A PRACTICAL CRITERION FOR POSITIVITY OF TRANSITION DENSITIES

DAVID P. HERZOG AND JONATHAN C. MATTINGLY

Abstract. We establish a simple criterion for locating points where the transition density of a degenerate diffusion is strictly positive. Throughout, we assume that the diffusion satisfies a stochastic differential equation (SDE) on $\mathbf{R}^d$ with additive noise and polynomial drift. In this setting, we will see that it is often that case that local information of the flow, e.g. the Lie algebra generated by the vector fields defining the SDE at a point $x \in \mathbf{R}^d$, determines where the transition density is strictly positive. This is surprising in that positivity is a more global property of the diffusion. This work primarily builds on and combines the ideas of Ben Arous and Léandre [2] and Jurdjevic and Kupka [6].

1. Introduction

The goal of this paper is to develop an easily applicable framework for locating points where the probability density of a degenerate diffusion is strictly positive. We will focus on the setting where the diffusion satisfies a stochastic differential equation (SDE) on $\mathbf{R}^d$ where each component of the drift is a polynomial in the standard Euclidean coordinates and the noise is additive. Our methods reduce finding points of positivity to computing a certain collection of constant vector fields generated by taking iterated commutators of the vector fields defining the SDE. This is convenient since a similar computation is typically used to show that the diffusion has a smooth probability density function $p_t(x, y)$ with respect to Lebesgue measure dy. While the existence of a smooth density is decided locally, we show that in some settings the bracket computation also determines the more global property of where the density is strictly positive. Additionally, uncovering sufficiently large regions of positivity is useful for proving unique ergodicity.

While methods already exist for proving positivity of transition densities, most require knowledge of attainable sets via controls. Here we have structured our assumptions to require as little global control information as possible. In particular, our results prove smoothness of the densities, the needed control statements, and positivity, all with one set of primarily local assumptions.

Although our general framework is limited to SDEs with polynomial drift and additive noise, working within such boundaries is reasonable in many applications. In particular, to illustrate the utility of our results, we will apply them to a collection of examples, each with quite different structure. Moreover, for the equations considered, either new results will be obtained or existing results will be improved upon.

The ideas used in this note build on a number existing works. Beyond the now classical theory of Hörmander [4] on hypoelliptic operators in the "sum of squares" form, we use the associated probabilistic techniques of Malliavin calculus

[12]. We also use a number of ideas from geometric control theory [7]. Moreover, we modify the idea that odd powered polynomial vector fields are "good" (due to their time reversal properties) and even powered polynomial vector fields are "bad" [6]. Similar ideas were critical in the work of Romito [14]. We also integrate into our results the powerful ideas of Ben Arous and Léandre [2] for proving positivity of densities of random variables over a Wiener space. Our hope is that by bringing these ideas together and adapting them to our specific context, we will provide a useful tool for many applied equations.

The layout of this paper is as follows. In Section 2, we introduce notation and terminology and state the main general results of the paper. In Section 3, we apply our results to specific examples. Section 4 contains heuristic discussions of why the main results hold and are natural. We also include an "non-example", that is an example where the main results fail to apply yet the corresponding density has regions of positivity (in space and time), and illustrate how to adapt the general theory in such cases. Additionally, Section 4 contains the proof of the main results as stated in Section 2.

Acknowledgements

The authors would like to thank Avanti Athreya, Richard Durrett, Tiffany Kolba, James Nolen, and Jan Wehr for helpful conversations on the topic of this paper. DPH would also like to thank Martin Hairer for suggesting the paper [6], from which his understanding of these ideas began and lead to the current collaboration. We would also like to acknowledge partial support of the NSF through grant DMS-08-54879 and the Duke University Dean's office.

2. Notation, Terminology and Main Results

Throughout, we study stochastic differential equations on $\mathbf{R}^d$ of the following form

$$dx_t \quad = \quad X_0(x_t)\,dt + \sum_{j=1}^{r} X_j\,dW_t^j \tag{2.1}$$

where X_0 is a *polynomial vector field*; that is, $X_0 = \sum_{j=1}^d X_0^j(x)\partial_{x_j}$ is such that each map $x \mapsto X_0^j(x)$ is a polynomial in the standard Euclidean coordinates, $X_1, \dots, X_r$ are constant vector fields; that is, they do not depend on the base point, and $W_t^1, W_t^2, \dots, W_t^r$ are standard independent real Wiener processes defined on a probability space $(\Omega, \mathscr{F}, \mathbf{P})$.

To deal with the issue of finite-time explosion in (2.1), we will need to stop the process x_t prior to the time of explosion. Thus for $n \in \mathbf{N}$, let $B_n(0)$ denote the open ball of radius n centered at the origin in $\mathbf{R}^d$, and define the stopping times $\tau_n = \inf\{t > 0 : x_t \notin B_n(0)\}$ and $\tau_\infty = \lim_{n\uparrow\infty} \tau_n$. Our results will be stated for the stopped processes $x_{t\wedge\tau_n}$, $n \in \mathbf{N}$. Of course, $x_{t\wedge\tau_n}$ coincides with x_t for all times $t \leq \tau_n$.

For vector fields $V = \sum_{j=1}^d V^j(x)\frac{\partial}{\partial x_j}$ and $W = \sum_{j=1}^d W^j(x)\frac{\partial}{\partial x_j}$, let $\mathrm{ad}^0 V(W) = W$,

$$\mathrm{ad}^1 V(W) = [V, W] := \sum_{j=1}^d \left(\sum_{k=1}^d V^k(x)\frac{\partial W^j(x)}{\partial x_k} - W^k(x)\frac{\partial V^j(x)}{\partial x_k} \right)\frac{\partial}{\partial x_j}.$$

Inductively, for $m \geq 2$ we let $\text{ad}^m V(W) = \text{ad}V\text{ad}^{m-1}V(W)$. For a set of vector fields $\mathcal{G}$ on $\mathbf{R}^d$, span($\mathcal{G}$) denotes the $\mathbf{R}$-linear span of $\mathcal{G}$ and

$$\text{cone}_{\geq 0}(\mathcal{G}) = \{\textstyle\sum_{i=1}^{j} \lambda_i V_i \,:\, \lambda_i \geq 0,\, V_i \in \mathcal{G}\}.$$

We call $x \in \mathbf{R}^d$ an *equilibrium point* of a set of vector fields $\mathcal{G}$ if $V(x) = 0$ for some $V \in \mathcal{G}$. If V is a constant vector field with constant value $v \in \mathbf{R}^d$ and W is a polynomial vector field, then we may define a map from $\mathbf{R}$ into $\mathbf{R}^d$ given by $\lambda \mapsto (W^j(\lambda v))$. Note that since W is a polynomial vector field, $(W^j(\lambda v))$ is a vector of polynomials in λ. Let $n(V, W)$ be the maximal degree among these polynomials (For purposes below, we assume that the zero polynomial has neither even nor odd degree). We call $n(V, W)$ the *relative degree* of V and W.

We now introduce the set of constant vector fields $\mathcal{C}$ which will play a fundamental role throughout the paper. It will be defined as the subset of constant vector fields in a larger set of vector fields which we now introduce. To initialize the inductive procedure let $\mathcal{G}_0 = \text{span}\{X_1, \dots, X_r\}$ and

$$\begin{aligned}
\mathcal{G}_1^{\text{o}} &= \mathcal{G}_0 \cup \{\text{ad}^{n(V,X_0)}V(X_0) \,:\, V \in \mathcal{G}_0, n(V, X_0) \text{ odd}\},\\
\mathcal{G}_1^{\text{e}} &= \{\text{ad}^{n(V,X_0)}V(X_0) \,:\, V \in \mathcal{G}_0,\, n(V, X_0) \text{ even}\},\\
\mathcal{G}_1 &= \text{span}(\mathcal{G}_1^{\text{o}}) + \text{cone}_{\geq 0}(\mathcal{G}_1^{\text{e}}).
\end{aligned}$$

For $j \geq 1$, we define $\mathcal{G}_{j+1}^{\text{o}}$, $\mathcal{G}_{j+1}^{\text{e}}$, $\mathcal{G}_{j+1}$ inductively as

$$\begin{aligned}
\mathcal{G}_{j+1}^{\text{o}} &= \mathcal{G}_j^{\text{o}} \cup \{\text{ad}^{n(V,W)}V(W) \,:\, V \in \mathcal{G}_j^{\text{o}} \text{ constant},\, W \in \mathcal{G}_j,\, n(V, W) \text{ odd}\},\\
\mathcal{G}_{j+1}^{\text{e}} &= \mathcal{G}_j^{\text{e}} \cup \{\text{ad}^{n(V,W)}V(W) \,:\, V \in \mathcal{G}_j^{\text{o}} \text{ constant},\, W \in \mathcal{G}_j,\, n(V, W) \text{ even}\},\\
\mathcal{G}_{j+1} &= \text{span}(\mathcal{G}_j^{\text{o}}) + \text{cone}_{\geq 0}(\mathcal{G}_j^{\text{e}}).
\end{aligned}$$

Let $\mathcal{C}^{\text{o}}$ denote the set of constant vector fields in $\cup_j \mathcal{G}_j^{\text{o}}$ and $\mathcal{C}^{\text{e}}$ denote the set of constant vector fields in $\cup_j \mathcal{G}_j^{\text{e}}$. Finally, define

$$\mathcal{C} = \text{span}(\mathcal{C}^{\text{o}}) + \text{cone}_{\geq 0}(\mathcal{C}^{\text{e}}). \tag{2.2}$$

Remark 2.3. Throughout, we will often identify a constant vector field on $\mathbf{R}^d$ with the vector in $\mathbf{R}^d$ which defines it. For example, depending on the context, $\mathcal{C}^{\text{o}}$ will be used to denote either the set of vector fields $\mathcal{C}^{\text{o}}$ defined above or the set of vectors $v \in \mathbf{R}^d$ such that $v = V(x)$ for some $V \in \mathcal{C}^{\text{o}}$.

Remark 2.4. The primary assumption we will make is that $\mathcal{C}$ is d-dimensional. This is equivalent to assuming that $\mathcal{C}$ spans the entire tangent space at all points $x \in \mathbf{R}^d$ as $\mathcal{C}$ contains only constant vector fields. Since $\mathcal{C}$ is contained in the Lie algebra generated by

$$X_1, \dots, X_r, [X_1, X_0], \dots, [X_r, X_0],$$

it follows by Hörmander's hypoellipticity theorem [4] that for every $n \geq 1$, $x \in B_n(0)$ and every Borel set $A \subset B_n(0)$

$$\mathbf{P}_x\{x_{t\wedge\tau_n} \in A\} = \int_A p_t^n(x, y)\, dy$$

for some nonnegative function $p_t^n(x, y)$ which is defined and smooth on $(0, \infty) \times B_n(0) \times B_n(0)$. Here we recall that $B_n(0)$ is the open ball of radius n centered at the origin in $\mathbf{R}^d$. Certainly, the transition kernel of $x_{t\wedge\tau_n}$ contains a singular component concentrated on the boundary of $B_n(0)$. However, this is invisible to sets contained in $B_n(0)$ since $B_n(0)$ is open.

We now state the main general result of the paper.

