STRONGLY DEGENERATE TIME INHOMOGENEOUS SDES: DENSITIES AND SUPPORT PROPERTIES. APPLICATION TO HODGKIN-HUXLEY TYPE SYSTEMS.

R. HÖPFNER, E. LÖCHERBACH, AND M. THIEULLEN

Abstract. In this paper we study the existence of densities for strongly degenerate stochastic differential equations (SDEs) whose coefficients depend on time and are not globally Lipschitz. In these models neither local ellipticity nor the strong Hörmander condition is satisfied. In this general setting we show that continuous transition densities indeed exist in all neighborhoods of points where the weak Hörmander condition is satisfied. We also exhibit regions where these densities remain positive. We then apply these results to stochastic Hodgkin-Huxley models with periodic input as a first step towards the study of ergodicity properties of such systems in the sense of [31]-[32].

1. Introduction

This paper belongs to a series of three articles (see also [23], [24]) in which we carry a probabilistic study of multidimensional strongly degenerate and time inhomogeneous random systems and their ergodic properties, with a view towards statistical inference in neuroscience. An important step on the road to ergodicity is to show that such systems possess Lebesgue densities and to address properties of the support of their law. This is the topic of the present paper. We establish that densities exist, are smooth and strictly positive, at least on suitable parts of the state space. The coefficients of our stochastic differential equations (SDEs) depend on time and are not globally Lipschitz. The noise is degenerate. In the main application we have in mind the noise is actually one dimensional and present in only few components of the system.

In order to prove existence of densities, is has become classical to use Malliavin Calculus and the Hörmander condition (cf. [36]). Hörmander sufficient condition ensures that the diffusion in the random system is actually strong enough even if the noise is visible only on a restricted number of components. It is satisfied when the Lie algebra generated by the coefficients of the SDE has full dimension and can be found under two forms: the strong form involving the diffusion coefficients only, and the weak form possibly including the drift coefficient. In our case, we can only hope for the weak Hörmander condition to be satisfied. Moreover in general this condition will hold only locally. SDEs satisfying local Hörmander condition with locally smooth coefficients have been considered recently in a time homogeneous setting (cf. [2], [3], [14], [18]). In these works the local Hörmander condition is ensured by a local ellipticity assumption (hence these papers deal with the strong form of this condition). However in our framework time homogeneity fails and we must work with the weak form of Hörmander condition, which holds only locally. In this general setting we show that smooth transition densities indeed exist in all neighborhoods of points where the weak Hörmander condition

Date: January 25, 2016.

2010 *Mathematics Subject Classification.* 60 J 60, 60 J 25, 60 H 07.

Key words and phrases. Degenerate diffusion processes, time inhomogeneous diffusion processes, Malliavin calculus, local Hörmander condition, support theorem, Hodgkin-Huxley system .

is satisfied. We also prove that these densities are lower semi continuous (lsc) w.r.t. the starting point even if our system does not enjoy the Feller property. In order to do so we extend a localization argument and estimates of the Fourier transform introduced in [2], [3], [14] and [18].

Our motivation is to describe with probabilistic tools the long time behavior of a neuron embedded in a network in order to be able to estimate either the parameters of the model or the underlying network activity or the characteristics of the spike trains generated by the neuron. The network activity is present via the synaptic stimulation the neuron receives through its dendritic tree. We describe the neuron by the Hodgkin-Huxley model which is very well known in physiology. This system is notoriously mathematically difficult and may exhibit a collection of different behaviors when submitted to a deterministic periodic input. The synaptic stimulation we consider is a random input carrying a deterministic and periodic signal. We are interested in ergodicity for the process composed of the neuron on the one hand and the input it receives on the other hand. This results in a five dimensional (5D) time inhomogeneous random system driven by a one dimensional Brownian motion present in the first and last component only. This system belongs to the general class of SDEs mentioned above.

Because of this original motivation we find it natural to introduce an intermediate family of models that we call *SDEs with internal variables and random input* which lies in between the general class of SDEs and our specific 5D-stochastic Hodgkin-Huxley. This family includes all conductance based models with synaptic input relevant for modeling in neuroscience (the Hodgkin-Huxley system is a conductance-based model). It is also relevant in biology and physics since it describes on a macroscopic scale the limit of a population of individuals of different types represented by the internal variables, coupled by a global variable. In the microscopic model, each individual can occupy two states, active or inactive. In neuro-physiology, individuals can be ion channels of different types, see e.g. [25]: here the active state corresponds to open channels, and the inactive state to closed channels. The transition rates between these two states depend on the global variable only. The deterministic system on which the SDEs are built is obtained as the limit of a sequence of Piecewise Deterministic Markov Processes when the number of individuals goes to infinity (cf. [11], [40], [17], [38]) in the sense of the Law of Large Numbers. When we consider SDEs with internal variables and random input we consider that the population is infinite, we neglect the intrinsic noise related to finite size effects and we focus on the external noise received from the environment. For instance, when we model a neuron, the individuals in the population are the ion channels and the global variable is the membrane potential (see [38]).

For systems in this family we provide an explicit discussion of the Hörmander condition that we later illustrate numerically in the last section devoted to the 5D-stochastic Hodgkin-Huxley model. Then we exhibit regions where densities, if they exist, remain positive. These regions are related to neighborhoods of equilibrium points of the underlying deterministic system. The particular structure of SDEs with internal variables and random input, namely some linearity in the internal equations, plays a key role for the control argument that we use. We also show that with positive probability, the solution of these systems can imitate any deterministic evolution resulting from an arbitrary input, on an arbitrary interval of time. This 'chamaeleon property', stated in Theorem 5, is one of the main results of our paper. It is used in [24] to prove that the 5D-stochastic Hodgkin-Huxley system emits an infinite number of spikes in the long run almost surely.

We finally apply the general results obtained for SDEs with internal variables and random input to the 5D-stochastic Hodgkin-Huxley model. In particular, the present paper shows that the 5D-stochastic Hodgkin-Huxley model possesses smooth Lebesgue densities. The Hörmander condition is satisfied at certain equilibrium points. Therefore, depending on the starting point, the 5D-stochastic Hodgkin-Huxley model possesses strictly positive densities in small neighborhoods of such equilibrium points. We use this result in two companion papers [23] and [24], where we address periodic ergodicity and prove limit theorems. To the best of our knowledge, no other probabilistic study has been presented in the literature before. There are some simulation studies (see e.g. [39] and [45]), but not much seems to be known mathematically.

The present paper paper is organized as follows. In Section 2 we first present our general SDEs and assumptions. Section 3 is devoted to proving our first main result, stated in Theorem 1, which shows the local existence of smooth densities for time inhomogeneous systems with locally Lipschitz coefficients. In Section 4 we introduce SDEs with internal variables and random input. We explicit the weak Hörmander condition and address the positivity of densities for such systems. Theorem 5 in Section 4 states the 'chamaeleon property', our second main result. Section 5 of the paper is devoted to the 5D-stochastic Hodgkin-Huxley model (with some reminders on the deterministic system). In the appendix (section 6) we provide complementary proofs.

2. The setting

In this section we describe the systems of SDEs to be considered in our paper. Given integers $m > 1$ and $l < m$, we write $x = (x_1, \ldots, x_m)$ for generic elements of $I\!R^m$. Let

$$\sigma(x) = \begin{pmatrix} \sigma_{1,1}(x) & \ldots & \sigma_{1,l}(x) \\ \vdots & & \vdots \\ \sigma_{m,1}(x) & \ldots & \sigma_{m,l}(x) \end{pmatrix} \quad \text{and} \quad b(t,x) = \begin{pmatrix} b_1(t,x) \\ \vdots \\ b_m(t,x) \end{pmatrix}.$$

We suppose that σ is measurable from $I\!R^m$ to $I\!R^{m\otimes l}$ and that $b(t,x)$ is a smooth function from $[0,\infty[\times I\!R^m$ to $I\!R^m$. For all $x \in I\!R^m$, we consider the SDE

$$X_{i,t} = x_i + \int_0^t b_i(s, X_s)ds + \sum_{k=1}^{l} \int_0^t \sigma_{i,k}(X_s)dW_s^k, \; t \geq 0, \; i = 1, \ldots, m, \tag{1}$$

and assume throughout this paper that a unique strong solution exists (at least up to some lifetime). Here, $W^1, \ldots, W^l$ are independent one-dimensional Brownian motions. Thus the system (1), an $m-$dimensional SDE driven by $l-$dimensional Brownian motion for $l < m$, is strongly degenerate. We write P_x for the probability measure under which the solution $X = (X_t)_{t\geq 0}$ of (1) starts at x. Note that the time dependence is in the drift only. We assume that (1) satisfies the following assumptions.

(H1) There exists an increasing sequence of compacts $K_n \subset K_{n+1}$ of the form $K_n = [a_n, b_n] = \prod_{i=1}^m [a_{n,i}, b_{n,i}]$ where $a_n = (a_{n,1}, \ldots, a_{n,m})$, such that for any $x \in \bigcup_n K_n$, the unique strong solution to (1) starting from x at time 0 satisfies that $T_n := \inf\{t : X_t \notin K_n\} \to \infty$ almost surely as $n \to \infty$.

(H2) The coefficients of (1) are locally smooth. Namely we assume that for all n, $\sigma_{i,k} \in C_b^\infty(K_n, I\!R)$ for all $1 \leq i \leq m, 1 \leq k \leq l$. Moreover, we suppose that for every multi-index $\beta \in \{0, \ldots, m\}^l, l \geq 1$, $\quad b(t,x) + \partial_\beta b(t,x)$ is bounded on $[0,T] \times K_n$ for all $T > 0$. Here $\partial_\beta = \frac{\partial^l}{\partial x_{\beta_1} \ldots \partial x_{\beta_l}}$ and we identify x_0 with t.

Notice that as a consequence of assumption **(H1)** we could choose as state space of the process $(X_t, t \geq 0)$ the set $E := \bigcup_n K_n$. We will do this in some parts of the paper, e.g. in Sections 4 and 5.

Example 1. *Consider the two dimensional damping Hamiltonian system*

$$X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix} \quad , \quad \left\{ \begin{array}{lcl} dX_{1,t} & = & g(X_t)dt \\ dX_{2,t} & = & \tilde{\sigma}(X_t)dW_t + V'(t, X_{1,t})dt \end{array} \right\}$$

evolving in a time dependent potential $V(t,\cdot)$. Here, W is a one dimensional Brownian motion. We suppose that $g, \tilde\sigma$ and V' are such that our conditions **(H1)** *and* **(H2)** *are satisfied. Systems of the above form are widely studied in the literature. They serve as models of physical systems subjected to random perturbations, see e.g.* [12] *where a chain of rotators is considered whose ends are coupled to stochastic heat baths. In the last years, the necessity of performing inference about unknown parameters within such models gave rise to several papers within the statistical literature. We refer the reader to* [42] *and to* [9] *and the references therein. A first step towards statistical inference is the existence, at least locally, of a transition density. This is the topic of the next section.*

3. Existence and smoothness of densities for (1)

Classically, one proves that the solution of an SDE admits a smooth density via Malliavin Calculus, imposing the Hörmander condition. In most of the cases it is assumed that the coefficients of the SDE are C^∞, bounded, with bounded derivatives of any order and that the Hörmander condition is satisfied all over the state space. However, in our case, the coefficients of (1) are not globally Lipschitz. Regarding the Hörmander condition there are actually two possibilities: either to work under the strong Hörmander condition or under the weak one which is a less stringent assumption. The strong degeneracy of (1) imposes to work under the weak form of Hörmander condition, which moreover may hold only locally. In addition the drift coefficient depends on time. Hence we have to apply local arguments in a time inhomogeneous setting.

3.1. **Local Hörmander condition in a time dependent setting.** In this section we state our local weak Hörmander condition. The first step is to rewrite (1) in Stratonovich form. This amounts to replace the drift $b(t,x)$ by

$$\tilde b_i(t,x) := b_i(t,x) - \frac{1}{2}\sum_{k=1}^{l}\sum_{j=1}^{m}\sigma_{j,k}(x)\frac{\partial \sigma_{i,k}}{\partial x_j}(x), \quad 1\le i\le m\ ,\ \ x\in \mathbb{R}^m\ , \tag{2}$$

which is still time inhomogeneous. Now we have to take care of time dependence in the drift of (1). Let us consider the vector fields (or linear differential operators of order one) A_0 and $A_1,\ldots,A_l$ on $[0,+\infty[\times \mathbb{R}^m$ whose coefficients are given by $\tilde b$ and $\sigma_1,\ldots,\sigma_l$, where σ_k denotes the $k-$th column of the matrix σ, for any $1\le k\le l$:

$$A_0 = \frac{\partial}{\partial t} + \sum_{i=1}^{m}\tilde b_i(t,x)\frac{\partial}{\partial x_i} = \frac{\partial}{\partial t} + \tilde b, \quad A_k = \sum_{i=1}^{m}\sigma_{i,k}(x)\frac{\partial}{\partial x_i}, 1\le k\le l.$$

A_0 and A_k can be identified respectively with the $(m{+}1)-$dimensional function $A_0(t,x) = (\tilde b_0,\ldots,\tilde b_m)$ where $\tilde b_0 = 1$ and $A_k(t,x) = (\sigma_{0,k},\sigma_{1,k},\ldots,\sigma_{m,k})$ where $\sigma_{0,k} = 0$. Actually there is a one-to-one correspondance between vector fields $\mathcal{T}(t,x) = \mathcal{T}_0(t,x)\frac{\partial}{\partial t} + \sum_{i=1}^m \mathcal{T}_i(t,x)\frac{\partial}{\partial x_i}$ and $(m{+}1)-$dimensional functions $(\mathcal{T}_0,\mathcal{T}_1,\ldots,\mathcal{T}_m)$.

