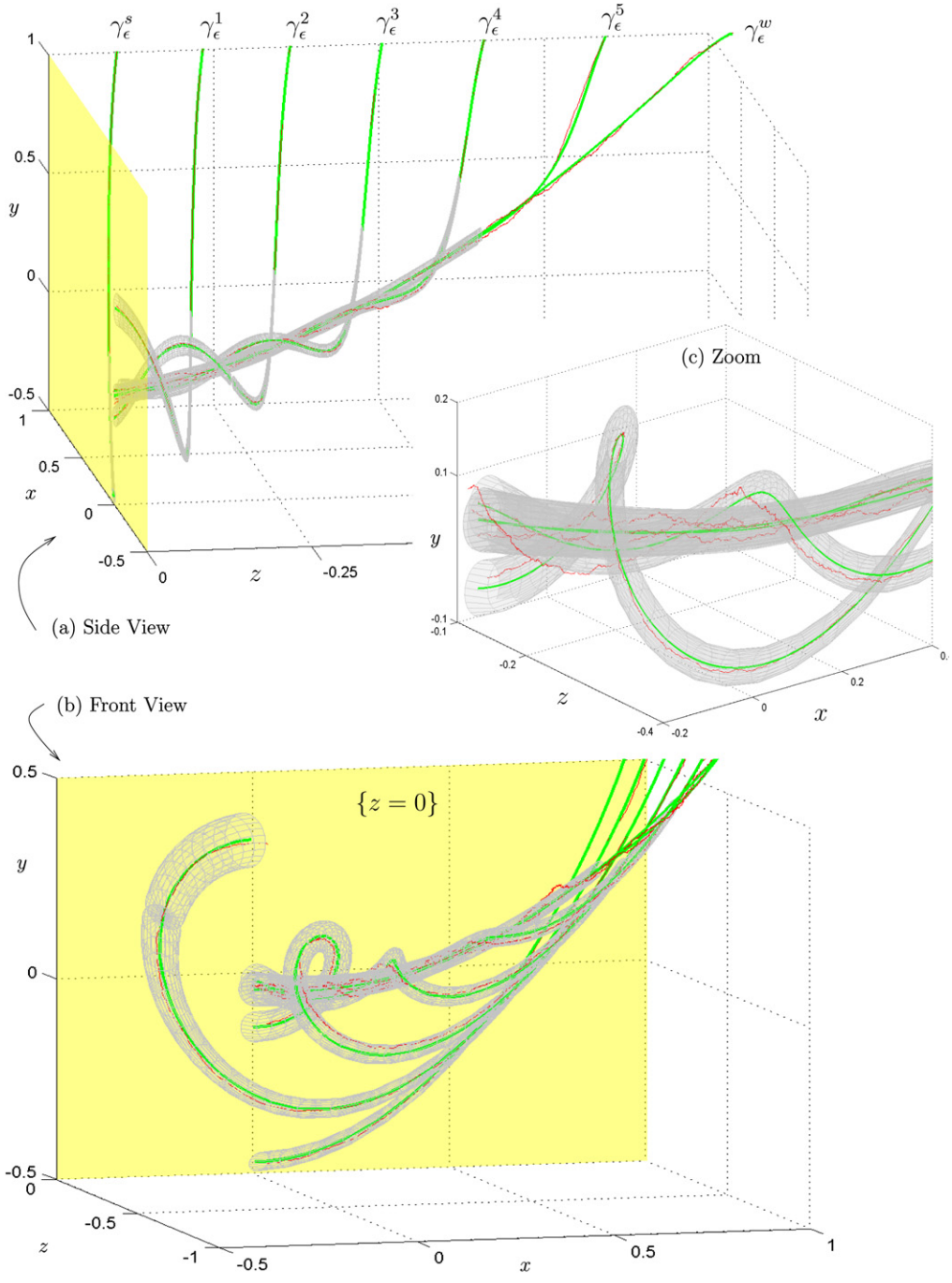

Fig. 7. Computation of the canards and covariance tubes near a folded node in system (6.4). A detailed description of the figure can be found in Section 7.1. (For interpretation of colors in this figure, the reader is referred to the web version of this article.)

see how the canards are organized on Σ^0; the maximal canard tubes for $\gamma_\epsilon^{s,1,2}$ do not interact while all other tubes interact near the weak canard. Fig. 7(c) shows a zoom that illustrates the twisting and also shows how the sample paths are indeed "trapped" inside the covariance tubes with very high probability.

7.2. Early jumps

To visualize the effect of early jumps we consider a folded node with global returns given by

$$\begin{aligned} \mathrm{d}x &= \frac{1}{\epsilon}\big(y - x^2 - x^3\big)\,\mathrm{d}s + \frac{\sigma}{\sqrt{\epsilon}}\,\mathrm{d}W_s^{(1)}, \\ \mathrm{d}y &= \big[-(\mu+1)x - z\big]\,\mathrm{d}s + \sigma'\,\mathrm{d}W_s^{(2)}, \\ \mathrm{d}z &= \left[\frac{\mu}{2} + ax + bx^2\right]\mathrm{d}s, \end{aligned} \tag{7.1}$$

which is a slight modification of a model system for folded node MMOs [30]. The critical manifold is cubic-shaped (or S-shaped) and given by

$$C_0 = \big\{(x,y,z)\in\mathbb{R}^3\colon\ y = x^2 + x^3\big\} = C_0^{a,-} \cup L_- \cup C^r \cup L^+ \cup C^{a,+}$$

where $C_0^{a,+} = C_0 \cap \{x < -2/3\}$, $C_0^r = C_0 \cap \{-2/3 < x < 0\}$, $C_0^{a,+} = C_0 \cap \{x > 0\}$, $L_- = C_0 \cap \{x = -2/3\}$ and $L_+ = C_0 \cap \{x = 0\}$. The parameters $(a,b)\in\mathbb{R}^2$ help to adjust the global return mechanism. If a, b are $\mathcal{O}(1)$ then they do not influence the local behavior of a folded node at the origin $(x,y,z) = (0,0,0)$. Fig. 8 shows the effect of early jumps after passage through a folded node region. Parameters for the simulation are:

$$\epsilon = 0.01, \qquad \mu = 0.143, \qquad a = 0.2, \qquad b = -1.1, \qquad \sigma = 0.005, \qquad \sigma' = 0. \tag{7.2}$$

In Fig. 8(a) a deterministic trajectory (thick blue curve) has been computed for $\sigma = 0$. Then an SDE sample path for (7.1) has been started on a point (green dot) of the deterministic solution and integrated forward. We define a cross-section

$$\Sigma^J := \big\{(x,y,z)\in\mathbb{R}^3\colon\ x = -0.3\big\}.$$

Escapes of sample paths from the folded node region are recorded on Σ^J. The next two returns are also shown as points (violet and red) on the cross-section Σ^J. It is clearly visible from Fig. 8(a) that the SDE sample path jumps before the deterministic solution. Note that this causes the path to get re-injected into the folded node region after a large excursion at a point slightly different from the deterministic solution. Hence the global return mechanism can potentially act as a control mechanism for the noise. To investigate the early jumps further we show in Fig. 8(b) a probability density

$$p(y,z) \quad \text{on } \Sigma^J.$$

The density has been computed by recording the intersections with Σ^J after passage through the folded node for 4000 sample paths that have been integrated for a time $s\in[0,20]$. The corresponding deterministic point measure p_{det} has been indicated as well. The density $p(y,z)$ clearly shows that paths are expected to jump before the deterministic solution if we consider the z-coordinate distance from the folded node. We also see that the density $p(y,z)$ is quite confined and shows a multi-modal structure. This structure can be explained from the fact that sample paths exit early but between different exit points they can make additional deterministic

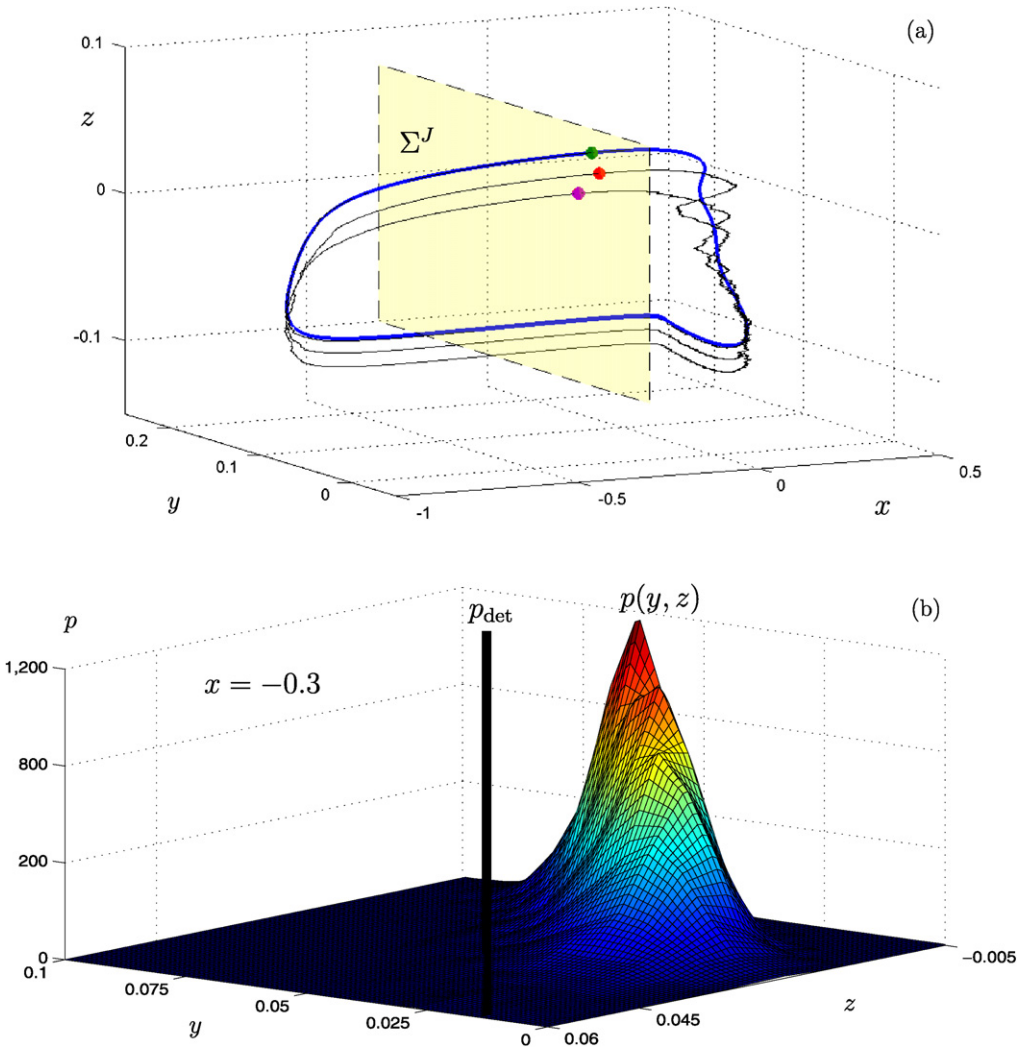

Fig. 8. (a) Phase space plot of simulations for (7.1). The deterministic solution (blue) and a stochastic sample path (black) are shown; parameters are given by (7.2) and for the deterministic solution we have $\sigma = 0 = \sigma'$. Intersections with the cross-section Σ^J (yellow) are shown as thick dots; the SDE sample path is started on the deterministic solution and on Σ^J (green dot). The next two intersections are shown as well (violet and red dots). (b) Probability density $p(y, z)$ for 4000 sample paths on Σ^J of sample path escapes from the folded node. The deterministic point-mass density is indicated as a bar p_{det} (black). (For interpretation of colors in this figure, the reader is referred to the web version of this article.)

small oscillations. A different number of these oscillations corresponds to the different maxima of $p(y, z)$.

8. Final remarks

In Section 6 we stated our results on the relation between the noise level, the parameter μ and the location of paths measured along the z-coordinate. Note that we proved and stated our results in zoomed-in (or re-scaled/blown-up) coordinates removing the ϵ-dependence. In particular, we worked in a neighborhood of the folded node that is of size $\mathcal{O}(\sqrt{\epsilon})$ in original coordinates. To obtain the results in original coordinates one has to apply a zoom-out (or blow-down) transformation. First, we

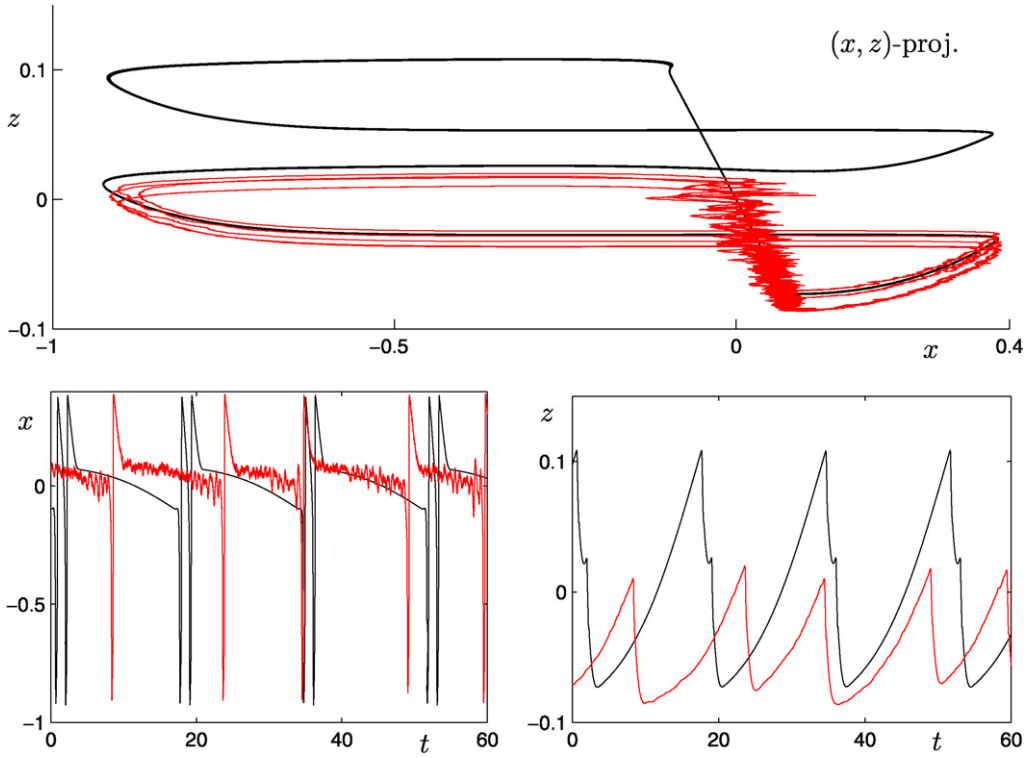

Fig. 9. Simulation for the model system (7.1) with parameter values (8.1). The upper plot shows a projection of a deterministic solution (black) and a stochastic sample path (red) into the (x, z)-plane. The lower two plots show the associated time series. (For interpretation of colors in this figure, the reader is referred to the web version of this article.)

replace $(x, y, z, \sigma, \sigma')$ by $(\bar{x}, \bar{y}, \bar{z}, \bar{\sigma}, \bar{\sigma}')$ (recall: we dropped the overbars for notational convenience). Then the identity

$$(\bar{x}, \bar{y}, \bar{z}, \bar{\sigma}, \bar{\sigma}', \mu) = (\epsilon^{-1/2}x, \epsilon^{-1}y, \epsilon^{-1/2}z, \epsilon^{-3/4}\sigma, \epsilon^{-3/4}\sigma', \mu)$$

provides the required zoom-out transformation. This implies e.g. that the interaction of canards in Corollary 6.3 is given, in original coordinates, by relations of the form

$$\sigma \approx \epsilon^{3/4}\mu^{1/4}\mathrm{e}^{-c_0(2k+1)^2\mu}$$

or that sample paths are likely to escape for

$$z \gg \epsilon^{1/2}\sqrt{\frac{\mu}{\kappa}\left|\log\sigma - \frac{3}{4}\log\epsilon\right|}$$

as shown in Theorem 6.4. Obviously one also has to translate the assumptions in a similar way e.g. $\bar{\sigma} \ll \mu^{3/4}$ becomes $\sigma \ll (\mu\epsilon)^{3/4}$.