Theorem 2.5. *Suppose that $\mathcal{C}$ is d-dimensional and let $\{y_1, \ldots, y_d\} \subset \mathcal{C}$ be a basis of $\mathcal{C}$ such that $\{y_1, \ldots, y_k\} \subset \mathcal{C}^{\mathrm{o}}$ and $\{y_{k+1}, \ldots, y_d\} \subset \mathcal{C}^{\mathrm{e}}$. For $x \in \mathbf{R}^d$, define the set*

$$\mathcal{D}(x) = \big\{x\big\} + \big\{\textstyle\sum_{i=1}^k \alpha_i y_i + \sum_{j=k+1}^d \lambda_j y_j \,:\, \alpha_i \in \mathbf{R},\, \lambda_j > 0\big\}.$$

and suppose that $x, z \in \mathbf{R}^d$ are such that $z \in \mathcal{D}(x)$.

(a) For all $T > 0$ there exist $t \in (0, T)$ and $N \in \mathbf{N}$ such that

$$p_t^n(x, z) > 0 \quad \textit{for all} \quad n \geq N.$$

(b) If there exists an equilibrium point $y \in \mathbf{R}^d$ of $\mathcal{G} = \{X_0 + \sum_{j=1}^r u_j X_j \,:\, u_j \in \mathbf{R}\}$ such that $y \in \mathcal{D}(x)$ and $z \in \mathcal{D}(y)$, then for all $T > 0$ there exists $N \in \mathbf{N}$ such that

$$p_t^n(x, z) > 0 \quad \textit{for all} \quad t \geq T,\, n \geq N.$$

Remark 2.6. Suppose that $\mathcal{C}$ is d-dimensional and that x_t is *non-explosive*; that is, for every $x \in \mathbf{R}^d$

$$\mathbf{P}_x\{\tau_\infty < \infty\} = 0.$$

Then x_t has a probability density function $p_t(x, y)$ with respect to Lebesgue measure dy which is smooth on $(0, \infty) \times \mathbf{R}^d \times \mathbf{R}^d$. Moreover, all conclusions of Theorem 2.5 hold with $p_t^n(x, z)$ replaced by $p_t(x, z)$.

Remark 2.7. Even if $\mathcal{C}$ is d-dimensional, it is still possible that the set $\mathcal{D}(x)$ cannot be chosen to be the entire space $\mathbf{R}^d$. See Example 3.4 in Section 3.

Remark 2.8. It is worth emphasizing that $y \in \mathbf{R}^d$ can be an equilibrium without being an equilibrium point of the drift vector field X_0. For example, if $X_0(y_1, y_2) = (g(y_1, y_2)(1 - y_2), f(y_2, y_1))$ for some scalar functions f, g and $X_1 = (0, 1)$ then all points of the form $(y_1, 1)$ are equilibrium points since $X(y_1, 1) + uX_1 = (0, 0)$ if $u = -f(y_1, 1)$.

Using the results of Theorem 2.5, we will also show:

Theorem 2.9. *Suppose that $\mathcal{C}$ is d-dimensional and x_t is non-explosive. Let $\mathcal{D}(x)$ be as in the statement of Theorem 2.5. Then there is at most one invariant probability measure corresponding to the Markov process x_t defined by (2.1). Moreover, if such an invariant probability measure μ exists, then $\mu(dx) = m(x)\,dx$ for some smooth, non-negative function m and if $x \in \operatorname{supp}(\mu)$ then for all $z \in \mathcal{D}(x)$, $m(z) > 0$.*

3. Examples

Before proving the main results, we apply them to specific examples to show their utility. A "non-example", that is an example where Theorem 2.5 is not applicable, is given in the next section in Remark 4.11 as it fits in better with the discussion there.

Example 3.1. As a first example, we consider the Langevin dynamics on $\mathbf{R}^{2d}$, $d \geq 1$,

$$\begin{aligned} dx_t &= [-\gamma x_t - \nabla F(y_t)]\, dt + \sum_{j=1}^{d} \sigma_j \, dW_t^j \\ dy_t &= x_t \, dt \end{aligned} \tag{3.2}$$

where $x_t, y_t \in \mathbf{R}^d$, $\gamma > 0$ is a constant, $F \in C^\infty(\mathbf{R}^d : \mathbf{R})$, $\sigma_j \in \mathbf{R}^d$ and the W_t^j are independent standard Wiener processes. So that solutions to (3.2) do not explode in finite time, we assume that F satisfies the one-sided Lipschitz condition and concavity and growth assumptions of Condition 3.1 of [9]. A prototypic example of a potential which satisfies these assumptions is $F(y) = \frac{1}{4}|y|^4 - \frac{1}{2}|y|^2$.

As a consequence of Theorem 2.5, we now prove:

Corollary 3.3. *If* $\operatorname{span}\{\sigma_1, \ldots, \sigma_d\} = \mathbf{R}^d$, *then for all* $(x, y), (x', y') \in \mathbf{R}^{2d}$ *and* $t > 0$

$$p_t((x, y), (x', y')) > 0.$$

Proof. Let $\mathbf{0} = (0, 0, \ldots, 0) \in \mathbf{R}^d$ and let $\mathcal{G} = \{X_0 + \sum_{j=1}^d u_j X_j \, : \, u_j \in \mathbf{R}\}$ where

$$X_0(x, y) = \begin{pmatrix} -\gamma x + \nabla F(y) \\ x \end{pmatrix} \qquad \text{and} \qquad X_j(x, y) = \begin{pmatrix} \sigma_j \\ \mathbf{0} \end{pmatrix}.$$

We begin by computing $\mathcal{C}$ (defined in the introduction) corresponding to equation (3.2). Since $n(X_0, X_j) = 1$ for all j, we see that

$$\mathcal{G}_1^{\circ} \supset \{[X_j, X_0] \, : \, j = 1, 2, \ldots, d\}$$

and

$$[X_j, X_0](x, y) = \begin{pmatrix} -\gamma \sigma_j \\ \sigma_j \end{pmatrix}.$$

Hence, in particular, $\mathcal{C} \supset \{X_j, [X_j, X_0] \, : \, j = 1, 2, \ldots, d\}$. Since the vectors $\sigma_1, \ldots, \sigma_d$ are linearly independent, it follows that $\mathcal{C}$ has a basis. Additionally, since $\mathcal{C}^{\circ} \supset \{X_j, [X_j, X_0] \, : \, j = 1, 2, \ldots, d\}$ we can choose a basis so that $\mathcal{D}(x, y) = \mathbf{R}^{2d}$ for all $(x, y) \in \mathbf{R}^{2d}$. To finish proving the result, we claim that the origin $(\mathbf{0}, \mathbf{0}) \in \mathbf{R}^{2d}$ is an equilibrium point of $\mathcal{G}$. Indeed, since

$$X_0(\mathbf{0}, \mathbf{0}) + \sum_{j=1}^{d} u_j X_j(\mathbf{0}, \mathbf{0}) = \begin{pmatrix} -\nabla F(\mathbf{0}) \\ \mathbf{0} \end{pmatrix} + \begin{pmatrix} \sum_{j=1}^d u_j \sigma_j \\ \mathbf{0} \end{pmatrix}$$

and the σ_j form a basis of $\mathbf{R}^d$, we may choose real numbers $u_j \in \mathbf{R}$ such that

$$X_0(\mathbf{0}, \mathbf{0}) + \sum_{j=1}^{d} u_j X_j(\mathbf{0}, \mathbf{0}) = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}.$$

In light of Remark 2.6, applying Theorem 2.5 (b) finishes the proof of Corollary 3.3. □

Example 3.4. Let $a_1, a_2 \in \mathbf{R}$, $\alpha_2 > \alpha_1 > 0$, and $\epsilon > 0$. With motivations from turbulent transport of inertial particles, the stochastic differential equation on $\mathbf{R}^2$

given by

$$dx_t = (a_1 x_t - \alpha_1 x_t^2 + y_t^2)\,dt$$
$$dy_t = (a_2 y_t - \alpha_2 x_t y_t)\,dt + \epsilon\, dW_t^2 \tag{3.5}$$

is considered in [3]. Here, we strengthen the results of Section 4 of this work. A more hands on application of some of the ideas of this note were applied to a specific case of this example in Section 11 of [1]. Applying Theorem 2.1 of [3], we first note that (x_t, y_t) is non-explosive.

We now prove:

Corollary 3.6. *Suppose that* $(x, y) \in \mathbf{R}^2$ *satisfies*

$$x < \frac{a_1 - |a_1|}{2\alpha_1} \quad \text{or} \quad x \geq \frac{a_1 + |a_1|}{2\alpha_1}.$$

Then for all $t > 0$ *and* $(x', y') \in \mathbf{R}^2$ *with* $x' > x$

$$p_t((x, y), (x', y')) > 0.$$

Otherwise if $(x, y) \in \mathbf{R}^2$ *satisfies*

$$\frac{a_1 - |a_1|}{2\alpha_1} \leq x \leq \frac{a_1 + |a_1|}{2\alpha_1},$$

then for all $t > 0$ *and* $(x', y') \in \mathbf{R}^2$ *with* $x' > \frac{a_1 + |a_1|}{2\alpha_1}$

$$p_t((x, y), (x', y')) > 0.$$

Remark 3.7. It is important to point out that Corollary 3.6 is not sharp. For example if $a_1 = a_2 = 0$, $\alpha_1 = 1$ and $\alpha_2 = 2$, it was shown in Section 11 of [1] that, in addition to the result above, for all $(x, y), (x', y') \in \mathbf{R}^2$ with $x' > 0$

$$p_t((x, y), (x', y')) > 0$$

for all $t > 0$ sufficiently large. The weakness of our result is due to the fact that Theorem 2.5 does not fully exploit the flow along X_0 in favor of making general statements for any positive time. However, Corollary 3.6 is more than sufficient to prove unique ergodicity in equation (3.5). Nevertheless, it is not hard to bootstrap from Corollary 3.6 to obtain the full (sharp) result proved in [1].

Proof. As in the previous example, we begin by computing the set $\mathcal{C}$ corresponding to equation (3.5). Let

$$\mathcal{G} = \left\{ X_0 + uX_1 \,:\, u \in \mathbf{R} \right\}$$

where $X_0 = (a_1 x - \alpha_1 x^2 + y^2)\partial_x + (a_2 y - \alpha_2 xy)\partial_y$ and $X_1 = \partial_y$. Since $n(X_0, X_1) = 2$, we find that $\text{ad}^2 X_1(X_0) = 2\partial_x \in \mathcal{G}_1^{\text{e}}$. Let

$$\mathcal{D}(x, y) = \{(x, y) + u(0, 1) + \lambda(1, 0) \,:\, u \in \mathbf{R},\, \lambda > 0\}.$$

As opposed to the previous example, the set $\mathcal{D}(x, y)$ is not the entire space. Hence we must make sure we have enough equilibrium points in the right locations.