The Lie bracket of two vector fields $\mathcal{T}(t,x) = \mathcal{T}_0(t,x)\frac{\partial}{\partial t} + \sum_{i=1}^m \mathcal{T}_i(t,x)\frac{\partial}{\partial x_i}$ and $\mathcal{V}(t,x) = \mathcal{V}_0(t,x)\frac{\partial}{\partial t} + \sum_{i=1}^m \mathcal{V}_i(t,x)\frac{\partial}{\partial x_i}$ is defined as

$$[\mathcal{T},\mathcal{V}]_i := \sum_{j=0}^{m}(\mathcal{T}_j\frac{\partial \mathcal{V}_i}{\partial x_j} - \mathcal{V}_j\frac{\partial \mathcal{T}_i}{\partial x_j}) = (\mathcal{T}_0\frac{\partial \mathcal{V}_i}{\partial t} - \mathcal{V}_0\frac{\partial \mathcal{T}_i}{\partial t}) + \sum_{j=1}^{m}(\mathcal{T}_j\frac{\partial \mathcal{V}_i}{\partial x_j} - \mathcal{V}_j\frac{\partial \mathcal{T}_i}{\partial x_j}).$$

In particular since σ in (1) does not depend on time, for all $1 \le k \le l$,

$$[A_k, \mathcal{V}]_i = \sum_{j=1}^{m} (\sigma_{j,k} \frac{\partial \mathcal{V}_i}{\partial x_j} - \mathcal{V}_j \frac{\partial \sigma_{i,k}}{\partial x_j}),$$

where no time derivative appears. On the contrary a time derivative is present in $[A_0, \mathcal{V}]$ since

$$[A_0, \mathcal{V}]_i = (\frac{\partial \mathcal{V}_i}{\partial t} - \mathcal{V}_0 \frac{\partial \tilde{b}_i}{\partial t}) + \sum_{j=1}^{m} (\tilde{b}_j \frac{\partial \mathcal{V}_i}{\partial x_j} - \mathcal{V}_j \frac{\partial \tilde{b}_i}{\partial x_j}).$$

Notice that whenever $\mathcal{V}_0$ vanishes identically, $[A_k, \mathcal{V}]_0 \equiv 0$ since $\sigma_{0,k}$ is zero and $[A_0, \mathcal{V}]_0 \equiv 0$ as well since $\tilde{b}_0$ is constant equal to 1. In this case the vector fields $[A_k, \mathcal{V}]$ and $[A_0, \mathcal{V}]$ belong to the m-dimensional space generated by the $\frac{\partial}{\partial x_i}$, $1 \le i \le m$. In particular $[A_k, A_0]$ belong to this latter space as well as all $A_k, 1 \le k \le l$, by definition.

Given $A_1, \dots, A_l$ and A_0 we can build two Lie algebras. On one hand the Lie algebra Λ generated by the set $\{A_0, A_1, \dots, A_l\}$ including the drift vector. On the other hand, we may define a set $\mathcal{L}_N$ of vector fields by 'initial condition' $A_1, \dots, A_l \in \mathcal{L}_N$ and at most N iteration steps

$$L \in \mathcal{L}_N \implies [L, A_0], [L, A_1], \dots, [L, A_l] \in \mathcal{L}_N \ , \tag{3}$$

for any fixed $N \in I\!N$. Notice that $\mathcal{L}_N$ does not contain the drift vector A_0, but the construction allows to take brackets with A_0 in further steps. Write $\mathcal{L}_N^*$ for the closure of $\mathcal{L}_N$ under Lie brackets and $\mathrm{LA}(\mathcal{L}_N)$ for the linear hull of $\mathcal{L}_N^*$, i.e. the Lie algebra spanned by $\mathcal{L}_N$. Finally, we write $\mathcal{L} = \mathrm{LA}(\bigcup_N \mathcal{L}_N)$. As just noticed the dimension of $\mathcal{L}(t,x)$ cannot exceed m whatever $(t,x) \in [0, +\infty[\times I\!R^m$. However the dimension of $\Lambda(t,x)$ can be equal to $m+1$. Actually the following result holds.

Proposition 1. *For all* $(t,x) \in [0, +\infty[\times I\!R^m$, $\dim \Lambda(t,x) = \dim \mathcal{L}(t,x) + 1$.

Before giving the proof of this proposition we state the local weak Hörmander condition we are going to work with. Recall $E = \bigcup_n K_n \subset I\!R^m$ from condition **(H1)**.

(LWH) We say that the Hörmander condition is satisfied at (t, y_0) if there exist $r \in]0, t[$ and $R > 0$ such that $B_{5R}(y_0) \subset E$ and $\dim \Lambda(s,y) = m+1$, $\forall (s,y) \in [t-r, t] \times B_{3R}(y_0)$ (local weak Hörmander condition).

Proof of Proposition 1. For a fixed integer N, consider the Lie algebra $\mathrm{LA}(\mathcal{L}_N)$ spanned by $\mathcal{L}_N$. Construct also iteratively the set Λ_N such that it contains $A_1, \dots, A_l$ and A_0 (initialization) and is stable by Lie brackets with A_0 and $A_k, 1 \le k \le l$ (iteration) of order up to N. Then define $\mathrm{LA}(\Lambda_N)$ as the Lie algebra spanned by Λ_N. The difference in the initialization between $\mathcal{L}_N$ and Λ_N plays a key role. $\Lambda_N \setminus \mathcal{L}_N$ consists of A_0 and the descendance of A_0 in the sense of iterated Lie brackets (3). This implies that $\Lambda_N \subset \{A_0\} \cup \mathcal{L}_N \cup -\mathcal{L}_N$ where we denote by $-\mathcal{L}_N$ the set $\{-L; L \in \mathcal{L}_N\}$. Notice that $\mathcal{L}_N \cup -\mathcal{L}_N$ belongs to the m-dimensional space generated by the $\frac{\partial}{\partial x_i}$, $1 \le i \le m$, and A_0 is the only vector field with non trivial coordinate in the direction of $\frac{\partial}{\partial t}$. As a consequence, for all N,

$$\dim Vect(\Lambda_N) = 1 + \dim Vect(\mathcal{L}_N), \qquad \text{whence} \qquad \dim Vect(\bigcup_N \Lambda_N) = 1 + \dim Vect(\bigcup_N \mathcal{L}_N).$$

Since $Vect(\bigcup_N \Lambda_N) = \Lambda$ is the Lie algebra generated by $\{A_0, A_1, \dots, A_l\}$ and $Vect(\bigcup_N \mathcal{L}_N) = \mathcal{L}$, this implies the result. □

In the sequel we will check **(LWH)** at (t, y_0) by successive computations of Lie brackets looking for $r \in]0, t[$ and $N \in I\!N$ such that $\dim \mathrm{LA}(\mathcal{L}_N)(s, y_0) = m$, $\forall s \in [t-r, t]$.

Remark 1. *We had to state our local weak Hörmander condition in a time inhomogeneous frame. Such a time inhomogeneous situation has already been considered in* [8] *- however supposing that the weak Hörmander condition holds globally.* [8] *recalls also an example from* [44] *which points out the necessity to incorporate the operator* $\frac{\partial}{\partial t}$ *to the original framework via extension of the coefficient* $\tilde{b}$ *into* A_0 *described above.*

3.2. **Local densities for (1).** Let us recall that an $I\!R^m$- valued random vector Z admits a density with respect to Lebesgue measure or is absolutely continuous *on an open set* $O \subset I\!R^m$, if for some function $p \in L^1(O)$,

$$E(f(X)) = \int f(y)p(y)dy,$$

for any continuous and bounded function $f \in C_b(I\!R^m)$ satisfying $supp(f) \subset O$.

Theorem 1. *Assume that (1) satisfies* **(H1)** *and* **(H2)**. *Assume moreover that* **(LWH)** *is satisfied at* (t, y_0). *Fix* $x \in I\!R^m$ *and denote by* $(X_t, t \geq 0)$ *the strong solution of (1) starting from* x. *Then the random variable* X_t *admits an infinitely differentiable density on* $B_R(y_0)$, *where* R *is given in* **(LWH)**.

Note that this density might be $\equiv 0$ near y_0; so far it is not granted that the process at time t visits such neighborhoods for $y_0 \in \mathrm{int}(E)$ for arbitrary choice of a starting point $x \in I\!R^m$ with positive probability.

Theorem 2. *Let us keep the assumptions and notations of Theorem 1 and for* x *in* $I\!R^m$ *denote by* $p_{0,t}(x, \cdot)$ *the density of* X_t *on* $B_R(y_0)$. *For any fixed* $y \in B_R(y_0)$, *the map* $I\!R^m \ni x \mapsto p_{0,t}(x, y)$ *is lower semi-continuous.*

Given the assumptions **(H1)-(H2)** on (1), we have to use localization arguments in order to prove these theorems. Localization arguments have been used in [30] and [14], however in a time homogeneous framework. Moreover [14] works under the condition of local ellipticity which fails to hold for (1). We prove below that [30] and [14] can be extended to a time inhomogeneous SDE satisfying only the local weak Hörmander condition **(LWH)**.

Proof of Theorem 1. In this proof we rely on the following criterion based on Fourier transform which ensures existence and regularity of Lebesgue densities. Let μ be a probability measure on $I\!R^m$ and $\hat{\mu}$ its Fourier transform defined by $\hat{\mu}(\xi) := \frac{1}{(2\pi)^m} \int_{I\!R^m} e^{i<y,\xi>} \mu(y)dy$. If $\hat{\mu}$ is integrable, then μ is absolutely continuous and a continuous version of its density is given by

$$p(y) = \frac{1}{(2\pi)^m} \int_{I\!R^m} e^{-i<\xi,y>} \hat{\mu}(\xi) d\xi. \tag{4}$$

If moreover

$$\int_{I\!R^m} |\xi|^k |\hat{\mu}(\xi)| d\xi < \infty \tag{5}$$

holds for all $k \in I\!N$, then p is of class C^∞.

We use this criterion in our situation in the following way. First, consider $R > 0$ provided by **(LWH)** and let N be such that $\dim \mathrm{LA}(\mathcal{L}_N)(s, y) = m, \ \forall s \in [t - r, t]$ and $y \in B_{3R}(y_0)$. Denote by Φ a localizing function in $C_b^\infty(I\!R^m)$ satisfying $1_{B_R(0)} \leq \Phi \leq 1_{B_{2R}(0)}$. Fix x, t and T with $t \leq T$. We put $m_0 := E_x(\Phi(X_t - y_0))$. If $m_0 = 0$, then it is trivially true that X_t has a density on $B_R(y_0)$. Indeed, in this case the density is simply $\equiv 0$ on $B_R(y_0)$. If $m_0 > 0$, then we prove below that the probability measure ν defined by

$$\int f(y)\nu(dy) := \frac{1}{m_0} E_x \left(f(X_t)\Phi(X_t - y_0) \right), \tag{6}$$

for all $f \in C_b(\mathbb{R}^m)$, is such that for $\hat{\nu}(\xi) = \frac{1}{m_0} E_x\left(e^{i<\xi, X_t>}\Phi(X_t - y_0)\right)$, $|\xi|^k|\hat{\nu}(\xi)|$ is integrable for any fixed $k \in \mathbb{N}$.

In the following, we will work for fixed $k \in \mathbb{N}$. The main step is to prove (12)-(13) below. Although the form of (12)-(13) is classical, we have to make sure that they hold in our time inhomogeneous framework. Let $\psi \in C_b^\infty(\mathbb{R}^m)$ such that

$$\psi(y) = \begin{cases} y & \text{if } |y| \leq 4R \\ 5R\frac{y}{|y|} & \text{if } |y| \geq 5R \end{cases}$$

and $|\psi(y)| \leq 5R$ for all y. Let $\bar{b}(t,y) = b(t, y_0 + \psi(y - y_0))$ and $\bar{\sigma}(y) = \sigma(y_0 + \psi(y - y_0))$ be the localized coefficients of (1). Assumption **(H2)** ensures that $\bar{b}$ and $\bar{\sigma}$ are $C_b^\infty-$extensions (w.r.t. x) of $b_{|B_{4R}(y_0)}$ and $\sigma_{|B_{4R}(y_0)}$ with $\bar{b}$ and its derivatives bounded on $[0,T]$. Let $\bar{X}$ satisfy the SDE

$$d\bar{X}_{i,s} = \bar{b}_i(s, \bar{X}_s)ds + \bar{\sigma}_i(\bar{X}_s)dW_s, \quad s \leq T, \quad 1 \leq i \leq m, \tag{7}$$

and $\bar{X}_{i,0} = X_0 = x$. If $x \in B_{4R}(y_0)$, the processes $\bar{X}$ and X coincide up to the first exit time of $B_{4R}(y_0)$. For a fixed δ in $]0, t/2 \wedge r \wedge 1[$, where r is provided by **(LWH)**, define $\tau_1 := \inf\{s \geq t-\delta; X_s \in B_{3R}(y_0)\}$ and $\tau_2 := \inf\{s \geq \tau_1; X_s \notin B_{4R}(y_0)\}$. The set $\{\Phi(X_t - y_0) > 0\}$ is equal to the union

$$\{\Phi(X_t - y_0) > 0; t - \delta = \tau_1 < t < \tau_2\} \cup \left\{\Phi(X_t - y_0) > 0; \sup_{0 \leq s \leq \delta} |\bar{X}_{\tau_1, \tau_1 + s}(X_{\tau_1}) - X_{\tau_1}| \geq R\right\},$$

where $\bar{X}_{u,v}(z)$ denotes the value at time v of the solution of (7) satisfying $\bar{X}_u = z$ at time u when $u \leq v$ (classical notation for flows). Note that $\Phi(X_t - y_0) > 0$ implies $X_t \in B_{2R}(y_0)$. Using the Markov property in τ_1, we obtain the following expression of $\hat{\nu}$,

$$\begin{aligned} m_0\hat{\nu}(\xi) &= E_x\left(e^{i<\xi,X_t>}\Phi(X_t - y_0)1_{\{\Phi(X_t - y_0)>0\}}\, 1_{\{\sup_{0\leq s\leq\delta} |\bar{X}_{\tau_1,\tau_1+s}(X_{\tau_1}) - X_{\tau_1}| \geq R\}}\right) \\ &+ E_x\left(e^{i<\xi,X_t>}\Phi(X_t - y_0)1_{\{\Phi(X_t - y_0)>0\}}\, 1_{\{t-\delta=\tau_1<t<\tau_2\}}\right). \end{aligned}$$

We are looking for upper bounds of $|\hat{\nu}(\xi)|$ to check whether $|\xi|^k|\hat{\nu}(\xi)|$ is integrable. The latter identity reads $m_0\hat{\nu}(\xi) = A + B$. We will see shortly that the important contribution comes from $|B|$. To control $|A|$ we use the classical estimate

$$P_x\left(\Phi(X_t - y_0) > 0; \sup_{0\leq s\leq\delta} |\bar{X}_{\tau_1,\tau_1+s}(X_{\tau_1}) - X_{\tau_1}| \geq R\right) \leq C(T,q,m,b,\sigma)R^{-q}\delta^{q/2}. \tag{8}$$

It is valid for all $q > 0$ and holds uniformly in x. The constant $C(T,q,m,b,\sigma)$ depends on the supremum norms of $\bar{b}$ and $\bar{\sigma}$, hence by construction, on the supremum norms of σ (resp. b) on $B_{5R}(y_0)$ (resp. $B_{5R}(y_0) \times [0,T]$). Notice that the right hand side of (8) follows from

$$E_x\left(\sup_{u: s\leq u\leq t} |\bar{X}_{i,u} - \bar{X}_{i,s}|^q\right) \leq C(T,q,m,b,\sigma)(t-s)^{q/2}, \quad \text{for all } 0 \leq s \leq t \leq T. \tag{9}$$