Another important point is that we focused on the detailed analysis near the folded node and did not consider different types of global return mechanisms. In Fig. 9 we show the interaction between

global returns and noise-induced early jumps on oscillatory patterns. The simulation in Fig. 9 has been carried out using the model system (7.1) with parameter values

$$\epsilon = 0.01, \qquad \mu = 0.029, \qquad a = -0.1, \qquad b = -0.5, \qquad \sigma = 0.005, \qquad \sigma' = 0.005. \quad (8.1)$$

The deterministic solution in Fig. 9 is an MMO with pattern 2^s where the number of small oscillations is difficult to count from the numerical results as μ is already extremely small. The SDE sample paths jump significantly earlier to $C_\epsilon^{a,-}$ than the deterministic solution as expected from our results. However, we see that the deterministic solution makes an additional large-amplitude oscillation (LAO) given by the passage

$$C_\epsilon^{a,-} \to \text{jump near } L_- \to C_\epsilon^{a,+} \to \text{jump near } L_+ \to C_\epsilon^{a,-}.$$

In particular, the early jumps change the number of LAOs in the MMO pattern from $L = 2$ to $L = 1$. The noise also influences the number of SAOs but the crucial point is that it can also have a global effect. Hence we end up with a stochastic-resonance-type mechanism for MMO patterns.

Furthermore, one could think about extending our results to capture the effect of a global return map on the escape density after the passage through a folded node region; see also the brief discussion in Section 7.2. Let Σ^J denote a cross-section on which we record the escape from the folded node region and let Σ^R denote a cross-section to the deterministic flow slightly before the re-entry to the folded node region; see also [81]. Then we have a global return map induced by the deterministic flow

$$M : \Sigma^J \to \Sigma^R. \quad (8.2)$$

If we are given a probability density p on Σ^J, we can then consider the induced density $p \circ M^{-1}/|\det DM \circ M^{-1}|$ on Σ^R. This should allow us to calculate a distribution for different MMO patterns, i.e. we can hope to assign a probability to each combination L^s after one return from Σ^J to Σ^J. Although this approach seems possible it is beyond the local analysis we focused on here. A solution of this problem crucially depends on the form of the global returns which are described by the map M.

Acknowledgments

The authors would like to thank Mathieu Desroches for inspiring discussions. N.B. and C.K. thank the CRC 701 at University of Bielefeld, B.G. thanks the MAPMO at Université d'Orléans and C.K. thanks Cornell University for hospitality and financial support.

Appendix A. Proof of Theorem 4.3 (canonical form)

We are going to need the following lemma for the proof of Theorem 4.3.

Lemma A.1. *Consider two non-autonomous vector fields $F, G : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$ for $(z, X) \in \mathbb{R} \times \mathbb{R}^N$. Suppose both are continuous in z, C^1 in X and defined on an open set $\mathcal{D}$ containing $(0, X_0)$. Suppose for all (z, X) in $\mathcal{D}$ we have*

$$\big\| F(z, X) - G(z, X) \big\| < \mu.$$

Let K be a Lipschitz constant for F in X, which is uniform in z. Suppose $X(z)$ and $Y(z)$ solve $\frac{dX}{dz} = F(z, X)$ and $\frac{dY}{dz} = G(z, Y)$ with $X(0) = X_0 = Y(0)$, then

$$\big\| X(z) - Y(z) \big\| \leqslant \frac{\mu}{K}\big(e^{K|z|} - 1\big). \quad (A.1)$$

Proof. A direct Gronwall lemma argument suffices; for details see [61, pp. 399–400]. □

The proof of Theorem 4.3 proceeds in several steps that aim to bring Eq. (4.8) into diagonal form [18,27].

Proof of Theorem 4.3. As a first step we apply a rescaling

$$u = \exp\left[\frac{1}{2\mu}\int_0^z \operatorname{Tr}\bigl(A(s)\bigr)\,\mathrm{d}s\right]u^{(0)} = \mathrm{e}^{z^2/\mu}u^{(0)}$$

where $u^{(0)} = (u_1^{(0)}, u_2^{(0)})$ are new coordinates. This yields

$$\mu\frac{\mathrm{d}u^{(0)}}{\mathrm{d}z} = A_0(z)u^{(0)}, \quad \text{with } A_0(z) = \begin{pmatrix} 2z & 2 \\ -2(\mu+1) & -2z \end{pmatrix} \tag{A.2}$$

where now $\operatorname{Tr}(A_0(z)) = 0$. Therefore the principal solution has determinant 1 and is area preserving. We are going to show that the solution of (A.2) is a rotation up to a small error using a sequence of z-dependent coordinate transformations $S_j(z)$. We set

$$u^{(0)} = S_0(z)u^{(1)}, \quad \text{where } S_0(z) = \begin{pmatrix} \frac{-z-\mathrm{i}\omega(z)}{1+\mu} & \frac{-z+\mathrm{i}\omega(z)}{1+\mu} \\ 1 & 1 \end{pmatrix}.$$

Then $S_0^{-1}A_0S_0$ is diagonal with entries $\pm 2\mathrm{i}\omega(z)$. We get

$$\mu\frac{\mathrm{d}u^{(1)}}{\mathrm{d}z} = A_1(z)u^{(1)},$$

where the new matrix A_1 is given by

$$A_1(z) = S_0^{-1}A_0S_0 - \mu S_0^{-1}\frac{\mathrm{d}S_0}{\mathrm{d}z} = \begin{pmatrix} 2\mathrm{i}\omega(z) + \mu\frac{\mathrm{i}-\omega'(z)}{2\omega(z)} & \mu\frac{\mathrm{i}+\omega'(z)}{2\omega(z)} \\ \mu\frac{-\mathrm{i}+\omega'(z)}{2\omega(z)} & -2\mathrm{i}\omega(z) + \mu\frac{-\mathrm{i}-\omega'(z)}{2\omega(z)} \end{pmatrix}.$$

As above we want to have a zero trace so we compute

$$\frac{1}{2\mu}\int_0^z \operatorname{Tr}\bigl(A_1(s)\bigr)\,\mathrm{d}s = -\frac{1}{2}\int_0^z \frac{\omega'(s)}{\omega(s)}\,\mathrm{d}s = -\frac{1}{2}\log\omega(z) + \text{const.}$$

Hence we consider the scaling

$$u^{(1)} = S_1(z)u^{(2)} = \frac{1}{\sqrt{\omega(z)}}u^{(2)}$$

which yields the equation

$$\mu\frac{\mathrm{d}u^{(2)}}{\mathrm{d}z} = A_2(z)u^{(2)}, \quad \text{with } A_2(z) = \begin{pmatrix} \mathrm{i}\omega_2(z) & \mu\bar{\rho}_2(z) \\ \mu\rho_2(z) & -\mathrm{i}\omega_2(z) \end{pmatrix}$$

where overbar denotes complex conjugate and

$$\omega_2(z) = 2\omega(z) + \frac{\mu}{2\omega(z)}, \qquad \rho_2(z) = \frac{-\mathrm{i} + \omega'(z)}{2\omega(z)}. \tag{A.3}$$

The next transformation

$$u^{(2)} = S_2(z)u^{(3)}$$

yields

$$\mu \frac{\mathrm{d}u^{(3)}}{\mathrm{d}z} = A_3(z)u^{(3)}, \qquad A_3(z) = S_2^{-1}\left[A_2 S_2 - \mu \frac{\mathrm{d}S_2}{\mathrm{d}z}\right]. \tag{A.4}$$

Instead of a given transformation we now impose the form of the matrices

$$S_2(z) = \begin{pmatrix} 1 & \mu\bar{v}(z) \\ \mu v(z) & 1 \end{pmatrix}, \qquad A_3(z) = \begin{pmatrix} \rho_1(z) & 0 \\ 0 & \bar{\rho}_1(z) \end{pmatrix} \tag{A.5}$$

so that $S_2 = \mathrm{Id} + \mathcal{O}(\mu)$ and A_3 is diagonal. Substituting (A.5) into (A.4) leads to the equations

$$\begin{aligned} 0 &= \mathrm{i}\omega_2(z) + \mu^2 \bar{\rho}_2(z) v - \rho_1(z), \\ \mu \frac{\mathrm{d}v}{\mathrm{d}z} &= \rho_2(z) - \mathrm{i}\omega_2(z) v - v\rho_1(z) \end{aligned}$$

and their complex conjugates. The first equation determines $\rho_1(z)$, and thus the second one becomes

$$\mu \frac{\mathrm{d}v}{\mathrm{d}z} = -2\mathrm{i}\omega_2(z) v - \mu^2 \bar{\rho}_2(z) v^2 + \rho_2(z). \tag{A.6}$$

If we can show that (A.6) has a bounded solution for times $|z| < 1$ then our prescribed coordinate change S_2 exists. Now let $a(z) = 2z + \operatorname{Re}\rho_1(z) = 2z + \mathcal{O}(\mu^2)$ and define $\alpha(z, z_0)$ by (4.24). A last transformation

$$u^{(3)} = S_3(z)\tilde{u} = \mathrm{e}^{-\alpha(z,0)/\mu} \frac{1}{1+\mathrm{i}} \begin{pmatrix} \mathrm{i} & 1 \\ 1 & \mathrm{i} \end{pmatrix} \tilde{u} \tag{A.7}$$

brings the equation into canonical form (4.21), with $\varpi(z) = \operatorname{Im}\rho_1(z) = 2\omega(z) + \mathcal{O}(\mu)$. Composing all transformations, we get $u = S(z)\tilde{u}$, where $S(z) = \mathrm{e}^{z^2/\mu} S_0(z) S_1(z) S_2(z) S_3(z)$ is indeed of the form (4.20).

To prove the existence of bounded solutions for (A.6) note that $\rho_2(z)$ and $\omega_2(z)$ are bounded away from zero and that they also have bounded norms for $|z| < 1$. Now set

$$F(z, v) := -\frac{2\mathrm{i}\omega_2(z)}{\mu} v + \frac{\rho_2(z)}{\mu}, \qquad G(z, v) := -\frac{2\mathrm{i}\omega_2(z)}{\mu} v - \mu\bar{\rho}_2(z) v^2 + \frac{\rho_2(z)}{\mu}.$$

Note that F is a vector field with bounded solutions on an $\mathcal{O}(1)$ time scale. Considering F, G as real planar vector fields on $\mathbb{R}^2$ we can apply Lemma A.1 to conclude that (A.6) admits bounded solutions on time intervals of length 1. □

Appendix B. Proof of Theorem 6.1 (covariance matrix)

In this appendix, we discuss the equation

$$\mu \frac{\mathrm{d}V}{\mathrm{d}z} = A\big(x_z^{\mathrm{det}}\big)V + V A\big(x_z^{\mathrm{det}}\big)^T + \big(F^0\big)\big(F^0\big)^T \tag{B.1}$$

describing the evolution of $V(z) = \{V_{ij}(z)\}_{i,j\in\{1,2\}} := \sigma^{-2}\operatorname{Cov}(z)$, the covariance matrix of the linearized variational equation around a deterministic solution of (6.4). Before proving Theorem 6.1 on the small-μ asymptotics of the solutions, we provide different approaches yielding information on the behavior of $V(z)$. A formal method based on iterative computations of a slow manifold is developed in Section B.1 to understand the asymptotics of (B.1). In Section B.2 we provide rigorous bounds using a Lyapunov function. In Section B.3 we refine the previous results by a transformation to real canonical form and results about delayed Hopf bifurcation, thereby proving the main theorem.

Proposition B.1. *Let*

$$v(z) = \begin{pmatrix} v_1(z) \\ v_2(z) \\ v_3(z) \end{pmatrix}, \quad \text{where } \begin{cases} v_1(z) = V_{11}(z), \\ v_2(z) = V_{22}(z), \\ v_3(z) = V_{12}(z) = V_{21}(z) \end{cases} \tag{B.2}$$

denote the two variances and the covariance. Then $v(z)$ satisfies the ODE

$$\mu \frac{\mathrm{d}v}{\mathrm{d}z} = \underbrace{\begin{pmatrix} -8x(z) & 0 & 4 \\ 0 & 0 & -4(\mu+1) \\ -2(\mu+1) & 2 & -4x(z) \end{pmatrix}}_{=:B(z)} v + \underbrace{\begin{pmatrix} 2 \\ 2\rho^2 \\ 0 \end{pmatrix}}_{=:E} \tag{B.3}$$

where we have abbreviated $x^{\mathrm{det}}(z) =: x(z)$.

Proof. Using the definitions from (6.6) and Eq. (B.1) we get

$$\begin{aligned} \mu \frac{\mathrm{d}V}{\mathrm{d}z} &= A\big(x_z^{\mathrm{det}}\big)V + V A\big(x_z^{\mathrm{det}}\big)^T + \big(F^0\big)\big(F^0\big)^T \\ &= \begin{pmatrix} -4x_z^{\mathrm{det}} & 2 \\ -2(\mu+1) & 0 \end{pmatrix} V + V \begin{pmatrix} -4x_z^{\mathrm{det}} & -2(\mu+1) \\ 2 & 0 \end{pmatrix} + \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2}\rho \end{pmatrix}\begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2}\rho \end{pmatrix} \\ &= \begin{pmatrix} 4V_{12} - 8V_{11}x_z^{\mathrm{det}} + 2 & 2V_{22} - 4V_{12}x_z^{\mathrm{det}} - 2V_{11}(1+\mu) \\ 2V_{22} - 4V_{12}x_z^{\mathrm{det}} - 2V_{11}(1+\mu) & -4V_{12}(1+\mu) + 2\rho^2 \end{pmatrix}. \end{aligned}$$

Therefore the result follows. □

Our goal is to analyze (B.3) for a given maximal canard solution $x(z)$. Observe that $B(z)$ has eigenvalues

$$-4x(z), \quad -4x(z) \pm \mathrm{i}\sqrt{1 - x(z)^2 + \mu} = -4x(z) \pm \mathrm{i}\omega_0\big(-x(z)\big).$$

We assume that z_0, z_1 are chosen so that

$$-1 < z_0 < 0 < z_1 < 1 \quad \text{and} \quad 1 - x(z)^2 + \mu > 0 \quad \forall z \in [z_0, z_1]. \tag{B.4}$$

In particular, the assumptions (B.4) are satisfied for any maximal canard solution approaching the folded node region from the slow manifold C_ϵ^a for some $z_0 < 0 < z_1$ of order 1 and μ sufficiently small.

B.1. Iteration and asymptotics

Notation B.2. Henceforth, we write $x(z,\mu) \asymp y(z,\mu)$ if

$$c_- y(z,\mu) \leqslant x(z,\mu) \leqslant c_+ y(z,\mu) \tag{B.5}$$

holds for all z, where $c_\pm$ are positive constants independent of z and μ.

A formal derivation for the asymptotics as $\mu \to 0$ for (B.3) can be carried out using an iterative scheme [93,18]. We set $v(z) = V_0^*(z) + V_1(z)$ where

$$V_0^*(z) := -B(z)^{-1}E = -\frac{1}{4(1+\mu)x(z)}\begin{pmatrix} 1+\mu+\rho^2 \\ (1+\mu)^2 + (1+4x(z)^2+\mu)\rho^2 \\ 2\rho^2 x(z) \end{pmatrix} \tag{B.6}$$

defines the critical manifold for (B.3) when viewed as a slowly time-dependent system. We get

$$\mu\frac{\mathrm{d}V_1}{\mathrm{d}z} = B(z)V_1 + \mu E_1(z), \qquad E_1(z) = -\frac{\mathrm{d}}{\mathrm{d}z}V_0^*(z).$$

The same change procedure also works for any $n \geqslant 1$ by setting

$$v(z) = \sum_{j=0}^{n} \mu^n V_j^*(z) + V_{n+1}(z). \tag{B.7}$$

Then $V_{n+1}(z)$ satisfies the equation

$$\mu\frac{\mathrm{d}V_{n+1}}{\mathrm{d}z} = B(z)V_{n+1} + \mu^{n+1}E_{n+1} \tag{B.8}$$

where V_n^* and E_n are given inductively by

$$V_{n+1}^*(z) = B(z)^{-1}\frac{\mathrm{d}}{\mathrm{d}z}V_n^*(z), \qquad E_{n+1}(z) = \frac{\mathrm{d}}{\mathrm{d}z}\big[B(z)^{-1}E_n(z)\big].$$

Remark. Observe that (B.7) is the asymptotic expansion for the slow manifold of (B.3). The iterative scheme we use here is very convenient for slowly time-dependent systems. Many other methods to calculate slow manifolds for general fast–slow systems have been explored; see [114,115] and references therein.