Consider the polynomial equation

$$a_1 x - \alpha_1 x^2 + y^2 = 0$$
$$a_2 y - \alpha_2 xy + u = 0$$

where $u \in \mathbf{R}$. Clearly, any pair $(x, y) \in \mathbf{R}^2$ satisfying the above equations for some $u \in \mathbf{R}$ is an equilibrium point of $\mathcal{G}$. In particular, we may solve $a_1 x - \alpha_1 x^2 + y^2 = 0$ producing

$$x = \frac{a_1 \pm \sqrt{a_1^2 + 4\alpha_1 y^2}}{2\alpha_1}.$$

Since we may pick $u = \alpha_2 xy - a_2 y$, we therefore deduce that all points $(x, y) \in \mathbf{R}^2$ such that either

$$x \geq \frac{a_1 + |a_1|}{2\alpha_1} \qquad \text{or} \qquad x \leq \frac{a_1 - |a_1|}{2\alpha_1}$$

are equilibrium points for the control system $\mathcal{G}$. Hence Remark 2.6 now implies Corollary 3.6. □

Example 3.8. Let $\nu > 0$ be a constant. We now study Galerkin truncations of the following randomly forced two-dimensional viscous Burgers' equation

$$\partial_t u(\mathbf{x}, t) + (u(\mathbf{x}, t) \cdot \nabla_{\mathbf{x}}) u(\mathbf{x}, t) = \nu \Delta_{\mathbf{x}} u(\mathbf{x}, t) + \xi(\mathbf{x}, t) \tag{3.9}$$

with periodic boundary conditions on the torus $\mathbb{T}^2 = [0, 2\pi]^2$. Here, we assume that there is no mean flow and that ξ is a Gaussian process which is white in time and colored in space. To emphasize, we do not require the divergence free condition $\nabla \cdot u = 0$; hence, (3.9) is not the 2D Navier Stokes equation. Moreover, we do not restrict ourselves to gradient solutions as is often done when considering the multidimensional Burgers equation. In the dynamics (3.9), we are precisely interested how the divergence free forcing spreads to the non-divergence free (gradiant-like directions). Since one does not have global solutions in this setting, here we must make use of the stopped processes.

Let us now be more precise. Writing

$$\sum_{0 \neq \mathbf{k} \in \mathbf{Z}^2} u_{\mathbf{k}}(t) e^{-i\langle \mathbf{k}, \mathbf{x} \rangle},$$

where $\langle \cdot, \cdot \rangle$ denotes the dot product, and fixing a positive integer $N \geq 2$, we consider the following stochastic differential equation on $\mathbf{C}^{2((2N+1)^2-1)}$

$$\begin{aligned} du_{\mathbf{k}} \quad = \quad & [iF_{\mathbf{k}}^N(u) - \nu |\mathbf{k}|^2 u_{\mathbf{k}}]\, dt + \frac{\mathbf{k}^{\perp}}{|\mathbf{k}|^2}(\sigma_{\mathbf{k}}\, dB_t^{\mathbf{k},(1)} + i\sigma'_{\mathbf{k}}\, dB_t^{\mathbf{k},(2)}) \\ & + \frac{\mathbf{k}}{|\mathbf{k}|^2}(\gamma_{\mathbf{k}}\, dW_t^{\mathbf{k},(1)} + i\gamma'_{\mathbf{k}}\, dW_t^{\mathbf{k},(2)}) \end{aligned} \tag{3.10}$$

where

- $u_{\mathbf{k}} \in \mathbf{C}^2$;
- the equation is over all indices $\mathbf{k} \in H_N = \big\{\mathbf{k} \in \mathbf{Z}^2 \setminus \{(0,0)\} \,:\, \|\mathbf{k}\|_\infty \leq N\big\}$;
- $$F_{\mathbf{k}}^N(u) = \sum_{\mathbf{l}, \mathbf{k}-\mathbf{l} \in H_N} \langle u_{\mathbf{l}}, \mathbf{k} - \mathbf{l} \rangle u_{\mathbf{k}-\mathbf{l}};$$
- $\sigma_{\mathbf{k}}, \sigma'_{\mathbf{k}}, \gamma_{\mathbf{k}}, \gamma'_{\mathbf{k}} \in \mathbf{R}$;
- $\mathbf{k}^{\perp} = (k_1, k_2)^{\perp} = (-k_2, k_1)$;
- $\{B_t^{\mathbf{k},(1)}, B_t^{\mathbf{k},(2)}, W_t^{\mathbf{k},(1)}, W_t^{\mathbf{k},(2)}\}_{\mathbf{k} \in H_N}$ is a set of independent Brownian motions.

To further illuminate the discussion, we first split the equation into incompressible and compressible directions. To this end, write

$$\begin{aligned} u_{\mathbf{k}} &= w_{\mathbf{k}} \frac{\mathbf{k}^\perp}{|\mathbf{k}|^2} + q_{\mathbf{k}} \frac{\mathbf{k}}{|\mathbf{k}|^2} \\ F_{\mathbf{k}}(u) &= F_{\mathbf{k}}^\perp(w,q) \frac{\mathbf{k}^\perp}{|\mathbf{k}|^2} + F_{\mathbf{k}}^{\|}(w,q) \frac{\mathbf{k}}{|\mathbf{k}|^2} \end{aligned}$$

where $w_{\mathbf{k}}, q_{\mathbf{k}} \in \mathbf{C}$. In particular, equation (3.10) now becomes

$$\begin{aligned} dw_{\mathbf{k}} &= [-\nu|\mathbf{k}|^2 w_{\mathbf{k}} + iF_{\mathbf{k}}^\perp(w,q)]\, dt + \sigma_{\mathbf{k}}\, dB_t^{\mathbf{k},(1)} + i\sigma'_{\mathbf{k}}\, dB_t^{\mathbf{k},(2)} \\ dq_{\mathbf{k}} &= [-\nu|\mathbf{k}|^2 q_{\mathbf{k}} + iF_{\mathbf{k}}^{\|}(w,q)]\, dt + \gamma_{\mathbf{k}}\, dW_t^{\mathbf{k},(1)} + i\gamma'_{\mathbf{k}}\, dW_t^{\mathbf{k},(2)} \end{aligned} \tag{3.11}$$

for some $F_{\mathbf{k}}^\perp$, $F_{\mathbf{k}}^{\|}$ to be computed in a moment. Note that (3.11) evolves on $\mathbf{C}^{2((2N+1)^2-1)} = \mathbf{C}^{8N(N+1)}$ for all $t < \tau_\infty$.

We will now use Theorem 2.5 to prove the following result:

Theorem 3.12. *Suppose that*

$$\{\mathbf{k} \in H_N \,:\, \sigma_{\mathbf{k}} \neq 0, \sigma'_{\mathbf{k}} \neq 0\} \supset \{\mathbf{k} \in H_N \,:\, \|\mathbf{k}\|_\infty = 1\}.$$

Then for all $(w,q), (w',q') \in \mathbf{C}^{8N(N+1)}$ *and* $T > 0$, *there exists* $N \in \mathbf{N}$ *large enough so that*

$$p_t^n((w,q),(w',q')) > 0 \quad \textit{for all } t \geq T,\, n \geq N.$$

Remark 3.13. It is interesting to note that, even if the process (w_t, q_t) is assumed to be incompressible initially; that is, $(w_0, q_0) = (w, 0) \in \mathbf{C}^{8N(N+1)}$, a small amount of low mode forcing ensures that any mixture of incompressible and compressible states becomes instantaneously possible. As we will see in the proof below, this cannot happen if we do not force the incompressible directions. In particular, if we assume that the process (w_t, q_t) is initially compressible; that is, $(w_0, q_0) = (0, q)$ and $\sigma_{\mathbf{k}} = \sigma'_{\mathbf{k}} = 0$ for all $\mathbf{k} \in H_N$, then $w_t \equiv 0$ for all $t \geq 0$.

Proof of Theorem 3.12. We will first write out and symmetrize the nonlinear terms $F_{\mathbf{k}}^\perp$ and $F_{\mathbf{k}}^{\|}$. Using the relations $\langle \mathbf{k}^\perp, \mathbf{l}\rangle = -\langle \mathbf{k}, \mathbf{l}^\perp\rangle$ and $\langle \mathbf{k}^\perp, \mathbf{l}^\perp\rangle = \langle \mathbf{k}, \mathbf{l}\rangle$, we find that

$$\begin{aligned} F_{\mathbf{k}}^\perp(w,q) &= \sum_{\mathbf{l},\, \mathbf{k}-\mathbf{l} \in H_N} w_{\mathbf{l}} w_{\mathbf{k}-\mathbf{l}} \frac{\langle \mathbf{l}^\perp, \mathbf{k}\rangle \langle \mathbf{k}-\mathbf{l}, \mathbf{k}\rangle}{|\mathbf{l}|^2 |\mathbf{k}-\mathbf{l}|^2} + w_{\mathbf{l}} q_{\mathbf{k}-\mathbf{l}} \frac{\langle \mathbf{l}^\perp, \mathbf{k}\rangle^2}{|\mathbf{l}|^2 |\mathbf{k}-\mathbf{l}|^2} \\ &\quad + \sum_{\mathbf{l},\, \mathbf{k}-\mathbf{l} \in H_N} q_{\mathbf{l}} w_{\mathbf{k}-\mathbf{l}} \frac{\langle \mathbf{l}, \mathbf{k}-\mathbf{l}\rangle \langle \mathbf{k}-\mathbf{l}, \mathbf{k}\rangle}{|\mathbf{l}|^2 |\mathbf{k}-\mathbf{l}|^2} - q_{\mathbf{l}} q_{\mathbf{k}-\mathbf{l}} \frac{\langle \mathbf{l}, \mathbf{k}-\mathbf{l}\rangle \langle \mathbf{l}, \mathbf{k}^\perp\rangle}{|\mathbf{l}|^2 |\mathbf{k}-\mathbf{l}|^2} \end{aligned}$$

and

$$\begin{aligned} F_{\mathbf{k}}^{\|}(w,q) &= \sum_{\mathbf{l},\, \mathbf{k}-\mathbf{l} \in H_N} -w_{\mathbf{l}} w_{\mathbf{k}-\mathbf{l}} \frac{\langle \mathbf{l}^\perp, \mathbf{k}\rangle^2}{|\mathbf{l}|^2 |\mathbf{k}-\mathbf{l}|^2} + w_{\mathbf{l}} q_{\mathbf{k}-\mathbf{l}} \frac{\langle \mathbf{l}^\perp, \mathbf{k}\rangle \langle \mathbf{k}-\mathbf{l}, \mathbf{k}\rangle}{|\mathbf{l}|^2 |\mathbf{k}-\mathbf{l}|^2} \\ &\quad \sum_{\mathbf{l},\, \mathbf{k}-\mathbf{l} \in H_N} -q_{\mathbf{l}} w_{\mathbf{k}-\mathbf{l}} \frac{\langle \mathbf{l}, \mathbf{k}-\mathbf{l}\rangle \langle \mathbf{l}^\perp, \mathbf{k}\rangle}{|\mathbf{l}|^2 |\mathbf{k}-\mathbf{l}|^2} + q_{\mathbf{l}} q_{\mathbf{k}-\mathbf{l}} \frac{\langle \mathbf{l}, \mathbf{k}-\mathbf{l}\rangle \langle \mathbf{k}-\mathbf{l}, \mathbf{k}\rangle}{|\mathbf{l}|^2 |\mathbf{k}-\mathbf{l}|^2}. \end{aligned}$$

