

Positive Harris recurrence for degenerate diffusions with internal variables and randomly perturbed time-periodic input

Simon Holbach*

August 5, 2019

Abstract

We consider a multidimensional time-homogeneous dynamical system and add a randomly perturbed time-dependent deterministic signal to some of its components, giving rise to a high-dimensional system of stochastic differential equations, which is driven by possibly very low-dimensional noise. Equations of this type commonly occur in biology when modelling neurons or in statistical mechanics for certain Hamiltonian systems. We provide verifiable conditions on the original deterministic dynamical system under which the solution to the respective stochastic system features a point in the interior of its state space, which can be proved to be attainable by deterministic control arguments, and at which a local Hörmander condition holds. Together with a Lyapunov condition, it follows that the corresponding process is positive Harris recurrent.

Keywords: degenerate diffusion, periodic ergodicity, Harris recurrence, local Hörmander condition, deterministic control, dynamical system

AMS 2010 subject classification: 60J60, 60J25

1 Introduction and main results

Let $U \subset \mathbb{R}^N$ and $V \subset \mathbb{R}^L$. For some functions $f: U \times V \rightarrow \mathbb{R}^N$ and $g: U \times V \rightarrow \mathbb{R}^L$ and a signal $S: [0, \infty) \rightarrow \mathbb{R}^N$, consider the deterministic dynamical system

$$(1) \quad \begin{aligned} \dot{x} &= f(x, y) + S, \\ \dot{y} &= g(x, y), \end{aligned}$$

where \dot{x} and \dot{y} are the time-derivatives of the time-dependent variables $x: [0, \infty) \rightarrow U$ and $y: [0, \infty) \rightarrow V$, respectively.

This system is divided into two groups of variables: The N components of x , whose dynamics depend directly on the signal, and the L components of y , which are affected by the signal only indirectly through the influence of x . Intuitively speaking, we can think of (1) as a dynamical system with no intrinsic time-inhomogeneity, which then receives an additional time-dependent *external input* in some of its variables, while the remaining variables merely describe an interior mechanism. This is why we sometimes refer to x as the *adjustable variable(s)* and y as the *internal variable(s)*. Note that the only source of time-inhomogeneity is indeed the signal – if the system receives constant external input $S \equiv c \in \mathbb{R}^N$ (or none at all, i.e. $c = 0$), it is homogeneous in time. Systems of this kind frequently arise in the context of neuroscience and statistical mechanics (see Examples 1.1 and 1.2 below).

We construct a stochastic model by following the idea that the signal is not actually received in its original shape, but is subject to random perturbations by external noise (i.e. noise that is independent of the rest of the system). Using integral notation, we write (1) as

$$(2) \quad \begin{aligned} dx(t) &= f(x, y)(t)dt + S(t)dt, \\ dy(t) &= g(x, y)(t)dt. \end{aligned}$$

*Fakultät für Mathematik, Universität Bielefeld, Postfach 10 01 31, 33501 Bielefeld, Germany, e-mail: sholbach@math.uni-bielefeld.de. Research was partially supported by the CRC 1283 at Universität Bielefeld. Some of the results featured in this article have been part of the author's PhD thesis [24] at Johannes Gutenberg-Universität Mainz.

and in order to perturb the signal, it seems natural to substitute the term $S(t)dt$ in (2) with the increment dZ_t of a process taking values in some $W \subset \mathbb{R}^N$ and satisfying an SDE of the type

$$dZ_t = [S(t) + b(Z_t)]dt + \sigma(Z_t)dW_t,$$

where W is an M -dimensional standard Brownian Motion, while $b: W \rightarrow \mathbb{R}^N$ and $\sigma: W \rightarrow \mathbb{R}^{N \times M}$ are suitable drift and diffusion coefficients. Note that this SDE can be viewed as a generalized Ornstein–Uhlenbeck type process with time-dependent mean-reversion level (think of $b(Z_t) = -\beta Z_t$ with $\beta \in (0, \infty)$). A particularly prominent special case is the classical signal in noise model (take $M = N = 1$, $b \equiv 0$, and $\sigma \equiv 1$), which arises in a wide variety of fields including communication, radiolocation, seismic signal processing, or computer-aided diagnosis and has been the subject of extensive study.

Perturbing $S(t)$ randomly in this way leads to the stochastic dynamical system

$$(3) \quad \begin{aligned} dX_t &= f(X_t, Y_t)dt + dZ_t, \\ dY_t &= g(X_t, Y_t)dt, \\ dZ_t &= [S(t) + b(Z_t)]dt + \sigma(Z_t)dW_t, \end{aligned}$$

with state space

$$E := U \times V \times W \subset \mathbb{R}^{N+L+N}.$$

This system can be thought of as degenerate in the following sense: Firstly, the equation for Y does not incorporate the driving Brownian Motion W explicitly, making it rather unclear which effect noise has on these components. Secondly, the dimension M of the driving Brownian Motion can (and will usually) be much lower than the dimension $N + L + N$ of the system. This is why we call a stochastic process satisfying a system of SDEs of the type (3) a *degenerate diffusion with internal variables and randomly perturbed time-inhomogeneous deterministic input*.

We now have three groups of variables: The entirely autonomous external input governed by dZ_t (the "noisy signal"), the components of X that are directly adjusted by the noisy signal, and the components of the internal variable Y , whose dynamics are only indirectly affected by noise, since the respective differential equations incorporate neither Z nor the driving Brownian Motion W explicitly. Note that for this reason Y is conditionally deterministic given X and has continuously differentiable trajectories.

The system (3) is a generalization of the one introduced in equation (18) of [27], which is a probabilistic version of a class of dynamical systems that are well-known in the mathematical modelling of neurons (see Example 1.1 below). In [23] (which can be viewed as a statistical companion article to the present one), we study the problem of estimating parameters that are present in the signal S (in particular its periodicity). Before we explain the focus of the current article, let us introduce some major examples.

Example 1.1. Let $N = 1$, $L = 3$, $U = \mathbb{R}$, $V = [0, 1]^3$ and consider the coefficient functions

$$f(x, y) = -36y_1^4(x + 12) - 120y_2^3y_3(x - 120) - 0.3(x - 10.6)$$

and

$$g(x, y) = \begin{pmatrix} \alpha_1(x)(1 - y_1) - \beta_1(x)y_1 \\ \alpha_2(x)(1 - y_2) - \beta_2(x)y_2 \\ \alpha_3(x)(1 - y_3) - \beta_3(x)y_3 \end{pmatrix}$$

with

$$\begin{aligned} \alpha_1(x) &= \begin{cases} \frac{0.1 - 0.01x}{\exp(1 - 0.1x) - 1}, & x \neq 10, \\ 0.1, & \text{else,} \end{cases} & \beta_1(x) &= 0.125 \exp(-x/80), \\ \alpha_2(x) &= \begin{cases} \frac{2.5 - 0.1x}{\exp(2.5 - 0.1x) - 1}, & x \neq 25, \\ 1, & \text{else,} \end{cases} & \beta_2(x) &= 4 \exp(-x/18), \\ \alpha_3(x) &= 0.07 \exp(-x/20), & \beta_3(x) &= \frac{1}{\exp(3 - 0.1x) + 1}. \end{aligned}$$

for all $(x, y) = (x, y_1, y_2, y_3)^\top \in \mathbb{R} \times [0, 1]^3$. The corresponding dynamical system (1) is known as the *Hodgkin–Huxley system* and it was first introduced by Hodgkin and Huxley in 1952 (see [21], note however that we use the slightly different model constants from [30]) with the aim of describing the initiation and propagation of action potentials in the cell membrane of a neuron in response to an external stimulus. While x is the membrane potential itself (usually labelled V in the literature), the internal variables y_1 , y_2 , and y_3 (commonly denoted by

n , m , and h) correspond to the ionic mechanism underlying its evolution. The two predominant ion currents in the cell membrane are import of sodium Na^+ and export of potassium K^+ through the membrane. Each of the internal variables signifies the probability that a specific type of gate in the respective ion channel is open at a given time. It is for this reason that n , m , and h are often called gating variables. In the context of this model, the signal S represents the dendritic input that the neuron receives from a large number of other neurons, transported by an even larger number of synapses located on the respective dendritic tree. The resulting "total dendritic input" can then be thought of as an average of interdependent and repeating similar currents, which is why S is usually assumed to be periodic (or even constant). When modelling neurons, particular interest lies in the typical spiking behaviour of the membrane potential, a feature that is commonly agreed upon to be adequately described by the Hodgkin–Huxley model. For a more detailed modern introduction, interpretation, and an in-depth comparison with other neuron models, see for example [30] and [10].

Adding noise in the sense of (3) by choosing $\sigma \in C^\infty(\mathbb{W})$ and $b(Z_t) = -\beta Z_t$ with $\beta \in (0, \infty)$, we acquire the so-called *stochastic Hodgkin–Huxley model (with mean reverting Ornstein–Uhlenbeck type input)*. It was first introduced and studied by Höpfner, Löcherbach, and Thieullen in the series of the three papers [26], [27], and [28]. The constant β is determined by the so-called time constant of the membrane, which represents spontaneous voltage decay not related to the input. For many types of neurons, the time constant is known from experiments (see [11]). A degree of freedom lies in the choice of the diffusion coefficient σ , which reflects the nature of the influence of noise. In the past, mean reverting Ornstein–Uhlenbeck type equations with various volatilities have been used to model the membrane potential itself (see for example [32] or [25]), and in a sense our stochastic Hodgkin–Huxley model can be viewed as a refinement of this kind of model. If σ is Lipschitz continuous, existence of a unique non-exploding strong solution follows from the same arguments as in [26, Proposition 1] and [27, Proposition 2].

In [28], the authors prove that if the external equation is of classical Ornstein–Uhlenbeck type

$$dZ_t = [S(t) - \gamma Z_t]dt + \sigma dW_t$$

with constant $\sigma \in (0, \infty)$ and $\mathbb{W} = \mathbb{R}$, or (when the signal S is non-negative) of Cox–Ingersoll–Ross type¹

$$dZ_t = [a + S(t) - \gamma Z_t]dt + \sqrt{Z_t}dW_t$$

with $2a \in (1, \infty)$ and $\mathbb{W} = [0, \infty)$, the solution to the stochastic Hodgkin–Huxley system is positive Harris recurrent (see [28, Theorem 2.7]). Moreover, they quantify the typical spiking behaviour: Almost surely, there are infinitely many spikes but also infinitely many periods of the signal in which no spike occurs (see [28, Theorem 2.8]). Harris recurrence then enables them to prove a Glivenko–Cantelli type theorem for the interspike intervals (see [28, Theorem 2.9]).

Analogously, one can introduce stochastic versions of simpler neuron models such as the FitzHugh–Nagumo model (see [30, equations (4.11) and (4.12)]) or the Morris–Lecar model (see [37] or, for a modern version, [45]).

Example 1.2. Systems of coupled oscillators are particularly intuitive Hamiltonian systems and several different stochastic models have recently been subject to research (see e.g [18], [8], [6], [44], [7]). We illustrate the connection to our model with a simple example.

Let us think of three rotors, each given by their angle $q_i(t) \in \mathbb{R}$ and momentum $p_i(t) \in \mathbb{R}$ at the time $t \in [0, \infty)$ for each $i \in \{1, 2, 3\}$. Assuming their respective masses to be all equal to 1 and not taking into account units, the laws of classical mechanics imply

$$(4) \quad \dot{q}_i = p_i \quad \text{for all } i \in \{1, 2, 3\}.$$

We suppose that these rotors are coupled in row, i.e.

$$(5) \quad \begin{aligned} \dot{p}_1 &= w_1(q_2 - q_1) - u_1(q_1), \\ \dot{p}_2 &= -[w_1(q_2 - q_1) + w_3(q_2 - q_3)] - u_2(q_2), \\ \dot{p}_3 &= w_3(q_2 - q_3) - u_3(q_3), \end{aligned}$$

where $w_1, w_2, w_3: \mathbb{R} \rightarrow \mathbb{R}$ and $u_1, u_2, u_3: \mathbb{R} \rightarrow \mathbb{R}$ are related to interaction potentials and pinning potentials, respectively. A classical model is the one that arises if we let one or both of the outer rotors receive external torques and interact with Langevin type heat baths. In order to give a mathematical description of this, we fix

¹Note that a Cox–Ingersoll–Ross type equation for Z is not contained in our model, since we require its state space to be the full euclidean space and the diffusion coefficient to be defined everywhere on it. This is only of major importance for Section 3, though.

$i \in \{1, 3\}$ for the remainder of this paragraph. Applying an external time-dependent torque $S_i: [0, \infty) \rightarrow \mathbb{R}$ to the i -th rotor means expanding the equation for p_i to

$$dp_i = [w_i(q_2 - q_i) - u_i(q_i)] dt + S_i dt,$$

which turns (4) and (5) into a system like (1). On top of that, we want to add interaction with a heat bath, i.e. for a temperature $\tau_i \in (0, \infty)$ and a dissipation constant $\delta_i \in (0, \infty)$, the equation for p_i is further expanded to

$$\begin{aligned} dp_i &= [w_i(q_2 - q_i) - u_i(q_i)] dt + S_i dt - \delta_i p_i dt + \sqrt{2\delta_i \tau_i} dW_t^{(i)} \\ &= [w_i(q_2 - q_i) - u_i(q_i) - \delta_i p_i] dt + \left[S_i dt + \sqrt{2\delta_i \tau_i} dW_t^{(i)} \right], \end{aligned}$$

where the last term in parentheses is the total sum of external influences. Following the spirit of (3), we may replace this term with the increments of a more general random perturbation of the torque: We take

$$dp_i = [w_i(q_2 - q_i) - u_i(q_i) - \delta_i p_i] dt + dZ_t^{(i)}$$

with

$$dZ_t^{(i)} = \left[S_i(t) + b_i(Z_t^{(i)}) \right] dt + \sigma_i(Z_t^{(i)}) dW_t^{(i)}$$

for some diffusion coefficient $\sigma_i: \mathbb{R} \rightarrow \mathbb{R}$ and a drift $b_i: \mathbb{R} \rightarrow \mathbb{R}$. What we end up with is indeed a degenerate diffusion with internal variables and randomly perturbed time-inhomogeneous deterministic input as in (3). If only the first rotor in the chain receives an external input, the dimensions are $M = N = 1$ and $L = 5$, $\mathbf{E} = \mathbb{R}^7$. If both of the outer rotors receive an external input, the dimensions are $M = N = 2$ and $L = 4$, $\mathbf{E} = \mathbb{R}^8$.

In the present text, this example will mainly serve as a hint at the potential of studying the general system (3) beyond the realm of neuroscience – which the theory displayed in this article is predominantly inspired by. Only certain parts of our results are applicable to this stochastic rotor system, but there are enough connections to spawn optimism for future work that may unify some of the approaches used in different settings.

Let us stress that the differences between the eight-dimensional system above with input on both ends of the chain and the six-dimensional system that is discussed in [8] mainly lie in the role of the external equations and in time-inhomogeneity: Forcing the latter into our notational framework, the equations for X and Y would remain unaltered, while the external equations would read

$$dZ_t^{(i)} = S_i dt + \sqrt{2\delta_i \tau_i} dW_t^{(i)}, \quad i \in \{1, 2\},$$

where the coefficient functions do not depend on $Z_t^{(i)}$. Therefore, including these variables separately into the system is rendered obsolete. In addition to that, with the external torque in [8] being constant, their entire system is homogeneous in time – in contrast to ours. Another important difference is that while the system in [8] lets the angle take values in the torus $\mathbb{R} \bmod 2\pi$, the state space for (3) has to be a proper subset of \mathbb{R}^{N+L+N} (without any topological identifications). Therefore, the system we describe here describes oscillators but not rotors in a strict sense.

For the six-dimensional system from [8], the authors' main result (Theorem 1.3) states the existence of smooth transition densities, unique existence of an invariant measure with a smooth density, and finally establishes a subgeometric convergence rate of the semi-group to the invariant measure.

Example 1.3. For a convenient toy example, let $\mathbf{U} = \mathbf{W} = \mathbb{R}^N$ with $N \in \{1, 2\}$ and $\mathbf{V} = \mathbb{R}^L$. Let

$$f(x, y) = -(1 + |y|^2)(|x|^2 - 1)x = -(1 + |y|^2) \frac{1}{4} \nabla_x (|x|^4 - |x|^2) \quad \text{for all } x \in \mathbb{R}^N, y \in \mathbb{R}^L.$$

This means that the x -variables are subject to a double well potential (for $N = 1$) or a mexican hat potential (for $N = 2$) where the slope is determined by the current y -position, but is bounded away from zero.

For $L = 1$, let

$$(6) \quad g(x, y) = -(1 + |x|^2)y + \sin(|x|^2 - 1) \quad \text{for all } x \in \mathbb{R}^N, y \in \mathbb{R}.$$

By the first summand, the y -variable experiences exponential decay with a rate that depends on x , but is again bounded away from zero. The second summand represents some bounded x -dependent source that vanishes at the bottom of the wells or of the hat.

A different choice for g in the case $N = 1$ (but for any L) is given by

$$(7) \quad g(x, y) = (h(x) - y_1, y_1 - y_2, \dots, y_{L-1} - y_L)^\top \quad \text{for all } x \in \mathbb{R}, y \in \mathbb{R}^L,$$

where $h: \mathbb{R} \rightarrow \mathbb{R}$ is a bounded smooth function with $h(x) = x^2$ for all $x \in (-2, 2)$. Here, the dependence of each component on the variables takes a cascade structure.

For the external equation, we take $\mathbb{W} = \mathbb{R}^N$, and

$$(8) \quad dZ_t = [S(t) - \beta Z_t]dt + \sigma(Z_t)dW_t$$

where $\beta \in \mathbb{R}^{N \times N}$ is strictly positive definite and all entries of $\sigma: \mathbb{R}^N \rightarrow \mathbb{R}^{N \times M}$ are bounded.