Proposition B.3. *Assume that the deterministic maximal canard solution $x(z)$ satisfies* (B.4)*. Then the asymptotic expansion* (B.7) *of $v(z)$ for $n \geqslant 0$ has components of order*

$$\mu^n V_{n,1}^*(z) \asymp \mu^n V_{n,2}^*(z) \asymp \frac{\mu^n}{|z|^{2n+1}}, \qquad \mu^n V_{n,3}^*(z) \asymp \frac{\mu^n}{|z|^{2n}}. \tag{B.9}$$

Proof. First observe that the symmetry (3.12) implies that $x(0) = 0$ for any maximal canard. Using this fact and the slow flow (3.10) we find that $x(z)$ must have a Taylor expansion with non-vanishing linear term, i.e.

$$x(z) = x_1 z + x_2 z^2 + \cdots \tag{B.10}$$

with $x_1 < 0$. The proof of (B.9) then proceeds by induction as follows: The base step $n = 0$ holds by formulas (B.6) and (B.10). The induction step from $n-1$ to n follows from direct differentiation

$$\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{z^{2n-1}}\right) = \frac{1-2n}{z^{2n}},$$

and the calculation of $B(z)^{-1}$

$$B(z)^{-1} = -\begin{pmatrix} \frac{1}{8x(z)} & \frac{1}{8(1+\mu)x(z)} & 0 \\ \frac{1+\mu}{8x(z)} & \frac{1+\mu}{8x(z)} - \frac{x(z)}{2(1+\mu)} & \frac{1}{2} \\ 0 & \frac{1}{4(1+\mu)} & 0 \end{pmatrix}$$

almost immediately; we just have to observe the block structure of $B(z)^{-1}$. □

Proposition B.3 is a formal asymptotic result. The asymptotic series (B.7) becomes "disordered" for $|z| = \mathcal{O}(\sqrt{\mu})$ because in this case all the terms for the coordinates v_1 and v_2 are of order $1/\sqrt{\mu}$ while all terms for v_3 are of order 1. Therefore we conjecture that

$$v_1 = \mathcal{O}\left(\frac{1}{\sqrt{\mu}}\right), \qquad v_2 = \mathcal{O}\left(\frac{1}{\sqrt{\mu}}\right), \qquad v_3 = \mathcal{O}(1), \tag{B.11}$$

for $-\sqrt{\mu} < z_0 \leqslant z \leqslant \sqrt{\mu}$.

B.2. Lyapunov function

The results in this section are not as sharp as the results obtained by coordinate changes in Section B.3 but they are obtained by a completely different technique which is of interest on its own in the context of folded nodes. We are going to require an auxiliary result to analyze that will be used below in the proof of Proposition B.5.

Lemma B.4. *Consider the linear non-autonomous differential equation on $\mathbb{R}$ given by*

$$\mu \frac{\mathrm{d}X}{\mathrm{d}z} = k_1 z X + k_2 \frac{\mu^n}{(-z)^{2n}} \tag{B.12}$$

where $k_{1,2} = \mathcal{O}(1)$ are two positive constants, $\mu > 0$, and either $n = 0$ and $z \geqslant z_0$ or $n \geqslant 1$ and $z_0 \leqslant z < 0$. Then

$$X(z) \asymp \begin{cases} \mu^n |z|^{-(2n+1)} & \text{for } z_0 + \mathcal{O}(\mu|\log\mu|) \leqslant z \leqslant -\sqrt{\mu}, \\ \mu^{n-1}|z|^{-(2n-1)} & \text{for } -\sqrt{\mu} \leqslant z < 0 \text{ if } n \geqslant 1, \\ \mu^{-1/2} & \text{for } -\sqrt{\mu} \leqslant z \leqslant \sqrt{\mu} \text{ if } n = 0. \end{cases}$$

Proof. The solution of (B.12) can be written as

$$X(z) = X(z_0)\mathrm{e}^{k_1(z^2-z_0^2)/2\mu} + k_2 I_n(z) \tag{B.13}$$

where

$$I_n(z) = \mu^{n-1}\int_{z_0}^{z} \mathrm{e}^{k_1(z^2-t^2)/2\mu}\,\frac{1}{(-t)^{2n}}\,\mathrm{d}t.$$

For $z \in [z_0, -\sqrt{\mu}\,]$ the leading-order asymptotics of (B.13) is given by $I_n(z)$. Using integration by parts we get an upper bound

$$\begin{aligned} k_1 I_n(z) &= k_1\mu^{n-1}\int_{z_0}^{z}\left(-\frac{\mu}{k_1 t}\,\frac{1}{(-t)^{2n}}\right)\left(-\frac{k_1 t}{\mu}\mathrm{e}^{k_1(z^2-t^2)/2\mu}\right)\mathrm{d}t \\ &= \mu^n \frac{1}{(-t)^{2n+1}}\mathrm{e}^{k_1(z^2-t^2)/2\mu}\bigg|_{z_0}^{z} - \mu^n\int_{z_0}^{z}\frac{\mathrm{d}}{\mathrm{d}t}\left[\frac{1}{(-t)^{2n+1}}\right]\mathrm{e}^{k_1(z^2-t^2)/2\mu}\,\mathrm{d}t \\ &= \frac{\mu^n}{(-z)^{2n+1}} - \frac{\mu^n}{(-z_0)^{2n+1}}\mathrm{e}^{k_1(z^2-z_0^2)/2\mu} - (2n+1)I_{n+1}(z) \\ &\leqslant \frac{\mu^n}{(-z)^{2n+1}}. \end{aligned} \tag{B.14}$$

The lower bound follows by inserting the upper bound for $I_{n+1}(z)$ in (B.14). Here the condition $z \geqslant z_0 + \mathcal{O}(\mu|\log\mu|)$ is needed to make the term $\mathrm{e}^{k_1(z^2-z_0^2)/2\mu}$ small. This implies in particular that $X(-\sqrt{\mu}) \asymp 1/\sqrt{\mu}$.

Finally, to describe the behavior for $-\sqrt{\mu} < z < 0$, we replace z_0 by $-\sqrt{\mu}$ in (B.13). Then all exponential terms are of order 1, and the integral can be estimated directly. □

Proposition B.5. *Suppose* (B.4) *holds and let $x(z)$ be a maximal canard solution. Then solutions to the variational equation* (B.3) *remain bounded by $\mathcal{O}(1/(|z| + \sqrt{\mu}\,))$ for $z_0 \leqslant z \leqslant \sqrt{\mu}$.*

Proof. Throughout the proof we are going to introduce several positive constants $c_j = \mathcal{O}(1)$ for $j \in \mathbb{N}$ whose actual value does not influence the asymptotic result. As a first step we want to find a symmetric matrix $M(z)$ such that

$$B(z)^T M(z) + M(z)B(z) = -x(z)\,\mathrm{Id}\,. \tag{B.15}$$

This can simply be accomplished by solving the six algebraic equations (B.15). We find that

$$M(z)|_{\mu=0} = \frac{1}{64(1+3x^2)}\begin{pmatrix} 7+12x^2 & 1+12x^2 & -12x \\ 1+12x^2 & 7+64x^2+48x^4 & -16x(1+3x^2) \\ -12x & -16x(1+3x^2) & 4(3+18x^2) \end{pmatrix}$$

where we have abbreviated $x = x(z)$. Since (B.10) holds for maximal canards it is straightforward to check that the matrix $M(z)$ is positive definite, uniformly in μ and z. Therefore it defines a family of quadratic forms

$$Y_n := V_n^T M(z) V_n \tag{B.16}$$

where V_n is defined by (B.7). The quadratic form (B.16) satisfies

$$c_1\|V_n\|^2 \leqslant Y_n(z) \leqslant c_2\|V_n\|^2$$

for some constants $c_1, c_2 > 0$. Essentially Y_n will act as a Lyapunov function to bound $\|V_n\|$. To show this we compute the derivative. Using (B.15) and (B.8), we get

$$\begin{aligned}\mu\frac{\mathrm{d}Y_n}{\mathrm{d}z} &= \mu\frac{\mathrm{d}V_n^T}{\mathrm{d}z}M(z)V_n + \mu V_n^T M(z)\frac{\mathrm{d}V_n}{\mathrm{d}z} + \mu V_n^T\frac{\mathrm{d}M}{\mathrm{d}z}V_n\\ &= -x(z)V_n^T V_n + \mu V_n^T\frac{\mathrm{d}M}{\mathrm{d}z}V_n + \mu^n\big[E_n(z)^T M(z)V_n + V_n^T M(z)E_n(z)\big]. \end{aligned}\tag{B.17}$$

Since $\|E_n\| = \mathcal{O}(|z|^{-2n})$ and $\|\frac{\mathrm{d}M}{\mathrm{d}z}\|$ is bounded we can find constants $c_3, c_4 > 0$ such that (B.17) implies

$$\mu\frac{\mathrm{d}Y_n}{\mathrm{d}z} \leqslant c_3\big(-x(z)+\mu\big)Y_n + c_4\frac{\mu^n}{(-z)^{2n}}\sqrt{Y_n}.$$

Setting $Y_n = Z_n^2$, we find that the last inequality is equivalent to

$$\mu\frac{\mathrm{d}Z_n}{\mathrm{d}z} \leqslant c_5\big(-x(z)+\mu\big)Z_n + c_6\frac{\mu^n}{(-z)^{2n}}.\tag{B.18}$$

Using (B.10) and Lemma B.4 we obtain that for z_0 of order -1,

$$Z_n(z) \leqslant c_7\frac{\mu^n}{|z|^{2n+1}} \quad \text{for } z_0 \leqslant z < 0.\tag{B.19}$$

Since Z_n is equivalent to $\|V_n(z)\|$, this shows that (B.8) is indeed an asymptotic expansion in powers of μ/z^2 for $z \leqslant -\sqrt{\mu}$, and in particular all components of $v(-\sqrt{\mu})$ are of order $1/\sqrt{\mu}$. To complete the proof up to time $z = \sqrt{\mu}$, we simply apply (B.19) in the particular case $n = 0$ (that is, for $V_0 = v$ and $E_0 = E$). □

In view of our conjecture (B.11) the bound on the covariance provided by Proposition B.5 is probably not sharp since we have not shown that $v_3 = \mathcal{O}(1)$.

B.3. Delayed Hopf bifurcation

To obtain a sharp bound on the covariance we consider a similar coordinate change idea as in Section 4.2. This procedure will give a variational equation for the covariance that has desirable symmetry properties.

Lemma B.6. *There exists a linear coordinate change $\xi^0 = S(z)\tilde{\zeta}^0$ transforming the linearized SDE* (6.6) *into*

$$\mathrm{d}\tilde{\zeta}_z^0 = \frac{1}{\mu}\tilde{A}(z)\tilde{\zeta}_z^0\,\mathrm{d}z + \frac{\sigma}{\sqrt{\mu}}\tilde{F}(z)\,\mathrm{d}W_z,\tag{B.20}$$

where $\tilde{A}(z)$ is in canonical form

$$\tilde{A}(z) = \begin{pmatrix} a(z) & \varpi(z)\\ -\varpi(z) & a(z)\end{pmatrix},\tag{B.21}$$

with $a(z) = -2x(z) + \mathcal{O}(\mu)$ and $\varpi(z) = 2\omega(z) + \mathcal{O}(\mu)$. The matrix $\tilde{F}(z)$ is positive definite, with eigenvalues bounded below and above uniformly in z.

Proof. It suffices to apply the coordinate change $\zeta_z^0 = S(z)\tilde{\zeta}_z^0$ constructed in the proof of Theorem 4.3 (with an obvious modification due to the fact that $x(z)$ is not necessarily given by the weak canard). The new diffusion coefficient is then given by $\tilde{F}(z) = S(z)^{-1}F^0$. □

A computation analogous to the one in the proof of Proposition B.1 then yields

Lemma B.7. *The covariance matrix of ζ_z^0 is given by $\sigma^2\tilde{V}(z)$, where the matrix elements of $\tilde{V}(z)$ satisfy the system*

$$\mu\frac{\mathrm{d}\tilde{v}}{\mathrm{d}z} = \underbrace{\begin{pmatrix} 2a(z) & 0 & 2\varpi(z) \\ 0 & 2a(z) & -2\varpi(z) \\ -\varpi(z) & \varpi(z) & 2a(z) \end{pmatrix}}_{=:\tilde{B}(z)}\tilde{v} + \tilde{E} \tag{B.22}$$

where $\tilde{E} \in \mathbb{R}^3$ with $\mathcal{O}(1)$-components.

It is already apparent from the form of (B.22) that the analysis of the variational equation simplifies. We can now prove Theorem 6.1, which we restate as follows for convenience.

Theorem B.8 *(Theorem 6.1). Suppose* (B.4) *holds. Then the solution $v = (v_1, v_2, v_3)$ for* (B.3) *satisfies the following asymptotics as $\mu \to 0$*

$$v_1 \asymp \frac{1}{|z| + \sqrt{\mu}}, \qquad v_2 \asymp \frac{1}{|z| + \sqrt{\mu}}, \qquad v_3 = \mathcal{O}(1), \qquad (v_1 - v_2) = \mathcal{O}(1) \tag{B.23}$$

for $z_0 + \mathcal{O}(\mu|\log\mu|) \leqslant z \leqslant \sqrt{\mu}$.