After considering the effect of the mapping $(\mathbf{l},\mathbf{k}-\mathbf{l})\mapsto(\mathbf{k}-\mathbf{l},\mathbf{l})$ on each of the terms above, we may write

$$\begin{aligned} F^{\perp}_{\mathbf{k}}(w,q) &= \sum_{\mathbf{l},\,\mathbf{k}-\mathbf{l}\in H_N} w_{\mathbf{l}}w_{\mathbf{k}-\mathbf{l}}\frac{\langle \mathbf{l}^\perp,\mathbf{k}\rangle}{2}\left(\frac{1}{|\mathbf{l}|^2}-\frac{1}{|\mathbf{k}-\mathbf{l}|^2}\right)+w_{\mathbf{l}}q_{\mathbf{k}-\mathbf{l}}\frac{\langle \mathbf{k}-\mathbf{l},\mathbf{k}\rangle}{|\mathbf{k}-\mathbf{l}|^2}\\ F^{\|}_{\mathbf{k}}(w,q) &= \sum_{\mathbf{l},\,\mathbf{k}-\mathbf{l}\in H_N} -w_{\mathbf{l}}w_{\mathbf{k}-\mathbf{l}}\frac{\langle \mathbf{l}^\perp,\mathbf{k}\rangle^2}{|\mathbf{l}|^2|\mathbf{k}-\mathbf{l}|^2}+w_{\mathbf{l}}q_{\mathbf{k}-\mathbf{l}}\frac{\langle \mathbf{l}^\perp,\mathbf{k}\rangle\langle \mathbf{k}-\mathbf{l},\mathbf{k}+\mathbf{l}\rangle}{|\mathbf{l}|^2|\mathbf{k}-\mathbf{l}|^2}\\ &\quad+\sum_{\mathbf{l},\,\mathbf{k}-\mathbf{l}\in H_N} q_{\mathbf{l}}q_{\mathbf{k}-\mathbf{l}}\frac{\langle \mathbf{l},\mathbf{k}-\mathbf{l}\rangle}{2}\frac{|\mathbf{k}|^2}{|\mathbf{l}|^2|\mathbf{k}-\mathbf{l}|^2}. \end{aligned}$$

The assertion made in the previous remark now follows easily from these expressions since if $\sigma_{\mathbf{k}}=\sigma'_{\mathbf{k}}=0$ for all $\mathbf{k}\in H_N$ and $w_0=0$, then $w_t=(w_{\mathbf{k}}(t))_{\mathbf{k}\in H_N}\equiv 0$ for all times t.

To prove Theorem 3.12, we do as in the previous two examples and start by computing $\mathcal{C}$ corresponding to (3.11). Define

$$\mathcal{G}=\left\{X_0+\sum_{\mathbf{k}\in FD_I} u_{\mathbf{k}}X_{\mathbf{k}}+v_{\mathbf{k}}Y_{\mathbf{k}}\;:\;u_{\mathbf{k}},v_{\mathbf{k}}\in\mathbf{R}\right\}$$

where

$$\begin{aligned} X_0 &= \sum_{\mathbf{k}\in H_N}\big[-\nu|\mathbf{k}|^2w_{\mathbf{k}}+iF^{\perp}_{\mathbf{k}}(w,q)\big]\frac{\partial}{\partial w_{\mathbf{k}}}+\big[-\nu|\mathbf{k}|^2q_{\mathbf{k}}+iF^{\|}_{\mathbf{k}}(w,q)\big]\frac{\partial}{\partial q_{\mathbf{k}}}\\ &\quad+\sum_{\mathbf{k}\in G_N}\big[-\nu|\mathbf{k}|^2\bar w_{\mathbf{k}}-iF^{\perp}_{\mathbf{k}}(\bar w,\bar q)\big]\frac{\partial}{\partial \bar w_{\mathbf{k}}}+\big[-\nu|\mathbf{k}|^2\bar q_{\mathbf{k}}-iF^{\|}_{\mathbf{k}}(\bar w,\bar q)\big]\frac{\partial}{\partial \bar q_{\mathbf{k}}} \end{aligned}$$

and

$$X_{\mathbf{k}}=\frac{\partial}{\partial w_{\mathbf{k}}}+\frac{\partial}{\partial \bar w_{\mathbf{k}}},\qquad\qquad Y_{\mathbf{k}}=i\frac{\partial}{\partial w_{\mathbf{k}}}-i\frac{\partial}{\partial \bar w_{\mathbf{k}}}.$$

Notice that $n(X_0,X_{\mathbf{j}})=1$ for all $\mathbf{j}\in\{\mathbf{k}\in H_N\;:\;\sigma_{\mathbf{k}}\neq 0,\,\sigma'_{\mathbf{k}}\neq 0\}$ since there are no diagonal terms in the nonlinear part of X_0. In particular,

$$[X_{\mathbf{j}},X_0]\in\mathcal{G}_1^{\circ}\;\text{ for all }\;\mathbf{j}\in\{\mathbf{k}\in H_N\;:\;\sigma_{\mathbf{k}}\neq 0,\,\sigma'_{\mathbf{k}}\neq 0\}.$$

Moreover, one can compute these commutators to see that

$$\begin{aligned} [X_{\mathbf{j}},X_0] &= -\nu|\mathbf{j}|^2\frac{\partial}{\partial w_{\mathbf{j}}}-\nu|\mathbf{j}|^2\frac{\partial}{\partial \bar w_{\mathbf{j}}}\\ &\quad+i\sum_{\mathbf{k}\in H_N}\left[w_{\mathbf{k}-\mathbf{j}}\langle \mathbf{j}^\perp,\mathbf{k}\rangle\left(\frac{1}{|\mathbf{j}|^2}-\frac{1}{|\mathbf{k}-\mathbf{j}|^2}\right)+\,q_{\mathbf{k}-\mathbf{j}}\frac{\langle \mathbf{k}-\mathbf{j},\mathbf{k}\rangle}{|\mathbf{k}-\mathbf{j}|^2}\right]\frac{\partial}{\partial w_{\mathbf{k}}}\\ &\quad-i\sum_{\mathbf{k}\in H_N}\left[\bar w_{\mathbf{k}-\mathbf{j}}\langle \mathbf{j}^\perp,\mathbf{k}\rangle\left(\frac{1}{|\mathbf{j}|^2}-\frac{1}{|\mathbf{k}-\mathbf{j}|^2}\right)+\bar q_{\mathbf{k}-\mathbf{j}}\frac{\langle \mathbf{k}-\mathbf{j},\mathbf{k}\rangle}{|\mathbf{k}-\mathbf{j}|^2}\right]\frac{\partial}{\partial \bar w_{\mathbf{k}}}\\ &\quad+i\sum_{\mathbf{k}\in H_N}\left[-2w_{\mathbf{k}-\mathbf{j}}\frac{\langle \mathbf{j}^\perp,\mathbf{k}\rangle^2}{|\mathbf{j}|^2|\mathbf{k}-\mathbf{j}|^2}+q_{\mathbf{k}-\mathbf{j}}\frac{\langle \mathbf{j}^\perp,\mathbf{k}\rangle\langle \mathbf{k}-\mathbf{j},\mathbf{k}+\mathbf{j}\rangle}{|\mathbf{j}|^2|\mathbf{k}-\mathbf{j}|^2}\right]\frac{\partial}{\partial q_{\mathbf{k}}}\\ &\quad-i\sum_{\mathbf{k}\in H_N}\left[-2\bar w_{\mathbf{k}-\mathbf{j}}\frac{\langle \mathbf{j}^\perp,\mathbf{k}\rangle^2}{|\mathbf{j}|^2|\mathbf{k}-\mathbf{j}|^2}+\bar q_{\mathbf{k}-\mathbf{j}}\frac{\langle \mathbf{j}^\perp,\mathbf{k}\rangle\langle \mathbf{k}-\mathbf{j},\mathbf{k}+\mathbf{j}\rangle}{|\mathbf{j}|^2|\mathbf{k}-\mathbf{j}|^2}\right]\frac{\partial}{\partial \bar q_{\mathbf{k}}}. \end{aligned}$$

Note also that for all $\mathbf{j},\mathbf{m}\in\{\mathbf{k}\in H_N\;:\;\sigma_{\mathbf{k}}\neq 0,\,\sigma'_{\mathbf{k}}\neq 0\}$ such that $\mathbf{j}+\mathbf{m}\in H_N$

$$n(X_{\mathbf{m}},[X_{\mathbf{j}},X_0])=n(Y_{\mathbf{m}},[X_{\mathbf{j}},X_0])=1.$$

Hence for all $\mathbf{j}, \mathbf{m} \in \{\mathbf{k} \in H_N \,:\, \sigma_{\mathbf{k}} \neq 0,\, \sigma'_{\mathbf{k}} \neq 0\}$ with $\mathbf{j}+\mathbf{m} \in H_N$, $[X_{\mathbf{m}}, [X_{\mathbf{j}}, X_0]] \in \mathcal{G}_2^{\mathrm{o}}$ and $[Y_{\mathbf{m}}, [X_{\mathbf{j}}, X_0]] \in \mathcal{G}_2^{\mathrm{o}}$. Computing these commutators we find that

$$[X_{\mathbf{m}}, [X_{\mathbf{j}}, X_0]] = \langle \mathbf{j}^{\perp}, \mathbf{m}\rangle \left(\frac{1}{|\mathbf{j}|^2} - \frac{1}{|\mathbf{m}|^2}\right) Y_{\mathbf{j}+\mathbf{m}} - 2\frac{\langle \mathbf{j}^{\perp}, \mathbf{m}\rangle^2}{|\mathbf{j}|^2|\mathbf{m}|^2}\widetilde{Y}_{\mathbf{j}+\mathbf{m}} \tag{3.14}$$

and

$$[Y_{\mathbf{m}}, [X_{\mathbf{j}}, X_0]] = -\langle \mathbf{j}^{\perp}, \mathbf{m}\rangle \left(\frac{1}{|\mathbf{j}|^2} - \frac{1}{|\mathbf{m}|^2}\right) X_{\mathbf{j}+\mathbf{m}} + 2\frac{\langle \mathbf{j}^{\perp}, \mathbf{m}\rangle^2}{|\mathbf{j}|^2|\mathbf{m}|^2}\widetilde{X}_{\mathbf{j}+\mathbf{m}} \tag{3.15}$$

where

$$\widetilde{X}_{\cdot} = \frac{\partial}{\partial q_{\cdot}} + \frac{\partial}{\partial \bar{q}_{\cdot}}, \qquad \widetilde{Y}_{\cdot} = i\frac{\partial}{\partial q_{\cdot}} - i\frac{\partial}{\partial \bar{q}_{\cdot}}.$$

We will now use the above computations to prove that

$$\left\{X_{\mathbf{j}}, Y_{\mathbf{j}}, \widetilde{X}_{\mathbf{j}}, \widetilde{Y}_{\mathbf{j}} \,:\, \|\mathbf{k}\|_{\infty} \leq k\right\} \subset \mathcal{C}^{\mathrm{o}}$$

for all $k = 1, 2, \ldots, N$ by induction on k. It will then follow that $\mathcal{C}^{\mathrm{o}}$ spans the tangent space, and so we may pick $\mathcal{D}(w,q) = \mathbf{C}^{8N(N+1)}$ for all $(w,q) \in \mathbf{C}^{8N(N+1)}$.