The objective of this article is to find verifiable conditions on (1) and the way in which the external equation of (3) perturbs the signal, under which positive Harris recurrence of (X, Y, Z) can be established in spite of its apparent degeneracy. Positive Harris recurrence is the foundation of a wide class of limit theorems for Markov processes, most prominently the Ratio Limit Theorem and Orey's Ergodic Theorem. Classical references for the theory of recurrence in the sense of Harris include [19], [43], and [39] in the discrete-time setting and [2] in the continuous-time case. Of course, in order to make use of the respective theory, one has to look for some time-homogeneous substructure, which suggests the basic assumption that the signal function S should be periodic. In this setting, one can modify known tools and provide a strategy to prove positive Harris recurrence (see Theorem 2.2 below, c.f. [33], [28]). Our main results provide conditions on the deterministic system (1) under which this strategy can be applied to the stochastic system (3). As one of the respective ingredients, Section 3 shows:

Result 1. If T and T^* are incommensurable and (x^*, y^*) lies on the trajectory of an attractive T^* -periodic solution to the deterministic system (1) with appropriate periodic input signal S , the solution (X, Y, Z) to the stochastic system (3) with *any* T -periodic signal S will almost surely visit every neighborhood of (x^*, y^*, z^*) with $z^* \in \mathbb{W}$, even if time is restricted to multiples of T . The only assumption on the external equation is pointwise surjectivity of $\sigma(\cdot) \in \mathbb{R}^{N \times M}$, no assumption is needed that noise is small in any sense.

This is obtained via control arguments relying on the Support Theorem for diffusions. The main corresponding result in precise terms are Theorems 3.6, 3.11, and 3.16. The second main ingredient is the content of Section 4 and can be summarized as follows:

Result 2. If σ is nice, simple and verifiable conditions on the derivatives of f and g imply a local Hörmander condition, which in turn yields local transition densities for (X, Y, Z) .

This is the content of Section 4. We give essentially two different sets of conditions on the derivatives of f and g (Theorems 4.10 and 4.17), both of which have a clear interpretation and can be used for applications. The third main ingredient for the strategy arising from Theorem 2.2 is a Lyapunov condition that is more reliant on ad hoc arguments and that is briefly touched upon in Remark 2.4.

2 Precise setting and general technique

For the sake of better readability, we will often use the short notation

$$\Phi = (\varphi, z) = (x, y, z) \in \mathbf{U} \times \mathbf{V} \times \mathbf{W} = \mathbf{E}.$$

With $\varphi(t) = (x(t), y(t))$ for all $t \in [0, \infty)$, we can then rewrite (1) as

$$(9) \quad \dot{\varphi}(t) = F_S(t, \varphi(t)),$$

where

$$(10) \quad F_S: [0, \infty) \times \mathbf{U} \times \mathbf{V} \rightarrow \mathbb{R}^{N+L}, \quad (t, \varphi) \mapsto \begin{pmatrix} f(\varphi) + S(t) \\ g(\varphi) \end{pmatrix}.$$

If it exists, we write $\varphi[\varphi_0, S](t)$ for the unique solution of (9) at time $t \in [0, \infty)$ with starting condition $\varphi(0) = \varphi_0 \in \mathbf{U} \times \mathbf{V}$. We also write

$$(11) \quad F(\varphi) := F_{S=0}(t, \varphi) = \begin{pmatrix} f(\varphi) \\ g(\varphi) \end{pmatrix} \quad \text{for all } (t, \varphi) \in [0, \infty) \times \mathbf{U} \times \mathbf{V}.$$

While we will often work with different signals that are fed into the deterministic system (9), for the random system we fix one reference signal

$$S_0: [0, \infty) \rightarrow \mathbb{R}^N.$$

Similarly as above, if we set $\Phi_t := (X_t, Y_t, Z_t)$ for all $t \in [0, \infty)$, we can express the diffusion equation (3) as

$$(12) \quad d\Phi_t = B(t, \Phi_t)dt + \Sigma(\Phi_t)dW_t$$

with drift

$$(13) \quad B: [0, \infty) \times \mathbf{E} \rightarrow \mathbb{R}^{N+L+N}, \quad (t, \Phi) \mapsto \begin{pmatrix} f(\varphi) + S_0(t) + b(z) \\ g(\varphi) \\ S_0(t) + b(z) \end{pmatrix},$$

and diffusion matrix

$$(14) \quad \Sigma: \mathbf{E} \rightarrow \mathbb{R}^{(N+L+N) \times M}, \quad \Phi \mapsto \begin{pmatrix} \sigma(z) \\ 0_{L \times M} \\ \sigma(z) \end{pmatrix}.$$

We fix a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and on it an M -dimensional Brownian Motion W .

Remark 2.1. The most practically relevant results are acquired in the case $M = N$, but distinguishing notationally between the space dimension and the dimension of the driving Brownian motion can make some arguments a little clearer and easier to read.

The following basic assertions will be assumed to hold throughout this article.

- (A0) For all deterministic starting points $\Phi_0 \in \mathbf{E}$, the equation (12) has a unique non-explosive strong solution $(\Phi_t)_{t \in [0, \infty)}$ on (Ω, \mathcal{A}) under \mathbb{P} .
- (A1) S_0 is periodic with periodicity $T \in (0, \infty)$.
- (A2) All of the coefficient functions F, G, S_0, b , and σ have derivatives of any order with respect to any of their variables.
- (A3) There is a strictly increasing sequence $(G_n)_{n \in \mathbb{N}}$ of bounded open convex subsets of \mathbf{E} with $\bigcup_{n \in \mathbb{N}} \overline{G_n} = \mathbf{E}$ such that for any $\Phi_0 \in \mathbf{E}$, the following properties hold \mathbb{P} -almost surely:
 - (a) If $\Phi_0 \in \partial \mathbf{E}$, then $\inf\{t \in (0, \infty) \mid \Phi_t \in \text{int}(\mathbf{E})\} = 0$.
 - (b) If $\Phi_0 \in \overline{G_n} \setminus G_{n+1}$ for some $n \in \mathbb{N}$, then $\inf\{t \in (0, \infty) \mid \Phi_t \in G_{n+1}\} = 0$.
 - (c) We have $\inf\{t \in (0, \infty) \mid \Phi_t \notin \overline{G_n}\} \xrightarrow{n \rightarrow \infty} \infty$.

Note that in view of (A3), \mathbf{U}, \mathbf{V} and \mathbf{W} are necessarily convex, σ -compact sets with non-empty interior. The condition (A3) can be checked for the stochastic Hodgkin–Huxley model (see [28, page 533]), and in the oscillator system from Example 1.2 it becomes trivial since $\mathbf{E} = \mathbb{R}^{N+L+N}$, which is also the case in Example 1.3.

Under (A1), the entire drift of (3) is T -periodic in time, and consequently the corresponding transition semi-group $(P_{s,t})_{t > s \geq 0}$ has the property

$$(15) \quad P_{s+kT, t+kT} = P_{s,t} \quad \text{for all } t > s \geq 0 \text{ and } k \in \mathbb{N}.$$

This implies that the T -grid chain $(\Phi_{nT})_{n \in \mathbb{N}_0}$ is a discrete time Markov chain with no time-inhomogeneity. As such, its recurrence properties can be studied with classical methods. Our general method is summarized in the following theorem, which is in its essence a time-inhomogeneous variant of [33, Theorem 2.5] (compare [35, Theorem 15.0.1]) which is also the foundation of [28, Theorem 2.2].

Theorem 2.2. *Grant the following assumptions:*

- (I) *There is a Lyapunov function for the T -grid chain, i.e. there are a measurable function $V: \mathbf{E} \rightarrow [1, \infty)$ and a compact set $K \subset \mathbf{E}$ such that*

$$V(\Phi) \xrightarrow{|\Phi| \rightarrow \infty} \infty, \quad \sup_K P_{0,T} V < \infty, \quad \text{and} \quad \inf_{\mathbf{E} \setminus K} (1 - P_{0,T}) V > 0.$$

(II) There is a point $\Phi^* \in \text{int}(\mathbf{E})$ that is T -attainable, i.e. for any starting point $\Phi_0 \in \mathbf{E}$ and any $\varepsilon > 0$ there is an $n = n(\Phi_0, \varepsilon) \in \mathbb{N}$ such that

$$P_{0,nT}(\Phi_0, B_\varepsilon(\Phi^*)) > 0.$$

(III) There is an open neighborhood \mathcal{U}^* of Φ^* such that for all $n \in \mathbb{N}$ the kernel $P_{0,nT}$ admits a lower semi-continuous local Lebesgue-density $p_{0,nT}: \mathbf{E} \times \mathcal{U}^* \rightarrow [0, \infty)$, i.e.

$$P_{0,nT}(\Phi_0, B) = \int_B p_{0,nT}(\Phi_0, \Psi) d\Psi \quad \text{for all } \Phi_0 \in \mathbf{E} \text{ and measurable sets } B \subset \mathcal{U}^*.$$

Then the T -grid chain $(\Phi_{nT})_{n \in \mathbb{N}_0}$ is positive Harris recurrent.

Condition (I) is classical and can often be treated with ad hoc arguments for the specific system. We only deal with it explicitly in Remark 2.4 below. For (II) and (III), we follow a strategy that is also used in [28]: In Section 3, we present stability properties of (9) under which we can prove (II) by constructing suitable control paths and applying the Support Theorem. In Section 4, we investigate properties of the coefficient functions that allow to check (III) via a local variant of Hörmander's condition.

Remark 2.3. 1.) It is well-known that if the T -grid chain is positive Harris recurrent, then the same is true for the $C([0, T]; \mathbf{E})$ -valued path-segment chain $((\Phi_t)_{t \in [(n-1)T, nT]})_{n \in \mathbb{N}}$ and the $[0, T] \times \mathbf{E}$ -valued time-space process $(t \bmod T, \Phi_t)_{t \in [0, \infty)}$.

2.) Note that by Lemma 3.1 below, (II) is slightly weaker than "attainability in a sense of deterministic control" as in [28, Definition 2.1].

3.) In [28], Theorems 1 and 2 from [27] are used to check (III). In their original statements, these theorems treat regularity of $p_{0,nT}$ only *separately* in \mathbf{E} and \mathcal{U}^* , but it is crucial for the proof of Theorem 2.2 that lower semi-continuity holds *jointly* in $\mathbf{E} \times \mathcal{U}^*$. Since $\mathcal{U}^* \ni \Psi \mapsto p_{0,nT}(\Phi, \Psi)$ is equi-continuous with respect to $\Phi \in \mathbf{E}$ (as noted at the end of the proof of [27, Theorem 1]), this can be fixed in that case.

4.) Note that the Lyapunov condition (I) is slightly weaker than its counterpart in [33, Theorem 2.5] where

$$(16) \quad P_{0,T}V \leq \alpha V + \beta \quad \text{for some } \alpha \in (0, 1) \text{ and } \beta \in (0, \infty),$$

which does not only yield positive Harris recurrence but even exponential ergodicity.

Remark 2.4. In [24], we give variants of classical drift conditions under which a Lyapunov function with the property (16) can be constructed. Possible choices for the external equation include classical multi-dimensional Ornstein-Uhlenbeck processes (see [24, Example 2.15] for more details). The relevant conditions on f and g cover the stochastic Hodgkin-Huxley model in particular, meaning that the conclusion of [28, Theorem 2.7] is immediately improved from just positive Harris recurrence to exponential ergodicity (which is not commented on in [28], even though the Lyapunov function given there in the proofs of Propositions 2.5 and 2.6 is not essentially different from the one in [24, Example 2.16]). For the toy model from Example 1.3, [24, Theorem 2.11] yields the existence of a Lyapunov function with the property (16). However, these methods fail for the oscillator model from Example 1.2. For the rotor system from [8] with no external equations, the authors construct a Lyapunov function yielding stretched exponential ergodicity (see [8, Proposition 2.2 and Section 4]). The key to their result lies in understanding the slow dynamics of the momentum p_2 of the middle rotor and using appropriate averaging techniques.

In order to make the strategy more transparent, we would like to give a short proof of Theorem 2.2 condensing the reasoning used in [28] to adjust [33, Theorem 2.5] to the time-periodic setting. We also use this occasion to fix some minor gaps and inaccuracies in [28]. Let us stress, though, that we make no claims of originality.

Proof of Theorem 2.2. The main idea is to combine (II) and (III) in order to prove that the set K from (I) is small in the sense of Nummelin and then apply [39, Theorem 3.7(v) and Proposition 5.10]. As the dependence on the initial condition in (II) is problematic for this approach, we get rid of it by switching to the resolvent kernel

$$R := \sum_{n=1}^{\infty} 2^{-n} P_{0,nT} = \sum_{n=1}^{\infty} 2^{-n} P_{0,T}^n$$

corresponding to the resolvent chain, i.e. the Markov chain arising from sampling the grid chain at independent geometric times. The grid chain is positive Harris recurrent if and only if the resolvent chain is, and we shift our focus to the latter for the rest of this proof.

By (II), we have

$$R(\Phi_0, B_\varepsilon(\Phi^*)) > 0 \quad \text{for all } \Phi_0 \in \mathbf{E} \text{ and } \varepsilon > 0,$$

and it is easy to see that (I) remains valid with R in place of $P_{0,T}$, while (III) yields (via Fatou's Lemma) that R admits a lower semi-continuous local density $r: \mathbf{E} \times \mathcal{U}^* \rightarrow [0, \infty)$.² The above implies

$$\int_{\mathcal{U}^*} r(\Phi^*, \Psi) d\Psi = R(\Phi^*, \mathcal{U}^*) > 0$$

and thus the existence of $\Psi^* \in \mathcal{U}^*$ with $r(\Phi^*, \Psi^*) > 0$. By lower semi-continuity there are open neighborhoods $C \subset \mathbf{E}$ of Φ^* and $D \subset \mathbf{E}$ of Ψ^* such that r is bounded away from zero on $C \times D$. This can be translated into the minorization condition

$$R \geq \alpha 1_C \otimes \mathcal{U}_D,$$

where $\alpha > 0$ and \mathcal{U}_D denotes the uniform distribution on D . By attainability of Φ^* and by lower-semi-continuity (and again Fatou's Lemma) we have

$$\beta := \inf_{\Phi_0 \in K} R(\Phi_0, C) > 0.$$

Since K is visited infinitely often by the resolvent chain as a consequence of (I), Borel-Cantelli now yields that so is C . As a byproduct of this, there automatically exists an irreducibility (even invariant) measure for R that charges C (compare [38, Proof of Theorem 3]). Taken all together, C is a small set that is visited infinitely often, and hence [39, Theorem 3.7(v)] yields that the resolvent chain is Harris recurrent. We denote its unique (up to a multiplicative constant) invariant measure by μ . As K is visited infinitely often, it follows from [39, Proposition 2.4] that $\mu(K) > 0$. Since

$$R^2 \geq R \circ (\alpha 1_C \otimes \mathcal{U}_D) \geq \beta \alpha 1_K \otimes \mathcal{U}_D,$$

this means that K is also a small set with respect to R and hence (I) and [39, Proposition 5.10] yield that the resolvent chain is positive Harris recurrent. \square

3 Deterministic control

In this section, we will present a technique that uses stability properties of the deterministic dynamical system (1) in order to prove condition (II) of Theorem 2.2 via the Support Theorem for diffusions (see Lemma 3.1 below).

We work under the following assumption.

(C1) Non-degeneracy of Z : At each point $z \in \mathbb{W}$, the linear mapping that is represented by $\sigma(z) \in \mathbb{R}^{N \times M}$ is surjective. Thus, $\sigma(z)$ has a right inverse, and we simply denote it by $\sigma^{-1}(z) \in \mathbb{R}^{M \times N}$.

Note that surjectivity of the linear mapping $\sigma(z): \mathbb{R}^M \rightarrow \mathbb{R}^N$ implies $M \geq N$, so (C1) guarantees that the external equation for Z is not degenerate itself (compare condition (H2) in Section 4). Also note that $\sigma^{-1}(z)$ depends continuously on $z \in \mathbb{W}$, as $\sigma(\cdot)$ is continuous and so is the function that maps a surjective matrix to its right inverse.

Since we want to apply the Support Theorem, let

$$(17) \quad \tilde{b}: \mathbb{W} \rightarrow \mathbb{R}^N, \quad z \mapsto b(z) - \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^M \sigma_{i,j}(z) \begin{pmatrix} \partial_{z_i} \sigma_{1,j}(z) \\ \vdots \\ \partial_{z_i} \sigma_{N,j}(z) \end{pmatrix},$$

denote the Stratonovich version of the drift b with respect to σ , i.e. \tilde{b} is the drift coefficient such that Z satisfies

$$dZ_t = [S_0(t) + \tilde{b}(Z_t)] dt + \sigma(Z_t) \circ dW_t,$$

where " $\circ dW_t$ " indicates that we interpret the stochastic integral in the sense of Stratonovich instead of Itô. Hence, if we define \tilde{B} in the same way as B in (13), only replacing b with \tilde{b} , we obviously get the Stratonovich drift for the entire system (3), i.e.

$$d\Phi_t = \tilde{B}(t, \Phi_t) dt + \Sigma(\Phi_t) \circ dW_t.$$

²Note that [28, Corollary 4.2] claims without proof that smoothness in the second variable is also passed on from $p_{0,nT}$ to r . It is not clear why this should be true in general, but luckily this is irrelevant for the proof of Theorem 2.2.

Lemma 3.1. *A point $\Phi^* \in \text{int}(\mathbf{E})$ is T -attainable in the sense of condition (II) of Theorem 2.2 if and only if for any starting point $\Phi_0 \in \mathbf{E}$ there are $\dot{h} \in \mathbb{L}_{\text{loc}}^2([0, \infty); \mathbb{R}^M)$ and a solution $\Psi \in C([0, \infty); \mathbf{E})$ to the deterministic integral equation*

$$(18) \quad d\Psi(t) = \tilde{B}(t, \Psi(t))dt + \Sigma(\Psi(t))\dot{h}(t)dt \quad \text{for all } t \in [0, \infty),$$

such that

$$(19) \quad \Phi_0 = \Psi(0) \quad \text{and} \quad \Phi^* \in \overline{\Psi(\mathbb{N}T)}.$$

A solution to (18) is called a control path.