Proof. We work in the coordinates provided by Lemma B.7. Summing the first two equations of (B.22) we get

$$\mu\frac{\mathrm{d}}{\mathrm{d}z}(\tilde{v}_1 + \tilde{v}_2) = 2a(z)(\tilde{v}_1 + \tilde{v}_2) + \tilde{e}_1 + \tilde{e}_2 \tag{B.24}$$

and we already know from Proposition B.5 (resp. Lemma B.4) that this yields

$$(\tilde{v}_1 + \tilde{v}_2)(z) \asymp \frac{1}{|z| + \sqrt{\mu}}$$

for $z_0 + \mathcal{O}(\mu|\log\mu|) \leqslant z \leqslant \sqrt{\mu}$. The difference of the first two equations in (B.22) and the third equation can be combined as

$$\mu\frac{\mathrm{d}}{\mathrm{d}z}\begin{pmatrix} \tilde{v}_1 - \tilde{v}_2 \\ \tilde{v}_3 \end{pmatrix} = \begin{pmatrix} 2a(z) & 4\varpi(z) \\ -\varpi(z) & 2a(z) \end{pmatrix}\begin{pmatrix} \tilde{v}_1 - \tilde{v}_2 \\ \tilde{v}_3 \end{pmatrix} + \begin{pmatrix} \tilde{e}_1 - \tilde{e}_2 \\ \tilde{e}_3 \end{pmatrix}. \tag{B.25}$$

Considering (B.25) as a fast–slow system with slow variable z we find that the critical manifold is given by the equation

$$\begin{pmatrix} (\tilde{v}_1 - \tilde{v}_2)^* \\ \tilde{v}_3^* \end{pmatrix} = -\begin{pmatrix} 2a(z) & 4\varpi(z) \\ -\varpi(z) & 2a(z) \end{pmatrix}^{-1}\begin{pmatrix} \tilde{e}_1 - \tilde{e}_2 \\ \tilde{e}_3 \end{pmatrix},$$

which is order 1. Observe that (B.25) undergoes a delayed (or dynamic) Hopf bifurcation at $z = 0$. Thus Neishtadt's theorem on delayed Hopf bifurcations [93] applies, and shows that solutions of the variational equation satisfy

$$(\tilde{v}_1 - \tilde{v}_2)(z) = (\tilde{v}_1 - \tilde{v}_2)^*(z) + \mathcal{O}(\mu), \qquad \tilde{v}_3(z) = v_3^*(z) + \mathcal{O}(\mu)$$

for $z_0 + \mathcal{O}(\mu|\log\mu|) \leqslant z \leqslant \mathcal{O}(1)$. Now the result (B.23) follows from $V(z) = S(z)\tilde{V}(z)S(z)^T$, by writing $\tilde{V}(z)$ as the sum of a leading term proportional to the identity matrix and a remainder of order 1. □

Appendix C. Proof of Theorem 6.2 (staying in covariance tubes)

Applying the transformation of Lemma B.6 to the nonlinear equation (6.12) and dropping the tildes yields the system

$$\mathrm{d}\zeta_z = \frac{1}{\mu}\big[A(z)\zeta_z + b(\zeta_z, z)\big]\,\mathrm{d}z + \frac{\sigma}{\sqrt{\mu}}F(z)\,\mathrm{d}W_z, \tag{C.1}$$

where

$$A(z) = \begin{pmatrix} a(z) & \varpi(z) \\ -\varpi(z) & a(z) \end{pmatrix}, \tag{C.2}$$

and $b(\zeta, z) = \mathcal{O}(\|\zeta\|^2)$. The solution of (C.1) with initial condition $\zeta_{z_0} = 0$ can be written as

$$\zeta_z = \frac{\sigma}{\sqrt{\mu}}\int_{z_0}^{z} U(z,s)F(s)\,\mathrm{d}W_s + \frac{1}{\mu}\int_{z_0}^{z} U(z,s)b(\zeta_s, s)\,\mathrm{d}s =: \zeta_z^0 + \zeta_z^1, \tag{C.3}$$

where $U(z,s)$ denotes the principal solution of the linear time-dependent system $\mu\dot{\zeta} = A(z)\zeta$. Owing to the particular form of $A(z)$, we have the explicit expression

$$U(z,s) = \mathrm{e}^{-\alpha(z,s)/\mu}\begin{pmatrix} \cos(\varphi(z,s)/\mu) & \sin(\varphi(z,s)/\mu) \\ -\sin(\varphi(z,s)/\mu) & \cos(\varphi(z,s)/\mu) \end{pmatrix}, \tag{C.4}$$

where

$$\alpha(z,s) = \int_s^z -a(u)\,\mathrm{d}u, \qquad \varphi(z,s) = \int_s^z \varpi(u)\,\mathrm{d}u. \tag{C.5}$$

Note in particular that since $-a(z) \asymp x(z) \asymp -z$ near $z = 0$, we have $\alpha(z,s) \asymp s^2 - z^2$.

For a two-by-two matrix M, let $\|M\|$ denote its L^2-operator norm, i.e., $\|M\|^2$ is the largest eigenvalue of MM^T.

Lemma C.1. *Let*

$$\Theta(z) = \frac{1}{\mu}\int_{z_0}^{z} \big\|U(z,s)\big\|\,\mathrm{d}s. \tag{C.6}$$

Then

$$\Theta(z) = \mathcal{O}\biggl(\frac{1}{|z| + \sqrt{\mu}}\biggr) \tag{C.7}$$

for all $z_0 \leqslant z \leqslant \sqrt{\mu}$.

Proof. Since $U(z,s)U(z,s)^T = \mathrm{e}^{-2\alpha(z,s)/\mu}\,\mathrm{Id}$, we have

$$\Theta(z) = \frac{1}{\mu}\int_{z_0}^{z} \mathrm{e}^{-2\alpha(z,s)/\mu}\,\mathrm{d}s, \tag{C.8}$$

and the result follows as in the proof of Lemma B.4. □

The next lemma provides bounds on the norms of $\bar{V}(z)$ and $\bar{V}(z)^{-1}$.

Lemma C.2. *Let*

$$K_+(z)^2 = \bigl\|\bar{V}(z)\bigr\|, \qquad K_-(z)^2 = \bigl\|\bar{V}(z)^{-1}\bigr\|. \tag{C.9}$$

Then

$$K_+(z)^2 = \mathcal{O}\biggl(\frac{1}{|z| + \sqrt{\mu}}\biggr), \qquad K_-(z)^2 = \mathcal{O}\bigl(|z| + \sqrt{\mu}\bigr) \tag{C.10}$$

for all $z_0 \leqslant z \leqslant \sqrt{\mu}$.

Proof. Note that

$$\bar{V}(z)\bar{V}(z)^T = \begin{pmatrix} \bar{v}_1 & \bar{v}_3 \\ \bar{v}_3 & \bar{v}_2 \end{pmatrix}^2 = \begin{pmatrix} \bar{v}_1^2 + \bar{v}_3^2 & (\bar{v}_1 + \bar{v}_2)\bar{v}_3 \\ (\bar{v}_1 + \bar{v}_2)\bar{v}_3 & \bar{v}_2^2 + \bar{v}_3^2 \end{pmatrix} \tag{C.11}$$

has eigenvalues given by

$$\frac{1}{2}\bigl[\bar{v}_1^2 + \bar{v}_2^2 + 2\bar{v}_3^2 \pm (\bar{v}_1 + \bar{v}_2)\sqrt{(\bar{v}_1 - \bar{v}_2)^2 + 4\bar{v}_3^2}\,\bigr]. \tag{C.12}$$

The larger eigenvalue is equal to $K_+(z)^4$, while the smaller one is equal to $K_-(z)^{-4}$. The result thus follows from the bounds obtained in Theorem B.8. □

We can now prove a local version of Theorem 6.2, on a small interval $[s,t] \subset [z_0, z]$.

Proposition C.3. *Fix times* $z_0 \leqslant s < t \leqslant z$ *such that* $\alpha(t,s) \leqslant \mathcal{O}(\mu)$. *Then for any* $0 < \gamma < 1$,

$$\begin{aligned}
&\mathbb{P}\Bigl\{\sup_{s \leqslant u \leqslant t} \bigl\langle \zeta_u, \bar{V}(u)^{-1}\zeta_u\bigr\rangle \geqslant r^2\Bigr\} \\
&\quad\leqslant \frac{1}{1-\gamma}\exp\biggl\{-\frac{\gamma r^2}{2\sigma^2}\biggl[1 - \mathcal{O}\biggl(\bigl(|s| + \sqrt{\mu}\bigr)\frac{t-s}{\mu}\biggr) - \mathcal{O}\biggl(\frac{r}{(|t| + \sqrt{\mu})^{3/2}}\biggr)\biggr]\biggr\}.
\end{aligned} \tag{C.13}$$

Proof. The proof is adapted from [25, Section 5.2], and we use almost the same notations as there. Let $\Upsilon_u = U(s,u)\zeta_u$. Then

$$\bigl\langle \zeta_u, \bar{V}(u)^{-1}\zeta_u\bigr\rangle = \bigl\langle \Upsilon_u, \underbrace{U(u,s)^T \bar{V}(u)^{-1} U(u,s)}_{=:Q_s(u)^2}\Upsilon_u\bigr\rangle = \bigl\| Q_s(u)\Upsilon_u\bigr\|^2. \tag{C.14}$$

We can again decompose $\Upsilon_u = \Upsilon_u^0 + \Upsilon_u^1$, where

$$\begin{aligned}\Upsilon_u^0 &= \frac{\sigma}{\sqrt{\mu}}\int_{z_0}^{u} U(s,v)F(v)\,\mathrm{d}W_v,\\ \Upsilon_u^1 &= \frac{1}{\mu}\int_{z_0}^{u} U(s,v)b(\zeta_v,v)\,\mathrm{d}v.\end{aligned} \tag{C.15}$$

The process Υ_u^0 is a Gaussian martingale. Lemma 5.1.8 in [25] can thus be applied and provides the bound

$$\mathbb{P}\Bigl\{\sup_{s\leqslant u\leqslant t}\bigl\| Q_s(t)\Upsilon_t^0\bigr\| \geqslant R_0\Bigr\} \leqslant \frac{1}{1-\gamma}\exp\Bigl\{-\gamma\frac{R_0^2}{2\sigma^2}\Bigr\}. \tag{C.16}$$

For this bound to be useful, we need to show that $Q_s(u)$ and $Q_s(t)$ are close to each other. Observe that

$$\mu\frac{\mathrm{d}}{\mathrm{d}u}Q_s(u)^{-2} = U(s,u)F(u)F(u)^T U(s,u)^T, \tag{C.17}$$

as a consequence of the definition of $U(s,u)$ and the differential equation satisfied by $\bar{V}(u)$. Integrating from u to t and multiplying on the left by $Q_s(u)^2$, we get

$$Q_s(u)^2 Q_s(t)^{-2} - \mathrm{Id} = Q_s(u)^2\frac{1}{\mu}\int_u^t U(s,v)F(v)F(v)^T U(s,v)^T\,\mathrm{d}v. \tag{C.18}$$

Now $\|U(s,v)\| \leqslant \mathcal{O}(1)$ owing to the assumption $\alpha(t,s) = \mathcal{O}(\mu)$. Thus the integral has order $t-u \leqslant t-s$. Furthermore,

$$\bigl\| Q_s(u)^2\bigr\| \leqslant \bigl\|U(u,s)\bigr\|^2 K_-(u)^2 = \mathcal{O}\bigl(K_-(u)^2\bigr). \tag{C.19}$$

As a consequence, we get

$$Q_s(u)^2 = Q_s(t)^2\biggl[\mathrm{Id} + \mathcal{O}\Bigl(K_-(s)^2\frac{t-s}{\mu}\Bigr)\biggr]. \tag{C.20}$$

Thus there exists an $R = r[1-\mathcal{O}(K_-(s)^2(t-s)/\mu)]$ such that

$$\mathbb{P}\Bigl\{\sup_{s\leqslant u\leqslant t}\bigl\langle\zeta_u,\bar{V}(u)^{-1}\zeta_u\bigr\rangle \geqslant r^2\Bigr\} \leqslant \mathbb{P}\Bigl\{\sup_{s\leqslant u\leqslant t\wedge\tau_{\mathcal{B}(r)}}\bigl\| Q_s(t)\Upsilon_u\bigr\| \geqslant R\Bigr\}. \tag{C.21}$$

For any decomposition $R = R_0 + R_1$ with $R_0, R_1 > 0$, we can bound the above probability by $P_0 + P_1$, where

$$P_i = \mathbb{P}\Big\{ \sup_{s \leqslant u \leqslant t \wedge \tau_{\mathcal{B}(r)}} \big\| Q_s(t) \Upsilon_u^i \big\| \geqslant R_i \Big\}, \quad i = 0, 1. \tag{C.22}$$

P_0 has already been estimated in (C.16). We want to choose R_1 in such a way that $P_1 = 0$. For any $u \leqslant t \wedge \mathcal{B}(r)$, we have

$$\begin{aligned} \big\| Q_s(t) \Upsilon_u^1 \big\| &\leqslant \text{const} \sup_{u \in [s,t]} K_-(u) \frac{1}{\mu} \int_{z_0}^{u} \big\| U(s,v) b(\zeta_v) \big\| \, \mathrm{d}v \\ &\leqslant \text{const} \sup_{u \in [s,t]} K_-(u) \Theta(u) \sup_{v \in [z_0, t \wedge \tau_{\mathcal{B}(r)}]} \|\zeta_v\|^2 \\ &\leqslant \text{const} \sup_{u \in [s,t]} K_-(u) \Theta(u) \sup_{v \in [z_0,t]} K_+(v)^2 r^2 \\ &\leqslant \text{const} \big(\sqrt{\mu} + |t| \big)^{-3/2} r^2. \end{aligned} \tag{C.23}$$

We can thus achieve $P_1 = 0$ by setting R_1 to be a sufficiently large constant times $(\sqrt{\mu} + |t|)^{-3/2} r^2$. This determines R_0, and the result then follows from (C.16). □

We can now complete the proof of Theorem 6.2, which will follow directly from

Theorem C.4. *There exist constants $\Delta_0, r_0, \mu_0 > 0$ such that for all $0 < \Delta < \Delta_0$, all $0 < \sigma < r < r_0 \mu^{3/4}$, $0 < \mu < \mu_0$ and all $0 < \gamma < 1$,*

$$\mathbb{P}\{\tau_{\mathcal{B}(r)} < z\} \leqslant C_+(z, z_0) \exp\Big\{ -\gamma \frac{r^2}{2\sigma^2} \big[1 - \mathcal{O}(\Delta) - \mathcal{O}\big(r \mu^{-3/4} \big) \big] \Big\} \tag{C.24}$$

holds for all $z \leqslant \sqrt{\mu}$, where

$$C_+(z, z_0) = \frac{\text{const}}{(1-\gamma)\Delta\mu} \int_{z_0}^{z} x_s^{\text{det}} \, \mathrm{d}s. \tag{C.25}$$

Proof. Let $z_0 = s_0 < s_1 < \cdots < s_N = z$ be a partition of $[z_0, z]$. Then

$$\mathbb{P}\{\tau_{\mathcal{B}(r)} < s\} \leqslant \sum_{k=1}^{N} P_k, \tag{C.26}$$

where

$$P_k = \mathbb{P}\Big\{ \sup_{s_{k-1} \leqslant u \leqslant s_k} \big\langle \zeta_u, \bar{V}(u)^{-1} \zeta_u \big\rangle \geqslant r^2 \Big\} \tag{C.27}$$

can be estimated by Proposition C.3. We want to choose the partition in such a way that the error terms in P_k are bounded uniformly in k. A convenient choice is to define the s_k by

$$\begin{cases} \alpha(s_{k+1}, s_k) = \Delta\mu & \text{when } s_{k+1} < -\sqrt{\mu}, \\ s_{k+1} - s_k = \Delta\sqrt{\mu} & \text{when } |s_{k+1}| \leqslant \sqrt{\mu}. \end{cases} \tag{C.28}$$

Using the fact that $\alpha(t,s) \asymp |s+t|(t-s)$ and applying Proposition C.3, one indeed checks that

$$P_k \leqslant \frac{1}{1-\gamma} \exp\left\{ -\frac{\gamma r^2}{2\sigma^2} \big[1 - \mathcal{O}(\Delta) - \mathcal{O}\big(r\mu^{-3/4}\big) \big] \right\} \quad \forall k = 1, \dots, N, \tag{C.29}$$

where the error terms are uniform in k. It remains to estimate the number N of elements of the partition, which will give the prefactor C_+. In the case $z \leqslant -\sqrt{\mu}$, we simply have

$$\alpha(z, z_0) = N\Delta\mu \quad \Rightarrow \quad N = \left\lceil \frac{\alpha(z, z_0)}{\Delta\mu} \right\rceil. \tag{C.30}$$

In the case $-\sqrt{\mu} \leqslant z \leqslant \sqrt{\mu}$, we have

$$N = \left\lceil \frac{\alpha(z, z_0)}{\Delta\mu} \right\rceil + \left\lceil \frac{z - (-\sqrt{\mu})}{\Delta\sqrt{\mu}} \right\rceil, \tag{C.31}$$

and the result follows from the fact that $\alpha(z, -\sqrt{\mu}) \asymp \sqrt{\mu}(z - (-\sqrt{\mu}))$. □

Theorem 6.2 is just a reformulation of this result, in which we have chosen $\gamma = 1 - \sigma^2/r^2$.