Let us now estimate $|B|$. Thanks to the Markov property at time $t - \delta$,

$$\Big|B\Big| \leq \sup_{y\in B_{3R}(y_0)} |E_x\left(e^{i<\xi,\bar{X}_{t-\delta,t}(y)>}\Phi(\bar{X}_{t-\delta,t}(y) - y_0)\right)|, \tag{10}$$

which again holds uniformly in x. As in [14], we take advantage of the relationship between the exponential $e^{i<\xi,z>}$ and its partial derivatives with respect to each component of z. Namely $\partial_{z_j}^{(2+k)} e^{i<\xi,z>} = -i^k\xi_j^{2+k}e^{i<\xi,z>}$. We denote by $\partial_\beta = \partial_{z_1}^{(2+k)} \ldots \partial_{z_m}^{(2+k)}$ the composition of these partial derivatives and set $\|\xi\| := \prod_{\ell=1}^m |\xi_\ell|$. Then

$$|E_x\left(e^{i<\xi,\bar{X}_{t-\delta,t}(y)>}\Phi(\bar{X}_{t-\delta,t}(y) - y_0)\right)| \leq \|\xi\|^{-2-k}\Big|E_x\left(\partial_\beta e^{i<\xi,\bar{X}_{t-\delta,t}(y)>}\Phi(\bar{X}_{t-\delta,t}(y) - y_0)\right)\Big|. \tag{11}$$

From the integration by parts formula of Malliavin calculus we conclude that for some functional $\mathcal{H}_k$,

$$|E_x(e^{i<\xi,\bar{X}_{t-\delta,t}(y)>}\Phi(\bar{X}_{t-\delta,t}(y)-y_0))| \le \|\xi\|^{-2-k}E_x(|\mathcal{H}_k(\bar{X}_{t-\delta,t}(y),\Phi(\bar{X}_{t-\delta,t}(y)-y_0))|). \tag{12}$$

We show in the Appendix (Section 6) that

$$\|\mathcal{H}_k(\bar{X}_{t-\delta,t}(y),\Phi(\bar{X}_{t-\delta,t}(y)-y_0))\|_p \le C(r,k,p,R,m,l)\delta^{-m(k+2)k_N}. \tag{13}$$

The constant k_N depends on the order N of successive Lie brackets needed to span $I\!R^m$ at any point of $B_{3R}(y_0)$ according to **(LWH)**. We deduce from (8) and (13) that, for any $q \ge 1$ and any $0 < \delta < \frac{t}{2}\wedge r$,

$$m_0\|\xi\|^k|\hat{\nu}(\xi)| \le C(T,r,k,R,q,m,l)\ \left[\|\xi\|^kR^{-q}\delta^{q/2}+\|\xi\|^{-2}\delta^{-m(k+2)k_N}\right].$$

In order to bound $\|\xi\|^k|\hat{\nu}(\xi)|$ above by an integrable function we now exploit (as in [14]) the freedom that still remains in the choice of the pair (δ,q). Indeed, for a given ξ, we can choose (δ,q) such that $\|\xi\|^kR^{-q}\delta^{q/2}+\|\xi\|^{-2}\delta^{-m(k+2)k_N}$ tends to zero faster than $\|\xi\|^{-3/2}$ as $\|\xi\|\to\infty$ as follows:

$$\delta = t/2\wedge r\wedge 1\wedge \|\xi\|^{-\frac{1}{2m(k+2)k_N}},\quad q = 2m(k+2)k_N(2k+3).$$

Then $m_0\|\xi\|^k|\hat{\nu}(\xi)| \le C(T,r,k,R,q,m,l)\ \|\xi\|^{-\frac{3}{2}}$.

The above estimates hold for any fixed $k\in I\!N$. Therefore the solution of (1) starting from x admits the density

$$p_{0,t}(x,y) = \frac{1}{(2\pi)^m}\int_{I\!R^m} e^{-i<\xi,y>}E_x(e^{i<\xi,X_t>}\Phi(X_t-y_0))d\xi, \tag{14}$$

on $B_R(y_0)$. It remains to show that (5) holds for $\hat{\nu}(\xi)$ and for any $k\in I\!N$. We split the integrals in (5) in two parts, over the bounded set $I := \{\|\xi\|\le M\}$ and its complement I^c for some $M>0$. The modulus of the integrand is bounded on I. On I^c, we use the fact that $E_x(e^{i<\xi,X_t>}\Phi(X_t-y_0))$ coincides with $\hat{\nu}(\xi)$ and the inequality just established: $m_0\|\xi\|^k|\hat{\nu}(\xi)| \le C(T,r,k,R,q,m,l)\ \|\xi\|^{-\frac{3}{2}}$, which is integrable over I^c.

Let us finally notice that applying the above arguments with $k=0$, we see that the continuity of $p_{0,t}(x,y)$ in y is uniform in x since the upper bounds in (8) and (10), obtained for $k=0$, do not depend on x. This finishes our proof. □

Proof of Theorem 2. We keep the notations introduced in the proof of Theorem 1, in particular Φ and ν. In order to prove the lower semi-continuity w.r.t. x, it is enough to show that for fixed $y\in B_R(y_0)$, the function $p_{0,t}(\cdot,y)$ is the limit of an increasing sequence of continuous functions $x\mapsto p^{(n)}_{0,t}(x,y)$. We also use localization arguments here but now the approximating sequence is obtained by considering X before it exits each compact K_n (cf. **(H1)**). Note that continuous dependence on the starting point holds for each approximating process which enjoys the flow property whereas this property may fail to hold for X itself. So, given an integer n, let $b^{(n)}(t,x)$ and $\sigma^{(n)}(x)$ denote $C^\infty-$extensions (in x) of $b(t,\cdot_{|K_n})$ and $\sigma_{|K_n}$. Let $X^{(n)}$ be the solution of the localized version of (1) with coefficients $b^{(n)}$ and $\sigma^{(n)}$. The first exit time of K_n by X is denoted by T_n (cf **(H1)**). Using that $T_n\to\infty$, we can write for any $x\in K_n$ and any positive measurable function f,

$$m_0\int f(y)\nu(dy) = \lim_n\uparrow E_x\left(f(X_t)\Phi(X_t-y_0)1_{\{T_n>t\}}\right).$$

Then for all n, since $X^{(n)}_t = X_t$ on $\{T_n>t\}$ almost surely and Φ is non negative,

$$m_0\int f(y)\nu(dy) \ge E_x\left(f(X_t)\Phi(X_t-y_0)1_{\{T_n>t\}}\right) = E_x\left(f(X^{(n)}_t)\Phi(X^{(n)}_t-y_0)1_{\{T_n>t\}}\right).$$

We approximate $1_{\{T_n>t\}}$ by some continuous functional on $\Omega := C(I\!R_+, I\!R^m)$. The set Ω is endowed with the topology of uniform convergence on compacts. $\mathbb{P}^{(n)}_{0,x}$ denotes the law of $X^{(n)}$ on $(\Omega, \mathcal{B}(\Omega))$, starting from x at time 0. The family $\{\mathbb{P}^{(n)}_{0,x}, x \in I\!R^m\}$ has the Feller property, i.e. if $x_k \to x$, then $\mathbb{P}^{(n)}_{0,x_k} \to \mathbb{P}^{(n)}_{0,x}$ weakly as $k \to \infty$. Thanks to this property, $E_x\left(f(X^{(n)}_t)\Phi(X^{(n)}_t - y_0)\right)$ is continuous w.r.t. x. Define $M^n_t = \max_{s\le t} X^{(n)}_s$ and $m^n_t = \min_{s\le t} X^{(n)}_s$ coordinate-wise. Due to the structure of the compacts K_n (see assumption **(H1)**), we can construct $C^\infty-$functions φ^n, Φ^n such that $1_{[a_{n-1},\infty[} \le \varphi^n \le 1_{[a_n,\infty[}$ and $1_{]-\infty,b_{n-1}]} \le \Phi^n \le 1_{]-\infty,b_n]}$ (these inequalities have to be understood coordinate-wise). Then, since X_t equals $X^{(n)}_t$ up to time T_n,

$$\{T_{n-1} > t\} = \{a_{n-1} \le m^n_t \le M^n_t \le b_{n-1}\} \subset \{\varphi^n(m^n_t) > 0, \Phi^n(M^n_t) > 0\} \subset \{T_n > t\},$$

and for any $f \ge 0$,

$$E_x\left(f(X^{(n)}_t)\Phi(X^{(n)}_t - y_0)1_{\{T_n>t\}}\right) \quad \ge \quad E_x\left(f(X^{(n)}_t)\Phi(X^{(n)}_t - y_0)\Phi^n(M^n_t)\varphi^n(m^n_t)\right).$$

Define now a sub-probability measure ν_n by

$$m_0 \int f(y)\nu_n(dy) := E_x\left(f(X^{(n)}_t)\Phi(X^{(n)}_t - y_0)\Phi^n(M^n_t)\varphi^n(m^n_t)\right). \tag{15}$$

The new functional $\Phi(X^{(n)}_t - y_0)\Phi^n(M^n_t)\varphi^n(m^n_t)$ satisfies the same hypotheses as the former $\Phi(X^{(n)}_t - y_0)$. For any $f \ge 0$,

$$\int f(y)\nu_n(dy) \le \int f(y)\nu_{n+1}(dy) \uparrow \int f(y)\nu(dy) \text{ as } n \to \infty.$$

If we can show that ν_n possesses a density, that we shall denote by $m_0^{-1}p^{(n)}_{0,t}(x,y)$, the following inequalities will hold true

$$p^{(n)}_{0,t}(x,y) \le p^{(n+1)}_{0,t}(x,y) \le p_{0,t}(x,y) \quad \text{for all } n \ge 1, \tag{16}$$

for any fixed x, $\lambda(dy)-$almost surely. So in a next step we show that indeed ν_n possesses a density. In order to indicate explicitly the dependence on the starting point x, we introduce the notation $\gamma_n(x,\xi)$ for $\hat{\nu}_n(\xi)$ as follows,

$$\gamma_n(x,\xi) := \frac{1}{m_0} E_x\left(e^{i<\xi, X^{(n)}_t>}\Phi(X^{(n)}_t - y_0)\Phi^n(M^n_t)\varphi^n(m^n_t)\right),$$

and we apply the argument in the proof of Theorem 1. Inequalities (8) and (12)-(13) for $k = 0$ hold for $m_0\gamma_n(x,\xi)$ which also satisfies $m_0|\gamma_n(x,\xi)| \le C(T,r,R,q,m)\ \|\xi\|^{-\frac{3}{2}}$. Therefore $\xi \to \gamma_n(x,\xi)$ is integrable. Hence, $m_0\nu_n$ admits a density that we denote $p^{(n)}_{0,t}(x,y)$ given by

$$p^{(n)}_{0,t}(x,y) = \frac{m_0}{(2\pi)^m} \int_{I\!R^m} e^{-i<\xi,y>}\gamma_n(x,\xi)d\xi. \tag{17}$$

From the fact that $\gamma_n(x,\xi) \to \hat{\nu}(\xi)$ as $n \to \infty$ and that the upper bounds for $|\gamma_n|$ do not depend on n, we deduce that $p^{(n)}_{0,t}(x,y) \to p_{0,t}(x,y)$. Taking into account (16), we conclude that $p_{0,t}(x,y) = \lim_n \uparrow p^{(n)}_{0,t}(x,y)$.

It remains to show (by dominated convergence) that for any $y \in B_R(y_0)$, the map $x \mapsto p^{(n)}_{0,t}(x,y)$ is continuous. This is a consequence of the continuity of $\gamma_n(x,\xi)$ in x (which follows from the Feller property of $\mathbb{P}^{(n)}_{0,x}$ and the fact that all operations appearing in $\gamma_n(x,\xi)$ are continuous on Ω) and the fact that (8) and (12)-(13) (which we use here for $k = 0$) hold uniformly in x. □

Example 1 continued *Consider again the two dimensional diffusion X_t of Example 1, under the conditions* **(H1)** *and* **(H2)**. *Then X_t admits a smooth density locally around each point (x_1, x_2) satisfying $\tilde{\sigma}(x_1, x_2) \neq 0$ and $\frac{\partial g}{\partial x_2}(x_1, x_2) \neq 0$.*

4. Densities for SDEs with internal variables and random input

4.1. **SDEs with internal variables and random input.** As a particular subclass of systems (1), we introduce in (18) below a structure of SDEs that we call 'SDEs with internal variables and random input'. This structure contains the stochastic Hodgkin-Huxley systems which we shall study in Section 5. Finally, in (19) and (20), we fix notations for certain deterministic $(m-1)-$dimensional systems which in view of the control theorem do have some relation to systems (18).

In order to model a neuron embedded in a network from which it receives an input through its dendritic tree, and able to activate ion channels modeled by the internal variables $2, \ldots, m-1$, we consider systems of the type

$$\begin{aligned} dX_{1,t} &= F(X_{1,t}, \ldots, X_{m-1,t})dt + dX_{m,t}\,, \\ dX_{i,t} &= [-a_i(X_{1,t})X_{i,t} + b_i(X_{1,t})]dt, \;\; i = 2, \ldots, m-1\,, \\ dX_{m,t} &= b_m(t, X_{m,t})dt \; + \sigma(X_{m,t})dW_t\,. \end{aligned} \tag{18}$$

Note that the last component X_m follows an autonomous equation and represents random external input to the system.