To prove the claim when $k=1$, first substitute

$$(\mathbf{j}, \mathbf{m}) = ((1,0),(0,1)), ((1,0),(0,-1)), ((-1,0),(0,-1)), ((-1,0),(0,1))$$

into equations (3.14)-(3.15) to see that $\widetilde{X}_{(1,1)}$, $\widetilde{Y}_{(1,1)}$, $\widetilde{X}_{(1,-1)}$, $\widetilde{Y}_{(1,-1)}$, $\widetilde{X}_{(-1,-1)}$, $\widetilde{Y}_{(-1,-1)}$, $\widetilde{X}_{(-1,1)}$, $\widetilde{Y}_{(-1,1)} \in \mathcal{C}^{\mathrm{o}}$. Substituting

$$(\mathbf{j}, \mathbf{m}) = ((1,1),(0,-1)), ((1,1),(-1,0)), ((-1,1),(0,-1)), ((-1,-1),(1,0))$$

into the same equations and using the fact that $X_{\mathbf{k}}, Y_{\mathbf{k}} \in \mathcal{C}^{\mathrm{o}}$ for any $\|\mathbf{k}\|_{\infty} = 1$, we find by taking linear combinations that $\widetilde{X}_{(1,0)}$, $\widetilde{Y}_{(1,0)}$, $\widetilde{X}_{(0,1)}$, $\widetilde{Y}_{(0,1)}$, $\widetilde{X}_{(-1,0)}$, $\widetilde{Y}_{(-1,0)}$, $\widetilde{X}_{(0,-1)}$, $\widetilde{Y}_{(0,-1)} \in \mathcal{C}^{\mathrm{o}}$. This proves the initial statement in the inductive argument. Suppose now that for some $1 \leq k < N$

$$\left\{X_{\mathbf{j}}, Y_{\mathbf{j}}, \widetilde{X}_{\mathbf{j}}, \widetilde{Y}_{\mathbf{j}} \,:\, \mathbf{j} \in H_N, \|\mathbf{j}\|_{\infty} \leq k\right\} \subset \mathcal{C}^{\mathrm{o}}.$$

Note that if $\mathbf{m}, \mathbf{j} \in H_N$ are such that $\|\mathbf{m}\|_{\infty} \leq k$, $\|\mathbf{j}\|_{\infty} = 1$, then $[\widetilde{X}_{\mathbf{m}}, [X_{\mathbf{j}}, X_0]] \in \mathcal{C}^{\mathrm{odd}}$ and $[\widetilde{Y}_{\mathbf{m}}, [X_{\mathbf{j}}, X_0]] \in \mathcal{C}^{\mathrm{o}}$. Note moreover that

$$[\widetilde{X}_m, [X_{\mathbf{j}}, X_0]] = \frac{\langle \mathbf{m}, \mathbf{j}+\mathbf{m}\rangle}{|\mathbf{m}|^2} Y_{\mathbf{j}+\mathbf{m}} + \frac{\langle \mathbf{j}^{\perp}, \mathbf{m}\rangle\langle \mathbf{m}, \mathbf{m}+2\mathbf{j}\rangle}{|\mathbf{j}|^2|\mathbf{m}|^2}\widetilde{Y}_{\mathbf{j}+\mathbf{m}} \tag{3.16}$$

and

$$[\widetilde{Y}_m, [X_{\mathbf{j}}, X_0]] = -\frac{\langle \mathbf{m}, \mathbf{j}+\mathbf{m}\rangle}{|\mathbf{m}|^2} X_{\mathbf{j}+\mathbf{m}} - \frac{\langle \mathbf{j}^{\perp}, \mathbf{m}\rangle\langle \mathbf{m}, \mathbf{m}+2\mathbf{j}\rangle}{|\mathbf{j}|^2|\mathbf{m}|^2}\widetilde{X}_{\mathbf{j}+\mathbf{m}}. \tag{3.17}$$

We claim that if $\mathbf{m}, \mathbf{j} \in H_N$ are such that $|\mathbf{j}| \neq |\mathbf{m}|$ and $\langle \mathbf{j}^{\perp}, \mathbf{m}\rangle \neq 0$, then the pairs (3.14) and (3.16), (3.15) and (3.17), are independent. Indeed, if they are dependent under these assumptions, then

$$|\mathbf{j}|^2\langle \mathbf{m}, \mathbf{m}+\mathbf{j}\rangle = \frac{1}{2}(|\mathbf{j}|^2 - |\mathbf{m}|^2)\langle \mathbf{m}, \mathbf{m}+2\mathbf{j}\rangle$$

which is true if and only if

$$|\mathbf{j}|^2 + |\mathbf{m}|^2 + 2\langle \mathbf{m}, \mathbf{j}\rangle = 0.$$

Note that this equality is impossible since $|\mathbf{j}| \neq |\mathbf{m}|$. Therefore, to finish the inductive argument, it suffices to show that for all $\mathbf{k} \in H_N$ with $\|\mathbf{k}\|_{\infty} = k+1$, there exist $\mathbf{m}, \mathbf{j} \in H_N$ such that

- $\mathbf{m}+\mathbf{j}=\mathbf{k}$;
- $\|\mathbf{m}\|_\infty = k$, $\|\mathbf{j}\|_\infty = 1$, $|\mathbf{m}| \neq |\mathbf{j}|$, and $\langle \mathbf{j}^\perp, \mathbf{m}\rangle \neq 0$.

For those such $\mathbf{k}$ away from the axes and the lines $|y|=|x|$ in the (x,y)-plane, take $\mathbf{j}\in H_N$ to be the unique member of the set $\{(1,0),(0,1),(-1,0),(0,-1)\}$ such that $\|\mathbf{k}-\mathbf{j}\|_\infty = k$. Thus define $\mathbf{m}=\mathbf{k}-\mathbf{j}$ and note that $\mathbf{j}$ and $\mathbf{m}$ have different Euclidean lengths and $\langle \mathbf{j}^\perp, \mathbf{m}\rangle \neq 0$. Now suppose $\mathbf{k}$ is on one of the axes or the lines $|y|=|x|$. Then there exists $\mathbf{j}\in\{(1,0),(0,1),(-1,0),(0,-1)\}$ such that $\mathbf{m}=\mathbf{k}-\mathbf{j}$ belongs to the set of indices generated up to this point of sup norm length $k+1$. It is easy to check that, again, $\mathbf{j}$ and $\mathbf{m}$ have different Euclidean lengths and $\langle \mathbf{j}^\perp, \mathbf{m}\rangle \neq 0$. This finishes the proof of the inductive argument.

Now note that we may choose a basis of $\mathcal{C}$ such that

$$\mathcal{D}(w,q) = \mathbf{C}^{8N(N+1)}$$

for all $(w,q)\in\mathbf{C}^{8N(N+1)}$. Moreover, the origin is clearly an equilibrium point of $\mathcal{G}$. Because the issue of explosion is still evident, Theorem 2.5 implies that for every $(w,q),(w',q')\in\mathbf{C}^{8N(N+1)}$ and $T>0$, there exists $N\in\mathbf{N}$ large enough such that

$$p_t^n((w,q),(w',q'))>0 \quad \text{for all } t\geq T,\, n\geq N.$$

for all $t>0$. □

4. Proof of Main Results

The goal of this section is to prove Theorem 2.5 and Theorem 2.9. Theorem 2.9 will be a relatively straightforward consequence of Theorem 2.5, so we focus our attention first on proving Theorem 2.5.

To prove Theorem 2.5, we will use a slight modification of the condition for positivity of the density given by Ben Arous and Léandre [2] (see also [12]). The slight modification is necessary to remove the global Lipschitzian and boundedness conditions often assumed of the coefficients in the SDE.

To setup the statement of our slight modification, let $H_\cdot = \int_0^\cdot h_s\,ds$, $h\in L^2([0,\infty):\mathbf{R}^r)$, and $\Phi_\cdot^x(H)$ denote the maximally-defined solution (in time) of the equation

$$\Phi_s^x(H) = x + \int_0^s X_0(\Phi_u^x(H))\,du + \sum_{j=1}^r X_j \int_0^s h_u^j\,du. \tag{4.1}$$

$J_{s,t}^x(H)$ denotes the maximally-defined $d\times d$ matrix-valued solution of

$$J_{s,t}^x(H) = \mathrm{Id}_{d\times d} + \int_s^t DX_0(\Phi_u^x(H))J_{s,u}^x(H)\,du \tag{4.2}$$

where $\mathrm{Id}_{d\times d}$ is the identity matrix and D is the Jacobian. Define the Gramian matrix $M_t^x(H)$ by

$$(M_t^x(H))^{nk} = \sum_{m=1}^r \int_0^t (J_{s,t}^x(H)X_m)^n (J_{s,t}^x(H)X_m)^k\,ds. \tag{4.3}$$

Remark 4.4. Sometimes $M_t^x(H)$ is called the deterministic Malliavin covariance matrix. Formally replacing H with a Brownian motion W yields the standard (stochastic) Malliavin covariance matrix.

Lemma 4.5. *Fix $x, z \in \mathbf{R}^d$ and $t > 0$ and suppose that $H_{\cdot} = \int_0^{\cdot} h_s\, ds$, $h \in L^2([0,\infty) : \mathbf{R}^r)$, is such that $\Phi_s^x(H)$ is defined for all times $s \in [0,t]$ and $\Phi_t^x(H) = z$. If $M_t^x(H)$ is invertible, then*

$$p_t^n(x,z) > 0$$

for any integer $n \geq 1$ such that $\Phi_s^x(H) \subset B_n(0)$ for all $s \in [0,t]$.

We defer the proof of Lemma 4.5 until the Appendix, and focus our efforts in this section on exhibiting a control $H_{\cdot} = \int_0^{\cdot} h_s\, ds$, $h \in L^2([0,\infty) : \mathbf{R}^r)$, so that $\Phi_{\cdot}^x(H)$ has all of the properties stated in Lemma 4.5. The proof of the existence of such a control splits into two parts. First, in Section 4.1 we will use the enlargement techniques of Jurjevic and Kupka [5, 6, 7] to see which directions can be flowed along in small times by $\Phi_s^x(H)$ over the class of controls H defined above. Second, we will see that there are enough directions so that we can construct a sufficiently "twisty" control H, ensuring that $M_t^x(H)$ is invertible. The existence of an equilibrium point $y \in \mathbf{R}^d$ as in the statement of Theorem 2.5 allows us control over the time parameter.

4.1. **A Primer on Geometric Control Theory.** For $x \in \mathbf{R}^d$ and $t > 0$, let $A(x, \leq t)$ be the set of points $z \in \mathbf{R}^d$ such that for some time $t_0 \in (0,t]$ there exists $H_{\cdot} = \int_0^{\cdot} h_s\, ds$, $h \in L^2([0,\infty) : \mathbf{R}^r)$, for which $\Phi_s^x(H)$ is defined for all $s \in [0,t_0]$ and $\Phi_{t_0}^x(H) = z$. Recalling the set $\mathcal{C}$ defined in Section 2, here we will use the techniques [5, 6, 7] to prove the following result:

Lemma 4.6. *For all $x \in \mathbf{R}^d$ and all $t > 0$, $\{x\} + \mathcal{C} \subset \overline{A(x, \leq t)}$.*

We start by making some heuristic observations, arguing intuitively why we should expect Lemma 4.6 to be true. To make notation more legible, for any C^∞ vector field V on $\mathbf{R}^d$ let $\exp(tV)(x)$ denote the maximally-defined integral curve of V passing through x at $t = 0$.