Proof. This is an immediate consequence of the Support Theorem and the fact that S_0 is T -periodic. A variant of the Support Theorem that applies to diffusions on state spaces with the property (A3) is given in [28, Theorem 3.1]. \square

Remark 3.2. Note that if we write $\Psi = (u, v, w)$, the control system (18) in differential notation becomes

$$(20) \quad \begin{aligned} \dot{u} &= f(u, v) + S_0 + \tilde{b}(w) + \sigma(w)\dot{h}, \\ \dot{v} &= g(u, v), \\ \dot{w} &= S_0 + \tilde{b}(w) + \sigma(w)\dot{h}. \end{aligned}$$

The last equation is still autonomous and w can be forced to any trajectory $\varrho \in C^1([0, \infty); \mathbb{W})$ with $\varrho(0) = w(0)$. Indeed, if we choose

$$(21) \quad \dot{h} := \sigma^{-1}(\varrho) \left(\dot{\varrho} - S_0 - \tilde{b}(\varrho) \right) \in C([0, \infty); \mathbb{R}^M) \subset \mathbb{L}_{\text{loc}}^2([0, \infty); \mathbb{R}^M),$$

then $w = \varrho$ solves the third equation of (20). The equations for u and v on the other hand can then be rewritten as

$$\dot{\psi}(t) = F_{\dot{\varrho}}(t, \psi(t)) \quad \text{with } \psi := (u, v),$$

and therefore $\psi = \varphi[(u(0), v(0)), \dot{\varrho}]$. In other words, a suitable choice of \dot{h} allows us to view ψ as a solution to (9) with a given continuous signal $S = \dot{\varrho}$, while $w = \varrho$ is its anti-derivative.

Remark 3.3. As a consequence of Remark 3.2 and the Support Theorem, the first two components (X_t, Y_t) of the solution Φ_t of the stochastic system (12) with the *fixed* periodic signal S_0 can (up until some time $t_0 \in (0, \infty)$) approximately imitate the behavior of the solution of the deterministic system (9) with *any* (not necessarily periodic) signal $S \in C([0, \infty); \mathbb{R}^N)$ satisfying

$$\mathbb{W}_S := \left\{ z \in \mathbb{W} \mid z + \int_0^t S(s)ds \in \mathbb{W} \text{ for all } t \in [0, t_0] \right\} \neq \emptyset.$$

Indeed, if $\varphi_0 = (x_0, y_0) \in \mathbb{U} \times \mathbb{V}$ is such that $\varphi[(x_0, y_0), S](t) = (x, y)(t)$ exists for all $t \in [0, t_0]$, and if $\Phi_0 = (x_0, y_0, z_0)$ with $z_0 \in \mathbb{W}_S$, [28, Theorem 3.1] yields that

$$\mathbb{P} \left(|(X_t, Y_t) - (x, y)(t)| < \varepsilon \text{ for all } t \in [0, t_0] \right) > 0 \quad \text{for all } \varepsilon > 0.$$

Note that for $\mathbb{W} = \mathbb{R}^N$ we have $\mathbb{W}_S = \mathbb{R}^N$, regardless of the choice of the signal.

Remark 3.4. The observation in Remark 3.2 inspires the following basic scheme for acquiring a T -attainable point:

1. Identify an equilibrium or a T^* -periodic orbit (with $T^* \notin \mathbb{Q}T$) of the deterministic system (9) with $S \equiv 0$ to which its solution is locally attracted.
2. Define \dot{h} in such a way that the resulting input \dot{w} steers $\psi = (u, v)$ into the domain of attraction of this orbit. Meanwhile, w must be guided to a given $z^* \in \text{int}(\mathbb{W})$.
3. Let w rest at z^* such that from that time on $\dot{w} \equiv 0$ and ψ converges to the attractive periodic orbit.

This way we establish (φ^*, z^*) as a T -attainable point whenever φ^* lies on the orbit. The choice of $z^* \in \text{int}(\mathbb{W})$ turns out to be completely arbitrary for our technique and our setting (see Theorems 3.6 and 3.11 below).

We will apply the strategy from Remark 3.4 in different scenarios: We start with a rather simple setting in Theorem 3.6, where we can illustrate the general idea in a context that allows to minimize the technical aspects. The main result of this section is then presented in Theorem 3.11, which treats a different, more nuanced and localized setting that is inspired by the Hodgkin–Huxley model. Finally, Theorem 3.16 demonstrates how our technique can be adapted to periodic orbits of (9) that are induced by periodic signals $S = S_* \neq 0$. All of these theorems have in common that they do not rely on ad hoc arguments exploiting properties of the $N + L + N$ -dimensional control system (18) itself, but work entirely under assumptions only on the $N + L$ -dimensional deterministic system (9) (except for condition (C1) above).

Remark 3.5. We want to make two remarks on the scope of the strategy from Remark 3.4 beyond what is explicitly dealt with in the sequel.

1.) If $c \in \mathbb{R}^N$, replacing f with $f - c$ and $S = S_0$ with $S_0 + c$ leaves the equation (1) entirely unaltered. In (3) this change influences only the external equation for Z . Since the Z -variable has no intrinsic meaning, one can always apply this transformation without fundamentally changing the setup. In this sense, working with properties of the system with zero-input (instead of general constant input) is hardly any restriction for practical purposes.

2.) In principle, it is possible that there is a solution to (9) such that all of its components are periodic, but their individual periodicities are not the same. If they are linearly independent over \mathbb{Q} , an extension of the arguments from Section 3.4 may allow to use the same strategy anyway.

3.1 Globally attractive periodic orbit under zero-input

We want to use the approach outlined in Remark 3.4 in a relatively simple setting where the orbit is in fact globally attractive. Before we give the exact conditions, let us stress that we can handle both periodic orbits and single equilibrium points, which are included in the following as orbits with period 0.

(C2) Incommensurable periodic orbit: For some $T^* \in [0, \infty)$ with $T^* = 0$ or $T^* \notin \mathbb{Q}T$, the system (9) with $S \equiv 0$ has a T^* -periodic orbit $\mathcal{O}^* \subset \text{int}(\mathbf{U} \times \mathbf{V})$, i.e. for all $\varphi_0 \in \mathcal{O}^*$ we have

$$\varphi[\varphi_0, S \equiv 0]([0, \infty)) = \mathcal{O}^*, \quad \text{and} \quad T^* = \inf\{t \in (0, \infty) \mid \varphi[\varphi_0, S \equiv 0](t) = \varphi_0\}.$$

(C3) Global attractivity of \mathcal{O}^* : For any $\varphi_0 \in \mathbf{U} \times \mathbf{V}$, there is a unique global solution to (9) with $S \equiv 0$, and it satisfies

$$\text{dist}(\varphi[\varphi_0, S \equiv 0](t), \mathcal{O}^*) \xrightarrow{t \rightarrow \infty} 0.$$

(C4) Non-explosiveness for moderate input: For any $\varphi_0 \in \mathbf{U} \times \mathbf{V}$, there is a $\delta_0 = \delta_0(\varphi_0) > 0$ such that for any $S \in C_b^\infty([0, \infty); \mathbb{R}^N)$ with $\|S\|_\infty \leq \delta_0$, the system (9) has a unique global solution.

We then have the following result.

Theorem 3.6. *Assume that (C1) - (C4) hold. Then for any $\varphi^* \in \mathcal{O}^*$ and $z^* \in \text{int}(\mathbf{W})$, the point $\Phi^* = (\varphi^*, z^*) \in \text{int}(\mathbf{E})$ is T -attainable.*

Proof. Let $\Phi_0 = (\varphi_0, z_0) \in \mathbf{E}$ and write $\Psi = (\psi, w)$ as in Remark 3.2. Since \mathbf{W} is convex, we can use linear interpolation and a mollifier in order to construct a smooth function $\varrho \in C_b^\infty([0, \infty); \mathbf{W})$ with $\varrho(0) = z_0$ and $\|\dot{\varrho}\|_\infty \leq \delta_0$ such that $\varrho \equiv z^*$ on $[kT, \infty)$ for some $k = k(z_0, \delta_0) \in \mathbb{N}$. Defining \dot{h} as in (21), we can force w onto ϱ and then $\psi = (u, v)$ has to be the solution $\varphi[\varphi_0, \dot{\varrho}]$ to (9), which uniquely exists thanks to (C4). From time kT on, the pertaining signal $\dot{\varrho}$ rests at zero, so we have

$$(22) \quad \psi(kT + t) = \varphi[\psi(kT), S \equiv 0](t) \quad \text{for all } t \in [0, \infty),$$

and hence Lemma 3.19 below will imply that $\varphi^* \in \overline{\psi(\mathbb{N}T)}$. Since $w \equiv z^*$ on $[kT, \infty)$, this means that $\Phi^* = (\varphi^*, z^*)$ is in $\overline{\Psi(\mathbb{N}T)}$. Consequently, Φ^* is T -attainable by Lemma 3.1. \square

Remark 3.7. In general, it is possible that (C2) and (C4) are fulfilled and the global attractivity (C3) is violated in only one point $\tilde{\varphi}_0 \in \mathbf{U} \times \mathbf{V}$ that is an unstable equilibrium of the system (9) with $S \equiv 0$ (see Example 3.8 below). This leads to a problem in the proof of Theorem 3.6, since we might be very unlucky and $\psi(kT) = \tilde{\varphi}_0$. If there are $m \in \mathbb{N}$ and $S \in C_b^\infty([0, mT]; \mathbb{R}^N)$ with

$$\int_0^{mT} S(t)dt = 0 \quad \text{and} \quad \varphi[\tilde{\varphi}_0, S](mT) \neq \tilde{\varphi}_0,$$

this problem can be solved by replacing ϱ with

$$\tilde{\varrho}(t) = \begin{cases} \varrho(t), & \text{if } t \in [0, \infty) \setminus [kT, (k+m)T], \\ z^* + \int_0^t S(s)ds, & \text{if } t \in [kT, (k+m)T], \end{cases}$$

Then, (22) holds with k replaced with $\tilde{k} = k + m$ and the rest of the proof can be finished with no further adjustments.

Example 3.8. Choose $U = V = \mathbb{R}$ and

$$\begin{aligned} f(x, y) &= x - y - x(x^2 + y^2), \\ g(x, y) &= x + y - y(x^2 + y^2) \end{aligned}$$

for all $x, y \in \mathbb{R}$. We see that

$$[0, \infty) \ni t \mapsto \varphi[(1, 0), S \equiv 0](t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$$

establishes the 2π -periodic orbit $\mathcal{O}^* = \mathbb{S}^1$ for the corresponding system (9) with $S \equiv 0$. Looking at the phase diagram in Figure 1, it is evident that the only initial condition violating (C3) is the unstable equilibrium in $\tilde{\varphi}_0 = (0, 0)$, but that this can be resolved by using Remark 3.7

$$S(t) = \cos \frac{2\pi t}{mT} \quad \text{for all } t \in [0, mT]$$

with some $m \in \mathbb{N}$. Obviously, (C4) is unproblematic.

Example 3.9. In the oscillator model from Example 1.2, we may think of a situation in which the interaction potentials are bounded, while the pinning potentials pull the oscillators to the centre with a force that grows with each rotation (for example $u_1 = u_2 = u_3 = -\varphi(|\cdot|)$ for some non-negative and non-decreasing function φ with $\varphi(0) = 0$ and with a sufficiently steep slope). We do not give a rigorous proof, but it is intuitively clear that such a system will satisfy (C2) - (C4) with $\mathcal{O}^* = \{0\}$, $T^* = 0$.

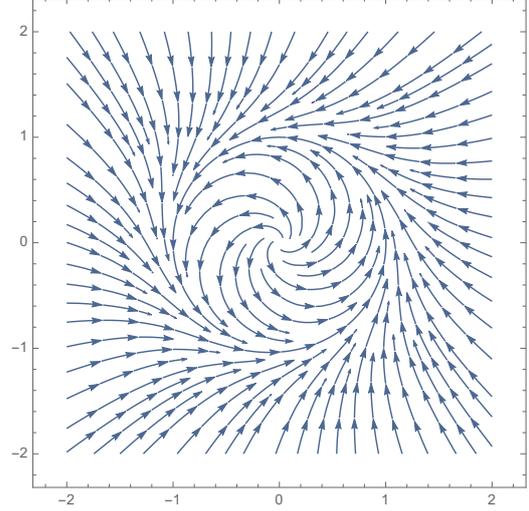


Figure 1

3.2 Locally attractive periodic orbit under zero-input

The main difficulty in the technique from Remark 3.4 is that the two tasks described in the second step may obviously compete with each other. In Theorem 3.6, we avoided this problem by simply assuming that the domain of attraction of the orbit is the entire space. In order to adapt our method to the case where the orbit is in fact only locally attractive, we subdivide this step into two steps:

- 2a. Define \dot{h} in such a way that the resulting input \dot{w} steers $\psi = (u, v)$ into the domain of attraction of this orbit.
- 2b. Move w to a given $z^* \in \text{int}(W)$ "gently enough" such that in the meantime ψ cannot leave the domain of attraction.

In order to realize Step 2a, we make some further assumptions on the interplay of x and y , which are inspired by the behaviour of the Hodgkin–Huxley model. For Step 2b, we have to make sure that the system is not too sensitive with respect to slight changes of the input. These ideas are made precise in the following set of assumptions. In order to grasp the essential idea, it may be helpful on first reading to think of the case $\mathcal{O}^* = \{(x^*, y^*)\}$, i.e. the periodic orbit is in fact just a single equilibrium point.

(C3') Local attractivity of \mathcal{O}^* : There is some $\varepsilon^* > 0$ such that for any initial value $\varphi_0 \in B_{\varepsilon^*}(\mathcal{O}^*)$ there is a unique global solution to (9) with $S \equiv 0$, and it satisfies

$$\text{dist}(\varphi[\varphi_0, S \equiv 0](t), \mathcal{O}^*) \xrightarrow{t \rightarrow \infty} 0.$$

(C4') Signal-dependent stability of φ^* : For all $\varepsilon > 0$, there is some $\delta(\varepsilon) \in (0, \varepsilon]$ such that for all $S \in C_b^\infty([0, \infty); \mathbb{R}^N)$ with $\|S\|_\infty < \delta(\varepsilon)$ and all $\varphi_0 \in B_{\delta(\varepsilon)}(\mathcal{O}^*)$ there is a unique global solution to (9), and it satisfies

$$\varphi[\varphi_0, S](t) \in B_\varepsilon(\mathcal{O}^*) \quad \text{for all } t \in [0, \infty).$$

(C5) Global attractivity in y for equilibrrious x : There is a continuous function

$$x^* : [0, \infty) \rightarrow \{x \in \mathbf{U} \mid (x, y) \in \mathcal{O}^* \text{ for some } y \in \mathbf{V}\}$$

such that for all $y_0 \in \mathbf{V}$ the initial value problem

$$\dot{y}(t) = g(x^*(t), y(t)), \quad y(0) = y_0,$$

has a unique solution $y[y_0, x^*] : [0, \infty) \rightarrow \mathbf{V}$, and it satisfies

$$\text{dist}((x^*(t), y[y_0, x^*](t)), \mathcal{O}^*) \xrightarrow{t \rightarrow \infty} 0.$$

(C6) Relocation to \mathcal{O}^* in x : For all $\varphi_0 = (x_0, y_0) \in \mathbf{U} \times \mathbf{V}$, there are $t_1 \in (0, \infty)$ and a function $\gamma \in C^1([0, \infty); \mathbf{U})$ with $\gamma(0) = x_0$ and $\gamma(t) = x^*(t)$ for all $t \in [t_1, \infty)$ (with x^* from (C5)) such that the differential equation

$$(23) \quad \dot{y}(t) = g(\gamma(t), y(t)), \quad y(0) = y_0,$$

has a unique local solution $\tilde{y}[y_0, \gamma] : [0, t_1] \rightarrow \mathbf{V}$.

Remark 3.10. Assumption (C4') is a stability condition that takes into account not only the starting point but also the signal: If we want the solution of (9) to stay close to \mathcal{O}^* , we can achieve this by starting not too far away from it, even if we slightly vary the input. If the input is in fact constantly zero, hypotheses (C3') and (C5) signify that \mathcal{O}^* is locally attractive, and its attraction is even global for y if x is already on a suitable trajectory on the projection of \mathcal{O}^* to \mathbf{U} . Condition (C6) makes sure that the x -variable can always be relocated to such a trajectory, while the corresponding equation for y still admits a local solution up to the time the x -variable hits x^* . Due to (C5), this solution can immediately be extended to a global solution. Thus, (C6) can equivalently be replaced by

(C6') Relocation to \mathcal{O}^* in x : For all $\varphi_0 = (x_0, y_0) \in \mathbf{U} \times \mathbf{V}$, there are $t_1 \in (0, \infty)$ and a function $S \in C([0, \infty); \mathbb{R}^N)$ such that there is a unique solution $(x, y) = \varphi[\varphi_0, S] : [0, \infty) \rightarrow \mathbf{U} \times \mathbf{V}$ to (1), and it satisfies $x = x^*$ on $[t_1, \infty)$.

If (C6') is satisfied, we can simply take γ as the corresponding trajectory of x and acquire (C6). If on the other hand (C6) holds, we can deduce (C6') by extending y to a global solution with (C5) and then taking $S = \dot{\gamma} - F(\gamma, y)$. While (C6') may be more intuitive, as it enhances the idea that one has to adjust x by giving the right input, (C6) is of a form that is usually easier to verify.

The set of conditions is now complete, and we can state the main result of this section.

Theorem 3.11. *Let (C1), (C2), (C3'), (C4'), (C5), and (C6) hold, and let $\mathbf{W} = \mathbb{R}^N$. Then for any $\varphi^* \in \mathcal{O}^*$ and $z^* \in \mathbb{R}^N$ the point $\Phi^* = (\varphi^*, z^*) \in \text{int}(\mathbf{E})$ is T -attainable.*

Proof. Let $\Phi_0 = (x_0, y_0, z_0) \in \mathbf{E}$ and write $\Psi = (u, v, w)$ as in (20). As indicated above, the construction of a suitable control path is much more delicate than in Theorem 3.6. Therefore, we divide its construction into several steps.