Appendix D. Proof of Theorem 6.4 (early jumps)

We consider again the equation for the difference ζ_z between stochastic sample paths and a deterministic reference solution, this time given by the weak canard. In canonical form, we have

$$\mathrm{d}\zeta_z = \frac{1}{\mu}\big[A(z)\zeta_z + b(\zeta_z, z) \big] \mathrm{d}z + \frac{\sigma}{\sqrt{\mu}} F(z)\, \mathrm{d}W_z, \tag{D.1}$$

where

$$A(z) = \begin{pmatrix} a(z) & \varpi(z) \\ -\varpi(z) & a(z) \end{pmatrix}, \qquad a(z) = 2z + \mathcal{O}\big(\mu^2\big), \tag{D.2}$$

and $b(\zeta, z) = \mathcal{O}(\|\zeta\|^2)$. The proof is split into several parts. In Section D.1, we show that sample paths are likely to leave a neighborhood of order slightly (that is, logarithmically) larger than σ/z of the weak canard in a time z of order $\sqrt{\mu|\log\mu|}$. Section D.2 analyzes the dynamics in a larger neighborhood of the weak canard, in which the drift term dominates. Section D.3 combines the two results to prove the main theorem.

D.1. Diffusion-dominated escape

We assume from now on that $\sigma \ll \mu^{3/4}$, because otherwise stochastic sample paths are no longer localized near deterministic solutions when $z = \sqrt{\mu}$. We define the set

$$\mathcal{S}(h) = \big\{ (\zeta, z) \colon z \geqslant \sqrt{\mu},\ \|\zeta\| < h\hat{\rho}(z) \big\}, \tag{D.3}$$

where

$$\hat{\rho}(z)^2 = \frac{\mathrm{Tr}(F(z)F(z)^T)}{4z}. \tag{D.4}$$

The following result is an adaptation of [22, Proposition 4.7] to the two-dimensional case.

Proposition D.1. *Let $h, \nu > 0$, with ν of order 1, satisfy the conditions*

$$\frac{\sigma}{h} \leqslant c_0 \quad \text{and} \quad \left(\frac{h}{\sigma}\right)^{2+\nu} \left[\log\left(1+\nu+\frac{h^2}{\sigma^2}\right)\right]^{1/2} \leqslant c_1 \frac{\mu^{3/4}}{\sigma} \tag{D.5}$$

for some $c_0, c_1 > 0$. If c_0 and c_1 are small enough, then there exist $T > 0$ and $C(\nu) > 0$ such that for any $(z_0, \zeta_0) \in \mathcal{S}(h)$ with $z_0 < T$,

$$\mathbb{P}^{(\zeta_0, z_0)}\{\tau_{\mathcal{S}(h)} \geqslant z\} \leqslant C(\nu)\left(\frac{h}{\sigma}\right)^{2\nu} \exp\left\{-\kappa(\nu)\frac{z^2 - z_0^2}{\mu}\right\} \tag{D.6}$$

holds for all $\sqrt{\mu} \leqslant z_0 \leqslant z \leqslant T$, where

$$\kappa(\nu) = \frac{2\nu}{1+\nu}\left[1 - \mathcal{O}\left(\frac{1}{\nu \log(h/\sigma)}\right)\right]. \tag{D.7}$$

We shall choose the value of the parameter ν later on, while h will be taken of the form $h = c\sigma|\log\sigma|$. Then condition (D.5) reduces to

$$\sigma|\log\sigma|^{2+\nu}\sqrt{\log|\log\sigma|} \leqslant \mathcal{O}\big(\mu^{3/4}\big), \tag{D.8}$$

which is slightly stronger than $\sigma \ll \mu^{3/4}$. The exponent $\kappa(\nu)$ in (D.7) becomes optimal in the limit $\nu \to \infty$, but condition (D.8) becomes more stringent as ν grows large. But we have to choose a finite ν anyway.

Proof of Proposition D.1. Let $\alpha(t,s) = \int_s^t a(u)\,\mathrm{d}u$. We define a partition $z_0 = s_0 < s_1 < \cdots < s_N = z$ of $[z_0, z]$ by

$$\alpha(s_k, s_{k-1}) = \Delta\mu, \quad k = 1, \dots, N-1, \tag{D.9}$$

where $\Delta > 0$ will be chosen later. The Markov property implies that

$$\mathbb{P}\{\tau_{\mathcal{S}(h)} \geqslant z\} \leqslant \prod_{k=1}^{N-1} P_k, \tag{D.10}$$

where

$$P_k = \sup_{\zeta:\, \|\zeta\| \leqslant h\hat{\rho}(s_{k-1})} \mathbb{P}^{\zeta, s_{k-1}}\left\{\sup_{s_{k-1} \leqslant s \leqslant s_k} \frac{\|\zeta_s\|}{\hat{\rho}(s)} < h\right\}. \tag{D.11}$$

We shall derive a uniform bound $q(\Delta)$ for all P_k, $1 \leqslant k \leqslant N-1$. Then (D.10) and the definition (D.9) of the partition imply

$$\mathbb{P}\{\tau_{\mathcal{S}(h)} \geqslant z\} \leqslant q(\Delta)^{-1} \exp\left\{-\frac{\alpha(z, z_0)}{\mu}\frac{\log q(\Delta)^{-1}}{\Delta}\right\}, \tag{D.12}$$

and the result will follow from an appropriate choice of Δ.

The process ζ_s starting at time s_{k-1} in ζ can be decomposed as $\zeta_s = \zeta_s^{k,0} + \zeta_s^{k,1}$, with

$$\zeta_s^{k,0} = U(s, s_{k-1})\zeta + \frac{\sigma}{\sqrt{\mu}} \int_{s_{k-1}}^{s} U(s,u)F(u)\,\mathrm{d}W_u,$$
$$\zeta_s^{k,1} = \frac{1}{\mu} \int_{s_{k-1}}^{s} U(s,u)b(\zeta_u,u)\,\mathrm{d}u, \tag{D.13}$$

where $U(s,u)$ is the principal solution of the linear system, given in (C.4). For any decomposition $h = H_0 - H_1$, we have

$$P_k \leqslant \sup_{\zeta:\, \|\zeta\| \leqslant h\hat{\rho}(s_{k-1})} \bigl[P_{k,0}(\zeta, H_0) + P_{k,1}(\zeta, H_1)\bigr], \tag{D.14}$$

where

$$P_{k,0}(\zeta, H_0) = \mathbb{P}^{\zeta, s_{k-1}}\biggl\{ \sup_{s_{k-1} \leqslant s \leqslant s_k} \frac{\|\zeta_s^{k,0}\|}{\hat{\rho}(s)} < H_0 \biggr\},$$
$$P_{k,1}(\zeta, H_1) = \mathbb{P}^{\zeta, s_{k-1}}\biggl\{ \sup_{s_{k-1} \leqslant s \leqslant s_k} \frac{\|\zeta_s^{k,1}\|}{\hat{\rho}(s)} \geqslant H_1,\ \sup_{s_{k-1} \leqslant s \leqslant s_k} \frac{\|\zeta_s\|}{\hat{\rho}(s)} < h \biggr\}. \tag{D.15}$$

We start by bounding $P_{k,0}(\zeta, H_0)$, using the endpoint estimate

$$P_{k,0}(\zeta, H_0) \leqslant \mathbb{P}^{\zeta, s_{k-1}}\bigl\{ \bigl\|\zeta_{s_k}^{k,0}\bigr\| < H_0 \hat{\rho}(s_k) \bigr\} \leqslant \frac{\pi H_0^2 \hat{\rho}(s_k)^2}{\sqrt{(2\pi)^2 \det \mathrm{Cov}(\zeta_{s_k}^{k,0})}}. \tag{D.16}$$

The last inequality follows from the fact that the random variable $\zeta_{s_k}^{k,0}$ is Gaussian, and we have bounded its density by the normalizing constant. We denote the diagonal matrix elements of $V = \sigma^{-2}\,\mathrm{Cov}(\zeta_{s_k}^{k,0})$ by v_1 and v_2, and the off-diagonal element by v_3. Then

$$\det \mathrm{Cov}\bigl(\zeta_{s_k}^{k,0}\bigr) = \sigma^4\bigl(v_1 v_2 - v_3^2\bigr) = \sigma^4\biggl[\frac{1}{4}\bigl((\mathrm{Tr}\,V)^2 - (v_1 - v_2)^2\bigr) - v_3^2\biggr]. \tag{D.17}$$

As already remarked in the proof of Theorem B.8, the quantities $v_1 - v_2$ and v_3 remain of order 1 up to some $z = T$ of order 1, as a consequence of Neishtadt's result on delayed Hopf bifurcations. In order to estimate the trace $\mathrm{Tr}\,V$, we use the fact that $\hat{\rho}(z)$ is decreasing in $(0, T]$ for T small enough, owing to the fact that $\mathrm{Tr}(F(z)F(z)^T)$ is bounded below by a positive constant, and has a derivative bounded in absolute value. Thus we have (cf. (B.24))

$$\begin{aligned}
\mathrm{Tr}\,V &= \frac{1}{\mu} \int_{s_{k-1}}^{s_k} \mathrm{e}^{2\alpha(s_k,u)/\mu}\,\mathrm{Tr}\bigl(F(u)F(u)^T\bigr)\,\mathrm{d}u \\
&= \mathrm{e}^{2\Delta} \int_{s_{k-1}}^{s_k} \frac{4u}{\mu} \mathrm{e}^{-2\alpha(u,s_{k-1})/\mu} \hat{\rho}(u)^2\,\mathrm{d}u \\
&\geqslant \hat{\rho}(s_k)^2\bigl[\mathrm{e}^{2\Delta} - 1\bigr].
\end{aligned} \tag{D.18}$$

Substituting (D.18) in (D.17) and then in (D.16) yields

$$P_{k,0}(\zeta, H_0) \leqslant \frac{H_0^2}{\sigma^2} \frac{1}{\mathrm{e}^{2\Delta}-1}\biggl[1+\mathcal{O}\biggl(\frac{1}{\mathrm{e}^{4\Delta}\hat{\rho}(s_k)^4}\biggr)\biggr]. \tag{D.19}$$

Next we estimate $P_{k,1}$. We first obtain the bound

$$\begin{aligned}\bigl\|\zeta^{k,1}_{s\wedge\tau_{\mathcal{S}(h)}}\bigr\| &\leqslant \frac{1}{\mu}\int_{s_{k-1}}^{s\wedge\tau_{\mathcal{S}(h)}} \bigl\|U(s,u)\bigr\|\bigl\|b(\zeta_u,u)\bigr\|\,\mathrm{d}u\\ &\leqslant \mathrm{const}\,\frac{1}{\mu}\int_{s_{k-1}}^{s\wedge\tau_{\mathcal{S}(h)}} \mathrm{e}^{\alpha(s,u)/\mu}h^2\hat{\rho}(u)^2\,\mathrm{d}u\\ &\leqslant \mathrm{const}\,h^2\frac{\hat{\rho}(s_{k-1})^2}{2s_{k-1}}\mathrm{e}^{\Delta},\end{aligned} \tag{D.20}$$

where we have used the fact that the function $u\mapsto\hat{\rho}(u)^2/2u$ is decreasing, and bounded the integral of $(2u/\mu)\mathrm{e}^{\alpha(s,u)/\mu}$ by e^{Δ}. Using a Taylor expansion of $\hat{\rho}(s)^2$ and the definitions of $\hat{\rho}$ and of the partition, one finds

$$\frac{\hat{\rho}(s_{k-1})^2}{\hat{\rho}(s)^2} \leqslant 1+\frac{s-s_{k-1}}{\hat{\rho}(s)^2}\sup_{u\in[s_{k-1},s]}\bigl|\bigl(\hat{\rho}(u)^2\bigr)'\bigr| \leqslant 1+\mathcal{O}(\Delta) \tag{D.21}$$

for all $s\in[s_{k-1},s_k]$. Together with (D.20), this implies

$$\frac{\|\zeta^{k,1}_{s\wedge\tau_{\mathcal{S}(h)}}\|}{\hat{\rho}(s)} \leqslant \mathrm{const}\,h^2\frac{\hat{\rho}(s_{k-1})}{s_{k-1}}\sqrt{\Delta}\mathrm{e}^{\Delta} \leqslant \mathrm{const}\frac{h^2\sqrt{\Delta}\mathrm{e}^{\Delta}}{\mu^{3/4}} =: \frac{H_1}{2}, \tag{D.22}$$

which yields $P_{k,1}(\zeta,H_1)=0$. Substituting $H_0=h+H_1$ in (D.19) thus yields

$$P_k \leqslant q(\Delta) := \frac{h^2}{\sigma^2}\frac{1}{\mathrm{e}^{2\Delta}-1}\biggl[1+\mathcal{O}\biggl(\frac{1}{\mathrm{e}^{4\Delta}}\biggr)+\mathcal{O}\biggl(\frac{h\sqrt{\Delta}\mathrm{e}^{\Delta}}{\mu^{3/4}}\biggr)\biggr] \tag{D.23}$$

for $k=1,\dots,N-1$. Finally, we make the choice

$$\Delta = \frac{1+\nu}{2}\log\biggl(1+\nu+\frac{h^2}{\sigma^2}\biggr). \tag{D.24}$$

Bounding P_k above amounts to bounding $q(\Delta)^{-1}$ below. For this we write

$$\begin{aligned}q(\Delta)^{-1} &= \frac{\sigma^2}{h^2}\mathrm{e}^{2\Delta}\biggl[1-\mathcal{O}\bigl(\mathrm{e}^{-2\Delta}\bigr)-\mathcal{O}\bigl(\mathrm{e}^{-4\Delta}\bigr)-\mathcal{O}\biggl(\frac{h\sqrt{\Delta}\mathrm{e}^{\Delta}}{\mu^{3/4}}\biggr)\biggr]\\ &\geqslant \biggl(1+\nu+\frac{h^2}{\sigma^2}\biggr)^{\nu}\biggl[1-\mathcal{O}\biggl(\biggl(\frac{\sigma^2}{h^2}\biggr)^{1+\nu}\biggr)\\ &\quad -\mathcal{O}\biggl(\frac{h}{\mu^{3/4}}\biggl(\frac{h}{\sigma}\biggr)^{1+\nu}\log\biggl(1+\nu+\frac{h^2}{\sigma^2}\biggr)^{1/2}\biggr)\biggr].\end{aligned} \tag{D.25}$$

Note that the error term $\mathcal{O}(\mathrm{e}^{-4\Delta})$ is negligible and no longer appears in the last line. Now by assumption (D.5), for c_0 and c_1 small enough we get

$$q(\Delta)^{-1} \geqslant \frac{1}{2}\left(1+\nu+\frac{h^2}{\sigma^2}\right)^{\nu} \tag{D.26}$$

and

$$\frac{\log q(\Delta)^{-1}}{\Delta} \geqslant \frac{2\nu}{1+\nu} - \frac{2\log 2}{(1+\nu)\log(1+\nu+h^2/\sigma^2)} =: \kappa(\nu). \tag{D.27}$$

Note that $\kappa(\nu)$ is indeed of the form (D.7). The result thus follows from (D.12). □

D.2. Averaging and drift-dominated escape

We consider again Eq. (D.1), but this time for slightly larger values of $\|\zeta\|$. We start by transforming the system to polar coordinates.