We first see why we should expect the following containment to hold

$$\{x\} + \operatorname{span}\{X_1, \ldots, X_r\} \subset \overline{A(x, \leq t)} \tag{4.7}$$

for all $x \in \mathbf{R}^d$, $t > 0$. Let $x \in \mathbf{R}^d$, $\alpha \in \mathbf{R} \setminus \{0\}$ and $j \in \{1, \ldots, r\}$ be given. The key is to realize that for $\lambda > 0$ large and $t > 0$ small

$$\exp(t(X_0 + \alpha\lambda X_j))(x) \approx \exp(t\alpha\lambda X_j)(x)$$

This is because the behavior of the flow along $X_0 + \alpha\lambda X_j$ is initially dominated for small times by the flow along $\alpha\lambda X_j$ since λ is large. More precisely, taking $t = t'/\lambda$ for some $t' > 0$ fixed, one can show that as $\lambda \to \infty$

$$\exp(t(X_0 + \alpha\lambda X_j))(x) = \exp\Big(\frac{t'}{\lambda}(X_0 + \alpha\lambda X_j)\Big)(x) \to \exp(t'\alpha X_j)(x).$$

Since $x \in \mathbf{R}^d$, $\alpha \in \mathbf{R} \setminus \{0\}$ and $j \in \{1, 2, \ldots, r\}$ were assumed to be arbitrary, we now see why one should believe the containment (4.7) as one could repeat the same argument with αX_j replaced by an arbitrary linear combination of $X_1, \ldots, X_r$.

To see how some of the commutators in the definition of $\mathcal{C}$ arise, we start by "tweaking" the directions $X_1, \ldots, X_r$ obtained in the previous step by X_0; that is, we will first flow along X_j for $\alpha\lambda$ units of times and then flow along X_0 for $t > 0$

units of time. Again let $x \in \mathbf{R}^d$, $\alpha \in \mathbf{R}\backslash\{0\}$ and $j \in \{1, \dots, r\}$ be given. If $x_j \in \mathbf{R}^d$ is the constant value of X_j, we notice that for $t > 0$ small

$$\begin{aligned}(4.8)\qquad \exp(tX_0) \circ \exp(\alpha\lambda X_j)(x) &= \exp(tX_0)(x + \alpha\lambda x_j) \\ &= x + \alpha\lambda x_j + \int_0^t X_0(x + \alpha\lambda x_j + \mathcal{O}(s))\, ds.\end{aligned}$$

Letting $t = t'/\lambda^{n(X_j,X_0)}$, it follows that as $\lambda \to \infty$

$$(4.9)\qquad \int_0^t X_0(x + \alpha\lambda x_j + \mathcal{O}(s))\, ds \to \frac{\alpha^{n(X_j,X_0)}}{n(X_j,X_0)!}\mathrm{ad}^{n(X_j,X_0)}X_j(X_0)(x).$$

As much as we would like to obtain this potentially new direction by taking $\lambda \to \infty$ in (4.8), we cannot as $\alpha\lambda x_j$ blows up as $\lambda \to \infty$. To rid ourselves of this problem, we need to flow backwards along X_j for $\alpha\lambda$ units of time producing the relation

$$\begin{aligned}&\exp(-\alpha\lambda X_j) \circ \exp(tX_0) \circ \exp(\alpha\lambda X_j)(x) \\ &= x + \int_0^t X_0(x + \alpha\lambda x_j + \mathcal{O}(s))\, ds.\end{aligned}$$

Using the same scaling of time $t = t'/\lambda^{n(X_j,X_0)}$, we now see how the commutator on the righthand side of (4.9), hence in the definition of $\mathcal{G}_1^{\mathrm{e}}$ and $\mathcal{G}_1^{\mathrm{o}}$, arises.

Remark 4.10. Note that this computation explains why the separation of $\mathcal{C}$ into $\mathcal{C}^{\mathrm{o}}$ and $\mathcal{C}^{\mathrm{e}}$ is needed. If $n(X_j, X_0)$ is even and $\mathrm{ad}^{n(X_j,X_0)}X_j(X_0)$ is constant, then relation (4.9) implies that we may only flow along $\mathrm{ad}^{n(X_j,X_0)}X_j(X_0)$ for positive times. Additionally, in the subsequent iteration of this method we cannot necessarily flow backwards along this vector field producing yet another direction.

Remark 4.11. Following these observations, it is evident where and why Theorem 2.5 will fail to either produce optimal results or be applicable at all. The failure is precisely due to the fact that the set $\mathcal{C}$ only includes those constant vector fields which can be flowed along in small positive times. In particular, Theorem 3.7 does not account for cases where there is an unavoidable time delay needed to access certain points in space (as in the example highlighted in Remark 3.7), usually due the need to employ the drift vector field X_0. Moreover, Theorem 3.7 will not even apply in situations if there is a more serious absence of time reversibility preventing $\mathcal{C}$ from being d-dimensional. As an example, consider the following SDE on $\mathbf{R}^3$

$$\begin{aligned}dx_t &= -x_t y_t\, dt + dB_t \\ dy_t &= (x_t^2 - y_t z_t)\, dt \qquad (4.12) \\ dz_t &= (y_t^2 - z_t)\, dt.\end{aligned}$$

For this system, it is not hard to check that Hörmander's bracket condition is satisfied globally but

$$\mathcal{C} = \{\alpha\partial_x + \lambda\partial_y \,:\, \alpha \in \mathbf{R},\, \lambda \geq 0\}.$$

Hence, Theorem 2.5 does not apply since $\mathcal{C}$ has dimension $2 < 3$.

Even though our general result does not apply in this example, computing $\mathcal{C}$ is still useful in that Lemma 4.6 is true regardless if $\mathcal{C}$ is d-dimensional. If $\mathcal{C}$ is not d-dimensional, one can now proceed to find more points in the set $\overline{A(x, \leq t)}$ by using $\mathcal{C}$ and the specific nature of the drift vector field X_0. Then, given the existence of $H_\cdot = \int_0^\cdot h_s\, ds$, $h \in L^2([0,\infty) : \mathbf{R}^r)$ such that $\Phi_t^x(H) = z$, positivity of the transition density $p_t^n(x,z)$ for n large enough can then be shown by following a similar line of reasoning to Lemma 4.22 or Remark 4.27.

We now turn the previous heuristics into a proof of Theorem 4.6. Our proof will employ results from the reference [7], so we will first introduce some further notation and terminology to connect with the setup there.

We recall that for any C^∞ vector field V on $\mathbf{R}^d$, $\exp(tV)(x)$ denotes the maximally defined integral curve of V passing through x at time $t=0$. Let $\mathcal{H}$ be any set of C^∞ vector fields on $\mathbf{R}^d$. For $x\in\mathbf{R}^d$ and $t>0$, $A_{\mathcal{H}}(x,\leq t)$ denotes the set of $z\in\mathbf{R}^d$ such that there exist positive times $t_1,\dots,t_k$ and corresponding vector fields $V_1,\dots,V_k\in\mathcal{H}$ such that $t_1+\cdots+t_k\leq t$ and

$$\exp(t_kV_k)\circ\exp(t_{k-1}V_{k-1})\circ\cdots\exp(t_1V_1)(x)=z.$$

Because there will be many different sets of vector fields, here we will absolutely need to emphasize the dependence of these sets on $\mathcal{H}$.

Two sets of C^∞ $\mathbf{R}^d$-vector fields, $\mathcal{H}$ and $\mathcal{I}$, are called *equivalent*, denoted by $\mathcal{H}\sim\mathcal{I}$, if $\overline{A_{\mathcal{H}}(x,\leq t)}=\overline{A_{\mathcal{I}}(x,\leq t)}$ for all $x\in\mathbf{R}^d$ and all $t>0$. One can show, see [7], that if $\mathcal{H}\sim\mathcal{I}$ and $\mathcal{H}\sim\mathcal{J}$, then $\mathcal{H}\sim\mathcal{I}\cup\mathcal{J}$. In particular, if we define

$$\text{sat}(\mathcal{H})=\bigcup_{\mathcal{I}\sim\mathcal{H}}\mathcal{I},$$

then it also follows that $\text{sat}(\mathcal{H})\sim\mathcal{H}$. $\text{sat}(\mathcal{H})$ is called the *saturate* of $\mathcal{H}$.

Remark 4.13. It is often the case that $\text{sat}(\mathcal{H})$ contains more vector fields than $\mathcal{H}$ itself. Moreover, the saturate maintains identical accessibility properties in the sense ($\sim$) described above. This is convenient in that it allows one to use simpler vector fields to determine accessibility properties of the original set of vector fields $\mathcal{H}$. For example, even though the constant vector field X_j, $j\geq 1$, does not belong to

$$\mathcal{G}=\{X_0+\textstyle\sum_{j=1}^r u_jX_j\,:\,u_j\in\mathbf{R}\},$$

we used it above to generate more directions in $\overline{A(x,\leq t)}$ as done in the arguments following equation (4.8). Using a limiting procedure, however, one can justify that this is indeed permissible.

In the next two lemmas, we list operations which allow us to expand (up to equivalence) a set of vector fields $\mathcal{H}$.

Lemma 4.14. *$\mathcal{H}$ is equivalent to the closed convex hull of the set*

$$\{\lambda V\,:\,\lambda\in[0,1],\,V\in\mathcal{H}\}.$$

Here the closure is taken in the topology of uniform convergence with all derivatives on compact subsets of $\mathbf{R}^d$.

Proof. Apply Theorem 5 and Theorem 6 in Chapter 2 of [7]. □

To state the next lemma, let $\psi:\mathbf{R}^d\to\mathbf{R}^d$ be a diffeomorphism. For any $V\in\mathcal{H}$, we may define a vector field $\psi_*(V)$ by

$$\psi_*(V)(x)=D\psi(\psi^{-1}(x))V(\psi^{-1}(x))$$

where $D\psi$ is the Jacobian of ψ. A diffeomorphism $\psi:\mathbf{R}^d\to\mathbf{R}^d$ is called a *normalizer* of $\mathcal{H}$ if $\psi(x),\psi^{-1}(x)\in\overline{A_{\mathcal{H}}(x,\leq t)}$ for all $x\in\mathbf{R}^d$ and all $t>0$. The set of normalizers of $\mathcal{H}$ is denoted by $\text{Norm}(\mathcal{H})$.

Lemma 4.15.
$$\mathcal{H} \sim \bigcup_{\psi \in Norm(\mathcal{H})} \{\psi_*(V) \,:\, V \in \mathcal{H}\}.$$

Proof. Notice that by the lemma immediately after Definition 5 of Chapter 2 of [7], if ψ is a normalizer of $\mathcal{H}$ using our definition, then it is also a normalizer using the definition given in [7]. The result then follows after applying Theorem 9 in Chapter 2 of [7] and using the fact that the identity map is a normalizer. □

Remark 4.16. We will see in the proof of Lemma 4.6 that the limiting procedure used in our heuristic calculations is exactly of the type covered by Lemma 4.14. We will also see that the use of normalizers is very much in line with one's ability to flow along a constant vector field for positive or negative times (hence the ψ and ψ^{-1} in the definition of a normalizer).

Using repeated applications of Lemma 4.14 and Lemma 4.15, we now prove Lemma 4.6.