Phase 1: At first, we want to choose \dot{h} in such a way that u is forced to the trajectory $\gamma : [0, \infty) \rightarrow \mathbf{U}$ from (C6). Therefore, we take a function $(u_\#, v_\#, w_\#)$ with

$$u_\# = \gamma, \quad \dot{v}_\# = g(\gamma, v_\#), \quad \dot{w}_\# = \dot{\gamma} - f(\gamma, v_\#),$$

and $v_\#(0) = y_0, w_\#(0) = z_0$. Here, the first equation indeed just prescribes $u_\#$ to coincide with γ , while (C6) secures that the second equation admits a solution in $[0, t_1]$, and the last equation is then simply solved by integrating, as the right hand side does not contain the variable $w_\#$. Rearranging the first equation of the control system (20) for u , we see that if we set

$$(24) \quad \begin{aligned} \dot{h} &:= \sigma^{-1}(w_\#) \left(\dot{w}_\# - S_0 - \tilde{b}(w_\#) \right) \\ &= \sigma^{-1}(w_\#) \left(\dot{u}_\# - f(u_\#, v_\#) - S_0 - \tilde{b}(w_\#) \right) \in C([0, \infty); \mathbb{R}^M), \end{aligned}$$

the function $(u, v, w) := (u_\#, v_\#, w_\#)$ satisfies (20) for all $t \in [0, t_1]$. Thus we have constructed a control path up until the time t_1 , from which on the first variable u follows the curve x^* .

Phase 2: Since $u_{\sharp} = \gamma = x^*$ after time t_1 , the stability condition (C5) implies that v_{\sharp} can be extended to a global solution with

$$\text{dist}((u_{\sharp}(t), v_{\sharp}(t)), \mathcal{O}^*) \xrightarrow{t \rightarrow \infty} 0,$$

which allows us to set

$$t_2 := \inf \left\{ t > t_1 \mid (u_{\sharp}(t), v_{\sharp}(t)) \in B_{\frac{1}{2}\delta(\varepsilon^*)}(\mathcal{O}^*) \right\} \in [t_1, \infty)$$

with ε^* from (C3') and $\delta(\varepsilon^*)$ chosen according to (C4'). Thus, by the time t_2 we have driven (u_{\sharp}, v_{\sharp}) into $B_{\delta(\varepsilon^*)}(\mathcal{O}^*)$.

Phase 3: Up to time t_2 , we have constructed a useful candidate for \dot{h} and the corresponding solution $(u, v, w) = (u_{\sharp}, v_{\sharp}, w_{\sharp})$ to (20). Our next step will be to change the definition (24) of \dot{h} beyond time t_2 and extend this solution accordingly. Our aim is to drive w to z^* while keeping (u, v) close to \mathcal{O}^* . Since \dot{w} can be thought of as an alternate signal that is fed into the equation, assumption (C4') makes sure that we can achieve this, as long as we move w sufficiently slowly. Linear interpolation and classical mollification techniques allow us to take some $\varrho \in C_b^\infty([t_2, \infty); \mathbb{R}^N)$ with $\varrho(t_2) = w_{\sharp}(t_2)$ and $\|\dot{\varrho}\|_\infty < \delta(\varepsilon^*)$ such that for some $k \in \mathbb{N}$ with $kT > t_2$ we have $\varrho \equiv z^*$ on $[kT, \infty)$. On $[t_2, \infty)$, we change the definition of \dot{h} to

$$(25) \quad \dot{h} := \sigma^{-1}(\varrho) \left(\dot{\varrho} - S_0 - \tilde{b}(\varrho) \right) \in C([t_2, \infty); \mathbb{R}^M).$$

With this choice, we can obviously extend w by taking it equal to ϱ beyond time t_2 , and thus $\psi = (u, v)$ has to obey

$$(26) \quad \dot{\psi}(t) = F_{\dot{\varrho}}(t, \psi(t)) \quad \text{for all } t \in [t_2, \infty).$$

Since $\dot{\varrho}$ now plays the role of a signal whose absolute value is bounded by $\delta(\varepsilon^*)$, assumption (C4') yields that we can indeed extend ψ correspondingly and $\psi(kT)$ will still be in $B_{\varepsilon^*}(\mathcal{O}^*)$. Furthermore, $w \equiv z^*$ and $\dot{w} \equiv 0$ on $[kT, \infty)$.

Phase 4: After time kT , the equation (26) turns into $\dot{\psi} = F(\psi)$ and therefore

$$\psi(kT + t) = \varphi[\psi(kT), S \equiv 0](t) \quad \text{for all } t \in [0, \infty).$$

Since $\psi(kT) \in B_{\varepsilon^*}(\mathcal{O}^*)$, we can apply Lemma 3.19 and the proof is completed in the same way as the proof of Theorem 3.6. \square

Remark 3.12. The route we take in this proof is inspired by the one used in the proof of [28, Proposition 2.5], where the authors treat the stochastic Hodgkin–Huxley system. They do not take into account general periodic orbits but only a constant equilibrium. For the special case in which the autonomous equation for Z is of one-dimensional Cox–Ingersoll–Ross type, they can in fact abandon the need for σ to be well-defined on the entire space \mathbb{R} . This would make the first two phases of our proof more complicated, as one has to make sure that w does not leave the state space by turning negative (this is why we assume $W = \mathbb{R}^N$). The case that Z is of Ornstein–Uhlenbeck type – of which our model is a generalization – is treated mostly in the same way as here, albeit only for the dimension $M = N = 1$ and for constant diffusion coefficient. However, there is one key difference in the third phase of the control (which corresponds to "Part (V)" of the proof of [28, Proposition 2.5]): In the mentioned article, the control is defined such that u is forced to a trajectory that moves it very slightly away from the equilibrium point (x^*, y^*) . If this is done slowly enough, v is assumed to stay close to y^* , and w approaches z^* similarly as in our third phase. No rigorous prove is given that the deterministic Hodgkin–Huxley system actually shows this kind of behaviour. We think that in general it is a more intuitive approach to force not u but w to a specific trajectory, since we can interpret this as feeding a suitable signal into (1) that does not let (x, y) escape from the domain of attraction of the equilibrium. Condition (C4') is the key to making this possible. Note though that for the Hodgkin–Huxley model, we still rely on numerical simulations and merely give a rough intuitive argument that (C4') holds (see Example 3.13 below). However, for a simpler example, the conditions of Theorem 3.11 can be checked rigorously (for a detailed discussion, we refer the reader to [24, Example 2.29]).

In the following examples, we want to discuss properties of certain systems like (1) that allow for the application of Theorem 3.11. The basic condition (C1) only concerns σ , and we simply assume that it holds.

Example 3.13. We return to the stochastic Hodgkin–Huxley system, the stability of which has mainly been studied numerically. There is an open neighborhood \mathcal{C} of 0 such that every constant input $S \equiv c \in \mathcal{C}$ is injectively and continuously mapped to an equilibrium

$$(x_c^*, y_c^*) = \left(x_c^*, \frac{\alpha_1(x_c^*)}{(\alpha_1 + \beta_1)(x_c^*)}, \frac{\alpha_2(x_c^*)}{(\alpha_2 + \beta_2)(x_c^*)}, \frac{\alpha_3(x_c^*)}{(\alpha_3 + \beta_3)(x_c^*)} \right) \in \mathbb{R} \times (0, 1)^3,$$

which – as simulations suggest – is stable and locally attractive (see [28, pages 533 and 548]), i.e. for all $\varepsilon > 0$ there is a $\delta = \delta(c, \varepsilon) > 0$ such that for all $\varphi_0 \in B_\delta(x_c^*, y_c^*)$ we have

$$(27) \quad (x, y)[x_0, y_0, S \equiv c](t) \in B_\varepsilon(x_c^*, y_c^*) \quad \text{for all } t \in [0, \infty),$$

and there is some $\varepsilon^* > 0$ such that for all $\varphi_0 \in B_{\varepsilon^*}(x_c^*, y_c^*)$ we actually even have

$$\varphi[(x_0, y_0), S \equiv c](t) \xrightarrow{t \rightarrow \infty} (x_c^*, y_c^*).$$

Therefore, (C2) and (C3') are taken for granted, and we write $\varphi^* := (x_0^*, y_0^*)$. Simulations show that if started slightly beside the equilibrium, the solution is immediately pulled in its direction, suggesting that one can in fact take $\delta = \varepsilon$ in (27). This is also the reason why stability in the sense of (27) should remain valid when also slightly and slowly changing the input over time as we do with the function ϱ in the third control phase in the proof of Theorem 3.11: A small change of the input c simply corresponds to a small change of the equilibrium (x_c^*, y_c^*) towards which the trajectory is headed. This is essentially (C4'). The conditions (C5) and (C6) (with $x^* \equiv x_0^*$) follow easily from the fact that for a given trajectory of the membrane potential x , variation of constants yields that we can write

$$y_i(t) = y_i(0)e^{-\int_0^t (\alpha_i + \beta_i)(x(s))ds} + \int_0^t \alpha_i(x(s))e^{-\int_s^t (\alpha_i + \beta_i)(x(r))dr} ds$$

for every $i \in \{1, 2, 3\}$ and $t \in [0, \infty)$.

Example 3.14. We consider the system from Example 1.3 for $N = 1$ and the choice of g that is given in (7). For $S \equiv 0$, the points $\varphi^* = (\pm 1, 1, \dots, 1)$ fulfill (C2), (C3'), (C5) and (C6). The key to seeing this lies in considering x and y separately and using the classical Lyapunov method for each variable in a way that works (locally) uniformly in the other variable (see [24, Example 2.29] for a detailed calculation for a similar example). Condition (C4') can be checked with essentially the same arguments as well (ibidem).

Example 3.15. We take the functions f and g from Example 3.8 and outside of $B_2(0) \subset \mathbb{R}^2$ we change their dependence on x arbitrarily. Then $\mathcal{O}^* = \mathbb{S}^1$ is still a periodic orbit of the corresponding system (9) with $S \equiv 0$, but now only local attractivity is guaranteed in x . However, (C3'), (C4') and (C5) (with $x^* = \cos(\cdot)$) can be checked via similar arguments as in Example 3.14. Note that this separation of the variables automatically solves the problem we had with the unstable equilibrium at the origin in Example 3.8.

3.3 Globally attractive periodic solution under nonzero-input

The scheme outlined in Remark 3.4 may be refined in the following way:

1. Find a T_* -periodic signal S_* (with $T_* \notin \mathbb{Q}T$) with mean zero over one period such that the corresponding deterministic system (9) has a T_* -periodic solution (with $T^* \in \mathbb{Q}T_*$) that is locally attractive.
2. Define \hat{h} in such a way that the resulting input \dot{w} steers $\psi = (u, v)$ into the domain of attraction of this solution. At the same time, w must be guided to a given $z^* \in \mathbb{W}$.
3. Once w reaches z^* , choose \hat{h} such that $\dot{w} = S_*$. Then w will keep returning to z^* while ψ approaches the attractive periodic solution.

Under certain conditions on the interplay of the different occurring periodicities, we can then establish (φ^*, z^*) as a T -attainable point whenever φ^* is a suitable value of the periodic solution.

(C2*) Incommensurable periodic solution: There is a non-constant T_* -periodic signal $S_* \in C([0, \infty); \mathbb{R}^N)$ with

$$T_* \in (0, \infty) \setminus \mathbb{Q}T \quad \text{and} \quad \int_0^{T_*} S_*(t) dt = 0$$

such that for some $T^* \in \mathbb{Q}T_*$ and $\varphi_* \in \text{int}(\mathbf{U} \times \mathbf{V})$, the unique solution

$$\varphi^*(t) := \varphi[\varphi_*, S_*](t), \quad t \in [0, \infty),$$

to the corresponding system (9) is T^* -periodic and $\varphi^*([0, T^*]) \subset \text{int}(\mathbf{U} \times \mathbf{V})$.

(C3*) Global attractivity of φ^* : For any $\varphi_0 \in \mathbf{U} \times \mathbf{V}$, there is a unique global solution to (9) with $S = S_*$, and it satisfies

$$|\varphi[\varphi_0, S_*](t) - \varphi^*(t)| \xrightarrow{t \rightarrow \infty} 0.$$

(C4*) Non-explosiveness for moderately varied input: For any $\varphi_0 \in \mathbf{U} \times \mathbf{V}$, there is a $\delta_0 = \delta_0(\varphi_0) > 0$ such that for any $S \in C([0, \infty); \mathbb{R}^N)$ with $\|S - S_*\|_\infty \leq \delta_0$, the system (9) has a unique global solution.

We can then prove the following result.

Theorem 3.16. *Let $\mathbf{W} = \mathbb{R}^N$ and assume that (C1) and (C2*) - (C4*) hold. Then for any $t^* \in [0, T^*]$ and $z^* \in \mathbb{R}^N$ the point $\Phi^* = (\varphi^*(t^*), z^*) \in \text{int}(\mathbf{E})$ is T -attainable.*

Proof. Note that for any $t^* \in (0, T^*)$, the conditions (C2*) - (C4*) are also satisfied with $S_*(t^* + \cdot)$, $\varphi^*(t^*)$, and $\varphi^*(t^* + \cdot)$ in place of S_* , φ_* , and φ^* . This proof can therefore be reduced to the case $\varphi^*(t^*) = \varphi^*(0) = \varphi_*$.

Let $\Phi_0 = (\varphi_0, z_0) \in \mathbf{E}$, $z^* \in \mathbb{R}^N$ and write $\Psi = (\psi, w)$ as in Remark 3.2. Using linear interpolation, we can construct a smooth function $\kappa: [0, \infty) \rightarrow \mathbf{W}$ with $\|\kappa\|_\infty \leq \delta_0$ such that $\kappa \equiv 0$ on $[k_0T_*, \infty)$ for some $k_0 = k_0(z_0, \delta_0) \in \mathbb{N}$ and

$$\int_0^{k_0T_*} \kappa(s) ds = z^* - z_0.$$

Then, the continuously differentiable function

$$\varrho: [0, \infty) \rightarrow \mathbb{R}^N, \quad t \mapsto z_0 + \int_0^t (S_*(s) + \kappa(s)) ds$$

fulfills not only

$$(28) \quad \varrho(0) = z_0 \quad \text{and} \quad \varrho((k_0 + k)T_*) = z^* \quad \text{for all } k \in \mathbb{N},$$

but also

$$(29) \quad \|\dot{\varrho} - S_*\|_\infty \leq \delta_0 \quad \text{and} \quad \dot{\varrho} = S_* \quad \text{on } [k_0T_*, \infty).$$

Now, we can once again define \dot{h} as in (21) in order to force w onto ϱ and then $\psi = (u, v)$ has to be the solution $\varphi[\varphi_0, \dot{\varrho}]$ to (9), which uniquely exists thanks to (C4*). By (29), we have

$$|\psi(k_0T_* + t) - \varphi^*(t)| = |\varphi[\psi(k_0T_*), S_*](t) - \varphi^*(t)| \xrightarrow{t \rightarrow \infty} 0,$$

and since φ^* has a period that is a rational multiple of T_* , we can combine this with (28) and conclude that $\Phi^* \in \overline{\Psi(\mathbb{N}T_*)}$. Since $T_* \notin \mathbb{Q}T$, an argument similar to the one in Lemma 3.19 (but simpler) shows that also $\Phi^* \in \overline{\Psi(\mathbb{N}T)}$. \square

Example 3.17. Let $T \in \mathbb{Q}$, $\mathbf{U} = \mathbf{V} = \mathbf{W} = \mathbb{R}$ and

$$f(x, y) = -(1 + y^2)x, \quad g(x, y) = x - y \quad \text{for all } x, y \in \mathbb{R}.$$

The corresponding system (9) fulfills the conditions of Theorem 3.6 with $\mathcal{O}^* = \{(0, 0)\} \subset \mathbb{R}^2$, and hence any point on the line $\{(0, 0)\} \times \mathbb{R}$ is T -attainable. If we set

$$S_*(t) = \cos t, \quad \varphi^*(t) = \frac{1}{2} \begin{pmatrix} \cos t + \sin t \\ \sin t \end{pmatrix} \quad \text{for all } t \in [0, \infty),$$

the corresponding system (9) with $S = S_*$ fulfills the assumptions of Theorem 3.16, showing that any point on the tube $\varphi^*([0, 2\pi]) \times \mathbb{R}$ is also T -attainable.

Example 3.18. We revisit Example 1.3 once again, this time for $N = 2$ and for g as in (6). If $S \equiv 0$, each point in $(\mathbb{S}^1 \cup \{0\}) \times \{0\} \subset \mathbb{R}^{2+L}$ is an equilibrium of the system (9), and no further periodic orbits exist. Unfortunately, none of these equilibria are attractive in the sense of (C3) or (C3'), so Theorems 3.6 and 3.11 provide no help at all in finding a T -attainable point. However, for $T \in \mathbb{Q}$ the assumptions of Theorem 3.16 are satisfied with

$$S_*(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}, \quad \varphi^*(t) = \begin{pmatrix} \sin t \\ -\cos t \\ 0 \end{pmatrix} \quad \text{for all } t \in [0, \infty),$$

showing that indeed any point in $\varphi^*([0, 2\pi]) \times \mathbb{R}^2 = \mathbb{S}^1 \times \{0\} \times \mathbb{R}^2 \subset \mathbb{R}^{2+L+2}$ is T -attainable. The influence of the signal S_* can be interpreted as a constant rotation of the mexican hat with the exact speed needed to turn a curve of unstable equilibria into a globally asymptotically stable periodic solution.

3.4 Two Lemmas

Lemma 3.19. *Let (C2) hold and let $\varphi_0 \in \mathbf{U} \times \mathbf{V}$. Further assume one of the following conditions:*

(a) (C3) holds.

(b) (C3') holds and $\varphi_0 \in B_{\varepsilon^*}(\mathcal{O}^*)$.

Then $\mathcal{O}^* \subset \overline{\varphi[\varphi_0, S \equiv 0](\mathbb{N}T)}$.