We shall also consider, for smooth functions $t \mapsto I(t)$, deterministic $(m-1)-$dimensional systems (19) where $I(t)dt$ replaces $dX_{m,t}$ of (18) and thus acts as deterministic input to the system:

$$\begin{aligned} dz_{1,t} &= F(z_{1,t}, \ldots, z_{m-1,t})dt + I(t)dt\,, \\ dz_{i,t} &= [-a_i(z_{1,t})z_{i,t} + b_i(z_{1,t})]dt, \;\; i = 2, \ldots, m-1, \end{aligned} \tag{19}$$

and in particular, corresponding to zero input $I(\cdot) \equiv 0$, the system

$$\begin{aligned} dz_{1,t} &= F(z_{1,t}, \ldots, z_{m-1,t})dt\,, \\ dz_{i,t} &= [-a_i(z_{1,t})z_{i,t} + b_i(z_{1,t})]dt, \;\; i = 2, \ldots, m-1. \end{aligned} \tag{20}$$

Example 2 (FitzHugh-Nagumo with random external input)**.** *The well known FitzHugh-Nagumo system is an important model in neuroscience, see e.g.* [25] *and* [6]. *We follow* [25] *and consider a FitzHugh-Nagumo system driven by random external input of Ornstein-Uhlenbeck type. It is given by*

$$\begin{aligned} dX_{1,t} &= [-X_{2,t} + f(X_{1,t})]dt + dX_{3,t}, \\ dX_{2,t} &= [bX_{1,t} - cX_{2,t}]dt, \\ dX_{3,t} &= (\,S(t) - X_{3,t}\,)\,dt \; + \; \gamma\, dW_t, \end{aligned} \tag{21}$$

where $b, c, \gamma > 0$ and where the function f is a cubic polynomial $f(x) = x(a-x)(x-1)$. If $a > 0$, then the deterministic system $dz_{1,t} = [-z_{2,t} + f(z_{1,t})]dt; dz_{2,t} = [bz_{1,t} - cz_{2,t}]dt$ has a stable equilibrium. In (21)*, the first variable $X_{1,t}$ models the action potential of the membrane of a single neuron at time t. $X_{2,t}$ is a summary variable representing the states of the ion channels in the membrane. $X_{3,t}$ is a random external input of Ornstein-Uhlenbeck type, carrying a deterministic signal $S(t)$.* (21) *is an example for* (18) *with $m = 3$ and $a_2(x_1) = c$, $b_2(x_1) = bx_1$.*

If for all $2 \leq i \leq m-1$, b_i and $a_i - b_i$ are positive, then the system (20) can be interpreted as the limit of a sequence of stochastic systems, known under the terminology of 'stochastic hybrid systems', in the sense of the Law of Large Numbers or Fluid Limit (cf [38]). Stochastic hybrid systems describe a deterministic dynamics which is coupled with jump Markov processes. More precisely, consider a population of individuals of $m-1$ different types, with N individuals of each type. Each individual is in two states (active or inactive, corresponding to open or closed channels). The individuals are coupled by a global variable z_1: the transition rates between the two states depend on z_1 only and are given by $b_i(z_1)$ for a transition from 'inactive' to 'active' and by $a_i(z_1) - b_i(z_1)$ for a transition from 'active' to 'inactive' for individuals of type i. The variable $z_{i,t}$ denotes the proportion of active individuals at time t, and in the $N \to \infty-$limit, $z_{i,t}$ gives the probability that an individual of type i is active at time t.

Systems (20) arise in various modeling issues. We refer the reader e.g. to Section 2.2.2. of [19] where the time evolution of the concentration of a molecule X in presence or absence of a rare molecular species is described by a model of type (20).

The detailed form of the functions F and a_i, b_i in the Hodgkin-Huxley system (cf. [20] and [25]) will be provided in Section 5. In this system, well-known in neurophysiology, three types of 'agents' are considered which are responsible for opening or closing of K^+ and Na^+ ion channels. In this particular model, we have three equations for internal variables corresponding to $m = 5$. z_1 describes the membrane potential of the neuron, which can be observed. The z_i, $i = 2, 3, 4$ are the gating variables associated to specific ion channels located in the membrane, that are not observed. One may consider models which include still more types of ion channels admitting their specific number of different types of 'agents', hence the interest to consider models (18) with general m.

By the general assumptions associated with (1), the coefficients $F : \mathbb{R}^{m-1} \to \mathbb{R}$, $a_i, b_i : \mathbb{R} \to \mathbb{R}$ for $2 \le i \le m-1$ and $b_m : [0, \infty] \times \mathbb{R} \to \mathbb{R}$ are smooth. In what follows, we suppose that the coefficients of (18) are such that **(H1)** and **(H2)** are satisfied. If we assume moreover that $0 \le b_i(z_1) \le a_i(z_1)$ and $a_i(z_1) > 0$ for all $i = 2, \ldots, m-1$, for all $z_1 \in \mathbb{R}$, then

$$y_{i,\infty}(z_1) := \frac{b_i(z_1)}{a_i(z_1)},\ z_1 \in \mathbb{R}, \tag{22}$$

are equilibrium points of the internal equations when we keep the first variable fixed at constant value z_1. In particular, introducing

$$F_\infty(z_1) := F(z_1, y_{2,\infty}(z_1), \ldots, y_{m-1,\infty}(z_1)),\ z_1 \in \mathbb{R}, \tag{23}$$

any point $(z_1, y_{2,\infty}(z_1), \ldots, y_{m-1,\infty}(z_1))$ such that $F_\infty(z_1) = 0$ is an equilibrium point of the system (20).

The aim of the next two sections is to make the conditions of Theorem 1 explicit for systems (18). First, we define some determinant $\mathbf{D}(x)$ on the points x of the state space and prove the following: if $\mathbf{D}(y_0) \neq 0$, then **(LWH)** holds at (t, y_0) for every t. Since this determinant can be evaluated numerically we can check condition **(LWH)** at (t, y_0). Second, we prove a 'chamaeleon property' for the system (18): on given finite time intervals, the solution of (18) imitates deterministic evolutions (19) for (almost) any smooth deterministic input $t \to I(t)$ with positive probability, provided both systems (18) and (19) have the same starting point. For systems (18), we thus do have tools to verify condition **(LWH)** of Theorem 1 and to prove positivity of the density.

4.2. **Weak Hörmander condition for (18).** We assume that assumptions **(H1)**-**(H2)** are satisfied as well as the following additional assumption on the autonomous equation for X_m.

(H3) There exists an open interval $U \subset \mathbb{R}$ such that the $m-$th equation in system (18)

$$dX_{m,t} = b_m(t, X_{m,t})dt\ + \sigma(X_{m,t})dW_t$$

possesses a unique strong solution taking values in U, whenever $X_{m,0} \in U$. Moreover $\sigma(\cdot)$ is strictly positive on U and its restriction to every compact interval in U is of class C^∞.

The linearity of the equation for dX_i, $i \in \{2, \ldots, m-1\}$, with respect to X_i has an important consequence that we recall in the following proposition (the proof of this proposition is provided in the Appendix).

Proposition 2. *Fix $i \in \{2, \ldots, m-1\}$. Suppose that $X_{i,0} \in [0,1]$ a.s., and also $0 \le b_i(x) \le a_i(x)$, for all $x \in I\!R$, x denoting the first component of* (18). *Then $\forall t > 0$, $X_{i,t} \in [0,1]$ a.s.*

In view of Proposition 2 we assume for the rest of this section that $X_{i,0} \in [0,1]$ a.s., and $0 \le b_i(x) \le a_i(x)$, for all $x \in I\!R$, for all $i \in \{2, \ldots, m-1\}$. We define $E_m := I\!R \times [0,1]^{m-2} \times U$ where U is given by **(H3)** and take E_m as state space for systems (18).

Definition 1. *For any integer $k \ge 1$ denote by $\partial_{x_1}^{(k)}$ the partial derivative of order k w.r.t. x_1. For any $x \in I\!R^{m-1} \times U$ consider $J_1(x) := F(x_1, x_2, \cdot, \cdot, \cdot, x_{m-1})$ and $J_i(x) := -a_i(x_1)x_i + b_i(x_1)$, $2 \le i \le m-1$. We define $\mathbf{D}(x)$ as the determinant of the matrix $(\partial_{x_1}^{(k)} J_i(x)\,;(i,k) \in \{1, \ldots, m-1\}^2)$.*

Notice that the above determinant makes only use of the drift vector of the zero input system (20).

Theorem 3. *Suppose that* (18) *satisfies* **(H1)**–**(H3)**. *For $x = (x_1, \ldots, x_m) \in \text{int}(E_m)$, the condition $\mathbf{D}(x) \ne 0$ implies that* **(LWH)** *holds at (t,x) for all t.*

We stress that the condition $\mathbf{D}(x) \ne 0$ is a sufficient condition implying **(LWH)**. $\mathbf{D}(x)$ can be evaluated numerically for given sets of functions $F(x_1, \ldots, x_{m-1})$, a_i and $b_i, 2 \le i \le m-1$, defining (18), see e.g. Section 5.4 below.

The proof of Theorem 3 will be given below (through Proposition 3). First we state some important consequences and make some remarks. It is important to note that $\mathbf{D}(x)$ actually depends only on the $m-1$ first components of x. In particular if the $m-1$ first components of two points x and x' coincide, then $\mathbf{D}(x) = \mathbf{D}(x')$. This remark will be important in the sequel (see e.g. Proposition 2 below). Moreover the condition in Theorem 3 implies a version of **(LWH)** uniform w.r.t. time on every compact interval $[0,T]$. Let us now define the set

$$\mathcal{D} := \{(x_1, \ldots, x_m) \in \text{int}(E_m); \quad \mathbf{D}(x) \ne 0\}.$$

The set $\mathcal{D}$ is an open subset of E_m by continuity of $\mathbf{D}$ on $I\!R^{m-1} \times U$. The following statement is a direct consequence of Theorems 1 and 2 of Section 3.2.

Theorem 4. *Suppose that* (18) *satisfies* **(H1)**–**(H3)**. *Assume that $y_0 \in \mathcal{D}$ and take $R > 0$ such that $B_{3R}(y_0) \subset \mathcal{D}$. Then for any $x \in E_m$ and $t > 0$, the random variable X_t admits an infinitely differentiable density $p_{0,t}(x, \cdot)$ on $B_R(y_0)$. The map $y \in B_R(y_0) \mapsto p_{0,t}(x,y)$ is continuous, and for any fixed $y \in B_R(y_0)$, the map $x \in E_m \mapsto p_{0,t}(x,y)$ is lower semi-continuous.*

Corollary 1. *Grant the assumptions of Theorem 4. For all $x \in E_m$, the following holds true. If there exists $y \in \mathcal{D}$ and $t > 0$ such that $P_{0,t}(x,U) > 0$ for all sufficiently small neighborhoods U of y, then there exists $\delta > 0$ such that, if K_1 (resp. K_2) denotes the closure of $B_\delta(x)$ (resp. $B_\delta(y)$),*

$$\inf_{x' \in K_1} \inf_{y' \in K_2} p_{0,t}(x', y') > 0.$$

The difficulty in practice is to obtain more information on $\mathcal{D}$, in particular to know whether it coincides with $\text{int}(E_m)$. At least one would like to be able to specify open regions included in $\mathcal{D}$. In general, one can hope to achieve this goal only numerically unless the coefficients of the system are very simple. In Section 5 we provide details for a stochastic Hodgkin-Huxley model. The definition of $\mathcal{D}$ comes from a particular choice of successive Lie brackets where we look for the directions in space which propagate the noise at maximal possible speed according to the following intuition: the noise in (18) is most rapidly transported through X_1 and X_m, since they are the only components carrying Brownian noise explicitly. Accordingly, except for the first Lie bracket $[A_0, A_1]$ which involves the drift A_0, we always use the diffusion coefficient A_1 in order to compute the brackets of higher order. The corresponding development of the solution of (18) into iterated Ito integrals for small time steps δ, shows that the speed of the diffusion is of order $\delta^{\frac{1}{2}}$ in the direction of A_1, of order $\delta^{1+\frac{1}{2}}$ in the direction of $[A_0, A_1]$ and for the subsequent Lie brackets we add a factor $\frac{1}{2}$ to the exponent each time we use A_1, so that the speed of the diffusion is of order δ^2 in the direction of $[[A_0, A_1], A_1]$, of order $\delta^{1+3\times\frac{1}{2}}$ in the direction of $[[[A_0, A_1], A_1], A_1]$ and so on. We refer the reader to [35], in particular identity (12). Hence it is important to remember that belonging to $\mathcal{D}$ is only a sufficient condition for **(LWH)** to hold.

We now prove Theorem 3 starting with the following key proposition about the computation of Lie brackets in this specific case. The proof is a direct consequence of the definition of Lie bracket recalled in Section 3.1 and is left to the reader.

Proposition 3. *Consider on one hand φ, ψ and ρ smooth functions of x_m defined on U and on the other hand a family of smooth functions y_i, $1 \le i \le m-1$, defined on $I\!R^{m-1}$, which do not depend on x_m. Let Ξ and Y denote vector fiels on $[0,+\infty] \times I\!R^m$ of the following form,*

$$\begin{aligned} \Xi(t,x) &:= \varphi(x_m)\,(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_m}), \\ Y(t,x) &:= \rho(x_m) \sum_{i=1}^{m-1} y_i \,\frac{\partial}{\partial x_i} + \psi(x_m)\,(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_m}). \end{aligned}$$

The Lie bracket $[\Xi, Y]$ takes the form

$$\begin{aligned} [\Xi, Y](t,x) &= \varphi(x_m)\rho(x_m) \sum_{i=1}^{m-1} \partial_{x_1} y_i \,\frac{\partial}{\partial x_i} \\ &+ \varphi(x_m)\rho'(x_m) \sum_{i=1}^{m-1} y_i \frac{\partial}{\partial x_i} + (\varphi\psi' - \varphi'\psi)(x_m)\;(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_m}). \end{aligned}$$

Proof of Theorem 3. According to the notations of Section 3.1 with one dimensional driving Brownian motion, hence $l = 1$, we write $A_1 = \sigma(x_m)(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_m})$ and $A_0 = \frac{\partial}{\partial t} + \sum_{i=1}^m \tilde{b}_i \frac{\partial}{\partial x_i}$ where $\tilde{b}$ is given in (2). Let us consider the Lie brackets defined recursively by $L_1 := [A_1, A_0]$ and $L_{k+1} = [A_1, L_k]$. In order to illustrate the relationship between the L_k and the determinant $\mathbf{D}(x)$ introduced in Definition 1, we compute explicitly L_1 and L_2. We find first that

$$L_1 = \sum_{i=1}^m \sigma(x_m)(\frac{\partial \tilde{b}_i}{\partial x_1} + \frac{\partial \tilde{b}_i}{\partial x_m})\frac{\partial}{\partial x_i} - \sigma'(x_m)\tilde{b}_m(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_m}).$$

The drift $\tilde{b}$ in (18) satisfies $\frac{\partial \tilde{b}_m}{\partial x_1} \equiv 0$, $\frac{\partial \tilde{b}_i}{\partial x_m} \equiv 0$ for all $i \in \{2, \dots, m-1\}$. Moreover $\frac{\partial \tilde{b}_i}{\partial x_1} \equiv \partial_{x_1} J_i$ for all $i \in \{1, \dots, m-1\}$. Hence

$$L_1 = \sum_{i=1}^{m-1} \sigma(x_m)\partial_{x_1} J_i \,\frac{\partial}{\partial x_i} + \sigma(x_m)(\frac{\partial \tilde{b}_1}{\partial x_m}\frac{\partial}{\partial x_1} + \frac{\partial \tilde{b}_m}{\partial x_m}\frac{\partial}{\partial x_m}) - \sigma'(x_m)\tilde{b}_m(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_m}).$$