Proof of Lemma 4.6. Let $\mathcal{G} = \{X_0 + \sum_{j=1}^r u_j X_j \,:\, u_j \in \mathbf{R}\}$. First note that it suffices to show that if $V \in \mathcal{C}^{\mathrm{o}}$ and $W \in \mathcal{C}^{\mathrm{e}}$, then $\alpha V, \lambda W \in \mathrm{sat}(\mathcal{G})$ for all $\alpha \in \mathbf{R}$ and all $\lambda \geq 0$. The result would then follow by Lemma 4.14 since if $V_1, V_2, \ldots, V_k \in \mathcal{C}^{\mathrm{o}}$ and $W_1, W_2, \ldots, W_j \in \mathcal{C}^{\mathrm{e}}$, then
$$\sum_{l=1}^k \alpha_l V_l + \sum_{i=1}^j \lambda_i W_i \in \mathrm{sat}(\mathcal{G})$$
for all $\alpha_i \in \mathbf{R}$ and all $\lambda_i \geq 0$.

We first demonstrate that $\alpha X_j \in \mathrm{sat}(\mathcal{G})$ for all $\alpha \in \mathbf{R}$ and $j \in \{1, \ldots, r\}$. Indeed, by Lemma 4.14 we have
$$\alpha X_j = \lim_{\lambda \to \infty} \frac{1}{\lambda}(X_0 + \alpha\lambda X_j) \in \mathrm{sat}(\mathcal{G}).$$
By induction, it is enough to show that if V is a constant vector field with $\alpha V \in \mathrm{sat}(\mathcal{G})$ for all $\alpha \in \mathbf{R}$ and $W \in \mathrm{sat}(\mathcal{G})$ is a polynomial vector field, then
$$\frac{\alpha^{n(V,W)}}{n(V,W)!}\mathrm{ad}^{n(V,W)} V(W) \in \mathrm{sat}(\mathcal{G})$$
for all $\alpha \in \mathbf{R}$. To prove this result, we seek to apply Lemma 4.15. Since V is a constant vector field, let $v = V(x) \in \mathbf{R}^d$ denote its constant value. For $\alpha \in \mathbf{R}$, define a map $\psi_\alpha : \mathbf{R}^d \to \mathbf{R}^d$ by
$$\psi_\alpha(x) = x - \alpha v.$$
Note that, for each $\alpha \in \mathbf{R}$, ψ_α is a normalizer for $\mathcal{G}$. Hence, for each $\alpha \in \mathbf{R}$, Lemma 4.15 implies that $(\psi_\alpha)_*(W) \in \mathrm{sat}(\mathcal{G})$. Since $D\psi_\alpha$ is the identity matrix, notice that
$$(\psi_\alpha)_*(W)(x) = W(x + \alpha v).$$
Applying Lemma 4.14, we thus find that for all $\alpha \in \mathbf{R}$
$$V_{\alpha W} := \lim_{\lambda \downarrow 0} \frac{1}{\lambda^{n(V,W)}}(\psi_{\lambda\alpha})_*(W) \in \mathrm{sat}(\mathcal{G}).$$
To finish the proof, all we must see is that
$$V_{\alpha W} = \frac{\alpha^{n(V,W)}}{n(V,W)!}\mathrm{ad}^{n(V,W)} V(W).$$

Recalling that $v \in \mathbf{R}^d$ denotes the constant value of V, for $x \in \mathbf{R}^d$ fixed consider the function $F : \mathbf{R} \to \mathbf{R}^d$ defined by $\alpha \mapsto W(x + \alpha v)$. By induction, for $j \geq 1$

$$F^{(j)}(\alpha) = \mathrm{ad}^j V(W)(x + \alpha v).$$

where $F^{(j)}$ is the jth derivative of F with respect to α. Hence we obtain the formula

$$(\psi_\alpha)_* W(x) = F(\alpha) = \sum_{j=0}^{n(V,W)} \frac{\alpha^j}{j!} F^{(j)}(0) = \sum_{j=0}^{n(V,W)} \frac{\alpha^j}{j!} \mathrm{ad}^j V(W)(x)$$

since each component of $F(\alpha)$ is a polynomial in α with degree $\leq n(V, W)$. Hence we now see that

$$V_{\alpha W} = \lim_{\lambda \to \infty} \frac{1}{\lambda^{n(V,W)}} (\psi_{\alpha\lambda})_*(W) = \frac{\alpha^{n(V,W)}}{n(V, W)!} \mathrm{ad}^{n(V,W)} V(W),$$

completing the proof. □

Before proceeding onto the second part of the argument, we state the following lemma which we will need later.

Lemma 4.17. *Suppose that, for some $x \in \mathbf{R}^d$, the Lie algebra generated by $\mathcal{H}$ evaluated at x spans the tangent space. Then for all $t, \epsilon > 0$*

$$\mathrm{interior}(A_{\mathcal{H}}(x, \leq t + \epsilon)) \supset \mathrm{interior}(\overline{A_{\mathcal{H}}(x, \leq t)}).$$

Proof. See Theorem 2 of Chapter 3 in [7]. □

4.2. **Strict Positivity.** The next two lemmas will operate as an easy-to-check criterion assuring that, for a given control H, $M_t^x(H)$ is invertible. Though not necessary (see Remark 4.27), these results use the fact that $\mathcal{G}$ contains only polynomial vector fields. In particular, the special structure of zero sets of polynomials is employed in the following lemma.

Lemma 4.18. *Suppose that $\mathcal{C}$ is d-dimensional and let $\mathcal{H} = \cup_{m=1}^r \{X_m, [X_0, X_m]\}$. Then for any non-empty open $A \subset \mathbf{R}^d$ the set of points in $\mathbf{R}^d$ given by*

$$\bigcup_{x \in A} \{V(x) : V \in \mathcal{H}\} \tag{4.19}$$

is d-dimensional.

Proof. Suppose that the subspace spanned by the set in (4.19) has dimension $l \leq d$ and choose a basis $v_1, v_2, \ldots, v_l \in \mathbf{R}^d$ for this subspace. The goal is to show that $l = d$. Let $V_1, V_2, \ldots, V_l$ be the constant vector fields with constant values $v_1, v_2, \ldots, v_l$, respectively. Notice that every vector field V in the span of $\mathcal{H}$ is a polynomial vector field and satisfies the following equality on the open set A

$$V = p_1 V_1 + p_2 V_2 + \cdots + p_l V_l \tag{4.20}$$

for some polynomials $p_1, \ldots, p_l$. Since A is open and V is a polynomial vector field, (4.20) is valid everywhere on $\mathbf{R}^d$. Moreover, since vector fields of the form (4.20) are closed under commutators and linear combinations, we see that

$$\mathrm{span}(\mathcal{C}) \subset \mathrm{span}\{v_1, v_2, \ldots, v_l\}$$

Note that this finishes the proof since $\mathcal{C}$ is d-dimensional. □

To setup the statement of the next result, define $K_t^x(H) \subset \mathbf{R}^d$ as follows:

$$K_t^x(H) = \bigcup_{m=1}^{r} \big\{X_m(\Phi_s^x(H)), [X_0, X_m](\Phi_s^x(H)) \,:\, s \in (0,t)\big\}. \tag{4.21}$$

Lemma 4.22. *Suppose that $K_t^x(H)$ is d-dimensional. Then the associated matrix $M_t^x(H)$ is invertible.*

Proof. It suffices to show that $M_t^x(H)$ is positive definite. Assume, to the contrary, that $M_t^x(H)$ is not positive-definite and let $\langle\,\cdot\,,\,\cdot\,\rangle$ denote the inner product on $\mathbf{R}^d$. Then there exists $y \in \mathbf{R}^d \setminus \{0\}$ such that

$$0 = \langle M_t^x(H)y, y\rangle = \sum_{m=1}^{r} \int_0^t \langle J_{s,t}^x(H)X_m, y\rangle^2\, ds.$$

To get a contradiction, we seek to obtain a lower bound $\langle M_t^x(H)y, y\rangle$ which is positive using the equality above. To derive such a bound, first observe that for $s \le s_0 \le u_0 \le t_0 \le t$, $J_{s_0,t_0}^x(H) = J_{u_0,t_0}^x(H) J_{s_0,u_0}^x(H)$ and that the matrix $J_{s_0,t_0}^x(H)$ is invertible. Using these two facts, it is not hard to check that for $s \le s_0 \le t_0 \le t$

$$\begin{aligned} \partial_{s_0} J_{s_0,t_0}^x(H) &= -J_{s_0,t_0}^x(H) DX_0(\Phi_{s_0}^x(H)) \\ J_{t_0,t_0}^x(H) &= \mathrm{Id}_{d\times d}. \end{aligned} \tag{4.23}$$

Letting $|\cdot|$ denote the Euclidean norm on $\mathbf{R}^d$, we then see that for all $u \in (0,t)$, $\epsilon \in (0, \min(u, t-u))$

$$\begin{aligned} 0 = \langle M_t^x(H)y, y\rangle &\ge \int_0^t \langle J_{s,t}^x(H)X_m, y\rangle^2\, ds \\ &\ge \int_{u-\epsilon}^{u+\epsilon} \langle J_{s,t}^x(H)X_m, y\rangle^2\, ds \\ &= \int_{u-\epsilon}^{u+\epsilon} \langle J_{s,u}^x(H)X_m, (J_{u,t}^x(H))^* y\rangle^2\, ds \\ &\ge |(J_{u,t}^x(H))^* y|^2 \inf_{y\,:\,\|y\|=1} \int_{u-\epsilon}^{u+\epsilon} \langle J_{s,u}X_m, y\rangle^2\, ds. \end{aligned} \tag{4.24}$$

Since $|J_{u,t}^* y| > 0$ and the unit disk is compact in $\mathbf{R}^d$, it suffices to show that for all nonzero $y \in \mathbf{R}^d$ there exists $m \in \{1, 2, \dots, r\}$, $u \in (0,t)$, and $\epsilon \in (0, \min(u, t-u))$ such that

$$\int_{u-\epsilon}^{u+\epsilon} \langle J_{s,u}^x(H)X_m, y\rangle^2\, ds > 0. \tag{4.25}$$

Thus let $y \in \mathbf{R}^d$, $y \neq 0$, be arbitrary. By hypothesis, either $\langle X_m, y\rangle \neq 0$ for some $m \in \{1, \dots, r\}$ or $\langle [X_m, X_0](\Phi_{t_0}^x(H)), y\rangle \neq 0$ for some $m \in \{1, \dots, r\}$, $t_0 \in (0,t)$. Clearly, if $\langle X_m, y\rangle \neq 0$ for some $m \in \{1, 2, \dots, r\}$, then there is nothing to show by continuity and (4.25). Thus suppose that $\langle X_m, y\rangle = 0$ for all $m = 1, 2, \dots, r$ and pick $t_0 \in (0,t)$, $m \in \{1, 2, \dots, r\}$ such that

$$\langle z, y\rangle = \langle [X_m, X_0](\Phi_{t_0}^x(H)), y\rangle \neq 0.$$

Since $\langle X_m, y\rangle = 0$, using the definition of $J^x_{s,t_0}(H)$ twice we see that

$$\begin{aligned}\langle J^x_{s,t_0}(H)X_m, y\rangle &= \int_s^{t_0} \langle DX_0(\Phi^x_u(H))J^x_{s,u}(H)X_m, y\rangle\, du \\ &= \int_s^{t_0} \langle [X_m, X_0](\Phi^x_u(H)), y\rangle\, du \\ &\quad + \int_s^{t_0} \left\langle DX_0(\Phi^x_u(H)) \int_s^u DX_0(\Phi^x_v(H))J^x_{s,v}(H)X_m\, dv, y\right\rangle du.\end{aligned}$$

Therefore, for s sufficiently close to t_0, $\langle J^x_{s,t_0}(H)X_m, y\rangle \neq 0$. Hence continuity then implies for any $\epsilon \in (0, t_0)$

$$\int_{t_0-\epsilon}^{t_0} \langle J^x_{s,t_0}(H)X_m, y\rangle^2\, ds > 0,$$

finishing the proof. □

We now use the previous two results and Lemma 4.6 to prove Theorem 2.5.