Proof. Let $\varphi^* \in \mathcal{O}^*$. If $T^* = 0$ and hence the orbit contains only a single equilibrium point, this lemma's claim is trivial. Therefore we assume $T^* \in (0, \infty) \setminus \mathbb{Q}T$, write $\varphi := \varphi[\varphi_0, S \equiv 0]$ and let $\varepsilon > 0$. We want to show that there is some $m \in \mathbb{N}$ such that $\varphi(mT) \in B_{2\varepsilon}(\varphi^*)$. In both cases (a) and (b), φ converges to \mathcal{O}^* , so there is some $t_0 = t_0(\varepsilon) \in (0, \infty)$ with $\varphi([t_0, \infty)) \subset B_\varepsilon(\mathcal{O}^*)$. Since $\mathcal{O}^* = \varphi[\varphi^*, S \equiv 0]([0, T^*])$ is compact, F is bounded on $B_\varepsilon(\mathcal{O}^*)$. As $\dot{\varphi} = F(\varphi)$, this means that $\varphi|_{[t_0, \infty)}$ is Lipschitz continuous with a Lipschitz constant $L(\varepsilon) \in (0, \infty)$. Let

$$\mathcal{N} := \{\varphi^* + \xi \mid \xi \in \mathbb{R}^{N+L}, \xi \cdot F(\varphi^*) = 0\} \subset \mathbb{R}^{N+L}$$

denote the hyperplane that is orthogonal to the orbit \mathcal{O}^* in the point φ^* . This is indeed a well-defined hyperplane, since $T^* > 0$ and therefore $F(\varphi^*) \neq 0$. If T_n^* denotes the n -th time after t_0 at which φ passes through $\mathcal{N} \cap B_\varepsilon(\varphi^*)$, [34, Lemma 8.5.3 and its proof] together with the attraction property (C3) or (C3') yield that T_n^* is well-defined for sufficiently small $\varepsilon > 0$ and that we have

$$T_n^* - T_{n-1}^* \xrightarrow{n \rightarrow \infty} T^*.$$

Then Lemma 3.20 below provides $m, n \in \mathbb{N}$ with

$$|T_n^* - mT| < \frac{\varepsilon}{L(\varepsilon)}$$

and hence we have

$$|\varphi(mT) - \varphi^*| \leq L(\varepsilon) |mT - T_n^*| + |\varphi(T_n^*) - \varphi^*| < 2\varepsilon,$$

which completes the proof. \square

Lemma 3.20. *Let $T \in (0, \infty)$ and $T^* \in (0, \infty) \setminus \mathbb{Q}T$, and let $(T_n^*)_{n \in \mathbb{N}} \subset \mathbb{R}$ with*

$$(30) \quad T_n^* - T_{n-1}^* \xrightarrow{n \rightarrow \infty} T^*.$$

Then for any $\varepsilon > 0$, there are $m, n \in \mathbb{N}$ such that

$$|T_n^* - mT| < \varepsilon.$$

Proof. Thanks to a strengthened version of Kronecker's Approximation Theorem (compare [1, Chapter 7, Exercise 3]), we know that for each $\varepsilon > 0$ there is an $N(\varepsilon) \in \mathbb{N}$ such that for all $t \in [0, \infty)$ there are $n(t, \varepsilon) \in \{1, \dots, N(\varepsilon)\}$ and $m(t, \varepsilon) \in \mathbb{N}$ with the property

$$|t + n(t, \varepsilon)T^* - m(t, \varepsilon)T| < \varepsilon.$$

Now let $\varepsilon > 0$. Thanks to the assumption (30) there is an $n_0 \in \mathbb{N}$ such that

$$T_{n_0}^* > 0 \quad \text{and} \quad |(T_k^* - T_{k-1}^*) - T^*| < \frac{\varepsilon}{2N(\frac{\varepsilon}{2})} \quad \text{for all } k \geq n_0.$$

Then for all $n \geq n_0$ and $m \in \mathbb{N}$ we have

$$\begin{aligned} |T_n^* - mT| &= \left| \sum_{k=n_0+1}^n ((T_k^* - T_{k-1}^*) - T^*) + T_{n_0}^* + (n - n_0)T^* - mT \right| \\ &\leq (n - n_0) \frac{\varepsilon}{2N(\frac{\varepsilon}{2})} + |T_{n_0}^* + (n - n_0)T^* - mT| \end{aligned}$$

and choosing $n = n_0 + n(T_{n_0}^*, \frac{\varepsilon}{2})$ and $m = m(T_{n_0}^*, \frac{\varepsilon}{2})$ finishes the proof. \square

4 Local Hörmander condition

In this section, we will present conditions on the external equation and on the way X and Y interact through the functions f and g that allow to prove condition (III) of Theorem 2.2 via the so-called *local Hörmander condition* (see Definition 4.2 and Lemma 4.3 below).

Notation 4.1. Whenever it is convenient, we will denote elements of $\mathbb{R}^{1+N+L+N}$ by

$$(31) \quad \xi = (\xi_0, \dots, \xi_{N+L+N}) := (t, x_1, \dots, x_N, y_1, \dots, y_L, z_1, \dots, z_N) = (t, x, y, z) = (t, \varphi, z) = (t, \Phi)$$

and use the abbreviation

$$\hat{b}(t, z) := S_0(t) + \tilde{b}(z) \quad \text{for all } (t, z) \in [0, \infty) \times \mathbb{W},$$

where \tilde{b} is as it was defined in (17). For any $m, n \in \mathbb{N}$, we will also use the notation

$$A_{\cdot, k} := \begin{pmatrix} A_{1, k} \\ \vdots \\ A_{n, k} \end{pmatrix} \in \mathbb{R}^n \quad \text{for all } k \in \{1, \dots, m\}$$

for the k -th column of a matrix $A \in \mathbb{R}^{n \times m}$, and we will write 0_n for the zero vector in \mathbb{R}^n .

For vector fields $V_1, V_2 \in C^1(D; \mathbb{R}^n)$ on some $D \subset \mathbb{R}^n$, $n \in \mathbb{N}$, the *Lie bracket* is defined as the vector field

$$[V_1, V_2] := J_{V_2} V_1 - J_{V_1} V_2 \in C(D; \mathbb{R}^n),$$

where J_{V_i} denotes the respective Jacobian matrix. Since this operation is bilinear, we can and will view the space $C^\infty(D; \mathbb{R}^n)$ as an algebra that has the Lie bracket as its multiplication. For any $\mathcal{C} \subset C^\infty(D; \mathbb{R}^n)$, we write $\langle \mathcal{C} \rangle$ for the subalgebra generated by \mathcal{C} . For any subalgebra $\mathcal{L} \subset C^\infty(D; \mathbb{R}^n)$ containing \mathcal{C} , we write $[\mathcal{C}]_{\mathcal{L}}$ for the ideal in \mathcal{L} that is generated by \mathcal{C} .

The idea behind the Lie bracket is that by combining motions in the directions V_1 and V_2 one can effectively approximate motion in the directions $[V_1, V_2]$, $[V_1, [V_1, V_2]]$, \dots and ultimately in any direction in $\langle [V_1, V_2] \rangle$. If V_1 and V_2 occur as the drift or as a column of the diffusion matrix of a Stratonovich SDE, this can be exploited in order to determine along which directions its solution can evolve locally. A detailed heuristic explanation of this idea can be found in [17, Section 2].

If we want to apply this reasoning to our time-inhomogeneous setting, we have to consider the homogeneous $(1 + N + L + N)$ -dimensional time-space process $(t, \Phi_t)_{t \in [0, \infty)}$. It solves the Stratonovich SDE

$$d(t, \Phi_t) = V_0(t, \Phi_t)dt + \sum_{k=1}^M V_k(t, \Phi_t) \circ dW_t^{(k)}$$

with the vector fields $V_k: [0, \infty) \times \mathbf{E} \rightarrow \mathbb{R}^{1+N+L+N}$ defined by

$$(32) \quad V_0(t, \Phi) := \begin{pmatrix} 1 \\ \tilde{B}(t, \Phi) \end{pmatrix}, \quad V_k(t, \Phi) := \begin{pmatrix} 0 \\ \Sigma_{\cdot, k}(\Phi) \end{pmatrix} \quad \text{for all } k \in \{1, \dots, M\},$$

with B and Σ as in (13) and (14). The following Definition 4.2 formalizes the idea that, locally, the process $(\Phi_t)_{t \in [0, \infty)}$ should be able to move in any direction. Context-specific variants are classically used in order to prove the existence of transition densities, usually incorporating tools from the Malliavin calculus (see for example [17] and the references therein).

Definition 4.2. We set

$$\mathcal{C} := \{V_1, \dots, V_M\}, \quad \mathcal{L} := \langle \{V_0, \dots, V_M\} \rangle, \quad \text{and} \quad \mathcal{L}^* := [\mathcal{C}]_{\mathcal{L}}.$$

If there is some $\Phi \in \text{int}(\mathbf{E})$ such that

$$(33) \quad \text{span}\{V(t, \Phi) \mid V \in \mathcal{L}^*\} \simeq \mathbb{R}^{N+L+N} \quad \text{for all } t \in [0, \infty),$$

we say that the *local Hörmander condition holds at Φ* .

Lemma 4.3. *The local Hörmander condition at $\Phi^* \in \text{int}(\mathbf{E})$ implies condition (III) of Theorem 2.2.*

Proof. This follows from [27, Theorems 1 and 2] and Remark 2.3. □

For general diffusions, there is very little one can say about non-trivial sufficient conditions for the local Hörmander condition. One usually depends on ad hoc arguments that exploit the particular shape of the drift and the diffusion matrix. For systems of the special type (3) however, there are two fairly general scenarios that we can treat without confining us to entirely specific examples. Both are based on the idea that external noise (in the form of the Z -variable) is explicitly imported only into the X -variable, while the coefficient functions f and g have to transport and distribute its influence suitably among all of the X - and Y -variables.

- (I) **Star shape:** Every component depends on the X -variables in such a way that sufficient amounts of noise are able to spread from Z via X to the rest of the system.
- (II) **Cascade structure:** There is a chain of components such that with each step exactly one more of them is directly influenced by the previous one, and this chain ultimately runs through the entire system.

Of course, in both of these scenarios we also have to make sure that X and Z are not coupled in a degenerate way. However, since the random terms in their respective equations coincide, their interaction is basically coded in the difference of their drift coefficients, which is again simply given by the function f . It therefore seems natural that, as long as σ is nice enough for Z not to be degenerate itself, it should be possible for us to find conditions solely on f and g that are sufficient for the local Hörmander condition.

The star shape setting is inspired by the Hodgkin-Huxley system and its variants (see Section 4 of [27], compare Example 4.15 below), while the cascade structure setting is inspired by a model of interacting neurons that is studied in Section 5.3 of [12] (see Remark 4.18). In this section, both scenarios will be treated under more general assumptions that transcend the scope of these specific models.

Before we proceed in treating each of these situations separately in the Subsections 4.1 and 4.2 below, we will first collect some preparatory observations and calculations.

Lemma 4.4. *Let $\mathcal{U} \subset \mathbf{E}$ open and assume that for $(t, \Phi) \in [0, \infty) \times \bar{\mathcal{U}}$ the mappings*

$$(t, z) \mapsto A_i(t, z) \in \mathbb{R}^N, \quad \varphi \mapsto W_i(\varphi) \in \mathbb{R}^{N+L}, \quad (t, z) \mapsto a_i(t, z) \in \mathbb{R},$$

are continuously differentiable for $i \in \{1, 2\}$. Then for all $(t, \Phi) \in [0, \infty) \times \bar{\mathcal{U}}$ we have

$$\begin{aligned} 1. & \left[\begin{pmatrix} A_1(t, \cdot) \\ 0_L \\ A_1(t, \cdot) \end{pmatrix}, \begin{pmatrix} A_2(t, \cdot) \\ 0_L \\ A_2(t, \cdot) \end{pmatrix} \right] (\Phi) = \begin{pmatrix} [A_1(t, \cdot), A_2(t, \cdot)](z) \\ 0_L \\ [A_1(t, \cdot), A_2(t, \cdot)](z) \end{pmatrix}, \\ 2. & \left[a_1(t, \cdot) \begin{pmatrix} W_1 \\ 0_N \end{pmatrix}, a_2(t, \cdot) \begin{pmatrix} W_2 \\ 0_N \end{pmatrix} \right] (\Phi) = a_1(t, z) a_2(t, z) \begin{pmatrix} [W_1, W_2](\varphi) \\ 0_N \end{pmatrix}, \\ 3. & \left[a_1(t, \cdot) \begin{pmatrix} W_1 \\ 0_N \end{pmatrix}, \begin{pmatrix} A_1(t, \cdot) \\ 0_L \\ A_1(t, \cdot) \end{pmatrix} \right] (\Phi) = - \sum_{j=1}^N A_1^{(j)}(t, z) \left(a_1(t, z) \begin{pmatrix} \partial_{x_j} W_1(\varphi) \\ 0_N \end{pmatrix} + \partial_{z_j} a_1(t, z) \begin{pmatrix} W_1(\varphi) \\ 0_N \end{pmatrix} \right). \end{aligned}$$

Proof. All of these formulas follow immediately from the definition of the Lie bracket by straight forward calculations. □

Remark 4.5. Let us make some comments on the role of time, i.e. the (0)-components of the occurring vector fields on the one hand and the derivatives with respect to $t = \xi_0$ on the other hand.

1.) First, note that for all $k \in \{0, \dots, M\}$ the (0)-component of V_k is constant. Hence, for any vector field $W \in C^1([0, \infty) \times \mathbf{E}; \mathbb{R}^{1+N+L+N})$ we have

$$[V_k, W]^{(0)} = \sum_{i=0}^{N+L+N} V_k^{(i)} \partial_{\xi_i} W^{(0)},$$

which vanishes everywhere whenever $W^{(0)}$ is constant as well. In particular, this is the case for $W \in \{V_0, \dots, V_M\}$. Consequently, any vector field that is contained in \mathcal{L}^* will have a vanishing (0)-component. Hence, $N + L + N$ is the maximum dimension that can possibly be achieved in (33) despite the fact that V_0, \dots, V_M are $(N + L + N + 1)$ -dimensional.

2.) Let $V, W \in C^1([0, \infty) \times \mathbf{E}; \mathbb{R}^{1+N+L+N})$. As

$$[V, W] = V^{(0)} \partial_t W - W^{(0)} \partial_t V + \sum_{i=1}^{N+L+N} (V^{(i)} \partial_{\xi_i} W - W^{(i)} \partial_{\xi_i} V),$$

any possible influence of time derivatives of W is killed by a vanishing $V^{(0)}$ and vice versa. The first part of this remark then yields that no time derivatives occur explicitly in

$$(34) \quad [V, W] = \sum_{i=1}^{N+L+N} (V^{(i)} \partial_{\xi_i} W - W^{(i)} \partial_{\xi_i} V) \quad \text{for all } V, W \in \mathcal{L}^*,$$

while

$$(35) \quad [V_0, W] = \partial_t W + \sum_{i=1}^{N+L+N} (V_0^{(i)} \partial_{\xi_i} W - W^{(i)} \partial_{\xi_i} V_0) \quad \text{for all } W \in \mathcal{L}^*.$$

As $\partial_t V_k$ vanishes for all $k \in \{1, \dots, M\}$, time derivatives cannot occur before taking Lie brackets at least twice in the construction of \mathcal{L}^* .

The following notational convention is of utmost importance for understanding the rest of this section.

Notation 4.6. 1.) As noted under 1.) in Remark 4.5, the only relevant vector field to feature a non-vanishing (0)-component is V_0 – which is not contained in \mathcal{L}^* . In order to simplify our notation, in the sequel **we will therefore systematically omit this component**. More precisely, we identify every vector field

$$(36) \quad W: [0, \infty) \times \mathbf{E} \rightarrow \mathbb{R}^{1+N+L+N}, \quad (t, \Phi) \mapsto \begin{pmatrix} 0 \\ W^{(1)}(t, \Phi) \\ \vdots \\ W^{(N+L+N)}(t, \Phi) \end{pmatrix},$$

with the collection of vector fields on \mathbf{E} given by

$$(37) \quad W(t, \cdot): \mathbf{E} \rightarrow \mathbb{R}^{N+L+N}, \quad \Phi \mapsto \begin{pmatrix} W^{(1)}(t, \Phi) \\ \vdots \\ W^{(N+L+N)}(t, \Phi) \end{pmatrix}, \quad \text{for all } t \in [0, \infty).$$

Either of these objects will simply be denoted by W . Let us stress that both directions of this identification are actively used: on the one hand, we tacitly omit vanishing zero-components when a vector field is given as in (36), but on the other hand, we also think of each vector field that is given in terms of (37) as one on $[0, \infty) \times \mathbf{E}$ with a vanishing zero-component.

2.) If V and W are of this type, so is their Lie bracket $[V, W]$ (confer part 1.) of Remark 4.5). Thus, it is natural (and consistent) to identify

$$(38) \quad [V, W](t, \Phi) = [V(t, \cdot), W(t, \cdot)](\Phi) \in \mathbb{R}^{N+L+N}$$

for all $(t, \Phi) \in [0, \infty) \times \mathbf{E}$.

3.) As indicated above, V_0 is the only relevant vector field not of the type in (36) and (37), but we only need it when taking Lie brackets as in (35). Using the notation we just introduced, (35) can be rewritten as

$$(39) \quad [V_0, W] = \partial_t W + \left[\begin{pmatrix} \hat{b} \\ 0_L \\ \hat{b} \end{pmatrix} + \begin{pmatrix} F \\ 0_N \end{pmatrix}, W \right].$$

Note that this formula is completely consistent: $[V_0, W]^{(0)}$ and $\partial_t W^{(0)}$ vanish again and are therefore omitted, while the Lie bracket on the right hand side is to be understood as the Lie bracket of vector fields on \mathbf{E} (which carry an additional dependence on time) and thus has only $N + L + N$ components to begin with.

4.) Note also that with this notational convention there is no difference between V_k and $\Sigma_{\cdot, k}$ for $k \in \{1, \dots, M\}$.