Lemma D.2. *Consider a system of the form*

$$\begin{aligned} \mathrm{d}\xi &= \frac{1}{\mu} f_\xi(\xi,\eta,z)\,\mathrm{d}z + \frac{\sigma}{\sqrt{\mu}} F_\xi(z)\,\mathrm{d}W_z, \\ \mathrm{d}\eta &= \frac{1}{\mu} f_\eta(\xi,\eta,z)\,\mathrm{d}z + \frac{\sigma}{\sqrt{\mu}} F_\eta(z)\,\mathrm{d}W_z, \end{aligned} \tag{D.28}$$

where W_z denotes a two-dimensional Wiener process, and F_ξ and F_η are row vectors of dimension 2. Then in polar coordinates ($\xi = r\cos\varphi$, $\eta = r\sin\varphi$) the system becomes

$$\begin{aligned} \mathrm{d}r &= \frac{1}{\mu} f_r(r,\varphi,z)\,\mathrm{d}z + \frac{\sigma}{\sqrt{\mu}} F_r(\varphi,z)\,\mathrm{d}W_z, \\ \mathrm{d}\varphi &= \frac{1}{\mu}\frac{1}{r} f_\varphi(r,\varphi,z)\,\mathrm{d}z + \frac{\sigma}{\sqrt{\mu}}\frac{1}{r} F_\varphi(\varphi,z)\,\mathrm{d}W_z, \end{aligned} \tag{D.29}$$

where the new and old diffusion coefficients are related by

$$\begin{aligned} F_r &= F_\xi\cos\varphi + F_\eta\sin\varphi, \\ F_\varphi &= F_\eta\cos\varphi - F_\xi\sin\varphi, \end{aligned} \tag{D.30}$$

while the drift coefficients are given by

$$\begin{aligned} f_r &= f_\xi\cos\varphi + f_\eta\sin\varphi + \frac{\sigma^2}{2r} F_\varphi F_\varphi^T, \\ f_\varphi &= f_\eta\cos\varphi - f_\xi\sin\varphi - \frac{\sigma^2}{r} F_r F_\varphi^T. \end{aligned} \tag{D.31}$$

Proof. The formulas can be checked directly by applying Itô's formula to (D.29). □

Applying this result to (D.1), we obtain a system of the form

$$\begin{aligned}
\mathrm{d}r &= \frac{1}{\mu}\left[a(z)r + r^2 b_r(\varphi, z) + \mathcal{O}\left(\frac{\sigma^2}{r}\right)\right]\mathrm{d}z + \frac{\sigma}{\sqrt{\mu}} F_r(\varphi, z)\,\mathrm{d}W_z,\\
\mathrm{d}\varphi &= \frac{1}{\mu}\left[-\varpi(z) + r b_\varphi(\varphi, z) + \mathcal{O}\left(\frac{\sigma^2}{r^2}\right)\right]\mathrm{d}z + \frac{\sigma}{\sqrt{\mu}}\frac{1}{r} F_\varphi(\varphi, z)\,\mathrm{d}W_z.
\end{aligned} \tag{D.32}$$

Note that the functions b_r and b_φ do not depend on r, owing to the fact that the nonlinearity in the original equation is homogeneous of degree 2. Another important observation is that the average of b_r (and b_φ) over φ is zero. This follows again from homogeneity, combined with (D.31). This observation suggests to simplify (D.32) by an averaging transformation.

Proposition D.3. *There exists a function $w(\varphi, z)$, which is bounded, smooth, and 2π-periodic in φ, such that $\bar{r} = r + r^2 w(\varphi, z)$ satisfies the SDE*

$$\mathrm{d}\bar{r} = \frac{1}{\mu}\big[a(z)\bar{r} + \beta(\bar{r}, \varphi, z)\big]\mathrm{d}z + \frac{\sigma}{\sqrt{\mu}}\tilde{F}_r(\bar{r}, \varphi, z)\,\mathrm{d}W_z, \tag{D.33}$$

where

$$\begin{aligned}
\beta(\bar{r}, \varphi, z) &= \mathcal{O}\big(\bar{r}^3\big) + \mathcal{O}\big(\mu\bar{r}^2\big) + \mathcal{O}\left(\frac{\sigma^2}{\bar{r}}\right),\\
\tilde{F}_r(\bar{r}, \varphi, z) &= F_r(\varphi, z) + \mathcal{O}(\bar{r}).
\end{aligned} \tag{D.34}$$

Proof. Using Itô's formula and the fact that $r = \bar{r} - \bar{r}^2 w(\varphi, z) + \mathcal{O}(\bar{r}^3)$, we obtain

$$\begin{aligned}
\mathrm{d}\bar{r} = \frac{1}{\mu}\Bigg[a(z)\bar{r} + \bar{r}^2\left(2z w(\varphi, z) + b_r(\varphi, z) - \varpi(z)\frac{\partial w}{\partial \varphi} + \mu\frac{\partial w}{\partial z}\right) + \mathcal{O}\big(\bar{r}^3\big) + \mathcal{O}\left(\frac{\sigma^2}{\bar{r}}\right)\Bigg]\mathrm{d}z\\
+ \frac{\sigma}{\sqrt{\mu}}\left[F_r(\varphi, z) + \bar{r}\left(2w(\varphi, z)F_r(\varphi, z) + \frac{\partial w}{\partial \varphi}F_\varphi(\varphi, z)\right) + \mathcal{O}\big(\bar{r}^2\big)\right]\mathrm{d}W_z.
\end{aligned} \tag{D.35}$$

It is thus sufficient to show that the equation

$$\frac{\partial w}{\partial \varphi}(\varphi, z) = \frac{a(z)}{\varpi(z)} w(\varphi, z) + \frac{b_r(\varphi, z)}{\varpi(z)} \tag{D.36}$$

admits a bounded, 2π-periodic solution. Letting $c = c(z) = a(z)/\varpi(z)$, the general solution of (D.36) can be written

$$w(\varphi, z) = w(0, z)\mathrm{e}^{c\varphi} + \int_0^\varphi \mathrm{e}^{c(\varphi - \varphi')}\frac{b_r(\varphi', z)}{\varpi(z)}\,\mathrm{d}\varphi'. \tag{D.37}$$

Thus choosing

$$w(0,z) = \frac{\mathrm{e}^{2\pi c}}{1-\mathrm{e}^{2\pi c}} \int_0^{2\pi} \mathrm{e}^{-c\varphi} \frac{b_r(\varphi,z)}{\varpi(z)}\,\mathrm{d}\varphi, \tag{D.38}$$

the resulting $w(\varphi,z)$ is indeed 2π-periodic. Finally note that

$$\lim_{z\to 0} w(0,z) = \frac{1}{2\pi} \int_0^{2\pi} \varphi \frac{b_r(\varphi,0)}{\varpi(0)}\,\mathrm{d}\varphi, \tag{D.39}$$

showing that $w(\varphi,z)$ is also bounded as $z\to 0$. □

We now define the set

$$\mathcal{D}(\eta) = \bigl\{(\bar r,\varphi,z)\colon\ z \geqslant \sqrt{\mu},\ \bar r < \eta\sqrt{z}\,\bigr\}. \tag{D.40}$$

Then the nonlinear term β satisfies, on the set $\mathcal{R} = \mathcal{D}(\eta)\setminus\mathcal{S}(h)$,

$$\frac{|\beta(\bar r,\varphi,z)|}{\bar r z} \leqslant M\biggl(\frac{\bar r^2}{z} + \frac{\mu\bar r}{z} + \frac{\sigma^2}{\bar r^2 z}\biggr) \leqslant M'\biggl(\eta^2 + \eta\mu^{3/4} + \frac{1}{|\log\sigma|^2}\biggr) \tag{D.41}$$

for some constants M, M'. We can thus find, for any $0<\kappa<2$, an $\eta=\eta(\kappa)$ such that (for sufficiently small σ and μ)

$$a(z)\bar r + \beta(\bar r,\varphi,z) \geqslant \kappa z\bar r \tag{D.42}$$

holds in $\mathcal{R}$. Thus we have

$$\mathrm{d}\bar r = \frac{1}{\mu}\bigl[\kappa z\bar r + \tilde\beta(\bar r,\varphi,z)\bigr]\,\mathrm{d}z + \frac{\sigma}{\sqrt{\mu}}\tilde F_r(\bar r,\varphi,z)\,\mathrm{d}W_z, \tag{D.43}$$

where $\tilde\beta(\bar r,\varphi,z)\geqslant 0$ in $\mathcal{R}$, implying

$$\bar r_z \geqslant \bar r_{z_0}\mathrm{e}^{\kappa(z^2-z_0^2)/2\mu} + \frac{\sigma}{\sqrt{\mu}}\int_{z_0}^{z} \mathrm{e}^{\kappa(z^2-s^2)/2\mu}\tilde F_r(\bar r_s,\varphi_s,s)\,\mathrm{d}W_s \tag{D.44}$$

holds as long as the process stays in $\mathcal{R}$.

Proposition D.4. *There exists a constant $\kappa_2>0$ such that for any initial condition $(\bar r_0,\varphi_0,z_0)\in\mathcal{R}$,*

$$\mathbb{P}^{(\bar r_0,\varphi_0,z_0)}\{\tau_{\mathcal{R}} > z\} \leqslant 2\exp\biggl\{-\kappa_2\frac{z^2-z_0^2}{\mu|\log\sigma|}\biggr\}. \tag{D.45}$$

Proof. We introduce a partition $z_0<z_1<\cdots<z_N=z$ of $[z_0,z]$, given by

$$z_{k+1}^2 - z_k^2 = \gamma\mu|\log\sigma| \quad \text{for } 0\leqslant k<N = \biggl\lceil \frac{z^2-z_0^2}{\gamma\mu|\log\sigma|}\biggr\rceil. \tag{D.46}$$

The Markov property implies that

$$\mathbb{P}^{(\bar r_0,\varphi_0,z_0)}\{\tau_{\mathcal R} > z\} \leqslant \prod_{k=1}^{N-1} P_k, \tag{D.47}$$

where

$$P_k = \sup_{\bar r,\varphi:\ (\bar r,\varphi,z_k)\in\mathcal R} P_k(\bar r,\varphi),$$
$$P_k(\bar r,\varphi) = \mathbb{P}^{(\bar r,\varphi,z_k)}\{\tau_{\mathcal R} > z_{k+1}\}. \tag{D.48}$$

Inequality (D.44) (with z_0 replaced by z_k) shows that

$$\bar r_z \geqslant \mathrm{e}^{\kappa(z^2-z_k^2)/2\mu}\left[\bar r_{z_k} + \frac{\sigma}{\sqrt{\mu}} M_z^k\right] \tag{D.49}$$

holds for $z_k \leqslant z \leqslant \tau_{\mathcal R}$, where M_z^k is the martingale

$$M_z^k = \int_{z_k}^{z} \mathrm{e}^{-\kappa(s^2-z_k^2)/2\mu}\tilde F_r(\bar r_s,\varphi_s,s)\,\mathrm{d}W_s. \tag{D.50}$$

It follows that

$$\begin{aligned} P_k(\bar r,\varphi) &\leqslant \mathbb{P}\left\{\bar r + \frac{\sigma}{\sqrt{\mu}} M^k_{z_{k+1}} < \eta\sqrt{z_{k+1}}\,\mathrm{e}^{-\kappa(z_{k+1}^2-z_k^2)/2\mu}\right\} \\ &\leqslant \mathbb{P}\big\{M^k_{z_{k+1}} < -c|\log\sigma|\sqrt{\mu}\,\hat\rho(z_k) + \eta\sqrt{\mu}\,\sigma^{\gamma\kappa/2-1}\sqrt{z_{k+1}}\big\}, \end{aligned} \tag{D.51}$$

where we have used $\bar r \geqslant h\hat\rho(z_k) = c\sigma|\log\sigma|\hat\rho(z_k)$ and $\mathrm{e}^{-\kappa(z_{k+1}^2-z_k^2)/2\mu} = \sigma^{\kappa\gamma/2}$. Choosing $\gamma > 2/\kappa$, we can guarantee that the term $c|\log\sigma|\sqrt{\mu}\hat\rho(z_k)$ dominates.

Since the noise acting on the system is non-degenerate, we may assume the existence of constants $D_+ \geqslant D_- > 0$ such that

$$D_- \leqslant \tilde F_r(\bar r,\varphi,z)\tilde F_r(\bar r,\varphi,z)^T \leqslant D_+. \tag{D.52}$$

Thus the variance of $M^k_{z_{k+1}}$ is bounded above by

$$V_+ = \int_{z_k}^{z_{k+1}} D_+\mathrm{e}^{-\kappa(s^2-z_k^2)/\mu}\,\mathrm{d}s \leqslant D_+\left(-\frac{\mu}{2\kappa z_k}\right)\mathrm{e}^{-\kappa(s^2-z_k^2)/\mu}\Big|_{z_k}^{z_{k+1}} \leqslant \frac{D_+\mu}{2\kappa z_k}. \tag{D.53}$$

A Bernstein-type estimate (cf. Lemma D.8 in Appendix D.4) provides the bound

$$\mathbb{P}\big\{M^k_{z_{k+1}} < -x\big\} \leqslant \mathrm{e}^{-x^2/2V_+}. \tag{D.54}$$

Using (D.53) and (D.54) in (D.51) shows that we may assume $P_k < 1/2$, and the result follows from (D.47) and the definition of N. Note that $\kappa_2 = \log 2/\gamma < (\log 2/2)\kappa$. □

D.3. Laplace transforms

In order to combine the results from the two previous subsections, we will use a lemma based on Laplace transforms. In the following we let $\{x_t\}_{t\geqslant 0}$ be a time-homogeneous $\mathbb{R}^d$-valued Markov process with continuous sample paths. All subsets $A \subset \mathbb{R}^d$ considered below are assumed to have smooth boundary, and to be such that the first-exit time $\tau_A = \inf\{t \geqslant 0\colon\ x_t \notin A\}$ is almost surely finite. The *Laplace transform* of τ_A is the non-decreasing function

$$\mathbb{R} \ni \lambda \mapsto \mathbb{E}^x\big[\mathrm{e}^{\lambda\tau_A}\big] = 1 + \lambda \int_0^\infty \mathbb{P}^x\{\tau_A > t\}\mathrm{e}^{\lambda t}\,\mathrm{d}t \in [0,\infty]. \tag{D.55}$$

Note that $\mathbb{E}^x[\mathrm{e}^{\lambda\tau_A}] \leqslant 1$ for all $\lambda \leqslant 0$. Thus there exists a $\lambda_0 \geqslant 0$ such that $\mathbb{E}^x[\mathrm{e}^{\lambda\tau_A}] < \infty$ for all $\lambda < \lambda_0$.