We can further reduce the expression of L_1 using that the drift $\tilde b$ in (18) satisfies also $\frac{\partial \tilde b_1}{\partial x_m} \equiv \frac{\partial \tilde b_m}{\partial x_m}$. We obtain

$$L_1 = \sum_{i=1}^{m-1} \sigma(x_m)\partial_{x_1} J_i \frac{\partial}{\partial x_i} + \Big(\sigma(x_m)\frac{\partial \tilde b_m}{\partial x_m} - \sigma'(x_m)\tilde b_m\Big)(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_m}). \tag{24}$$

Proposition 3 applies to $\Xi = A_1$ and $Y = L_1$ with $\varphi(x_m) \equiv \rho(x_m) \equiv \sigma(x_m)$, $y_i \equiv \partial_{x_1} J_i$, for $i \in \{1, \ldots, m-1\}$, $\psi(x_m) \equiv \sigma(x_m)\frac{\partial \tilde b_m}{\partial x_m} - \sigma'(x_m)\tilde b_m$. Therefore, with this specific choice,

$$\begin{aligned} L_2 \;&=\; \sum_{i=1}^{m-1} \sigma(x_m)^2\, \partial^{(2)}_{x_1} J_i\, \frac{\partial}{\partial x_i} \\ &+\; \sum_{i=1}^{m-1} \sigma(x_m)\sigma'(x_m)\partial_{x_1} J_i\, \frac{\partial}{\partial x_i} + (\varphi\psi' - \varphi'\psi)(x_m)\, (\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_m}). \end{aligned} \tag{25}$$

Once again, identity (25) coupled with Proposition 3 enables us to work by iteration. We thus obtain the following expression for L_k, for any $k \geq 1$:

$$L_k = \sum_{i=1}^{m-1} \sigma(x_m)^k\, \partial^{(k)}_{x_1} J_i\, \frac{\partial}{\partial x_i} + \sum_{\ell=1}^{k-1}\sum_{i=1}^{m-1} \Phi_\ell(x_m)\, \partial^{(\ell)}_{x_1} J_i\, \frac{\partial}{\partial x_i} + \Phi(x_m)(\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_m}). \tag{26}$$

The explicit expression of Φ_ℓ, Φ are not necessary to conclude. Indeed let us identify these Lie brackets with the column vectors in $I\!R^m$ obtained with their coordinates on the basis $(\frac{\partial}{\partial x_i},\, i \in \{1, \ldots, m\})$. A sufficient condition for **(LWH)** to be satisfied is that the vector space generated by $(A_1, L_k,\, k \in \{1, \ldots, m-1\})$ coincides with $I\!R^m$. It is sufficient that the determinant formed with these vectors does not vanish. The definition of A_1 and formula (26) imply that this determinant coincides with the determinant obtained with the vectors A_1 and $\tilde L_k$, $k \in \{1, \ldots, m-1\}$ where $\tilde L_k := \sum_{i=1}^{m-1} \sigma(x_m)^k\, \partial^{(k)}_{x_1} J_i\, \frac{\partial}{\partial x_i}$. Since σ does not vanish on U (cf. **(H3)**, we conclude that a sufficient condition is that $\mathbf{D}(x)$ does not vanish. □

Example 2 continued *In the situation of Example 3, we have* $J_1(x_1, x_2) = -x_2 + f(x_1)$ *and* $J_2(x_1, x_2) = bx_1 - cx_2$. *It is easy to see that in this case* $\mathbf{D}(x) = -bf''(x_1) \neq 0$ *for all* $x_1 \neq \frac{a+1}{3}$ *by definition of* f. *Taking one more derivative, i.e. calculating the Lie bracket* $[A_1, [A_1, [A_1, A_0]]]$ *leads actually to the condition* $f'''(x) \neq 0$ *which is always true. This highlights that* $\mathbf{D}(x) \neq 0$ *is only a sufficient condition implying* **(LWH)**, *and the Fitzhugh-Nagumo system with random external input of Example 3 actually satisfies the weak Hörmander condition on the whole state space.*

4.3. **Positivity of densities for models** (18)**.** Once we have proved that densities exist for (18), even if only locally, we look for regions where they are positive. For this purpose we combine control arguments and the support theorem. We keep the notation $E_m = I\!R \times [0,1]^{m-2} \times U$ introduced in the previous section. We start by proving an accessibility result for (18) in Proposition 4 below, which holds without any assumption on the existence of densities and relies on some stability properties of the underlying deterministic system (20). We refer the reader to [5] for similar ideas in the framework of Piecewise Deterministic Markov Processes.

Let $(X_u)_{u\geq 0}$ be a solution of (18). We denote by $P_{0,t}(x, \cdot)$ the law of X_t when $X_0 = x$ a.s. Recall the structure of system (20).

Proposition 4. *Grant* **(H1)**–**(H3)** *and assume that* $U = I\!R$. *We keep moreover the assumptions of Proposition 2 and suppose that* $0 < b_i < a_i$ *for all* $i \in \{2, \ldots, m-1\}$. *Given an arbitrary real number* z_1, *consider* $\overline{z} := (z_1,\, y_{i,\infty}(z_1),\, i \in \{2, \ldots, m-1\})$ *in* $I\!R^{m-1}$, *where* $y_{i,\infty}(z_1) := \frac{b_i(z_1)}{a_i(z_1)}$ *is an* *equilibrium point for the $i-$th equation of* (20) *when we keep the first variable fixed at constant value z_1. For all $x \in E_m$ and any neighborhood $\mathcal{N}$ of $\overline{z}$ in $I\!R \times (0,1)^{m-2}$ there exists t_0 such that*

$$\forall t \geq t_0, \quad P_{0,t}(x, \mathcal{N} \times I\!R) > 0. \tag{27}$$

Proposition 5. *Let us keep the notations and assumptions of Proposition 4. Consider an arbitrary real number z_1 and the associated point $\overline{z} := (z_1,\, y_{i,\infty}(z_1),\; i \in \{2, \ldots, m-1\})$ in $I\!R^{m-1}$. Assume that $\mathbf{D}(\overline{z}, u) \neq 0$ for some $u \in I\!R$, where the determinant $\mathbf{D}$ has been introduced in Definition 1. Then for all $x \in E_m$, there is $t_0 > 0$ such that for all $t \geq t_0$ the following holds true. There exist $\overline{u} = \overline{u}(t) \in I\!R$ and $\delta = \delta(t) > 0$ such that, if K_1 (resp. K_2) denotes the closure of $B_\delta(x)$ (resp. $B_\delta(\overline{z}, \overline{u})$),*

$$\inf_{x' \in K_1} \inf_{y' \in K_2} p_{0,t}(x', y') > 0.$$

Remark 2. *Notice that for each $i \in \{2, \ldots, m-1\}$, the solution of $dy_t = (-a_i(z_1)y_t + b_i(z_1))dt$ with z_1 as a fixed parameter, converges to $y_{i,\infty}(z_1)$ when $t \to +\infty$ and that $y_{i,\infty}(z_1)$ is globally asymptotically stable. Proposition 4 holds in particular when $F(\overline{z}) = 0$. In this case $\overline{z}$ is an equilibrium point of (20) and – from* (32) *in the proof below – we can choose $\overline{u}(t) \equiv \overline{u}$ as a constant not depending on time.*

Proof of Proposition 4. Let $z_1 \in I\!R$ and the associated point $\overline{z} := (z_1,\, y_{i,\infty}(z_1),\; i \in \{2, \ldots, m-1\})$ in $I\!R^{m-1}$. As in the proof of Theorem 2 we write Ω for $C([0, \infty[, I\!R^m)$ and endow it with its canonical filtration $(\mathcal{F}_t)_{t \geq 0}$. Recall that $\mathbb{P}_{0,x}$ is the law of $(X_u)_{u \geq 0}$ starting from x at time 0. We first localize the system by a sequence of compacts (K_n) according to **(H1)** and let $T_n = \inf\{t : X_t \in K_n^c\}$ be the exit time of K_n. For a fixed n, let $b^{(n)}(t,x)$ and $\sigma^{(n)}(x)$ be $C_b^\infty-$extensions in x of $b(t, \cdot_{|K_n})$ and $\sigma_{|K_n}$ respectively and $X^{(n)}$ be the associated diffusion process (here we denote the coefficients of (18) by b and σ for short). For any integer $n \geq 1$ and starting point x, we write $\mathbb{P}_{0,x}^{(n)}$ for the law of $(X_u^{(n)})_{u \geq 0}$ on Ω satisfying $X_0^{(n)} = x$. We wish to find lower bounds for quantities of the form $\mathbb{P}_{0,x}(B)$ where $B = \{f \in \Omega : f(t) \in \mathcal{N} \times I\!R\} \in \mathcal{F}_t$, for any $t > 0$ given . We start with the following inequality which holds for any $t > 0$ and n:

$$\mathbb{P}_{0,x}(B) \geq \mathbb{P}_{0,x}(\{f \in B; T_n > t\}) = \mathbb{P}_{0,x}^{(n)}(\{f \in B; T_n > t\}). \tag{28}$$

In the sequel we show that for some integer n_0 and any fixed $x \in K_{n_0}$, the quantity $\mathbb{P}_{0,x}^{(n)}(\{f \in B; T_n > t\})$ is indeed positive provided that n is sufficiently large. We are therefore interested in the support of $\mathbb{P}_{0,x}^{(n)}$. Fix t and let $\mathcal{C} := \{\mathtt{h} : [0,t] \to I\!R : \mathtt{h}(s) = \int_0^s \dot{\mathtt{h}}(u)du, \forall\, s \leq t, \int_0^t \dot{\mathtt{h}}^2(u)du < \infty\}$ be the Cameron-Martin space. Given $\mathtt{h} \in \mathcal{C}$, consider $X(\mathtt{h}) \in I\!R^m$ the solution of the differential equation

$$X(\mathtt{h})_s = x + \int_0^s \sigma^{(n)}(X(\mathtt{h})_u)\dot{\mathtt{h}}(u)du + \int_0^s \tilde{b}^{(n)}(u, X(\mathtt{h})_u)du, \quad s \leq t. \tag{29}$$

If (29) were time homogeneous, the support theorem would imply that the support of $\mathbb{P}_{0,x}^{(n)}$ in restriction to $\mathcal{F}_t$ is the closure of the set $\{X(\mathtt{h}) : \mathtt{h} \in \mathcal{C}\}$ with respect to the uniform norm on $[0,t]$ (see e.g. [33] Theorem 3.5 or [4] Theorem 4). To conclude in our situation as well, it is enough to replace the $m-$dimensional process $X^{(n)}$ by the $(m+1)-$dimensional process $(t, X_t^{(n)})$ which is time-homogenous. In order to proceed further we construct a control $\mathtt{h}$ so that $X(\mathtt{h})$ remains in K_n during $[0,t]$ provided n is sufficiently large. We start by exploiting stability properties of the underlying deterministic system (20). The main idea of our proof is to choose a a smooth function $\gamma : I\!R \mapsto I\!R$ satisfying $\gamma(\tau) := z_1$ for all $\tau \geq 1$. Once γ is fixed, consider $y_s \in I\!R^{m-2}$ solving $dy_{i,s} = [-a_i(\gamma(s))y_{i,s} + b_i(\gamma(s))]ds,\; i = 2, \ldots, m-1$. Then for all $t > 1$,

$$y_{i,t} = y_{i,0}e^{-\int_0^t a_i(\gamma(s))ds} + \int_0^t b_i(\gamma(u))e^{-\int_u^t a_i(\gamma(r))dr}du = y_{i,1}e^{-a_i(z_1)(t-1)} + y_{i,\infty}(z_1)(1 - e^{-a_i(z_1)(t-1)}).$$

This formula expresses the fact that on $[1, +\infty[$, the coefficients $a_i(\gamma(s))$ (resp. $b_i(\gamma(s))$) are constant equal to $a_i(z_1)$ (resp. $b_i(z_1)$). It shows that $y_{i,t}$ – whenever z_1 acts as a fixed parameter – converges to $y_{i,\infty}(z_1)$ when $t \to +\infty$ and that $y_{i,\infty}(z_1)$ is globally asymptotically stable. Hence for any $\varepsilon > 0$ there exists $t_0 > 1$ such that $|y_{i,t} - y_{i,\infty}(z_1)| < \varepsilon$ for all $t \geq t_0$ and all $2 \leq i \leq m-1$. Now take ε so small that $B_\varepsilon(\overline{z}) \subset \mathcal{N}$. Then for all $t \geq t_0 > 1$, the vector $(\gamma(t), y_{2,t}, \ldots, y_{m-1,t})$ belongs to $B_\varepsilon(\overline{z})$ (remember that for $t > 1$, $\gamma(t)$ is fixed at z_1).
Fix an integer n_0 and x in K_{n_0}. We are now able to construct a control $h \in \mathcal{C}$ such that the solution of (29) remains in K_n during finite time intervals for all n large enough. Choose a function γ as above satisfying moreover $\gamma(0) = x_1$. Define $(Z_s)_{s\geq 0} \in \mathbb{R}^m$, the deterministic path starting from x such that

$$\begin{aligned} Z_{1,s} &= \gamma(s)\,, \\ dZ_{i,s} &= [-a_i(Z_{1,s})Z_{i,s} + b_i(Z_{1,s})]ds, \;\; i = 2, \ldots, m-1\,, \\ Z_{m,s} &= x_m - x_1 + \gamma(s) - \int_0^s F(Z_u)du. \end{aligned} \tag{30}$$

Conditionally on $Z_{1,s}, \ldots, Z_{m-1,s}$, the last component $Z_{m,s}$ plays the role of an integrated deterministic input $s \to \int_0^s I(u)du$ in (19).