5.) Let us further illustrate this notation by rewriting the formulas from Lemma 4.4 accordingly. With the identification of (36) and (37) and with the convention (38), they read

$$\begin{aligned} 1. \quad & \left[\begin{pmatrix} A_1 \\ 0_L \\ A_1 \end{pmatrix}, \begin{pmatrix} A_2 \\ 0_L \\ A_2 \end{pmatrix} \right] (\xi) = \begin{pmatrix} [A_1(t, \cdot), A_2(t, \cdot)](z) \\ 0_L \\ [A_1(t, \cdot), A_2(t, \cdot)](z) \end{pmatrix}, \\ 2. \quad & \left[a_1 \begin{pmatrix} W_1 \\ 0_N \end{pmatrix}, a_2 \begin{pmatrix} W_2 \\ 0_N \end{pmatrix} \right] (\xi) = a_1(t, z) a_2(t, z) \begin{pmatrix} [W_1, W_2](\varphi) \\ 0_N \end{pmatrix}, \\ 3. \quad & \left[a_1 \begin{pmatrix} W_1 \\ 0_N \end{pmatrix}, \begin{pmatrix} A_1 \\ 0_L \\ A_1 \end{pmatrix} \right] (\xi) = - \sum_{j=1}^N A_1^{(j)}(t, z) \begin{pmatrix} \partial_{x_j} W_1(\varphi) \\ 0_N \end{pmatrix} \partial_{z_j} a_1(t, z) \begin{pmatrix} W_1(\varphi) \\ 0_N \end{pmatrix} \end{aligned}$$

for all $\xi = (t, \varphi, z) \in [0, \infty) \times \bar{U}$.

Using Notation 4.6 and Lemma 4.4, we will now see which vector fields can be constructed by taking Lie brackets with V_0, \dots, V_M . The first formula in Lemma 4.4 yields

$$(40) \quad \mathcal{L}^* \supset \langle \{V_1, \dots, V_M\} \rangle = \left\{ \begin{pmatrix} w \\ 0_L \\ w \end{pmatrix} \mid w \in \langle \{\sigma_{\cdot, 1}, \dots, \sigma_{\cdot, M}\} \rangle \right\},$$

i.e. without incorporating the drift, we can span at most N dimensions.

Definition 4.7. We say that σ is non-degenerate in $z \in \mathbb{W}$ if

$$(41) \quad \text{span}\{w(z) \mid w \in \langle \{\sigma_{\cdot, 1}, \dots, \sigma_{\cdot, M}\} \rangle\} = \mathbb{R}^N.$$

If, for example, $\sigma(z)$ is surjective (as in condition (C1) in Section 3), it has to have N linearly independent columns, and hence we get non-degeneracy without even taking Lie brackets.

Let $(x, y, z) \in \mathbf{E}$ and assume that σ is non-degenerate in z . In order to span the remaining $N + L$ dimensions, we will always start by taking the Lie bracket of the drift V_0 with some column V_k of the diffusion matrix, $k \in \{1, \dots, M\}$. Using (39) and Lemma 4.4, we see that

$$(42) \quad [V_0, V_k](\xi) = \begin{pmatrix} [\hat{b}(t, \cdot), \sigma_{\cdot, k}](z) \\ 0_L \\ [\hat{b}(t, \cdot), \sigma_{\cdot, k}](z) \end{pmatrix} - \sum_{i=1}^N \sigma_{i, k}(z) \begin{pmatrix} \partial_{x_i} F(x, y) \\ 0_N \end{pmatrix}$$

and hence

$$(43) \quad [V_k, V_0](\xi) = -[V_0, V_k](\xi) = \begin{pmatrix} [\sigma_{\cdot, k}, \hat{b}(t, \cdot)](z) \\ 0_L \\ [\sigma_{\cdot, k}, \hat{b}(t, \cdot)](z) \end{pmatrix} + \sum_{i=1}^N \sigma_{i, k}(z) \begin{pmatrix} \partial_{x_i} F(x, y) \\ 0_N \end{pmatrix}.$$

In order to get a better idea of what we are dealing with here, let us see what happens when we take an extra Lie bracket with V_0 or some column V_l of the diffusion matrix, $l \in \{1, \dots, M\}$. Using (39), Lemma 4.4, and basic properties of the Lie bracket again, we see that

$$\begin{aligned}
[V_0, [V_0, V_k]](\xi) &= \partial_t[V_0, V_k](\xi) + \left[\begin{pmatrix} \hat{b} \\ 0_L \\ \hat{b} \end{pmatrix} + \begin{pmatrix} F \\ 0_N \end{pmatrix}, [V_0, V_k] \right](\xi) \\
&= \begin{pmatrix} [\hat{b}(t, \cdot), [\hat{b}(t, \cdot), \sigma_{\cdot, k}]](z) + \partial_t[\hat{b}(t, \cdot), \sigma_{\cdot, k}](z) \\ 0_L \\ [\hat{b}(t, \cdot), [\hat{b}(t, \cdot), \sigma_{\cdot, k}]](z) + \partial_t[\hat{b}(t, \cdot), \sigma_{\cdot, k}](z) \end{pmatrix} \\
(44) \quad &- \sum_{i=1}^N \left([\hat{b}(t, \cdot), \sigma_{\cdot, k}]^{(i)}(z) \begin{pmatrix} \partial_{x_i} F(x, y) \\ 0_N \end{pmatrix} + \sigma_{i, k}(z) \begin{pmatrix} [F, \partial_{x_i} F](x, y) \\ 0_N \end{pmatrix} \right) \\
&\quad + \sum_{j=1}^N \hat{b}^{(j)}(t, z) \left(\sigma_{i, k}(z) \begin{pmatrix} \partial_{x_j} \partial_{x_i} F(x, y) \\ 0_N \end{pmatrix} + \partial_{z_j} \sigma_{i, k}(z) \begin{pmatrix} \partial_{x_i} F(x, y) \\ 0_N \end{pmatrix} \right).
\end{aligned}$$

Similarly, we can calculate

$$(45) \quad [V_l, [V_k, V_0]](\xi) = \begin{pmatrix} \zeta_{k, l}(t, z) \\ 0_L \\ \zeta_{k, l}(t, z) \end{pmatrix} + \sum_{i, j=1}^N \sigma_{j, l}(z) \left(\sigma_{i, k}(z) \begin{pmatrix} \partial_{x_j} \partial_{x_i} F(x, y) \\ 0_N \end{pmatrix} + \partial_{z_j} \sigma_{i, k}(z) \begin{pmatrix} \partial_{x_i} F(x, y) \\ 0_N \end{pmatrix} \right),$$

where

$$(46) \quad \begin{aligned}
\zeta_{k_1}(t, z) &:= [\sigma_{\cdot, k_1}, \hat{b}(t, \cdot)](z), \\
\zeta_{k_1, \dots, k_n}(t, z) &:= [\sigma_{\cdot, k_n}, \zeta_{k_1, \dots, k_{n-1}}(t, \cdot)](z) \quad \text{for all } n \geq 2,
\end{aligned}$$

with any $k_1, k_2, \dots \in \{1, \dots, M\}$.

Having thus acquired a basic idea of the typical structure of higher iterations

$$[V_0, [\dots, [V_0, V_k]]], [V_{k_n}, [\dots, [V_{k_1}, V_0]]] \in \mathcal{L}^*,$$

we formulate the following two Lemmas that will help us treat these in detail later. Their proofs are straight forward, following the same line of arguments as in the above calculations of $[V_0, [V_0, V_k]]$ and $[V_l, [V_k, V_0]]$.

Lemma 4.8. *If there is an open $\mathcal{U} \subset \mathbf{E}$ such that $W \in C^1(\mathbf{E}, \mathbb{R}^{1+N+L+N})$ can locally be written as*

$$W(\xi) = \begin{pmatrix} A(t, z) \\ 0_L \\ A(t, z) \end{pmatrix} - \sum_{i=1}^n a_i(t, z) \begin{pmatrix} W_i(x, y) \\ 0_N \end{pmatrix} \quad \text{for all } \xi \in [0, \infty) \times \bar{\mathcal{U}},$$

for suitable $n \in \mathbb{N}$ and smooth functions A, a_i and W_i , then

$$\begin{aligned}
[V_0, W](\xi) &= \begin{pmatrix} [\hat{b}(t, \cdot), A(t, \cdot)](z) + \partial_t A(t, z) \\ 0_L \\ [\hat{b}(t, \cdot), A(t, \cdot)](z) + \partial_t A(t, z) \end{pmatrix} - \sum_{i=1}^n \partial_t a_i(t, z) \begin{pmatrix} W_i(x, y) \\ 0_N \end{pmatrix} \\
&- \sum_{j=1}^N A^{(j)}(t, z) \begin{pmatrix} \partial_{x_j} F(x, y) \\ 0_N \end{pmatrix} - \sum_{i=1}^n a_i(t, z) \begin{pmatrix} [F, W_i](x, y) \\ 0_N \end{pmatrix} \\
&- \sum_{i=1}^n \sum_{j=1}^N \hat{b}^{(j)}(t, z) \left(a_i(t, z) \begin{pmatrix} \partial_{x_j} W_i(x, y) \\ 0_N \end{pmatrix} + \partial_{z_j} a_i(t, z) \begin{pmatrix} W_i(x, y) \\ 0_N \end{pmatrix} \right)
\end{aligned}$$

for all $\xi \in [0, \infty) \times \bar{\mathcal{U}}$.

Lemma 4.9. *If there is an open $\mathcal{U} \subset \mathbf{E}$ such that $W \in C^1(\mathbf{E}, \mathbb{R}^{1+N+L+N})$ can locally be written as*

$$W(\xi) = \begin{pmatrix} A(t, z) \\ 0_L \\ A(t, z) \end{pmatrix} + \sum_{i=1}^n a_i(z) \begin{pmatrix} W_i(x, y) \\ 0_N \end{pmatrix} \quad \text{for all } \xi \in [0, \infty) \times \mathcal{U}.$$

for suitable $n \in \mathbb{N}$ and smooth functions A, a_i and W_i , then

$$[V_k, W](\xi) = \begin{pmatrix} [\sigma_{\cdot, k}, A(t, \cdot)](z) \\ 0_L \\ [\sigma_{\cdot, k}, A(t, \cdot)](z) \end{pmatrix} + \sum_{i=1}^n \sum_{j=1}^N \sigma_{j, k}(z) \left(a_i(z) \begin{pmatrix} \partial_{x_j} W_i(x, y) \\ 0_N \end{pmatrix} + \partial_{z_j} a_i(z) \begin{pmatrix} W_i(x, y) \\ 0_N \end{pmatrix} \right)$$

for all $\xi \in [0, \infty) \times \bar{\mathcal{U}}$ and any $k \in \{1, \dots, M\}$.

4.1 Star shape

The aim of this section is to prove the following theorem and two other variants of it that also include the case $N \neq M$ (see Theorem 4.13 and Corollary 4.14 below).

Theorem 4.10. *Let $\Phi = (x, y, z) \in \text{int}(\mathbf{E})$, $M = N$ and assume that $\sigma(z)$ is an invertible diagonal matrix. If for some $k_1, \dots, k_{N+L} \in \{1, \dots, N\}$ the vectors*

$$(47) \quad \partial_{x_{k_1}} \cdots \partial_{x_{k_n}} F(x, y), \quad n \in \{1, \dots, N+L\},$$

are linearly independent, then the local Hörmander condition holds at Φ .

Our proof strategy is to span all dimensions with vector fields that are constructed in the following fashion: Let $l \in \mathbb{N}$ and $\kappa = (k_1, \dots, k_l) \in \{1, \dots, M\}^l$ and define

$$L_{\kappa,1} := [V_{k_1}, V_0] \quad \text{and} \quad L_{\kappa,n} := [V_{k_n}, L_{\kappa,n-1}] \quad \text{for all } n \in \{2, \dots, l\}.$$

When this recursive procedure is finished, we have acquired the vector field $L_{\kappa,l} \in \mathcal{L}^*$. As this is the result of successively taking Lie brackets with vector fields that follow the "path" $\kappa = (k_1, \dots, k_l)$ through the columns of the diffusion matrix, we will use the notation $L_\kappa := L_{\kappa,l}$. Our goal is to find verifiable criteria under which for some sequence $\kappa_1, \dots, \kappa_{N+L}$ of such paths (of possibly different lengths) and for some

$$W_1, \dots, W_N \in \langle \{V_1, \dots, V_M\} \rangle,$$

the span of the vector fields

$$W_1, \dots, W_N, L_{\kappa_1}, \dots, L_{\kappa_{N+L}} \in \mathcal{L}^*,$$

evaluated at a suitable point $\xi \in [0, \infty) \times \mathbf{E}$ has dimension $N + L + N$. To this end, we will have to calculate L_κ for a general path

$$\kappa = (k_1, \dots, k_{l(\kappa)}) \in \bigcup_{l \in \mathbb{N}} \{1, \dots, M\}^l,$$

where we write $l(\kappa)$ for the path length of κ . Thanks to Lemma 4.9, these calculations are not too hard. Recall the definition of the functions ζ_{k_1, \dots, k_n} that was given in (46).

Lemma 4.11. *For any $\kappa = (k_1, \dots, k_l) \in \{1, \dots, M\}^l$, $n \in \{1, \dots, l\}$, and $\xi \in [0, \infty) \times \mathbf{E}$, we have*

$$(48) \quad L_{\kappa,n}(\xi) = \begin{pmatrix} \zeta_{k_1, \dots, k_n}(t, z) \\ 0_L \\ \zeta_{k_1, \dots, k_n}(t, z) \end{pmatrix} + \sum_{i_1, \dots, i_n=1}^N \sigma_{i_1, k_1}(z) \cdots \sigma_{i_n, k_n}(z) \begin{pmatrix} \partial_{x_{i_1}} \cdots \partial_{x_{i_n}} F(x, y) \\ 0_N \end{pmatrix} \\ + \sum_{\alpha \in \mathbb{N}_0^N, 1 \leq |\alpha|_1 \leq n-1} p_{\kappa, n, \alpha}(z) \begin{pmatrix} \partial_x^\alpha F(x, y) \\ 0_N \end{pmatrix},$$

where each coefficient function $p_{\kappa, n, \alpha}$ is a polynomial expression of the terms

$$\partial_z^\beta \sigma_{i, k_j} \quad \text{with } \beta \in \mathbb{N}_0^N, |\beta|_1 \leq |\alpha|_1, i \in \{1, \dots, N\}, j \in \{1, \dots, n\},$$

where $|\alpha|_1 := \alpha_1 + \dots + \alpha_N$ for all $\alpha \in \mathbb{N}_0^N$.

Proof. Fix some $\kappa = (k_1, \dots, k_l) \in \{1, \dots, M\}^l$. We want to prove this Lemma by induction, so at first we notice that the cases $n = 1$ and $n = 2$ were already checked in (43) and (45). Assume now that this Lemma's claim holds for some $n \in \{1, \dots, l-1\}$. Then Lemma 4.9 yields

$$L_{\kappa, n+1}(\xi) = \begin{pmatrix} \zeta_{k_1, \dots, k_{n+1}}(t, z) \\ 0_L \\ \zeta_{k_1, \dots, k_{n+1}}(t, z) \end{pmatrix} + \sum_{i_1, \dots, i_{n+1}=1}^N \sigma_{i_1, k_1}(z) \cdots \sigma_{i_{n+1}, k_{n+1}}(z) \begin{pmatrix} \partial_{x_{i_1}} \cdots \partial_{x_{i_{n+1}}} F(x, y) \\ 0_N \end{pmatrix} \\ + \sum_{i_1, \dots, i_{n+1}=1}^N \sigma_{i_{n+1}, k_{n+1}}(z) \partial_{z_{i_{n+1}}} [\sigma_{i_1, k_1}(z) \cdots \sigma_{i_n, k_n}(z)] \begin{pmatrix} \partial_{x_{i_1}} \cdots \partial_{x_{i_n}} F(x, y) \\ 0_N \end{pmatrix} \\ + \sum_{\alpha \in \mathbb{N}_0^N, 1 \leq |\alpha|_1 \leq n-1} \sum_{i=1}^N \sigma_{i, k_{n+1}}(z) p_{\kappa, n, \alpha}(z) \begin{pmatrix} \partial_{x_i} \partial_x^\alpha F(x, y) \\ 0_N \end{pmatrix} \\ + \sum_{\alpha \in \mathbb{N}_0^N, 1 \leq |\alpha|_1 \leq n-1} \sum_{i=1}^N \sigma_{i, k_{n+1}}(z) \partial_{z_i} p_{\kappa, n, \alpha}(z) \begin{pmatrix} \partial_x^\alpha F(x, y) \\ 0_N \end{pmatrix},$$

where the summands in the first line are already exactly of the shape we are aiming for. All of the other terms correspond to derivatives of F that are of the order n at most. By the induction hypothesis, each $p_{\kappa,n,\alpha}$ is a polynomial expression of the terms $\partial_z^\beta \sigma_{i,k_j}$ with $\beta \in \mathbb{N}_0^N$, $|\beta|_1 \leq |\alpha|_1$, $i \in \{1, \dots, N\}$, $j \in \{1, \dots, n\}$. Trivially, the same is then true for any partial derivative $\partial_{z_i} p_{\kappa,n,\alpha}$ and hence all of the respective coefficients in the above formula are polynomial expressions of $\partial_z^\beta \sigma_{i,k_j}$ with $i \in \{1, \dots, N\}$, $j \in \{1, \dots, n+1\}$ and $\beta \in \mathbb{N}_0^N$ with $|\beta|_1$ bounded by the order of the respective derivative of F . Thus, $L_{\kappa,n+1}(\xi)$ is indeed of the desired form, which completes the proof. \square

Remark 4.12. Having a closer look at the formula in the proof of Lemma 4.11, we note that the coefficients $p_{\kappa,n,\alpha}$ can be calculated by the following scheme:

1. For all $n \in \mathbb{N}$, define

$$p_{\kappa,n,0_N} := 0$$

and

$$p_{\kappa,n,\alpha} := \prod_{j=1}^n \sigma_{i_j,k_j}$$

for all $\alpha \in \mathbb{N}_0^N$ with $|\alpha|_1 = n$ and $\partial_x^\alpha = \partial_{x_{i_1}} \cdots \partial_{x_{i_n}}$.

2. Now we can recursively calculate

$$p_{\kappa,n,\alpha} = \sum_{i=1}^N \sigma_{i,k_n} (\partial_{z_i} p_{\kappa,n-1,\alpha} + p_{\kappa,n-1,\alpha-e_i} \cdot 1_{\alpha_i \neq 0})$$

for all $n \geq 2$ and $\alpha \in \mathbb{N}_0^N$ with $|\alpha|_1 \in \{1, \dots, n-1\}$, where e_i denotes the i -th canonical unit vector in \mathbb{R}^N .