Lemma D.5. *Choose nested bounded open sets $\mathcal{S}_1 \subset \mathcal{S}_2 \subset \mathcal{D} \subset \mathbb{R}^d$. Let $\mathcal{R} = \mathcal{D} \setminus \mathcal{S}_1$ and consider the Laplace transforms*

$$\begin{aligned} G_{\mathcal{D}}(\lambda) &= \sup_{x\in\mathcal{S}_1} \mathbb{E}^x\big[\mathrm{e}^{\lambda\tau_{\mathcal{D}}}\big],\\ G_{\mathcal{S}}(\lambda) &= \sup_{x\in\mathcal{S}_1} \mathbb{E}^x\big[\mathrm{e}^{\lambda\tau_{\mathcal{S}_2}}\big],\\ G_{\mathcal{R}}(\lambda) &= \sup_{x\in\partial\mathcal{S}_2} \mathbb{E}^x\big[\mathrm{e}^{\lambda\tau_{\mathcal{R}}}\big],\\ Q(\lambda) &= \sup_{x\in\partial\mathcal{S}_2} \mathbb{E}^x\big[1_{\{\tau_{\mathcal{S}_1^c} < \tau_{\mathcal{D}}\}}\mathrm{e}^{\lambda\tau_{\mathcal{R}}}\big]. \end{aligned} \tag{D.56}$$

Let λ be such that $G_{\mathcal{S}}(\lambda)$ and $G_{\mathcal{R}}(\lambda)$ are finite, and assume that $Q(\lambda)G_{\mathcal{S}}(\lambda) < 1$. Then $G_{\mathcal{D}}(\lambda)$ is also finite and satisfies

$$G_{\mathcal{D}}(\lambda) \leqslant \frac{G_{\mathcal{S}}(\lambda)G_{\mathcal{R}}(\lambda)}{1 - Q(\lambda)G_{\mathcal{S}}(\lambda)}. \tag{D.57}$$

Proof. For $x_0 \in \mathcal{S}_1$, one necessarily has $\tau_{\mathcal{S}_2} \leqslant \tau_{\mathcal{D}}$, and thus the strong Markov property implies

$$\begin{aligned} \mathbb{E}^{x_0}\big[\mathrm{e}^{\lambda\tau_{\mathcal{D}}}\big] &= \mathbb{E}^{x_0}\big[\mathrm{e}^{\lambda\tau_{\mathcal{S}_2}}\mathbb{E}^{x_{\tau_{\mathcal{S}_2}}}\big[\mathrm{e}^{\lambda\tau_{\mathcal{D}}}\big]\big]\\ &\leqslant \mathbb{E}^{x_0}\big[\mathrm{e}^{\lambda\tau_{\mathcal{S}_2}}\big] \sup_{x\in\partial\mathcal{S}_2} \mathbb{E}^x\big[\mathrm{e}^{\lambda\tau_{\mathcal{D}}}\big]. \end{aligned} \tag{D.58}$$

Similarly, for $x \in \partial\mathcal{S}_2$, since $\tau_{\mathcal{R}} = \tau_{\mathcal{S}_1^c} \wedge \tau_{\mathcal{D}}$ we have

$$\begin{aligned} \mathbb{E}^x\big[\mathrm{e}^{\lambda\tau_{\mathcal{D}}}\big] &= \mathbb{E}^x\big[1_{\{\tau_{\mathcal{S}_1^c} < \tau_{\mathcal{D}}\}}\mathrm{e}^{\lambda\tau_{\mathcal{S}_1^c}}\mathbb{E}^{x_{\tau_{\mathcal{S}_1^c}}}\big[\mathrm{e}^{\lambda\tau_{\mathcal{D}}}\big]\big] + \mathbb{E}^x\big[1_{\{\tau_{\mathcal{D}} < \tau_{\mathcal{S}_1^c}\}}\mathrm{e}^{\lambda\tau_{\mathcal{D}}}\big]\\ &\leqslant \mathbb{E}^x\big[1_{\{\tau_{\mathcal{S}_1^c} < \tau_{\mathcal{D}}\}}\mathrm{e}^{\lambda\tau_{\mathcal{R}}}\mathbb{E}^{x_{\tau_{\mathcal{S}_1^c}}}\big[\mathrm{e}^{\lambda\tau_{\mathcal{D}}}\big]\big] + \mathbb{E}^x\big[1_{\{\tau_{\mathcal{D}} < \tau_{\mathcal{S}_1^c}\}}\mathrm{e}^{\lambda\tau_{\mathcal{R}}}\big]\\ &\leqslant Q(\lambda)G_{\mathcal{D}}(\lambda) + G_{\mathcal{R}}(\lambda). \end{aligned} \tag{D.59}$$

(D.58) and (D.59) imply

$$G_{\mathcal{D}}(\lambda) \leqslant G_{\mathcal{S}}(\lambda)\big[Q(\lambda)G_{\mathcal{D}}(\lambda) + G_{\mathcal{R}}(\lambda)\big], \tag{D.60}$$

which yields the result. □

We will apply this lemma to sets $\mathcal{S}_1 = \mathcal{S}(h_1)$, $\mathcal{S}_2 = \mathcal{S}(h_2)$ and $\mathcal{D} = \mathcal{D}(\kappa)$, where $h_i = c_i\sigma|\log\sigma|$, $i = 1, 2$, with $0 < c_1 < c_2$. We introduce a new time $t = z^2$, and let x_t be the time-homogeneous Markov process $(\sqrt{t}, \zeta_{\sqrt{t}})$.

Proposition D.1 yields a control of $G_{\mathcal{S}}(\lambda)$ in the following way. The bound (D.6) translates in terms of the new process x_t as

$$\mathbb{P}^{x_0}\{\tau_{\mathcal{S}_2} \geqslant t\} \leqslant C_1 \mathrm{e}^{-\lambda_1 t}, \qquad x_0 = (z_0, \zeta_{z_0}), \tag{D.61}$$

where $C_1 = C(\nu)(h_2/\sigma)^{2\nu}$ and $\lambda_1 = \kappa(\nu)/\mu$. It follows that

$$G_{\mathcal{S}}(\lambda) \leqslant 1 + \lambda \int_0^\infty \mathrm{e}^{\lambda t} \sup_{x_0 \in \mathcal{S}_1} \mathbb{P}^{x_0}\{\tau_{\mathcal{S}_2} \geqslant t\}\,\mathrm{d}t \leqslant 1 + \frac{C_1\lambda}{\lambda_1 - \lambda} \tag{D.62}$$

holds for all $\lambda < \lambda_1 = \kappa(\nu)/\mu$. In a similar way, Proposition D.4 yields

$$G_{\mathcal{R}}(\lambda) \leqslant 1 + \frac{C_2\lambda}{\lambda_2 - \lambda} \tag{D.63}$$

for all $\lambda < \lambda_2 = \kappa_2/\mu|\log\sigma|$, where $C_2 = 2$. It remains to estimate $Q(\lambda)$. Let us first show that $Q(\lambda)$ can be bounded in terms of $Q(0)$.

Lemma D.6. *For all $\lambda < \lambda_2$, one has*

$$Q(\lambda) \leqslant \frac{C_2^{\lambda/\lambda_2}}{1 - \lambda/\lambda_2} Q(0)^{1-\lambda/\lambda_2}. \tag{D.64}$$

Proof. First note that for all $T \geqslant 0$,

$$\mathrm{e}^{\lambda\tau_{\mathcal{R}}} \leqslant \mathrm{e}^{\lambda T} + \lambda \int_T^\infty 1_{\{\tau_{\mathcal{R}} > t\}} \mathrm{e}^{\lambda t}\,\mathrm{d}t. \tag{D.65}$$

Plugging this into the definition of $Q(\lambda)$ yields

$$Q(\lambda) \leqslant \sup_{x \in \partial\mathcal{S}_2} \left[\mathbb{P}^x\{\tau_{\mathcal{S}_1^c} < \tau_{\mathcal{D}}\}\mathrm{e}^{\lambda T} + \lambda \int_T^\infty \mathbb{P}^x\{\tau_{\mathcal{R}} > t\}\mathrm{e}^{\lambda t}\,\mathrm{d}t \right]. \tag{D.66}$$

The first term on the right-hand side is bounded by $Q(0)\mathrm{e}^{\lambda T}$. The second one can be estimated with Proposition D.4, yielding

$$Q(\lambda) \leqslant Q(0)\mathrm{e}^{\lambda T} + \frac{C_2\lambda}{\lambda_2 - \lambda} \mathrm{e}^{-(\lambda_2 - \lambda)T}. \tag{D.67}$$

Optimizing over T, we find that the optimal bound is obtained when $\mathrm{e}^{\lambda_2 T} = C_2/Q(0)$, which yields (D.64). □

Finally, $Q(0)$ can be estimated in a similar way as in the proof of Proposition D.4.

Proposition D.7. *We have*

$$Q(0) = \sup_{x\in\partial\mathcal{S}_2} \mathbb{P}^x\{\tau_{\mathcal{S}_1^c} < \tau_{\mathcal{D}}\} \leqslant \sigma^{\kappa(c_2-c_1)^2|\log\sigma|z_0\hat\rho(z_0)^2/D_+}. \tag{D.68}$$

Proof. Inequality (D.44) and the fact that $\mathrm{e}^{\kappa z^2/2\mu}/\hat\rho(z)$ is increasing for sufficiently small z imply

$$\bar r_z - h_1\hat\rho(z) \geqslant \mathrm{e}^{\kappa(z^2-z_0^2)/2\mu}\left[(h_2-h_1)\hat\rho(z_0) + \frac{\sigma}{\sqrt{\mu}}M_z^0\right] \tag{D.69}$$

for $z \leqslant \tau_{\mathcal{R}}$, where M_z^0 is the martingale introduced in (D.50). The variance of M_z^0 being bounded by $D_+\mu/2\kappa z_0$ (cf. (D.53)), the Bernstein-type inequality of Lemma D.8 allows to write

$$\begin{aligned}\mathbb{P}\{\tau_{\mathcal{S}_1^c} < z\wedge\tau_{\mathcal{R}}\} &= \mathbb{P}\Big\{\inf_{z_0\leqslant s\leqslant z\wedge\tau_{\mathcal{R}}}\big(\bar r_s - h_1\hat\rho(z)\big) \leqslant 0\Big\}\\ &\leqslant \exp\left\{-\kappa\frac{(c_2-c_1)^2}{D_+}|\log\sigma|^2\right\} = \sigma^{\kappa(c_2-c_1)^2|\log\sigma|z_0\hat\rho(z_0)^2/D_+}.\end{aligned} \tag{D.70}$$

Note that the right-hand side of (D.70) does not depend on z. The result thus follows from taking the limit $z\to\infty$. □

Proof of Theorem 6.4. Since $\lambda_2 < \lambda_1$ for sufficiently small σ, we set $\lambda = (1-\theta)\lambda_2$ for a fixed $0<\theta<1$. For this λ we have

$$G_{\mathcal{R}}(\lambda) \leqslant 1 + \frac{C_2}{\theta}, \qquad G_{\mathcal{S}}(\lambda) \leqslant 1 + \frac{C_1\lambda_2}{\lambda_1-\lambda_2} = 1 + \mathcal{O}\left(\frac{1}{|\log\sigma|}\right). \tag{D.71}$$

Furthermore Lemma D.6 and Proposition D.7 yield

$$Q(\lambda) \leqslant \frac{C_2}{\theta}\sigma^{c\theta|\log\sigma|} \tag{D.72}$$

for some constant $c>0$. Thus Lemma D.5 can be applied to show that $G_{\mathcal{D}}(\lambda)$ is finite. Finally, by Markov's inequality,

$$\mathbb{P}^x\{\tau_{\mathcal{D}} \geqslant t\} = \mathbb{P}^x\big\{\mathrm{e}^{\lambda\tau_{\mathcal{D}}} \geqslant \mathrm{e}^{\lambda t}\big\} \leqslant \mathrm{e}^{-\lambda t}G_{\mathcal{D}}(\lambda), \tag{D.73}$$

which gives the theorem when translated back to the process ζ_z. □

D.4. A Gaussian tail estimate for martingales

Let W_z be the two-dimensional Brownian motion, and consider the martingale

$$M_z = \int_0^z g(X_s,s)\,\mathrm{d}W_s = \sum_{i=1}^n\int_0^z g_i(X_s,s)\,\mathrm{d}W_s^{(i)}, \tag{D.74}$$

where $g=(g_1,\dots,g_n)$ takes values in $\mathbb{R}^n$ and the process X_z is assumed to be adapted to the filtration generated by W_z. We will assume that the integrand satisfies

$$G_-(z)^2 \leqslant g(X_z, z)g(X_z, z)^T \leqslant G_+(z)^2 \tag{D.75}$$

almost surely, for deterministic functions $G_\pm(z)$, and that the integrals

$$V_\pm(z) = \int_0^z G_\pm(s)^2 \,\mathrm{d}s \tag{D.76}$$

are finite.

Lemma D.8. *For any $x > 0$,*

$$\mathbb{P}\Big\{ \sup_{0\leqslant s\leqslant z} M_s > x \Big\} \leqslant \mathrm{e}^{-x^2/2V_+(z)}. \tag{D.77}$$

Proof. Let

$$[M]_z = \int_0^z g(X_s, s)g(X_s, s)^T \,\mathrm{d}s \tag{D.78}$$

be the increasing process associated with M_z. Then for any $\gamma \in \mathbb{R}$ the Doléans exponential

$$\mathrm{e}^{\gamma M_z - \gamma^2 [M]_z/2} \tag{D.79}$$

is a martingale. It follows that

$$\begin{aligned}
\mathbb{P}\Big\{ \sup_{0\leqslant s\leqslant z} M_s > x \Big\} &= \mathbb{P}\Big\{ \sup_{0\leqslant s\leqslant z} \mathrm{e}^{\gamma M_s} > \mathrm{e}^{\gamma x} \Big\} \\
&\leqslant \mathbb{P}\Big\{ \sup_{0\leqslant s\leqslant z} \mathrm{e}^{\gamma M_s - \gamma^2 [M]_s/2} > \mathrm{e}^{\gamma x - \gamma^2 [M]_z/2} \Big\} \\
&\leqslant \mathbb{P}\Big\{ \sup_{0\leqslant s\leqslant z} \mathrm{e}^{\gamma M_s - \gamma^2 [M]_s/2} > \mathrm{e}^{\gamma x - \gamma^2 V_+(z)/2} \Big\} \\
&\leqslant \mathrm{e}^{-\gamma x + \gamma^2 V_+(z)/2} \mathbb{E}\big[\mathrm{e}^{\gamma M_z - \gamma^2 [M]_z/2} \big]
\end{aligned} \tag{D.80}$$

by Doob's submartingale inequality. Now the expectation in the last line is equal to 1, and the result follows by optimizing over γ, that is, choosing $\gamma = x/V_+(z)$. □

References

[1] M. Abramowitz, I.A. Stegun, Handbook of Mathematical Functions, 9th edition, Dover, 1965.
[2] Jean-Philippe Aguilar, Nils Berglund, The effect of classical noise on a quantum two-level system, J. Math. Phys. 49 (2008) 102102, 23 pp.
[3] Michael Allman, Volker Betz, Breaking the chain, Stochastic Process. Appl. 119 (8) (2009) 2645–2659.
[4] Michael Allman, Volker Betz, Martin Hairer, A chain of interacting particles under strain, Stochastic Process. Appl. 121 (9) (2011) 2014–2042.
[5] L. Arnold, Recent progress in stochastic bifurcation theory, in: IUTAM Symposium on Nonlinearity and Stochastic Structural Dynamics, Springer, 2001, pp. 15–27.
[6] L. Arnold, Random Dynamical Systems, Springer, 2003.
[7] V.I. Arnold, Encyclopedia of Mathematical Sciences: Dynamical Systems V, Springer, 1994.
[8] S.M. Baer, T. Erneux, Singular Hopf bifurcation to relaxation oscillations I, SIAM J. Appl. Math. 46 (5) (1986) 721–739.
[9] S.M. Baer, T. Erneux, Singular Hopf bifurcation to relaxation oscillations II, SIAM J. Appl. Math. 52 (6) (1992) 1651–1664.