Note that $(Z_s,\, s \in [0,t])$ is bounded and therefore remains in K_n for all n large enough. Now fix $t \geq t_0$ and consider a function $\mathtt{h}$ defined by

$$\dot{\mathtt{h}}(s) := \frac{\dot{\gamma}(s) - F(Z_s) - b_m(s, Z_{m,s}) + \frac{1}{2}\sigma(Z_{m,s})\sigma'(Z_{m,s})}{\sigma(Z_{m,s})}. \tag{31}$$

Since by assumption $\sigma(\cdot) > 0$ on $\mathbb{R}$ (we have assumed that $U = \mathbb{R}$), the expression (31) is well-defined. This assumption also provides that $\dot{\mathtt{h}} \in L^2([0,t])$, hence $\mathtt{h} \in \mathcal{C}$. Hence, with such a choice of $\mathtt{h}$, the solution $X(\mathtt{h})$ of equation (29) coincides with the solution Z of system (30). As explained previously, we can choose n such that $(Z_s,\, s \in [0,t])$ remains in K_n.
Consider now, for $\delta > 0$, the tubular neighborhood T_δ of $(Z_s,\, s \in [0,t])$ in Ω of size δ, namely the set $\{f \in \Omega : \sup_{s\leq t} |f(s) - Z_s| < \delta\}$. By the support theorem $\mathbb{P}^{(n)}_{0,x}(T_\delta) > 0$. Remember that we have chosen ϵ and t_0 in order to satisfy $T_\delta \subset \{f \in \Omega : f(t) \in B_\varepsilon(\overline{z}) \times \mathbb{R}\}$ as well as $B_\varepsilon(\overline{z}) \subset \mathcal{N}$. Choosing $\delta \leq \varepsilon/2$ such that $T_\delta \subset \{f \in \Omega : T_n(f) > t\}$, we conclude as announced that

$$P_{0,t}(x, \mathcal{N} \times \mathbb{R}) \geq P_x(X_t \in B_\varepsilon(\overline{z}) \times \mathbb{R}) \geq \mathbb{P}^{(n)}_{0,x}(T_\delta) > 0.$$

□

Proof of Proposition 5. The fact that $\mathbf{D}(\overline{z}, u) \neq 0$ for some $u \in \mathbb{R}$ implies that $\mathbf{D}(\overline{z}, \tilde{u}) \neq 0$ for all $\tilde{u} \in \mathbb{R}$, by Theorem 3. The attainability at time t is proven as in the proof of Proposition 4. For $x \in E_m$, γ_1 and t_0 as there, $t_0 > 1$, we define for $t \geq t_0$

$$\overline{u}(t) = x_m - x_1 + z_1 - \int_0^t F(Z_s)ds \in U = \mathbb{R}. \tag{32}$$

Then there is some $\delta(t) > 0$ such that $P_{0,t}(x, B_{\delta(t)}(\overline{z}, \overline{u}(t))) > 0$. Applying Corollary 1 to $y = (\overline{z}, \overline{u}(t)) \in \mathcal{D}$ finishes the proof. □

We are now able to prove the main result of this section which shows that, during any arbitrary long period, with positive probability, the stochastic system (18) is able to reproduce the behavior of $(dz_t, I(t)) \in \mathbb{R}^m$ where $z(t)$ is a solution of (19) with $I(t)$ an arbitrary smooth input applied to (20). Note that by comparing (18) and (19) we see that the $m-$the component X_m of the stochastic system (18) has to be compared to a deterministic control path (19) to which we add an $m-$th coordinate given by $t \to X_{m,0} + \int_0^t I(s)ds$.

Remember $B_\delta(x)$ denotes the open ball of radius δ centered at x. In the following, $U \subset I\!R$ is again the open interval of Assumption **(H3)**.

Theorem 5. *Suppose that* (18) *satisfies* **(H1)**–**(H3)**. *Fix $x \in E_m$ and $t > 0$. Let I be a smooth deterministic input such that $x_m + \int_0^s I(r)dr \in U$ for all $s \le t$. Define $\mathbb{X}^x_s := (\mathbb{Y}^{\tilde{x}}_s,\, x_m + \int_0^s I(r)dr,\, s \le t)$ where $\mathbb{Y}^{\tilde{x}}$ is the deterministic path solution of (19) starting from $\tilde{x} := (x_1, \dots, x_{m-1})$. We denote by $\mathbb{P}_{0,x}$ the law of the solution of (18) starting at x. Then for any $\varepsilon > 0$ there exists $\delta > 0$ such that for all $x'' \in B_\delta(x)$*

$$\mathbb{P}_{0,x''}\left(\left\{ f \in \Omega : \sup_{s \le t} |f(s) - \mathbb{X}^x_s| \le \varepsilon \right\}\right) > 0.$$

Proof of Theorem 5. We keep the notations introduced in the proof of Proposition 4. In the course of this proof we have shown that the support theorem applies to inhomogeneous diffusions like the one obtained after localizing (18). We will prove the positivity we are looking for through inequalities (28) and paths solving (29) for $h \in \mathcal{C}$, that remain in K_n during $[0,t]$ for n sufficiently large. So the system we work with is the localized one. Consider I a deterministic input such that $x_m + \int_0^s I(r)dr \in U$ for all $s \le t$. Define $\chi_{m,s} := x_m + \int_0^s I(r)dr$ for all $s \le t$ and

$$\dot{\mathtt{h}}(s) := \frac{I(s) - b_m(s, \chi_{m,s}) + \frac{1}{2}\sigma(\chi_{m,s})\sigma'(\chi_{m,s})}{\sigma(\chi_{m,s})}. \tag{33}$$

By definition $(\chi_{m,s}, s \le t)$ lies in a compact interval included in U. Then, the expression (33) is well-defined by assumption **(H3)**. This assumption also provides that $\dot{\mathtt{h}} \in L^2([0,t])$ hence $\mathtt{h} \in \mathcal{C}$. Moreover, with such a choice of $\mathtt{h}$, the controlled path $X(\mathtt{h})$, solution of (29), coincides with $(\mathbb{Y}^{\tilde{x}}_s,\, \chi_{m,s},\, s \le t)$ where $\mathbb{Y}^{\tilde{x}}$ is the deterministic path solution of (19) starting from $\tilde{x} = (x_1, \dots, x_{m-1})$. We can choose n large enough such that $(\mathbb{Y}^{\tilde{x}}_s,\, \chi_{m,s},\, s \in [0,t])$ remains in K_n. We write $\mathbb{X}^x_s$ for $(\mathbb{Y}^{\tilde{x}}_s,\, \chi_{m,s})$. Remember that $\Omega = C([0,\infty[; I\!R^m)$ and for $\delta > 0$, consider the tubular neighborhood T_δ of $\mathbb{X}^x$ on $[0,t]$ namely the set $\{f \in \Omega : \sup_{s \le t} |f(s) - \mathbb{X}^x_s| < \delta\}$. By the support theorem $\mathbb{P}^{(n)}_{0,x}(T_\delta) > 0$. Choose now δ such that $T_\delta \subset \{f \in \Omega : T_n(f) > t\}$. Taking T_δ as the set B in (28) yields the first statement of Theorem 5. The second one follows from the Feller property of $\mathbb{P}^{(n)}_{0,x}$ which enables us to extend the first statement to a small ball around x. □

We close this section with the following consequence of Theorem 5 from which we borrow the notations. We focus on equilibria for deterministic systems (19) under a particular choice of time-constant input:

$$\text{Associated to } z^* \in I\!R \times (0,1)^{m-1}, \text{ we have } I(\cdot) \equiv c^* := -F(z^*),$$

where we write $z^* := (z^*_1, y_{1,\infty}(z^*_1), \dots, y_{m-1,\infty}(z^*_1))$ for $z^*_1 \in I\!R$, and $F_\infty(z^*_1) = F(z^*)$ as in (22) and (23). The following improves on Proposition 5.

Corollary 2. *Assume that* (18) *satisfies* **(H1)**-**(H3)**, *and consider z^* and c^* as above. Assume that $\mathbf{D}(z^*, u) \neq 0$ for some $u \in U$. Consider $x_m \in U$ and $t > 0$ such that $x_m + c^* s \in U$ for all $0 \le s \le t$. Then for $x := (z^*, x_m)$ and $y := (z^*, x_m + c^* t)$, there exists $\delta > 0$ such that*

$$\inf_{x' \in B_\delta(x)} \inf_{y' \in B_\delta(y)} p_{0,t}(x', y') > 0\ .$$

Proof of Corollary 2. For z^* and $I(t) \equiv c^* = -F(z^*)$ as above, the deterministic path solution to (19) with starting point z^* is constant in time. Attainability of y at time t follows from Theorem 5. $\mathbf{D}(y) \neq 0$ follows from $\mathbf{D}(z^*, u) \neq 0$ for some $u \in U$, thus **(LWH)** holds at (t, y) for all $t > 0$ by Theorem 3. □

The following specializes to equilibria under zero input $I(\cdot) \equiv 0$:

Corollary 3. *Assume that* (18) *satisfies* **(H1)**-**(H3)**. *If equation (20) admits equilibria, i.e. points* $z^* \in I\!\!R \times (0,1)^{m-1}$ *such that* $F_\infty(z_1^*) = F(z^*) = 0$, *and if* $\mathbf{D}(z^*, u) \neq 0$ *for some* $u \in U$, *then the assertion of Corollary 2 holds with* $y = x := (z^*, x_m)$ *for arbitrary* $x_m \in U$ *and arbitrary* $t > 0$.

5. Application to physiology

In this section we apply the above results to a random system based on the Hodgkin-Huxley model well known in physiology. This random system belongs to the family of SDEs with internal variables and random input presented in Section 4. We start by some reminders on the deterministic Hodgkin-Huxley model that we call (HH) for short.

5.1. **The deterministic (HH) system.** The deterministic Hodgkin-Huxley model for the membrane potential of a neuron (cf [20] and [25]) has been extensively studied over the last decades. There seems to be a large agreement that it models adequately many observations made on the response to an external input, in many types of neurons. This model belongs to the family of conductance-based models. It features two types of voltage-gated ion channels responsible for the import of Na^+ and export of K^+ ions through the membrane. The time dependent conductance of a sodium (resp. potassium) channel depends on the state of four gates which can be open or closed; it is maximal when all gates are open. There are two types of gates m and h for sodium, one type n for potassium. The variables n_t, m_t, h_t describe the probability that a gate of corresponding type be open at time t. Then, the Hodgkin-Huxley equations with deterministic input I which may be time dependent, is the $4D$ system

$$
\begin{aligned}
dV_t &= I(t)\,dt - \left[\overline{g}_{\mathrm{K}}\, n_t^4\,(V_t - E_{\mathrm{K}}) + \overline{g}_{\mathrm{Na}}\, m_t^3\, h_t\,(V_t - E_{\mathrm{Na}}) + \overline{g}_{\mathrm{L}}\,(V_t - E_{\mathrm{L}})\right] dt, \\
dn_t &= [\,\alpha_n(V_t)\,(1-n_t) - \beta_n(V_t)\,n_t\,]\,dt, \\
dm_t &= [\,\alpha_m(V_t)\,(1-m_t) - \beta_m(V_t)\,m_t\,]\,dt, \\
dh_t &= [\,\alpha_h(V_t)\,(1-h_t) - \beta_h(V_t)\,h_t\,]\,dt,
\end{aligned}
\tag{34}
$$

where we adopt the notations and constants of [25]. The functions $\alpha_n, \beta_n, \alpha_m, \beta_m, \alpha_h, \beta_h$ take values in $(0,\infty)$ and are analytic, i.e. they admit a power series representation on $I\!\!R$. They are given as follows:

$$
\begin{aligned}
\alpha_n(v) &= \tfrac{0.1-0.01v}{\exp(1-0.1v)-1}, & \beta_n(v) &= 0.125\exp(-v/80), \\
\alpha_m(v) &= \tfrac{2.5-0.1v}{\exp(2.5-0.1v)-1}, & \beta_m(v) &= 4\exp(-v/18), \\
\alpha_h(v) &= 0.07\exp(-v/20), & \beta_h(v) &= \tfrac{1}{\exp(3-0.1v)+1}.
\end{aligned}
\tag{35}
$$

Moreover if we set $a_n := \alpha_n + \beta_n$, $b_n := \alpha_n$ and analogously for m and h, we see that (HH) can be written as a particular case of (20) with F given by

$$
\begin{aligned}
F(v,n,m,h) &= -[\overline{g}_{\mathrm{K}}\, n^4\,(v - E_{\mathrm{K}}) + \overline{g}_{\mathrm{Na}}\, m^3\, h\,(v - E_{\mathrm{Na}}) + \overline{g}_{\mathrm{L}}\,(v - E_{\mathrm{L}})] \\
&= -[36\, n^4\,(v+12) + 120\, m^3\, h\,(v-120) + 0.3\,(v-10.6)]\,.
\end{aligned}
\tag{36}
$$

The parameter $\overline{g}_{\mathrm{Na}}$ (resp. $\overline{g}_{\mathrm{K}}$) is the maximal conductance of a sodium (resp. potassium) channel while $\overline{g}_{\mathrm{L}}$ is the leak conductance. The parameters E_{K}, E_{Na}, E_{L} are called reversal potentials. Their values $\overline{g}_{\mathrm{K}} = 36$, $\overline{g}_{\mathrm{Na}} = 120$, $\overline{g}_{\mathrm{L}} = 0.3$, $E_{\mathrm{K}} = -12$, $E_{\mathrm{Na}} = 120$, $E_{\mathrm{L}} = 10.6$ are those of [25].

If the variable V is kept constant at $v \in I\!\!R$, the variables n_t, m_t, h_t converge when $t \to +\infty$ respectively towards

$$
n_\infty(v) := \frac{\alpha_n}{\alpha_n + \beta_n}(v)\ ,\ m_\infty(v) := \frac{\alpha_m}{\alpha_m + \beta_m}(v)\ ,\ h_\infty(v) := \frac{\alpha_h}{\alpha_h + \beta_h}(v)\ .
\tag{37}
$$

The Hodgkin-Huxley system exhibits a broad range of possible and qualitatively quite different behaviors, depending on the specific input I. In response to a periodic input, the solution of (34) displays a periodic behavior (regular spiking of the neuron on a long time window) only in special situations. Let us first mention that there exists some interval U such that time-constant input in U results in periodic behavior for the solution of (34) (see [41]). For an oscillating input, there exists some interval J such that oscillating inputs with frequencies in J yield periodic behavior (see [1]). Periodic behavior includes that the period of the output can be a multiple of the period of the input. However, the input frequency has to be compatible with a range of preferred frequencies of (34), a fact which is similarly encountered in biological observations (see [25]). Indeed there are also intervals $\widetilde{I}$ and $\widetilde{J}$ for which time-constant input in $\widetilde{I}$ or oscillating input at frequency $f \in \widetilde{J}$ leads to chaotic behavior. Using numerical methods [16] gives a complete tableau.