Let us now consider a finite sequence of paths

$$\kappa_1, \dots, \kappa_{N+L} \in \bigcup_{l \in \mathbb{N}} \{1, \dots, M\}^l$$

and the corresponding sequence of vector fields $L_{\kappa_1}, \dots, L_{\kappa_{N+L}} \in \mathcal{L}^*$ as defined in (48). Let $\xi = (t, x, y, z) \in [0, \infty) \times \mathbf{E}$ and assume that σ is non-degenerate in z , i.e. there are

$$w_1, \dots, w_N \in \langle \{\sigma_{\cdot,1}, \dots, \sigma_{\cdot,M}\} \rangle$$

such that $w_1(z), \dots, w_N(z) \in \mathbb{R}^N$ are linearly independent. Setting

$$W_1 := \begin{pmatrix} w_1 \\ 0_L \\ w_1 \end{pmatrix}, \dots, W_N := \begin{pmatrix} w_N \\ 0_L \\ w_N \end{pmatrix} \in \langle \{V_1, \dots, V_M\} \rangle,$$

we want to find a sufficient condition for

$$W_1(z), \dots, W_N(z), L_{\kappa_1}(\xi), \dots, L_{\kappa_{N+L}}(\xi)$$

to be linearly independent. We first note that for each $\kappa_n = (k_{n,1}, \dots, k_{n,l(\kappa_n)})$ the term

$$\zeta_{k_{n,1}, \dots, k_{n,l(\kappa_n)}}(t, z) \in \mathbb{R}^N$$

can of course be expressed as a linear combination of the linearly independent vectors $w_1(z), \dots, w_N(z) \in \mathbb{R}^N$, and consequently the first summand

$$\begin{pmatrix} \zeta_{k_{n,1}, \dots, k_{n,l(\kappa_n)}}(t, z) \\ 0_L \\ \zeta_{k_{n,1}, \dots, k_{n,l(\kappa_n)}}(t, z) \end{pmatrix} \in \mathbb{R}^{N+L+N}$$

of $L_{\kappa_n}(\xi)$ is a linear combination of $W_1(z), \dots, W_N(z) \in \mathbb{R}^{N+L+N}$. All of its other summands are zero in their last N components. Therefore, the vectors

$$W_1(z), \dots, W_N(z), L_{\kappa_1}(\xi), \dots, L_{\kappa_{N+L}}(\xi) \in \mathbb{R}^{N+L+N}$$

are linearly independent if and only if

$$\bar{L}_{\kappa_1}(\xi), \dots, \bar{L}_{\kappa_{N+L}}(\xi) \in \mathbb{R}^{N+L}$$

are, where

$$(49) \quad \begin{aligned} \bar{L}_{\kappa_n}(\xi) = & \sum_{i_1, \dots, i_{l(\kappa_n)}=1}^N \sigma_{i_1, k_{n,1}}(z) \cdots \sigma_{i_{l(\kappa_n)}, k_{n, l(\kappa_n)}}(z) \partial_{x_{i_1}} \cdots \partial_{x_{i_{l(\kappa_n)}}} F(x, y) \\ & + \sum_{\alpha \in \mathbb{N}_0^N, 1 \leq |\alpha| \leq l(\kappa_n) - 1} p_{\kappa_n, l(\kappa_n), \alpha}(z) \partial_x^\alpha F(x, y) \end{aligned}$$

for every $n \in \{1, \dots, N+L\}$. Note that these vector fields in fact no longer depend on time, which is why we can simply write

$$\bar{L}_{\kappa_n}(x, y, z) = \bar{L}_{\kappa_n}(\xi).$$

The following theorem collects the insight we have gained so far. Note that the signal (the only time-dependent object present) is not part of any of the occurring vector fields, which is why uniformity in time (as required for the local Hörmander condition) is entirely unproblematic.

Theorem 4.13. *Let $\Phi = (x, y, z) \in \text{int}(\mathbf{E})$ and assume that the following two conditions hold.*

- (i) σ is non-degenerate in z .
- (ii) There are

$$\kappa_1, \dots, \kappa_{N+L} \in \bigcup_{l \in \mathbb{N}} \{1, \dots, M\}^l$$

such that

$$\bar{L}_{\kappa_1}(\Phi), \dots, \bar{L}_{\kappa_{N+L}}(\Phi) \quad \text{are linearly independent.}$$

Then the local Hörmander condition holds at Φ .

Proof. As we have reasoned above, the assumptions of this theorem imply that there are $W_1, \dots, W_N \in \langle \{V_1, \dots, V_M\} \rangle \subset \mathcal{L}^*$ such that the span of

$$W_1(z), \dots, W_N(z), L_{\kappa_1}(t, \Phi), \dots, L_{\kappa_{N+L}}(t, \Phi)$$

has the dimension $N+L+N$ for all $t \in [0, \infty)$. Since by construction $L_{\kappa_1}, \dots, L_{\kappa_{N+L}}$ are also in \mathcal{L}^* , it follows that the local Hörmander condition holds at Φ . \square

In general, the assumptions of Theorem 4.13 can turn out to be still very hard to verify, since the vector fields in (49) have a rather complicated structure. However, we can present certain interesting situations in which they can be radically simplified.

Corollary 4.14. *Let $\Phi = (x, y, z) \in \text{int}(\mathbf{E})$ and $\kappa_1, \dots, \kappa_{N+L} \in \bigcup_{l \in \mathbb{N}} \{1, \dots, M\}^l$.*

- (a) *Assume that in a neighborhood of z , the diffusion matrix σ is constant and its only value is a surjective matrix. If the vectors*

$$\sum_{i_1, \dots, i_{l(\kappa_n)}=1}^N \sigma_{i_1, k_{n,1}} \cdots \sigma_{i_{l(\kappa_n)}, k_{n, l(\kappa_n)}} \partial_{x_{i_1}} \cdots \partial_{x_{i_{l(\kappa_n)}}} F(x, y), \quad n \in \{1, \dots, N+L\},$$

are linearly independent, then the local Hörmander condition holds at Φ .

- (b) *Let $M = N$ and assume that in a neighborhood of z , the diffusion matrix σ is constant and that its only value is an invertible diagonal matrix. If the vectors*

$$\partial_{x_{k_{n,1}}} \cdots \partial_{x_{k_{n, l(\kappa_n)}}} F(x, y), \quad n \in \{1, \dots, N+L\},$$

are linearly independent, then the local Hörmander condition holds at Φ .

Proof. Since (b) follows from (a), we only have to prove (a). As the diffusion matrix takes only one value $\sigma \in \mathbb{R}^{M \times N}$, condition (i) of Theorem 4.13 is in fact equivalent to surjectivity of this matrix. Furthermore, it follows from Remark 4.12 that for constant σ all of the coefficients $p_{\kappa_n, l(\kappa_n), \alpha}$ with $1 \leq |\alpha|_1 \leq l(\kappa_n) - 1$ vanish. Therefore, $\bar{L}_{\kappa_1}, \dots, \bar{L}_{\kappa_{N+L}}$ can be simplified to the respective expressions given in this Corollary. Their linear independence yields condition (ii) of Theorem 4.13, which can then be applied to complete the proof. \square

With Theorem 4.13, we can also handle the

Proof of Theorem 4.10. Once again, we want to apply Theorem 4.13. By invertibility of $\sigma(z)$, all of its columns are linearly independent, and hence condition (i) is trivially fulfilled.

In order to check condition (ii), we set

$$(50) \quad \kappa_n := (k_1, \dots, k_n) \quad \text{for every } n \in \{1, \dots, N+L\}.$$

The vector field \bar{L}_{κ_n} can be calculated by the formula (49), and since $\sigma(z)$ is a diagonal matrix, the first sum is reduced to just the one summand

$$\sigma_{k_1, k_1}(z) \cdots \sigma_{k_n, k_n}(z) \partial_{x_{k_1}} \cdots \partial_{x_{k_n}} F(x, y),$$

while the second sum is reduced to

$$\sum_{l=1}^{n-1} p_{\kappa_n, n, \alpha(l)}(z) \partial_{x_{k_1}} \cdots \partial_{x_{k_l}} F(x, y)$$

with $\alpha(l) := e_{k_1} + \dots + e_{k_l}$, as follows immediately from Remark 4.12. Invertibility of $\sigma(z)$ yields that $\sigma_{k_1, k_1}(z) \cdots \sigma_{k_n, k_n}(z)$ can never be zero, so none of the leading coefficients of $\bar{L}_{\kappa_n}(\Phi)$ vanish. Hence, starting with $n = 1$, we can successively add multiples of $\bar{L}_{\kappa_n}(\Phi)$ to $\bar{L}_{\kappa_{n+1}}(\Phi)$ in order to eliminate any summand in $\bar{L}_{\kappa_{n+1}}(\Phi)$ with derivatives of the order less than $n + 1$. The resulting vectors are clearly linearly independent if and only if $\bar{L}_{\kappa_1}(\Phi), \dots, \bar{L}_{\kappa_{N+L}}(\Phi)$ are. This leads to the conclusion that linear independence of

$$\sigma_{k_1, k_1}(z) \cdots \sigma_{k_n, k_n}(z) \partial_{x_{k_1}} \cdots \partial_{x_{k_n}} F(x, y), \quad n \in \{1, \dots, N+L\},$$

is sufficient for condition (ii) of Theorem 4.13. As leaving out scaling factors has no effect on linear independence, the proof is completed. \square

Example 4.15. For $N = M = 1$ and locally positive σ , Theorem 4.10 contains Theorem 3 of [27] as a special case. In particular, it covers the Hodgkin–Huxley model, as is calculated in detail in Section 5.3 of [27].

Example 4.16. In continuation of Example 3.18, we consider the toy model from Example 1.3 for $N = 2$, $L = 1$ and g as in (6). In Example 3.18, we showed in particular that (for $T \in \mathbb{Q}$) the point $\Phi = (1, 0, 0, z)$ is T -attainable for any $z \in \mathbb{R}^2$. Straight forward calculations show that Theorem 4.10 with $(k_1, k_2, k_3) = (1, 2, 2)$ yields the local Hörmander condition in Φ when $\sigma(z) \in \mathbb{R}^{2 \times 2}$ is an invertible diagonal matrix. If $M \neq N$ and for instance

$$\sigma \equiv \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{in some open set } \mathcal{U} \subset \mathbb{R}^2,$$

we can check the local Hörmander condition in Φ by taking $\kappa_1 = (1)$, $\kappa_2 = (1, 2)$ and $\kappa_3 = (2, 2)$ and applying Corollary 4.14. Since, as noted in Remark 2.4, there is also a Lyapunov function in this setting, this means that for either choice of σ , all of the assumptions of Theorem 2.2 are satisfied and the solution to the corresponding SDE (3) is positive Harris recurrent (even exponentially ergodic).

4.2 Cascade structure

For this section, we will stick to the case $M = N = 1$. Since σ is now \mathbb{R} -valued, we have

$$(51) \quad \mathbf{B} := \{z \in \mathbb{W} \mid \sigma \text{ is non-degenerate in } z \in \mathbb{W}\} = \{z \in \mathbb{W} \mid \sigma(z) \neq 0\} \subset \mathbb{W} \subset \mathbb{R}.$$

We tacitly assume that \mathbf{B} is non-empty, since otherwise our model would be devoid of any randomness. Moreover, we fix an open set $\mathbf{A} \subset \mathbf{U} \times \mathbf{V}$ and write

$$\mathbf{E}^* := \mathbf{A} \times \mathbf{B} \subset \mathbf{E} \subset \mathbb{R}^{1+L+1}.$$

Let us first state this section's main result.

Theorem 4.17. *Consider the following conditions.*

(H1) *In \mathbf{A} we have $\partial_x g_1, \partial_{y_1} g_2, \partial_{y_2} g_3, \dots, \partial_{y_{L-1}} g_L \neq 0$ and $\partial_{y_i} g_j = 0$ for all $i \in \{1, \dots, L\}, j \in \{i+2, \dots, L+1\}$.*

(H2) *In \mathbf{A} we have $\partial_x f \partial_x^2 g_1 \neq \partial_x^2 f \partial_x g_1$.*

If (H1) and (H2) are both valid, the local Hörmander condition holds at any $\Phi \in \mathbf{E}^$.*

Remark 4.18. A more compact formulation of the cascade condition (H1) using (11) and (31) would be that in \mathbf{A} for all $i \in \{1, \dots, L\}$ we have

$$\partial_{\xi_i} F_{i+1} \neq 0 \quad \text{and} \quad \partial_{\xi_i} F_j = 0 \quad \text{for all } j \in \{i+2, \dots, L+1\},$$

or in less strict but maybe more intuitive notation

$$\partial_{\xi_i} F = \begin{pmatrix} \square_i \\ \neq 0 \\ 0_{L-i} \end{pmatrix} \quad \text{for all } i \in \{1, \dots, L\},$$

where $\square_i \in \mathbb{R}^i$ indicates that we do not care about the values of these first i entries. Intuitively speaking, while the variable ξ_i may or may not influence the first i components of F , it definitely influences F_{i+1} and definitely does not immediately influence any components afterwards – hence the term cascade structure. This setting is mainly inspired by the diffusion given by (5.26) in [12], which features two cascades that are similar in nature to the scenario we study in this section. There, the authors discuss an approximating diffusion for a model of interacting neurons that are divided into two groups. In each of these groups, a current is passed on from one neuron to the next, while only at the ends of these chains neurons are subject to noise and interact with neurons from the other group.

Condition (H2) serves the purpose of decoupling X and Z , as will become clear in the proof of Theorem 4.17. Note that if (H1) already holds,

(H2') for all $(x, y) \in \mathbf{A}$ the vectors $\partial_x F(x, y)$ and $\partial_x^2 F(x, y)$ are linearly independent,

is equivalent to (H2).

In order to prove Theorem 4.17, we take Lie brackets in a way that is somewhat dual to the approach we used in Section 4.1. If we start with the vector field V_1 corresponding to the single column of the $(1+L+1) \times 1$ -dimensional diffusion matrix and then keep taking Lie brackets with the vector field V_0 corresponding to the drift, intuition suggests that the structure from (H1) will turn up there again in some sense, providing more and more linearly independent vectors. So, this time we set

$$(52) \quad L_1 := [V_0, V_1], \quad \text{and} \quad L_n := [V_0, L_{n-1}] \quad \text{for all } n \geq 2.$$

We have already calculated L_1 and L_2 in (42) and (44). Using Lemma 4.8, we are able to get a better grasp of what L_n looks like for larger n . This is the content of the following Lemma, which shows that the cascade structure in F is indeed passed on to L_1, \dots, L_L in some sense.

Lemma 4.19. *Let $n \in \{1, \dots, L\}$ and assume that the cascade condition (H1) holds. Then for all $\xi = (t, x, y, z) \in [0, \infty) \times \mathbf{E}^*$ we can write*

$$(53) \quad L_n(\xi) = \begin{pmatrix} A_n(t, z) \\ 0_L \\ A_n(t, z) \end{pmatrix} - \begin{pmatrix} \bar{L}_n(\xi) \\ 0 \end{pmatrix},$$

where

$$(54) \quad \bar{L}_n(\xi) := \sum_{i=1}^{m_n-1} a_{n,i}(t, z) W_{n,i}(x, y) + (-1)^{n+1} \sigma(z) W_{n,m_n}(x, y)$$

for suitable $m_n \in \mathbb{N}$ and smooth functions $a_{n,i}$, A_n , and $W_{n,i}$ with

$$W_{n,i} = \begin{pmatrix} \square_n \\ 0_{1+L-n} \end{pmatrix} \quad \text{for all } i \in \{1, \dots, m_n - 1\}, \quad W_{n,m_n} = \begin{pmatrix} \square_n \\ \neq 0 \\ 0_{1+L-(n+1)} \end{pmatrix}$$

everywhere in \mathbf{A} .

Proof. In analogy to Lemma 4.11, we prove this Lemma by induction. As seen in (42), a representation of L_n as in (53) holds for $n = 1$. We assume now that such a representation holds for some $n \in \{1, \dots, L-1\}$. Set $a_{n, m_n}(t, z) := (-1)^{n+1} \sigma(z)$ for all $(t, z) \in [0, \infty) \times \mathbf{B}$ and let $\xi = (t, x, y, z) \in [0, \infty) \times \mathbf{E}^*$. Then, according to Lemma 4.8, we obtain

$$\begin{aligned} L_{n+1}(\xi) &= \begin{pmatrix} [\hat{b}(t, \cdot), A_n(t, \cdot)](z) + \partial_t A_n(t, z) \\ 0_L \\ [\hat{b}(t, \cdot), A_n(t, \cdot)](z) + \partial_t A_n(t, z) \end{pmatrix} - \sum_{i=1}^{m_n} \partial_t a_{n,i}(t, z) \begin{pmatrix} W_{n,i}(x, y) \\ 0 \end{pmatrix} \\ &\quad - A_n(t, z) \begin{pmatrix} \partial_x F(x, y) \\ 0 \end{pmatrix} - \sum_{i=1}^{m_n} a_{n,i}(t, z) \begin{pmatrix} [F, W_{n,i}](x, y) \\ 0 \end{pmatrix} \\ &\quad - \sum_{i=1}^{m_n} \hat{b}(t, z) \left(a_{n,i}(t, z) \begin{pmatrix} \partial_x W_{n,i}(x, y) \\ 0 \end{pmatrix} + \partial_z a_{n,i}(t, z) \begin{pmatrix} W_{n,i}(x, y) \\ 0 \end{pmatrix} \right). \end{aligned}$$

The first summand is already of the desired type, and every other summand has a vanishing last component. Therefore, we are left with the task to study the term

$$(55) \quad \begin{aligned} \bar{L}_{n+1}(\xi) &:= \sum_{i=1}^{m_n} \partial_t a_{n,i}(t, z) W_{n,i}(x, y) + A_n(t, z) \partial_x F(x, y) + \sum_{i=1}^{m_n} a_{n,i}(t, z) [F, W_{n,i}](x, y) \\ &\quad + \sum_{i=1}^{m_n} \hat{b}(t, z) \left(a_{n,i}(t, z) \partial_x W_{n,i}(x, y) + \partial_z a_{n,i}(t, z) W_{n,i}(x, y) \right). \end{aligned}$$

We have to prove that we can extract one summand from this sum that is of the type $(-1)^{n+2} \sigma(z) W(x, y)$ with some

$$W(x, y) = \begin{pmatrix} \square_{n+1} \\ \neq 0 \\ 0_{1+L-(n+2)} \end{pmatrix},$$

while every other summand is of the type $\varphi(t, z) V(x, y)$ with some

$$V(x, y) = \begin{pmatrix} \square_{n+1} \\ 0_{1+L-(n+1)} \end{pmatrix}.$$

In order to do so, we will expand the expressions in (55) step for step and discuss the occurring terms one by one.