[10] Richard Bellman, Introduction to Matrix Analysis, McGraw–Hill, New York, 1960.
[11] E. Benoît, Systems lents-rapides dans $\mathbb{R}^3$ et leurs canards, in: Third Snepfenried Geometry Conference, vol 2, Soc. Math. France, 1982, pp. 159–191.
[12] E. Benoît, Enlacements de canards, C. R. Acad. Sci. Paris 300 (8) (1985) 225–230.
[13] E. Benoît, Canards et enlacements, Publ. Math. Inst. Hautes Etudes Sci. 72 (1990) 63–91.
[14] E. Benoît, J.L. Callot, F. Diener, M. Diener, Chasse au canards, Collect. Math. 31 (1981) 37–119.
[15] E. Benoît, C. Lobry, Les canards de $\mathbb{R}^3$, C. R. Math. Acad. Sci. Paris 294 (1982) 483–488.
[16] Roberto Benzi, Alfonso Sutera, Angelo Vulpiani, The mechanism of stochastic resonance, J. Phys. A 14 (11) (1981) L453–L457.
[17] N. Berglund, D. Landon, Mixed-mode oscillations and interspike interval statistics in the stochastic FitzHugh–Nagumo model, arXiv:1105.1278v2, 2011, submitted for publication.
[18] Nils Berglund, Adiabatic dynamical systems and hysteresis, PhD thesis, EPFL, 1998.
[19] Nils Berglund, Barbara Gentz, Beyond the Fokker–Planck equation: Pathwise control of noisy bistable systems, J. Phys. A 35 (9) (2002) 2057–2091.
[20] Nils Berglund, Barbara Gentz, The effect of additive noise on dynamical hysteresis, Nonlinearity 15 (3) (2002) 605–632.
[21] Nils Berglund, Barbara Gentz, Metastability in simple climate models: Pathwise analysis of slowly driven Langevin equations, Stoch. Dyn. 2 (2002) 327–356.
[22] Nils Berglund, Barbara Gentz, Pathwise description of dynamic pitchfork bifurcations with additive noise, Probab. Theory Related Fields 122 (3) (2002) 341–388.
[23] Nils Berglund, Barbara Gentz, A sample-paths approach to noise-induced synchronization: Stochastic resonance in a double-well potential, Ann. Appl. Probab. 12 (2002) 1419–1470.
[24] Nils Berglund, Barbara Gentz, Geometric singular perturbation theory for stochastic differential equations, J. Differential Equations 191 (2003) 1–54.
[25] Nils Berglund, Barbara Gentz, Noise-Induced Phenomena in Slow–Fast Dynamical Systems. A Sample-Paths Approach, Probab. Appl., Springer, London, 2006.
[26] Nils Berglund, Barbara Gentz, Stochastic dynamic bifurcations and excitability, in: Carlo Laing, Gabriel Lord (Eds.), Stochastic Methods in Neuroscience, Oxford University Press, 2009, pp. 64–93.
[27] Nils Berglund, Hervé Kunz, Memory effects and scaling laws in slowly driven systems, J. Phys. A 32 (1) (1999) 15–39.
[28] B. Braaksma, Singular Hopf bifurcation in systems with fast and slow variables, J. Nonlinear Sci. 8 (5) (1998) 457–490.
[29] M. Brons, T.J. Kaper, H.G. Rotstein, Introduction to focus issue – mixed mode oscillations: experiment, computation, and analysis, Chaos 18 (2008) 015101.
[30] M. Brøns, M. Krupa, M. Wechselberger, Mixed mode oscillations due to the generalized canard phenomenon, in: Fields Inst. Commun., vol. 49, 2006, pp. 39–63.
[31] Jean-Louis Callot, Francine Diener, Marc Diener, Le problème de la "chasse au canard", C. R. Acad. Sci. Paris Sér. A–B 286 (22) (1978) A1059–A1061.
[32] H. Degn, L.F. Olsen, J.W. Perram, Bistability, oscillation, and chaos in an enzyme reaction, Ann. New York Acad. Sci. 316 (1) (1979) 623–637.
[33] M. Desroches, J. Guckenheimer, C. Kuehn, B. Krauskopf, H. Osinga, M. Wechselberger, Mixed-mode oscillations with multiple time scales, SIAM Rev. 54 (2) (2012), in press.
[34] M. Desroches, B. Krauskopf, H.M. Osinga, The geometry of slow manifolds near a folded node, SIAM J. Appl. Dyn. Syst. 7 (4) (2008) 1131–1162.
[35] M. Desroches, B. Krauskopf, H.M. Osinga, Mixed-mode oscillations and slow manifolds in the self-coupled FitzHugh–Nagumo system, Chaos 18 (2008) 015107.
[36] M. Desroches, B. Krauskopf, H.M. Osinga, Numerical continuation of canard orbits in slow–fast dynamical systems, Nonlinearity 23 (3) (2010) 739–765.
[37] R.E. Lee DeVille, Eric Vanden-Eijnden, Cyrill B. Muratov, Two distinct mechanisms of coherence in randomly perturbed dynamical systems, Phys. Rev. E (3) 72 (3) (2005) 031105, 10 pp.
[38] C.T. Dickson, J. Magistretti, M.H. Shalisnky, E. Fransen, M.E. Hasselmo, A. Alonso, Properties and role of I_h in the pacing of subtreshold oscillations in entorhinal cortex layer II neurons, J. Neurophysiol. 83 (2000) 2562–2579.
[39] F. Diener, M. Diener, Nonstandard Analysis in Practice, Springer, 1995.
[40] Catherine Doss, Michèle Thieullen, Oscillations and random perturbations of a FitzHugh–Nagumo system, preprint, hal-00395284, 2009.
[41] F. Dumortier, Techniques in the theory of local bifurcations: Blow-up, normal forms, nilpotent bifurcations, singular perturbations, in: D. Schlomiuk (Ed.), Bifurcations and Periodic Orbits of Vector Fields, Kluwer, 1993, pp. 19–73.
[42] F. Dumortier, R. Roussarie, Canard Cycles and Center Manifolds, Mem. Amer. Math. Soc., vol. 121(577), 1996.
[43] W. Eckhaus, Relaxation oscillations including a standard chase on French ducks, in: Lecture Notes in Math., vol. 985, 1983, pp. 449–494.
[44] Neil Fenichel, Geometric singular perturbation theory for ordinary differential equations, J. Differential Equations 31 (1) (1979) 53–98.
[45] Ronald F. Fox, Stochastic resonance in a double well, Phys. Rev. A 39 (1989) 4148–4153.
[46] Mark I. Freidlin, Quasi-deterministic approximation, metastability and stochastic resonance, Phys. D 137 (2000) 333–352.
[47] Mark I. Freidlin, On stable oscillations and equilibriums induced by small noise, J. Stat. Phys. 103 (2001) 283–300.
[48] M.I. Freidlin, A.D. Wentzell, Random Perturbations of Dynamical Systems, Springer, 1998.
[49] L. Gammaitoni, E. Menichella-Saetta, S. Santucci, F. Marchesoni, C. Presilla, Periodically time-modulated bistable systems: Stochastic resonance, Phys. Rev. A 40 (1989) 2114–2119.
[50] Luca Gammaitoni, Peter Hänggi, Peter Jung, Fabio Marchesoni, Stochastic resonance, Rev. Modern Phys. 70 (1998) 223–287.

[51] J. Grasman, Asymptotic Methods for Relaxation Oscillations and Applications, Springer, 1987.
[52] J. Guckenheimer, Return maps of folded nodes and folded saddle-nodes, Chaos 18 (2008).
[53] J. Guckenheimer, Singular Hopf bifurcation in systems with two slow variables, SIAM J. Appl. Dyn. Syst. 7 (4) (2008) 1355–1377.
[54] J. Guckenheimer, R. Haiduc, Canards at folded nodes, Mosc. Math. J. 5 (1) (2005) 91–103.
[55] J. Guckenheimer, C. Kuehn, Homoclinic orbits of the FitzHugh–Nagumo equation: Bifurcations in the full system, SIAM J. Appl. Dyn. Syst. 9 (2010) 138–153.
[56] J. Guckenheimer, C. Scheper, A geometric model for mixed-mode oscillations in a chemical system, SIAM J. Appl. Dyn. Syst. 10 (1) (2011) 92–128.
[57] E. Hairer, G. Wanner, Solving Ordinary Differential Equations II, Springer, 1991.
[58] Samuel Herrmann, Peter Imkeller, The exit problem for diffusions with time-periodic drift and stochastic resonance, Ann. Appl. Probab. 15 (1A) (2005) 36–68.
[59] Samuel Herrmann, Peter Imkeller, Dierk Peithmann, Transition times and stochastic resonance for multidimensional diffusions with time periodic drift: a large deviations approach, Ann. Appl. Probab. 16 (4) (2006) 1851–1892.
[60] D.J. Highham, An algorithmic introduction to numerical simulation of stochastic differential equations, SIAM Rev. 43 (3) (2001) 525–546.
[61] M.W. Hirsch, S. Smale, R. Devaney, Differential Equations, Dynamical Systems, and an Introduction to Chaos, 2nd edition, Academic Press, 2003.
[62] P. Hitczenko, G.S. Medvedev, Bursting oscillations induced by small noise, SIAM J. Appl. Math. 69 (2009) 1359–1392.
[63] J.L. Hudson, M. Hart, D. Marinko, An experimental study of multiple peak periodic and nonperiodic oscillations in the Belousov–Zhabotinskii reaction, J. Chem. Phys. 71 (4) (1979) 1601–1606.
[64] P. Imkeller, I. Pavlyukevich, Model reduction and stochastic resonance, Stoch. Dyn. 2 (4) (2002) 463–506.
[65] E. Izhikevich, Neural excitability, spiking, and bursting, Internat. J. Bifur. Chaos Appl. Sci. Engrg. 10 (2000) 1171–1266.
[66] Kalvis M. Jansons, G.D. Lythe, Stochastic calculus: application to dynamic bifurcations and threshold crossings, J. Stat. Phys. 90 (1–2) (1998) 227–251.
[67] C. Jones, N. Kopell, Tracking invariant manifolds with differential forms in singularly perturbed systems, J. Differential Equations 108 (1994) 64–88.
[68] C.K.R.T. Jones, Geometric singular perturbation theory, in: Dynamical Systems, Montecatini Terme, 1994, Springer, 1995.
[69] Yuri Kabanov, Sergei Pergamenshchikov, Two-Scale Stochastic Systems: Asymptotic Analysis and Control, Appl. Math. (New York), Stoch. Model. Appl. Probab., vol. 49, Springer, Berlin, 2003.
[70] Olav Kallenberg, Foundations of Modern Probability, 2nd edition, Springer, 2002.
[71] P.E. Kloeden, E. Platen, Numerical Solution of Stochastic Differential Equations, Springer, 2010.
[72] E.F. Mishchenko, Yu.S. Kolesov, A.Yu. Kolesov, N.Kh. Rozov, Asymptotic Methods in Singularly Perturbed Systems, Plenum Press, 1994.
[73] M.T.M. Koper, Bifurcations of mixed-mode oscillations in a three-variable autonomous Van der Pol–Duffing model with a cross-shaped phase diagram, Phys. D 80 (1995) 72–94.
[74] M. Krupa, P. Szmolyan, Extending geometric singular perturbation theory to nonhyperbolic points – fold and canard points in two dimensions, SIAM J. Math. Anal. 33 (2) (2001) 286–314.
[75] M. Krupa, P. Szmolyan, Extending slow manifolds near transcritical and pitchfork singularities, Nonlinearity 14 (2001) 1473–1491.
[76] M. Krupa, P. Szmolyan, Geometric analysis of the singularly perturbed fold, in: Multiple-Time-Scale Dynamical Systems, in: IMA Vol. Math. Appl., vol. 122, 2001, pp. 89–116.
[77] M. Krupa, P. Szmolyan, Relaxation oscillation and canard explosion, J. Differential Equations 174 (2001) 312–368.
[78] M. Krupa, M. Wechselberger, Local analysis near a folded saddle-node singularity, J. Differential Equations 248 (12) (2010) 2841–2888.
[79] C. Kuehn, From first Lyapunov coefficients to maximal canards, Internat. J. Bifur. Chaos Appl. Sci. Engrg. 20 (5) (2010) 1467–1475.
[80] C. Kuehn, A mathematical framework for critical transitions: bifurcations, fast–slow systems and stochastic dynamics, arXiv:1101.2908, 2011, submitted for publication.
[81] C. Kuehn, On decomposing mixed-mode oscillations and their return maps, Chaos 21 (3) (2011) 033107.
[82] C. Kuehn, Multiple Time Scale Dynamics, 2012, book in preparation.
[83] R. Kuske, Probability densities for noisy delay bifurcations, J. Stat. Phys. 96 (3–4) (1999) 797–816.
[84] R. Kuske, S.M. Baer, Asymptotic analysis of noise sensitivity of a neuronal burster, Bull. Math. Biol. 64 (2002) 447–481.
[85] Carlo Laing, Gabriel J. Lord (Eds.), Stochastic Methods in Neuroscience, Oxford University Press, Oxford, 2010.
[86] André Longtin, Stochastic resonance in neuron models, J. Stat. Phys. 70 (1993) 309–327.
[87] André Longtin, Effect of noise on the tuning properties of excitable systems, Chaos Solitons Fractals 11 (2000) 1835–1848.
[88] Bruce McNamara, Kurt Wiesenfeld, Theory of stochastic resonance, Phys. Rev. A 39 (1989) 4854–4869.
[89] E.F. Mishchenko, N.Kh. Rozov, Differential Equations with Small Parameters and Relaxation Oscillations, Plenum Press, New York, 1980.
[90] E.F. Mishchenko, N.Kh. Rozov, Differential Equations with Small Parameters and Relaxation Oscillations, Plenum Press, 1980 (translated from Russian).
[91] C.B. Muratov, E. Vanden-Eijnden, Noise-induced mixed-mode oscillations in a relaxation oscillator near the onset of a limit cycle, Chaos 18 (2008) 015111.
[92] C.B. Muratov, E. Vanden-Eijnden, W. E, Self-induced stochastic resonance in excitable systems, Phys. D 210 (2005) 227–240.
[93] A.I. Neĭshtadt, Persistence of stability loss for dynamical bifurcations. I, Differ. Equ. 23 (1987) 1385–1391.

[94] C. Nicolis, G. Nicolis, Stochastic aspects of climatic transitions—additive fluctuations, Tellus 33 (3) (1981) 225–234.
[95] B. Øksendal, Stochastic Differential Equations, Springer, 2003.
[96] V. Petrov, S.K. Scott, K. Showalter, Mixed-mode oscillations in chemical systems, J. Chem. Phys. 97 (9) (1992) 6191–6198.
[97] H.G. Rotstein, M. Wechselberger, N. Kopell, Canard induced mixed-mode oscillations in a medial entorhinal cortex layer II stellate cell model, SIAM J. Appl. Dyn. Syst. 7 (4) (2008) 1582–1611.
[98] M. Scheffer, J. Bascompte, W.A. Brock, V. Brovkhin, S.R. Carpenter, V. Dakos, H. Held, E.H. van Nes, M. Rietkerk, G. Sugihara, Early-warning signals for critical transitions, Nature 461 (2009) 53–59.
[99] B. Schmalfuss, K.R. Schneider, Invariant manifolds for random dynamical systems with slow and fast variables, J. Dynam. Differential Equations 20 (1) (2008) 133–164.
[100] L.F. Shampine, M.W. Reichelt, The MatLab ODE suite, SIAM J. Sci. Comput. 18 (1) (1997) 1–22.
[101] Richard B. Sowers, Random perturbations of canards, J. Theoret. Probab. 21 (2008) 824–889.
[102] N.G. Stocks, R. Manella, P.V.E. McClintock, Influence of random fluctuations on delayed bifurcations: The case of additive white noise, Phys. Rev. A 40 (1989) 5361–5369.
[103] Jianzhong Su, Jonathan Rubin, David Terman, Effects of noise on elliptic bursters, Nonlinearity 17 (1) (2004) 133–157.
[104] J.B. Swift, P.C. Hohenberg, Guenter Ahlers, Stochastic Landau equation with time-dependent drift, Phys. Rev. A 43 (1991) 6572–6580.
[105] P. Szmolyan, M. Wechselberger, Canards in $\mathbb{R}^3$, J. Differential Equations 177 (2001) 419–453.
[106] Axel Timmermann, Hezi Gildor, Michael Schulz, Eli Tziperman, Coherent resonant millennial-scale climate oscillations triggered by massive meltwater pulses, J. Climate 16 (2003) 2569–2585.
[107] S.-K. Tin, On the dynamics of tangent spaces near normally hyperbolic manifolds and singularly perturbed boundary value problems, PhD thesis, Brown University, 1994.
[108] M.C. Torrent, M. San Miguel, Stochastic-dynamics characterization of delayed laser threshold instability with swept control parameter, Phys. Rev. A 38 (1988) 245–251.
[109] Henry C. Tuckwell, Stochastic Processes in the Neurosciences, SIAM, Philadelphia, PA, 1989.
[110] F. Verhulst, Methods and Applications of Singular Perturbations: Boundary Layers and Multiple Timescale Dynamics, Springer, 2005.
[111] M. Wechselberger, Existence and bifurcation of canards in $\mathbb{R}^3$ in the case of a folded node, SIAM J. Appl. Dyn. Syst. 4 (1) (2005) 101–139.
[112] M. Wechselberger, A propos de canards (apropos canards), preprint, 2010, 20 pp.
[113] N. Yu, R. Kuske, Y.X. Li, Stochastic phase dynamics and noise-induced mixed-mode oscillations in coupled oscillators, Chaos 18 (2008) 015112.
[114] A. Zagaris, H.G. Kaper, T.J. Kaper, Fast and slow dynamics for the computational singular perturbation method, Multiscale Model. Simul. 2 (4) (2004) 613–638.
[115] A. Zagaris, H.G. Kaper, T.J. Kaper, Two perspectives on reduction of ordinary differential equations, Math. Nachr. 278 (12) (2005) 1629–1642.