5.2. **(34) with random input.** It has been shown in [38] that conductance-based models like (34) are fluid limits of a sequence of Piecewise Deterministic Markov Processes. Such limit theorems enable to study the impact of *channel noise* (also called *intrinsic noise*) on latency coding. Our setting is different. The noise here is external coming from the network in which the neuron is embedded, through its dendritic system. This system has a complicated topological structure and carries a large number of synapses which register spike trains emitted from a large number of other neurons within the same active network. We model the cumulated dendritic input as a diffusion of mean-reverting type carrying a deterministic signal S. The resulting system that we consider is the following particular case of (18):

$$
\begin{aligned}
dV_t &= d\xi_t - \left[\overline{g}_{\mathrm{K}}\, n_t^4\,(V_t - E_{\mathrm{K}}) + \overline{g}_{\mathrm{Na}}\, m_t^3\, h_t\,(V_t - E_{\mathrm{Na}}) + \overline{g}_{\mathrm{L}}\,(V_t - E_{\mathrm{L}})\right] dt\,, \\
dn_t &= [\,\alpha_n(V_t)\,(1-n_t) - \beta_n(V_t)\, n_t\,]\, dt\,, \\
dm_t &= [\,\alpha_m(V_t)\,(1-m_t) - \beta_m(V_t)\, m_t\,]\, dt\,, \\
dh_t &= [\,\alpha_h(V_t)\,(1-h_t) - \beta_h(V_t)\, h_t\,]\, dt\,, \\
d\xi_t &= (\,S(t) - \xi_t\,)\,\tau dt + \gamma\, q(\xi_t)\,\sqrt{\tau} dW_t\,,
\end{aligned} \tag{38}
$$

parametrized in terms of τ (governing speed) and γ (governing spread). For instance ξ can be of Ornstein-Uhlenbeck (OU) type (then $U = I\!R$, $q(\cdot) \equiv 1$) or of Cox-Ingersoll-Ross (CIR) type (then $U = (-K, \infty)$, $q(x) = \sqrt{(x+K) \vee 0}$ for $x \in U$, and K is chosen in $]\frac{\gamma^2}{2} + \sup|S|, +\infty[$). Such a choice builds on the statistical study [21]. When the deterministic signal S is periodic, it is shown in [22] that ξ of OU type admits a periodically invariant regime under which the signal $S(\cdot)$ is related to expectations of ξ via the formula $s \to E_{\pi,0}(\xi_s) = \int_0^\infty S(s - \frac{r}{\tau})e^{-r} dr$. In the companion papers [23] and [24] we address the periodic ergodicity of the solution to (38). Ergodicity properties when ξ is of OU type are the topic of [23]. The case of CIR is covered in [24] where also limit theorems are proved. Below we will conduct a numerical study of **(LWH)** for (38), based on Theorem 3. In this theorem, the specific nature of ξ plays no role in the definition of the determinant $\mathbf{D}$ provided that the SDE satisfied by ξ satisfies assumption **(H3)**, cf. Proposition 6 below. Therefore, the results of this numerical study apply to general random (HH) where we replace the last line in (38) by $d\xi_t = b_5(t, \xi_t)dt + \sigma(\xi_t)dW_t$.

5.3. **Weak Hörmander condition for (38).**

5.3.1. *The determinant* $\mathbf{\Delta}$. Applying Theorem 3 and Definition 1 we have to consider points where the $4D$ determinant, whose columns are the partial derivatives of the coefficients of (34) with respect to the first variable v from order one to order four, does not vanish. Since in this case the function F given in (36) is linear in v, we obtain that $\partial_v^{(k)} F = 0$ for $k \in \{2, 3, 4\}$. Moreover $\partial_v F(v, n, m, h) =$

$-(\overline{g}_{\mathrm{K}}\, n^4 + \overline{g}_{\mathrm{Na}}\, m^3\, h + \overline{g}_{\mathrm{L}})$ never vanishes on $[0,1]^3$. So actually in this case, it is sufficient to consider a $3D$ determinant extracted from $\mathbf{D}$.

Proposition 6. *Assume that σ remains strictly positive on U. Let us introduce the notation $d_n(v,n) := -a_n(v)n + b_n(v)$ and analogous ones for m and h. Then* **(LWH)** *for (38) is satisfied at any point $(t,v,n,m,h,\zeta) \in [0,\infty[\times I\!R \times (0,1)^3 \times U$ where $\Delta(v,n,m,h) \neq 0$ with*

$$\mathbf{\Delta}(v,n,m,h) \;:=\; \det \begin{pmatrix} \partial_v^{(2)} d_n & \partial_v^{(3)} d_n & \partial_v^{(4)} d_n \\ \partial_v^{(2)} d_m & \partial_v^{(3)} d_m & \partial_v^{(4)} d_m \\ \partial_v^{(2)} d_h & \partial_v^{(3)} d_h & \partial_v^{(4)} d_h \end{pmatrix}. \tag{39}$$

Proposition 7. *The set of points in $(v,n,m,h,\zeta) \in I\!R \times (0,1)^3 \times U$ where $\mathbf{\Delta}$ does not vanish has full Lebesgue measure.*

Proof. We say that a set has full Lebesgue measure if its complement has Lebesgue measure zero. Firstly it can be shown numerically that indeed there exists points (v,n,m,h,ζ) such that $\mathbf{\Delta}(v,n,m,h) \neq 0$ (see Section 5.4 below). Moreover, for any fixed $v \in I\!R$, the function $(n,m,h) \mapsto \mathbf{\Delta}(v,n,m,h)$ is a polynomial of degree three in the variables n,m,h. In particular, for any fixed v, either $\mathbf{\Delta}(v,.,.,.)$ vanishes identically on $(0,1)^3$, or its zeros form a two-dimensional sub-manifold of $(0,1)^3$. Finally, since $\mathbf{\Delta}$ is a sum of terms

$$(\text{some power series in } v) \cdot n^{\varepsilon_n} m^{\varepsilon_m} h^{\varepsilon_h}$$

with epsilons taking values 0 or 1, it is impossible to have small open v-intervals where it vanishes identically on $(0,1)^3$. We conclude the proof by integrating over v and using Fubini's Theorem. □

Although the condition $\mathbf{\Delta} \neq 0$ is only a sufficient condition ensuring that **(LWH)** is satisfied locally, it is convenient since it is possible to evaluate $\mathbf{\Delta}(v,n,m,h)$ numerically. This is done in section 5.4 below.

5.4. **Numerical study of the determinant $\mathbf{\Delta}$.** We compute numerically the value of $\mathbf{\Delta}$ at points of the form $(v, n_\infty(v), m_\infty(v), h_\infty(v))$ as in (37). The function $F_\infty(v) := F(v, n_\infty(v), m_\infty(v), h_\infty(v))$ is strictly increasing at least on an interval $\mathcal{I}$ containing $\mathcal{I}_0 = (-15, +30)$ hence it defines a bijection between the constant input $I(t) = c$ in (34) and the solution of the equation $F_\infty(v) = c$ that we denote by v_c. Therefore for any $v \in \mathcal{I}$, the point $(v, n_\infty(v), m_\infty(v), h_\infty(v))$ is the equilibrium point of (34) submitted to the constant input $c = F_\infty(v)$. We use this fact below since it may be more convenient to work with v than with c even if classically one considers c as the parameter of interest. For instance the point $(0, n_\infty(0), m_\infty(0), h_\infty(0))$ corresponds to $c = F(0, n_\infty(0), m_\infty(0), h_\infty(0)) \approx -0.0534$. We found that $\mathbf{\Delta}(0, n_\infty(0), m_\infty(0), h_\infty(0)) < 0$ and moreover the function $v \mapsto \mathbf{\Delta}\,(v, n_\infty(v), m_\infty(v), h_\infty(v))$ has exactly two zeros on the interval $\mathcal{I}_0 = (-15, +30)$ located at $v \approx -11.4796$ and $v \approx +10.3444$. As a consequence, for all values of c belonging to $]F_\infty(-10), F_\infty(+10)[=]-6.15, 26.61[$, the determinant $\mathbf{\Delta}(v_c, n_\infty(v_c), m_\infty(v_c), h_\infty(v_c))$ remains strictly negative.

5.5. **Positivity regions for (38).** In this section we apply the results of section 4.3 to (38). Remember that by comparing (18) and (38) we see that $\xi_t - \xi_0$ corresponds to $\int_0^t I(s)ds$.

Consider first suitable constant $I(t) \equiv c$, fix $\zeta \in U$ and $t > 0$, and consider

$$x_c := (v_c, n_\infty(v_c), m_\infty(v_c), h_\infty(v_c), \zeta) \quad , \quad x'_c := (v_c, n_\infty(v_c), m_\infty(v_c), h_\infty(v_c), \zeta + ct) \tag{40}$$

where v_c is the unique solution of $F(v_c, n_\infty(v_c), m_\infty(v_c), h_\infty(v_c)) = c$ (see section 5.4). Let us denote by $P_{s,t}(\cdot,\cdot)_{s<t}$ the semigroup of the process $(X_t)_{t\geq 0}$ which satisfies (38). Then Theorem 5 and Corollary 2 read as follows.

Proposition 8. *Assume that $\zeta + cs \in U$ for all $0 \le s \le t$. Consider x_c and x'_c defined in (40).*
1. Then for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x'' \in B_\delta(x_c)$, $P_{0,t}(x'', B_\varepsilon(x'_c)) > 0$.
2. If moreover $\mathbf{\Delta}(v_c, n_\infty(v_c), m_\infty(v_c), h_\infty(v_c)) \neq 0$, then there exists $\delta > 0$ such that for $K_c = \overline{B}_\delta(x_c)$ and $K'_c = \overline{B}_\delta(x'_c)$,

$$\inf_{x \in K_c} \inf_{x' \in K'_c} p_{0,t}(x, x') > 0.$$

Remember that the assumption $\mathbf{\Delta}(v_c, n_\infty(v_c), m_\infty(v_c), h_\infty(v_c)) \neq 0$ ensures that **(LWH)** holds both at x_c and x'_c. We have checked numerically in section 5.4 that this assumption is satisfied for $c \in]-6.15, 26.61[$. Hence for this range of c the above proposition applies.

6. Appendix

6.1. **Simple properties of (18). Proof of Proposition 2.** Given the trajectory of X_1, the variation of constants method yields

$$X_{i,t} = X_{i,0} e^{-\int_0^t a_i(X_{1,s})ds} + \int_0^t b_i(X_{1,u}) e^{-\int_u^t a_i(X_{1,r})dr} du. \tag{41}$$

However note that (41) does not provide an explicit formula for $X_{i,t}$ since X_1 depends on X_i (the system is fully coupled). Writing $\int_0^t b_i(X_{1,u}) e^{-\int_u^t a_i(X_{1,r})dr} du = \int_0^t \frac{b_i(X_{1,u})}{a_i(X_{1,u})} a_i(X_{1,u}) e^{-\int_u^t a_i(X_{1,u})dr} du$, the assumptions on $a_i(\cdot)$ and $b_i(\cdot)$ imply that

$$0 \le X_{i,t} \le X_{i,0} e^{-\int_0^t a_i(X_{1,s})ds} + \int_0^t a_i(X_{1,u}) e^{-\int_u^t a_i(X_{1,r})dr} du. \tag{42}$$

By straightforward integration it follows that

$$0 \le X_{i,t} \le (X_{i,0} + e^{\int_0^t a_i(X_{1,r})dr} - 1) e^{-\int_0^t a_i(X_{1,s})ds} = 1 + (X_{i,0} - 1) e^{-\int_0^t a_i(X_{1,s})ds}.$$

The statement follows. □

6.2. **Proof of (13).** We keep the notations introduced in the proof of Theorem 1 as well as in section 3.1. In order to establish (13) we extend the argument of [14], Theorem 2.3. To sum up this argument we can say that by an iterative procedure on the Sobolev norms of $\mathcal{H}(\bar{X}_{t-\delta,t}(y), \Phi(\bar{X}_{t-\delta,t}(y) - y_0))$ (in the sense of Malliavin calculus) of different indices, it is proved that estimating these Sobolev norms amounts to estimate the Sobolev norms of $\bar{X}$ and of the inverse of the Malliavin covariance matrix $(\Gamma_{\bar{X}_t})_{i,j} :=< D\bar{X}_{i,t}, D\bar{X}_{j,t} >_{L^2[0,t]}$, $1 \le i, j \le m$, where D denotes Malliavin derivative. Since by a classical identity this inverse can be written using the inverse of $det\ \Gamma_{\bar{X}_t}$ and the coefficients of Γ itself, the key ingredient is to estimate the Sobolev norms of $\bar{X}$ and expressions of the form $E_z \left(|det\ \Gamma_{\bar{X}_t}|^{-p} \right)^{1/p}$. We show below that no difficulty comes from the Sobolev norms of $\bar{X}$ and we prove that for any $p \ge 1$ and $t \le 1$, for any $N \in I\!N$ and z such that $\dim \mathrm{LA}(\mathcal{L}_N)(s, z) = m$, $\forall s \in [0, t]$,

$$E_z \left(|det\ \Gamma_{\bar{X}_t}|^{-p} \right)^{1/p} \le C(p, m, N, z)\ t^{-m(1+N)}. \tag{43}$$

Formula (43) is the main step to obtain (13). Indeed it suffices to apply it to the process $\bar{X}_{t-\delta,t}$ on an interval of length δ instead of t in (43). A particular version of (43) obtained by taking $N = 0$ is proved in [14] where the restriction to $N = 0$ is possible due to the fact that local ellipticity is assumed to hold. However local ellipticity fails to hold in our framework. This is why we prove the general version of (43).

We proceed in three steps. In the first step we check that the usual upper bound for the Sobolev norms of $\bar{X}$ is still valid and at the end of this step we obtain an expression of a key term of $\Gamma_{\bar{X}_t}$ that involves the successive Lie brackets introduced in section 3.1. The scheme of this argument is classical (cf. [30]) but we have to take care of the time dependence in the drift. We describe its main points for the sake of completeness. In the second step we prove (43) where N is the order of the successive Lie brackets that we need to generate $I\!R^m$ according to **(LWH)**. When local ellipticity holds, the diffusion coefficients themselves generate $I\!R^m$ and it is not necessary to compute Lie brackets ($N = 0$). Finally, in the third step, we show how the arguments of the proof of Theorem 2.3 of [14] allow to obtain (13) from (43), with $t = \delta$. Since $\delta \le 1$, when stating the following estimates, we will always concentrate on the case $t \le 1$.