Due to our induction hypothesis, we get that

$$W_{n,i}(x, y), \partial_x W_{n,i}(x, y) \text{ are of the type } \begin{pmatrix} \square_{n+1} \\ 0_{1+L-(n+1)} \end{pmatrix} \text{ for all } i \in \{1, \dots, m_n\},$$

and due to (H1), the same is true for $\partial_x F(x, y)$. This is why it remains to look at

$$\sum_{i=1}^{m_n} a_{n,i}(t, z) [F, W_{n,i}](x, y) = \sum_{i=1}^{m_n} a_{n,i}(t, z) \sum_{j=1}^{1+L} \left(F^{(j)} \partial_{\xi_j} W_{n,i} - W_{n,i}^{(j)} \partial_{\xi_j} F \right) (x, y).$$

Thanks again to the induction hypothesis,

$$\partial_{\xi_j} W_{n,i}(x, y) = \begin{pmatrix} \square_{n+1} \\ 0_{1+L-(n+1)} \end{pmatrix} \text{ for all } i \in \{1, \dots, m_n\} \text{ and } j \in \{1, \dots, 1+L\},$$

which is of course still valid for the same terms multiplied by $F^{(j)}(x, y)$. Therefore, it remains to discuss

$$- \sum_{i=1}^{m_n} a_{n,i}(t, z) \sum_{j=1}^{1+L} W_{n,i}^{(j)}(x, y) \partial_{\xi_j} F(x, y).$$

Using the induction hypothesis once again, we know in particular that

$$W_{n,i}^{(j)}(x, y) = 0 \quad \text{for all } i \in \{1, \dots, m_n - 1\} \text{ and } j \in \{n+1, \dots, 1+L\}$$

and

$$W_{n,m_n}^{(j)}(x, y) = 0 \quad \text{for all } j \in \{n+2, \dots, 1+L\}.$$

Thus, it remains to look at

$$- \sum_{i=1}^{m_n-1} a_{n,i}(t, z) \sum_{j=1}^n W_{n,i}^{(j)}(x, y) \partial_{\xi_j} F(x, y) - (-1)^{n+1} \sigma(z) \sum_{j=1}^{n+1} W_{n,m_n}^{(j)}(x, y) \partial_{\xi_j} F(x, y).$$

Using (H1) again, we see that

$$\partial_{\xi_j} F(x, y) = \begin{pmatrix} \square_{n+1} \\ 0_{1+L-(n+1)} \end{pmatrix} \quad \text{for all } j \in \{1, \dots, n\},$$

which implies that it now remains to comment on the summand

$$(-1)^{n+2} \sigma(z) W_{n,m_n}^{(n+1)}(x, y) \partial_{\xi_{n+1}} F(x, y).$$

The two leading scalar factors do not vanish (by (51) and by the induction hypothesis), and

$$\partial_{\xi_{n+1}} F(x, y) = \begin{pmatrix} \square_{n+1} \\ \neq 0 \\ 0_{1+L-(n+2)} \end{pmatrix}$$

thanks to (H1). Thus, $\bar{L}_{n+1}(\xi)$ is in fact of the desired form with

$$W_{n+1,m_{n+1}}(x, y) := W_{n,m_n}^{(n+1)}(x, y) \partial_{\xi_{n+1}} F(x, y) = \begin{pmatrix} \square_{n+1} \\ \neq 0 \\ 0_{1+L-(n+2)} \end{pmatrix},$$

and the proof by induction is complete. \square

We can now turn to the

Proof of Theorem 4.17. Assumption (H1) and Lemma 4.19 immediately imply that for arguments from $[0, \infty) \times \mathbf{E}^*$ the vector fields $\bar{L}_1, \dots, \bar{L}_L$ are of the type

$$\bar{L}_n = \begin{pmatrix} \square_n \\ \neq 0 \\ 0_{1+L-(n+1)} \end{pmatrix} \quad \text{for all } n \in \{1, \dots, L\},$$

which in turn yields their linear independence. Lemma 4.19 also implies that L_n differs from

$$- \begin{pmatrix} \bar{L}_n(\xi) \\ 0 \end{pmatrix}$$

only by a multiple of the diffusion coefficient $V_1(z)$, which is nonzero since $z \in \mathbf{B}$. Combining these two facts lets us conclude that

$$V_1(z), L_1(\xi), \dots, L_L(\xi)$$

are linearly independent. In order to span the remaining dimension, we set

$$\begin{aligned} L_{L+1}(\xi) &:= [V_1, [V_1, V_0]](\xi) = -[V_1, L_1](\xi) \\ &= \begin{pmatrix} \zeta_{1,1}(t, z) \\ 0_L \\ \zeta_{1,1}(t, z) \end{pmatrix} + \sigma^2(z) \begin{pmatrix} \partial_x^2 f(x, y) \\ \partial_x^2 g_1(x, y) \\ 0_L \end{pmatrix} + \sigma(z) \sigma'(z) \begin{pmatrix} \partial_x f(x, y) \\ \partial_x g_1(x, y) \\ 0_L \end{pmatrix}, \end{aligned}$$

where we used (43) and the cascade condition. Using similar arguments as in the proof of Corollary 4.10, we see after subtracting a suitable linear combination of $V_0(z)$ and $L_1(\xi)$ that $L_{L+1}(\xi)$ is linearly independent from $V_1(z), L_1(\xi), \dots, L_L(\xi)$ whenever the decoupling condition (H2) is fulfilled. \square

Remark 4.20. Note that Lemma 4.11 provides a general representation of $L_{\kappa,n}$ while specific conditions on f and g come in only afterwards. In contrast to this, the calculations of L_n in Lemma 4.19 already rely explicitly on the condition (H1), without which the representation would become very bulky and rather unhelpful. This slightly different approach is also reflected in the fact that in the present section the assumptions on f and g are made not only in a single point but in the open neighborhood \mathbf{A} .

Example 4.21. For $M = N = 1$ and the choice of g that is given in (7), the system from Example 1.3 obviously fulfills (H1) and (H2) in a neighborhood \mathbf{A} of $(\pm 1, 1, \dots, 1)$, and \mathbf{B} is non-empty unless $\sigma \equiv 0$. In Example 3.15, we showed that $(\pm 1, 1, \dots, 1, z)$ is also T -attainable for any $z \in \mathbb{R}$. Since, as noted in Remark 2.4, there is also a Lyapunov function in this setting, this means that all of the assumptions of Theorem 2.2 are satisfied and the solution to the corresponding SDE (3) is positive Harris recurrent (even exponentially ergodic).

Example 4.22. Even though it might seem like the oscillator model from Example 1.2 should involve some kind of cascade structure (at least in the setting with one-sided input), calculating the derivatives of the respective function g shows that it actually does not (even after renumbering the variables). If we apply the strategy from (52) anyway and do calculations in analogy to Lemma 4.19, we can span $N + L = 1 + 5$ space directions in the one-sided case at points where $w'_1(q_2 - q_1)$ and $w'_3(q_2 - q_3)$ do not vanish. For the case of two-sided input, we have $N = 2$ and therefore we have to do the same construction twice: Once just as in (52) and once with V_1 replaced with V_2 . For the respective calculations, recall that Lemma 4.8 works for any $N \in \mathbb{N}$. Again, we can span $N + L = 2 + 4$ space directions at points where $w'_1(q_2 - q_1)$ and $w'_3(q_2 - q_3)$ do not vanish. For the time-homogeneous system from [8] without explicit external equations, one can allow those derivatives to vanish as long as some higher order of them does not, as the authors prove in [8, Lemma 5.3]. In our case however, time-inhomogeneity and the extra dependence on z in the drift make this more difficult. Decoupling of X and Z is even more problematic, since any dependence on the x -variable (i.e. p_1 or (p_1, p_3) , respectively) is linear, which rules out working with a condition like (H2). At the moment, we do not know how to provide a method that will work under assumptions that are neither too restrictive nor physically irrelevant.

Acknowledgements. The author would like to thank Reinhard Höpfner, Eva Löcherbach and Barbara Gentz for fruitful discussions and helpful remarks and suggestions.

References

- [1] T. M. APOSTOL: *Modular Functions and Dirichlet Series in Number Theory*. Springer, 1976.
- [2] J. AZÉMA, M. DUFLO, D. REVUZ: *Mesure invariante des processus de Markov récurrents*. Séminaire de Probabilités III. In: *Lecture Notes in Mathematics* Vol. 88 (1969).
- [3] R. BASS: *Diffusions and Elliptic Operators*. Springer, 1998.
- [4] K. S. COLE, H.A. ANTOSIEWICZ, P. RABINOWITZ: *Automatic Computation of Nerve Excitation*. In: *Journal of the Society for Industrial and Applied Mathematics* Vol. 3 (1955), pp. 153–172.
- [5] A. CRUDU, A. DEBUSSCHE, A. MULLER, O. RADULESCU: *Convergence of stochastic gene networks to hybrid piecewise deterministic processes*. In: *The Annals of Applied Probability* Vol. 22(5) (2012), pp. 1822–1859.
- [6] N. CUNEO, J. P. ECKMANN: *Non-Equilibrium Steady States for Chains of Four Rotors*. In: *Communications in Mathematical Physics* Issue 1 (2016), pp. 185–221.
- [7] N. CUNEO, J. P. ECKMANN, M. HAIRER, L. REY-BELLET: *Non-equilibrium steady states for networks of oscillators*. In: *Electronic Journal of Probability* Vol. 23 (2018), no. 55, pp. 1–28.
- [8] N. CUNEO, J. P. ECKMANN, C. POQUET: *Non-equilibrium steady state and subgeometric ergodicity for a chain of three coupled rotors*. In: *Nonlinearity* Vol. 28 (2015), pp. 2397–2421.
- [9] M. DESROCHES, J. GUCKENHEIMER, B. KRAUSKOPF, C. KUEHN, H. M. OSINGA, M. WECHSELBERGER: *Mixed-mode oscillations with multiple time scales*. In: *SIAM Review* Vol. 54(2) (2012), pp. 211–288.
- [10] A. DESTEXHE: *Conductance-based integrate and fire models*. In: *Neural Computation* Vol. 9 (1997), pp. 503–514.
- [11] S. DITLEVSEN, P. LÁNSKÝ: *Estimation of the input parameters in the Feller neuronal model*. In: *Physical Review E* Vol. 73 (2006), 061910.

- [12] S. DITLEVSEN, E. LÖCHERBACH: On oscillating systems of interacting neurons. In: *Stochastic Processes and their Applications* Vol. 127 (2017), pp. 1840–1869.
- [13] D. DOWN, S. P. MEYN, R. L. TWEEDIE: Exponential and Uniform Ergodicity of Markov Processes. In: *Annals of Probability* Vol. 23(4) (1995), pp. 1671–1691.
- [14] K. ENDLER: Periodicities in the Hodgkin–Huxley model and versions of this model with stochastic input. Master Thesis, Institute of Mathematics, University of Mainz (see under <http://ubm.opus.hbz-nrw.de/volltexte/2012/3083/>), 2012.
- [15] A. FAGGIONATO, D. GABRIELLI, M. RIBEZZI CRIVELLARI: Non-equilibrium Thermodynamics of Piecewise Deterministic Markov Processes. In: *Journal of Statistical Physics* Vol. 137 (2009), pp. 259–304.
- [16] J. GUCKENHEIMER, R. A. OLIVA: Chaos in the Hodgkin–Huxley model. In: *SIAM Journal on Applied Dynamical Systems* Vol. 1(1) (2002), pp. 105–114.
- [17] M. HAIRER: On Malliavin’s proof of Hörmander’s Theorem. In: *Bulletin des Sciences Mathématiques* Vol. 135 (2011), pp. 650–666.
- [18] M. HAIRER, J. C. MATTINGLY: Slow energy dissipation in anharmonic oscillator chains. In: *Communications on Pure and Applied Mathematics* Vol. 62 (2009), pp. 999–1032.
- [19] T. HARRIS: The existence of stationary measures for certain Markov processes. In: *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability* Vol. II (1956), pp. 113–124.
- [20] B. HASSARD: Bifurcation of periodic solutions of the Hodgkin–Huxley model for the squid giant axon. In: *Journal of Theoretical Biology* Vol. 71 (1978), pp. 401–420.
- [21] A. L. HODGKIN, A. F. HUXLEY: A Quantitative Description of Membrane Current And Its Application to Conduction And Excitation in Nerve. In: *Journal of Physiology* Volume 117, Vol. 4 (1952), pp. 500–544.
- [22] S. HOLBACH: Local asymptotic normality for shape and periodicity in the drift of a time inhomogeneous diffusion. In: *Statistical Inference for Stochastic Processes* Vol. 21 (2018), pp. 527–538.
- [23] S. HOLBACH: Local asymptotic normality for shape and periodicity of a signal in the drift of a degenerate diffusion with internal variables. arXiv:1903.03440 [math.PR].
- [24] S. HOLBACH: Recurrence and parameter estimation for degenerate diffusions with internal variables and randomly perturbed time-inhomogeneous deterministic input. Dissertation, Johannes Gutenberg-Universität Mainz, 2018, available online at <http://publications.ub.uni-mainz.de/theses/volltexte/2018/100002300/pdf/100002300.pdf>
- [25] R. HÖPFNER: On a set of data for the membrane potential in a neuron. In: *Mathematical Biosciences* Vol. 207 (2007), pp. 275–301.
- [26] R. HÖPFNER, E. LÖCHERBACH, M. THIEULLEN: Ergodicity for a Stochastic Hodgkin–Huxley Model Driven by Ornstein–Uhlenbeck Type Input. In: *Annales de l’Institut Henri Poincaré* Vol. 1 (2016), pp. 483–501.
- [27] R. HÖPFNER, E. LÖCHERBACH, M. THIEULLEN: Strongly degenerate time inhomogeneous SDEs: densities and support properties. Application to a Hodgkin–Huxley system with periodic input. In: *Bernoulli* Vol. 23(4A) (2017), 2587–2616.
- [28] R. HÖPFNER, E. LÖCHERBACH, M. THIEULLEN: Ergodicity and Limit Theorems for Degenerate Diffusions with Time Periodic Drift. Application to a Stochastic Hodgkin–Huxley Model. In: *ESAIM P&S* Vol. 20 (2016), pp. 527–554.
- [29] N. IKEDA, S. WATANABE: *Stochastic Differential Equations and Diffusion Processes*. North-Holland Library, 2nd edition, 1989.
- [30] E. M. IZHIKEVICH: *Dynamical Systems in Neuroscience*. The MIT Press, 2007.
- [31] R. Z. KHASMINSKII: *Stochastic Stability of Differential Equations*. Second English Edition, Springer, 2012.
- [32] P. LÁNSKÝ, L. SACERDOTE, F. TOMASSETTI: On the Comparison of Feller and Ornstein–Uhlenbeck Models for Neural Activity. In: *Biological Cybernetics* Vol. 73 (1995), pp. 457–465.

- [33] J. MATTINGLY, A. STUART, D. HIGHAM: Ergodicity of SDEs and approximations: locally Lipschitz vector fields and degenerate noise. In: *Stoch. Proc. Appl.* Vol. 101 (2002), pp. 185–232.
- [34] K. R. MEYER, G. R. HALL, D. OFFIN: Introduction to Hamiltonian Dynamical Systems and the N -Body Problem. Second Edition, Springer, 2009.
- [35] S. MEYN, R. TWEEDIE: Stability of Markovian processes I: criteria for discrete-time chains. In: *Advances in Applied Probability* Vol. 24 (1992), pp. 542–574.
- [36] A. MILLET, M. SANZ-SOLÉ: A simple proof of the support theorem for diffusion processes. In: *Séminaire de probabilités de Strasbourg* Vol. 28 (1994), pp. 36–48.
- [37] C. MORRIS, H. LECAR: Voltage oscillations in the barnacle giant muscle fiber. In: *Biophysical Journal* Vol. 35 (1981), pp. 193–213.
- [38] E. NUMMELIN: A splitting technique for Harris recurrent Markov chains. In: *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* Vol. 43 (1978), pp. 309–318.
- [39] E. NUMMELIN: General irreducible Markov chains and non-negative operators. Cambridge University Press, 1985.
- [40] K. PAKDAMAN, M. THIEULLEN, G. WAINRIB: Fluid limit theorems for stochastic hybrid systems with application to neuron models. In: *Advances in Applied Probability* Vol. 42 (2010), pp. 761–794.
- [41] J. PRÜSS, M. WILKE: Gewöhnliche Differentialgleichungen und dynamische Systeme. Birkhäuser, 2010.
- [42] J. REALPE-GOMEZ, T. GALLA, A. J. MCKANE: Demographic noise and piecewise deterministic Markov processes. In: *Physical Review E* Vol. 86 (2012), 011137.
- [43] D. REVUZ: Markov Chains. North-Holland, 1975.
- [44] L. REY-BELLET, L. E. THOMAS: Exponential convergence to non-equilibrium stationary states in classical statistical mechanics. In: *Communications in Mathematical Physics* Vol. 225 (2002), pp. 309–329.
- [45] J. RINZEL, B. ERMENTROUT: Analysis of neural excitability and oscillations. In: *Methods in neuronal modeling: from ions to networks, 2nd edition* (1998), pp. 251–291.
- [46] J. RUBIN, M. WECHSELBERGER: Giant squid-hidden canard: the 3D geometry of the Hodgkin–Huxley model. In: *Biological Cybernetics* Vol. 97(1) (2007), pp. 5–32.