Step 1. Let $\tilde{\bar{b}}_i(t,x) := \bar{b}_i(t,x) - \frac{1}{2}\sum_{k=1}^{l}\sum_{j=1}^{m} \bar{\sigma}_{j,k}(x)\frac{\partial \bar{\sigma}_{i,k}}{\partial x_j}(x),\ 1 \le i \le m$, be the Stratonovich drift for (7) and $\bar{A}_0 := \frac{\partial}{\partial t} + \tilde{\bar{b}}$, $\bar{A}_k := \bar{\sigma}_k, 1 \le k \le l$, the corresponding vector fields, where $\bar{\sigma}_k$ denotes the $k-$th column of the matrix $\bar{\sigma}$. Define $(Y_t)_{i,j} := \frac{\partial \bar{X}_{i,t}}{\partial x_j}, 1 \le i,j \le m$. Then Y satisfies the following linear SDE with bounded coefficients w.r.t. time and space,

$$Y_t = I_m + \int_0^t \partial\tilde{\bar{b}}(s,\bar{X}_s)Y_s ds + \sum_{k=1}^{l}\int_0^t \partial\bar{\sigma}_k(\bar{X}_s)Y_s \circ dW_s^k,$$

where I_m is the $m \times m-$unity matrix and $\partial\tilde{\bar{b}}$ and $\partial\bar{\sigma}_k$ are the $m \times m-$matrices having components $(\partial\tilde{\bar{b}})_{i,j}(t,x) = \frac{\partial \tilde{\bar{b}}_i}{\partial x_j}(t,x)$ and $(\partial\bar{\sigma}_k)_{i,j}(x) = \frac{\partial\bar{\sigma}_{i,k}}{\partial x_j}(x)$. In the above formula, $\circ dW_s$ denotes the Stratonovich integral. By means of Itô's formula, one shows that Y_t is invertible. Its inverse Z_t satisfies the linear SDE (again with bounded coefficients w.r.t. time and space) given by

$$Z_t = I_m - \int_0^t \partial\tilde{\bar{b}}(s,\bar{X}_s)Z_s ds - \sum_{k=1}^{l}\int_0^t \partial\bar{\sigma}_k(\bar{X}_s)Z_s \circ dW_s^k. \tag{44}$$

In this framework, the following estimates are classical (see e.g. [30]) and will be sufficient for our purpose. For all $0 \le s \le t \le T$, for all $p \ge 1$,

$$\sup_{s \le t} E(|(Z_s)_{i,j}|^p) \le C(T,p,m,\bar{b},\bar{\sigma}),\ 1 \le i,j \le m, \tag{45}$$

$$\sup_{r_1,\ldots,r_k \le t} E\left(|D_{r_1,\ldots,r_k}\bar{X}_{i,t}|^p\right) \le C(T,p,m,k,\bar{b},\bar{\sigma})t^{k/2}\left(t^{1/2}+1\right)^{(k+1)^2 p}, \tag{46}$$

where the constants $C(T,p,m,k,\bar{b},\bar{\sigma})$ depend only on the bounds of the space derivatives of $\bar{b}$ and $\bar{\sigma}$. Up to this point, the fact that the drift coefficient depends on time did not play an important role since all coefficients are bounded, uniformly in time.

Step 2. It is well known (see for example [36], page 110, formula (2.40)) that

$$\Gamma_{\bar{X}_t} = Y_t\left(\int_0^t Z_s\bar{\sigma}(\bar{X}_s)\bar{\sigma}^*(\bar{X}_s)Z_s^* ds\right)Y_t^*.$$

In order to prove (43) one has to evaluate the latter integral and therefore to control expressions of the form $Z_sV(s,\bar{X}_s)$, where $V(t,x)$ is a smooth function. This is done by iterating the formula,

$$\begin{aligned} Z_tV(t,\bar{X}_t) &= V(0,x) + \sum_{k=1}^{l}\int_0^t Z_s[\bar{\sigma}_k,V](s,\bar{X}_s)\circ dW_s^k + \int_0^t Z_s[\frac{\partial}{\partial t} + \tilde{\bar{b}},V](s,\bar{X}_s)ds \\ &= V(0,x) + \sum_{k=1}^{l}\int_0^t Z_s[\bar{A}_k,V](s,\bar{X}_s)\circ dW_s^k + \int_0^t Z_s[\bar{A}_0,V](s,\bar{X}_s)ds, \end{aligned} \tag{47}$$

starting with $V \equiv \bar{\sigma}_k$, for $1 \leq k \leq l$, where we identify functions with vector fields (cf. [36], formula (2.42)). Here, the fact that the drift coefficient is time dependent is important and gives rise to the extra term $\frac{\partial}{\partial t}$ within the second integral of the first line. In particular with $V \equiv \bar{\sigma}_\ell$ for a fixed $1 \leq \ell \leq k$, we obtain (cf. (1.9) of [8])

$$Z_t \bar{\sigma}_\ell(t, \bar{X}_t) = \bar{\sigma}_\ell(x) + \sum_{k=1, k \neq \ell}^{l} \int_0^t Z_s[\bar{A}_k, \bar{A}_\ell](s, \bar{X}_s) \circ dW_s^k + \int_0^t Z_s[\bar{A}_0, \bar{A}_\ell](s, \bar{X}_s) ds.$$

Iterating (47) we see that $Z_s \bar{\sigma}(\bar{X}_s)$ can be written as the sum of two terms. The first term is a finite sum of iterated Itô integrals where the integrands are $\bar{A}_k, 1 \leq k \leq l$, and the successive Lie brackets of order at most N obtained with $\bar{A}_k, 1 \leq k \leq l$ and $\bar{A}_0$. The second term is a remainder R_N (this is analogous to Theorem 2.12 of [30]). The most important feature is that the behavior of R_N depends only on the supremum norms of derivatives with respect to time and space of $\bar{b}$ and with respect to space of $\bar{\sigma}$. Based on (47), (43) follows by Theorem (2.17), estimate (2.18) of [30].

Step 3. Once (43) is established, (13) follows by a straightforward adaptation of the proof of Theorem 2.3 of [14], replacing the number of derivatives k there by $m(k+2)$ which is the number of derivatives to be considered in our context to handle ∂_β. For completeness let us note that (46) is the same bound as (2.17) in [14] whereas (43) plays the role of (2.20) in [14]. For t close to zero, the right-hand side of (2.20) in [14] is of order t^{-m} due to the local ellipticity condition, while our bound is of order $t^{-m(1+N)}$ due to our condition **(LWH)**. Plugging (43) and (46) in (2.25) of [14] (cf. the proof of (2.23)) replaces the r.h.s. obtained there for (2.23) by $O\left(t^{-(\frac{1}{2}+N)}\right)$ for small t. With such changes, the argument developed there goes through. In our framework we end up with $O\left(t^{-m\,(k+2)\,k_N}\right)$ as a control for (2.21) of [14], for small t as r.h.s., with some positive constant k_N depending on **(LWH)**. Here $k+2$ is the number of derivatives to be considered in our case. □

Acknowledgments

This work has been supported by the Agence Nationale de la Recherche through the project MANDy, Mathematical Analysis of Neuronal Dynamics, ANR-09-BLAN-0008-01 and is part of the project Labex MME-DII (ANR11-LBX-0023-01).

We thank Vlad Bally and Michel Benaïm for very stimulating discussions. We also thank two anonymous referees for helpful comments and suggestions.

References

[1] Aihara, K., Matsumoto, G., Ikegaya, Y.: Periodic and nonperiodic responses of a periodically forced Hodgkin-Huxley oscillator. J. Theoret. Biol. **109**, 249–269 (1984)

[2] Bally, V., Integration by parts formula for locally smooth laws and applications to equations with jumps I. Preprints Institut Mittag-Leffler, The Royal Swedish Academy of Sciences, 2007

[3] Bally, V., Clément, E. Integration by part formula and applications to stochastic equations with jumps. Probab. Theory. Rel. Fields **151**, 613–657 (2011)

[4] Ben Arous, G., Gradinaru, M., Ledoux, M., Hölder norms and the support theorem for diffusions. Annales de l'IHP **30**, 415-436 (1994)

[5] Benaïm, M., Le Borgne, S., Malrieu, F., Zitt, P.-A., Qualitative properties of certain piecewise deterministic Markov processes. Annales de l'IHP **51**, 1040–1075 (2015)

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

[7] Brodda, K., Höpfner, R., A stochastic model and a functional limit theorem for information processing in large systems of neurons. J. Math. Biol. **52**, 439-457 (2006)

[8] Cattiaux, P., Mesnager, L., Hypoelliptic non-homogeneous diffusions. Probab. Theory. Rel. Fields **123**, 453-483 (2002)
[9] Cattiaux, P., León, J., Prieur, C., Estimation for stochastic damping Hamiltonian systems under partial observations: I. Invariant density. Stoch. Proc. and their Appl. **124**, 1236–1260 (2014)
[10] Clark, J. M. C., An introduction to stochastic differential equations on manifolds. In: R. Brockett and D. Q. Mayne (edts), Geometric methods in Systems Theory, Reidel Pub. Comp., Dordrecht 1973, 131–149
[11] Crudu, A., Debussche, A., Muller, A., Radulescu, O., Convergence of stochastic gene networks to hybrid piecewise deterministic processes. The Annals of Applied Probability **22**, 1822–1859 (2012)
[12] Cuneo, N., Eckmann, J.-P., Poquet, C., Non-equilibrium steady state and subgeometric ergodicity for a chain of three coupled rotors. Nonlinearity **28**, 2397–2421 (2015)
[13] Davis, M., Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models. Journal of the Royal Statistical Society (B) **43**, 353–388 (1984)
[14] De Marco, S., Smoothness and Asymptotic Estimates of densities for SDEs with locally smooth coefficients and Applications to square-root diffusions. Annals of Applied Probability **21**, 1282–1321 (2011)
[15] Destexhe, A., Conductance-based integrate and fire models. Neural Comput. **9**, 503–514 (1997)
[16] Endler, K., Periodicities in the Hodgkin-Huxley model and versions of this model with stochastic input. Master Thesis, Institute of Mathematics, University of Mainz, 2012 (see under http://ubm.opus.hbz-nrw.de/volltexte/2012/3083/)
[17] Faggionato, A., Gabrielli, D., Ribezzi Crivellari, M., Non-equilibrium Thermodynamics of Piecewise Deterministic Markov Processes. J. Stat. Phys. **137**, 259–304 (2009)
[18] Fournier, N., Smoothness of the law of some one-dimensional jumping S.D.E.s with non-constant rate of jump. Electronic Journal of Probability **13**, 135–156 (2008)
[19] Goreac, D., Some topics in deterministic and stochastic control. Habilitation à diriger les recherches, Université Paris-Est Marne la Vallée, 2013 (see under https://tel.archives-ouvertes.fr/tel-00864555/document.)
[20] Hodgkin, A., Huxley, A., A quantitative description of ion currents and its applications to conduction and excitation in nerve membranes. J. Physiol. **117**, 500-544 (1952)
[21] Höpfner, R., On a set of data for the membrane potential in a neuron. Math. Biosci. **207**, 275–301 (2007)
[22] Höpfner, R., Kutoyants, Yu., Estimating discontinuous periodic signals in a time inhomogeneous diffusion. Statist. Inference Stoch. Proc. **13**, 193–230 (2010)
[23] Höpfner, R., Löcherbach, E., Thieullen, M., Ergodicity for stochastic Hodgkin-Huxley models driven by an Ornstein-Uhlenbeck process. Annales de l'IHP **52**, 483–501 (2016)
[24] Höpfner, R., Löcherbach, E., Thieullen, M., Ergodicity and limit theorems for degenerate diffusions with time periodic drift. Application to a stochastic Hodgkin-Huxley model. arXiv:1503.01648, 2015
[25] Izhikevich, E., Dynamical systems in neuroscience: the geometry of excitability and bursting. MIT Press 2009
[26] Ikeda, N., Watanabe, S., Stochastic differential equations and diffusion processes. North-Holland 1989
[27] Karatzas, J., Shreve, S., Brownian motion and stochastic calculus. Springer 1991
[28] Kunita, H., Lectures on Stochastic Flows And Applications. Delivered at the Indian Institute of Science, Bangalore, under the T.I.F.R.-I.I.Sc. programme in applications of mathematics. Notes by M. K. Ghosh. Lectures on Mathematics and Physics. Mathematics, 78. Tata Institute of Fundamental Research, Bombay. Springer 1986
[29] Kunita, H., Stochastic flows and stochastic differential equations. Cambridge University Press 1990
[30] Kusuoka, S., Stroock, D., Applications of the Malliavin calculus, Part II. J. Fac. Sci. Univ. Tokyo Sect. IA Math. **32**, 1–76 (1985)
[31] Meyn, S., P., Stability of Markovian Processes II: Continuous-Time Processes and Sampled Chains. Advances in Applied Probability **25**, 487–517 (1993)
[32] Meyn, S., P., Stability of Markovian Processes III: Foster-Lyapunov criteria for Continuous-Time Processes. Advances in Applied Probability **25**, 518–548 (1993)
[33] Millet, A., Sanz-Solé, M., A simple proof of the support theorem for diffusion processes. Séminaire de probabilités (Strasbourg), tome **28**, 36–48 (1994)
[34] Morris, C., Lecar, H., Voltage oscillations in the barnacle giant muscle fiber. Biophysics Journal **35**, 193-213 (1981)
[35] Nagel, A., Stein, E. M., Wainger, S., Balls and metrics defined by vector fields I: Basic properties. Acta Mathematica **155**, 103–147 (1985)
[36] Nualart, D., The Malliavin calculus and Related Topics. Springer 1995
[37] Nummelin, E., General irreducible Markov chains and non-negative operators. Cambridge University Press 1985
[38] Pakdaman, K., Thieullen, M., Wainrib, G., Fluid limit theorems for stochastic hybrid systems with application to neuron models. Adv. Appl. Probab. **42**, 761–794 (2010)

[39] Pankratova, E., Polovinkin, A., Mosekilde, E., Resonant activation in a stochastic Hodgkin-Huxley model: interplay between between noise and suprathreshold driving effects. Eur. Phys. J. B **45**, 391–397 (2005)
[40] Realpe-Gomez, J., Galla, T., McKane, A., J., Demographic noise and piecewise deterministic Markov processes. Physical Review E, **86** (2012)
[41] Rinzel, J., Miller, R., Numerical calculation of stable and unstable periodic solutions to the Hodgkin-Huxley equations. Mathem. Biosci. **49**, 27–59 (1980)
[42] Samson, A., Thieullen, M., Contrast estimator for completely or partially observed hypoelliptic diffusion. Stoch. Proc. and their Appl. **122** , 2521–2552 (2012).
[43] Stroock, D., Varadhan, S., On the support of diffusion processes with applications to the strong maximum principle. Proc. Sixth Berkeley Symp. Math. Statist. Prob. III, 333–359, Univ. California Press, Berkeley, 1972
[44] Taniguchi, S., Applications of Malliavin calculus to time-dependent systems of heat equations. Osaka J. Math., **22**, N. 2, 307-320 (1985)
[45] Yu, Y., Wang, W., Wang, J., Liu, F., Resonance-enhanced signal detection and transduction in the Hodgkin-Huxley neuronal systems. Physical Review E **63**, 021907 (2001)

R. Höpfner: Johannes Gutenberg-Universität Mainz

E-mail address: hoepfner@mathematik.uni-mainz.de

E. Löcherbach: Université de Cergy-Pontoise

E-mail address: eva.loecherbach@u-cergy.fr

M. Thieullen: Université Pierre et Marie Curie

E-mail address: michele.thieullen@upmc.fr