Proof of Theorem 2.5. We first prove Theorem 2.5 part (b) and then show how part (a) follows by a similar argument. Therefore suppose that $y \in \mathbf{R}^d$ is an equilibrium point of $\mathcal{G}$ and that $x, z \in \mathbf{R}^d$ are such that $y \in \mathcal{D}(x)$ and $z \in \mathcal{D}(y)$. By Lemma 4.5, our goal is to exhibit $H_\cdot = \int_0^\cdot h_s\, ds$, $h \in L^2([0,t] : \mathbf{R}^r)$, such that $\Phi^x_t(H) = z$ and $M^x_t(H)$ invertible. To ensure that $M^x_t(H)$ is invertible, we will build $H_\cdot$ in such a way so as to "twist" the path of $\Phi^x_\cdot(H)$ from x to z.

We first claim that there exist countably many non-empty disjoint open subsets U_l, $l \geq 0$, with the property that

$$U_{l+1} \subset \bigcup_{w \in U_l} \mathcal{D}(w) \tag{4.26}$$

for all $l \geq 0$. Suppose first that $\mathcal{D}(x) = \mathbf{R}^d$. Then it follows that $\mathcal{D}(x') = \mathbf{R}^d$ for all $x' \in \mathbf{R}^d$. Thus in this case simply let U_l be any partition of $\mathbf{R}^d$. If $\mathcal{D}(x) \neq \mathbf{R}^d$, then since $y \in \mathcal{D}(x)$ write

$$y = x + \textstyle\sum_{j=1}^k \alpha_j y_j + \sum_{j=k+1}^d \lambda_j y_j$$

for some $\alpha_j \in \mathbf{R}$ and $\lambda_j > 0$. Let $\lambda = \min_j \lambda_j > 0$ and define constants $\alpha_0 = 0$ and $\alpha_l = \sum_{k=1}^l 2^{-k}$, $l \geq 1$. Note that for $l \geq 0$ the sets

$$U_l \;=\; x + \operatorname{span}\{y_1, \ldots, y_k\} + \{\mu_{k+1}y_{k+1} + \cdots + \mu_d y_d \,:\, \mu_j \in (\alpha_l \lambda, \alpha_{l+1}\lambda)\}$$

are disjoint, open and satisfy (4.26). This finishes the proof of the claim.

By construction of the sets U_l, $l \geq 0$, and Lemma 4.18, there exist $x_{l+r} \in U_l$ such that

$$\bigcup_{m=1}^r \{x_1, \ldots, x_r, [X_m, X_0](x_{r+1}), \ldots, [X_m, X_0](x_{r+j})\}$$

is d-dimensional. Here, recall that $x_1, \ldots, x_r$ are the constant values of $X_1, \ldots, X_r$, respectively. Moreover, $x_{r+1} \in \mathcal{D}(x)$, $y \in \mathcal{D}(x_{j+r})$ and

$$x_{l+1+r} \in \mathcal{D}(x_{l+r})$$

for all $l = 1, 2, \ldots, j$.

We now show that we can build $H_\cdot$ so that the path $\Phi^x_\cdot(H)$ passes through each of these points prior to time $t > 0$ and so that $\Phi^x_t(H) = z$. Observe that Lemma

4.17 and Lemma 4.6 together imply $A(w, \leq s) \supset \mathcal{D}(w)$ for all $w \in \mathbf{R}^d$ and all $s > 0$. Hence by definition of $A(w, \leq s)$, there exist positive times $t_1, t_2, \ldots, t_{j+1}$ with $\sum_{l=1}^{j+1} t_l < \frac{t}{2}$ and corresponding $H_l(\cdot) = \int_0^{\cdot} h_l(s)\, ds$, $h_l \in L^2([0, t_l] : \mathbf{R}^r)$, such that $\Phi_{t_1}^x(H_1) = x_{r+1}$, $\Phi_{t_{l+1}}^{x_{r+l}}(H_{l+1}) = x_{r+l+1}$, $l = 1, \ldots, j-1$, and $\Phi_{t_{j+1}}^{x_{r+j}}(H_{j+1}) = y$. By piecing together the H_l's, this now gives us the path from x to y. For the rest of the path, we may also pick a positive time $t_{j+3} < \frac{t}{2}$ and $H_{j+3}(\cdot) = \int_0^{\cdot} h_{j+3}(s)\, ds$, $h_{j+3} \in L^2([0, t_{j+3}] : \mathbf{R}^r)$ such that $\Phi_{t_{j+3}}^y(H_{j+3}) = z$. Moreover, since y is an equilibrium point of $\mathcal{G}$, letting $t_{j+2} = t - (t_1 + \cdots + t_{j+1} + t_{j+3}) > 0$ there exists a control $H_{j+2}(\cdot) = \int_0^{\cdot} h_{j+2}(s)\, ds$, $h_{j+2} \in L^2([0, t_{j+2}] : \mathbf{R}^r)$ such that $\Phi_{t_{j+2}}^y(H_{j+2}) = y$. By Lemma 4.22, we now obtain the conclusion in part (b).

To prove part (a), simply let $z = y$ in the first argument and, for an arbitrary $T > 0$, choose $t < T$. Note that this now finishes the proof of Theorem 2.5. □

Remark 4.27. Without using the special structure of polynomial vector fields, one can prove Theorem 2.5 alternatively by choosing the path from x to y differently as follows. Define

$$D(x, y) = \begin{cases} \mathcal{D}(x) \setminus \mathcal{D}(y) & \text{if } \mathcal{D}(x) \neq \mathbf{R}^d \\ \mathbf{R}^d & \text{otherwise} \end{cases}$$

and let $y' \in D(x, y)$ be arbitrary. Since $D(x, y)$ is open, let $\delta > 0$ be such that $B_\delta(y') \subset D(x, y)$. By the support theorems [15, 16], there exists $s_1 \in (0, t/4)$ such that for all n large enough

$$\mathbf{P}_x\{s_1 < \tau_n,\, x_{s_1} \in B_\delta(y')\} > 0.$$

Now recall that $W_s = (W_s^1, \ldots, W_s^r)$ is an r-dimensional standard Wiener process defined on the probability space $(\Omega, \mathscr{F}, \mathbf{P})$. In this remark, we identify the set Ω with the space of continuous paths $C([0, \infty) : \mathbf{R}^r)$. Letting $M_t^x(W(\omega))$ denote the matrix $M_t^x(H)$ when $H_s = (W_s^1(\omega), \ldots, W_s^r(\omega))$, we note that by Malliavin's proof of Hörmander's theorem [8, 11]

$$\mathbf{P}_x\{s_1 < \tau_n,\, x_{s_1} \in B_\delta(y'),\, M_{s_1}^x(W) \text{ invertible}\} = \mathbf{P}_x\{s_1 < \tau_n,\, x_{s_1} \in B_\delta(y')\} > 0$$

for all n sufficiently large. Therefore, fix

$$\omega \in \{s_1 < \tau_n,\, x_{s_1} \in B_\delta(y'),\, M_{s_1}(W(\omega)) \text{ invertible}\}$$

and define $H_s = (W_s^1(\omega), \ldots, W_s^r(\omega))$ on the time interval $[0, s_1]$. Hence $\Phi_{s_1}^x(H) \in B_\delta(y')$. Since

$$y \in \bigcap_{w \in B_\delta(y')} \mathcal{D}(w),$$

pick $\tilde{H}$ such that for some $s_2 < \frac{t}{4}$

$$\Phi_{s_2}^{\Phi_{s_1}^x(H)}(\tilde{H}) = y.$$

We can complete our path from y to z in exactly the same way as in the proof of Theorem 2.5. Invertibility of the covariance matrix for our chosen control at time t follows immediately since $M_{s_1}^x(W(\omega))$ is invertible. See Theorem 8.1 in [10] for a similar argument.

Remark 4.28. Yet another way to prove Theorem 2.5 is to use a Feynman-Kac representation of the probability density function $p_t^n(x, z)$. Indeed fixing $n \in \mathbf{N}$

and $x \in B_n(0)$, observe that the time-reversed density $q_s^n(x,z) = p_{t-s}^n(x,z)$ solves the following PDE

$$\frac{\partial q_s^n}{\partial s} = -\mathcal{L}_z^* q_s^n \quad \text{on} \quad [0,t) \times B_n(0)$$

where $\mathcal{L}_z^*$ is the formal adjoint (in the z variable) of the Markov generator $\mathcal{L}$ corresponding to the diffusion x_t. Now consider the process y_t solving

$$dy_t = -X_0(y_t)\,dt - \sum_{j=1}^r X_j\,dW_t^j$$

and let $T_n = \inf\{t > 0 : |y_t| \geq n\}$. It then follows that we may write $p_t^n(x,z)$ as

$$p_t^n(x,z) = q_0^n(x,z) = \mathbf{E}_z e^{\int_0^{s\wedge T_n} f(y_u)\,du} q_{s\wedge T_n}(x, y_{s\wedge T_n})$$

for some $f \in C^\infty(\mathbf{R}^d : \mathbf{R})$. One can use now the expression above coupled with the support theorems [15, 16] applied to the time-reversed process y_t to bound $p_t^n(x,z)$ from below by a positive quantity.

We finish this section by proving Theorem 2.9 as a consequence of Theorem 2.5 (a).

Proof of Theorem 2.9. Let μ be an invariant probability measure for the Markov process x_t defined by (2.1). Again, since $\mathcal{C}$ is contained in the Lie algebra generated by $X_1, \ldots, X_r, [X_1, X_0], \ldots, [X_r, X_0]$ and $\mathcal{C}$ is d-dimensional, it follows by Hörmander's theorem [4] that $\mu(dx) = m(x)\,dx$ for some nonnegative function $m \in C^\infty(\mathbf{R}^d)$. Recall also that, for the same reasons, the Markov process x_t defined by (2.1) has a probability density function $p_t(x,y)$ with respect to Lebesgue measure on $\mathbf{R}^d$ which is smooth for $(t,x,y) \in (0,\infty) \times \mathbf{R}^d \times \mathbf{R}^d$. Since μ is an invariant probability measure, we have the following relation for almost every $z \in \mathbf{R}^d$ and $t > 0$

$$m(z) = \int_{\mathbf{R}^d} m(y) p_t(y,z)\,dy.$$

We now use this relation to prove the positivity assertion. Let $x \in \text{supp}(\mu)$. Hence $\mu(B_\delta(x)) > 0$ for all $\delta > 0$. By smoothness of the density m, for each $\delta > 0$ there exists $x_1 = x_1(\delta) \in B_\delta(x)$ such that $m(x_1) > 0$. Since m is smooth, in particular continuous, there exists $\gamma > 0$ such that $B_\gamma(x_1) \subset B_\delta(x)$ and $m(y) \geq \epsilon > 0$ for all $y \in B_\gamma(x_1)$. Hence for almost every $z \in \mathbf{R}^d$ we have

$$m(z) \geq \int_{B_\gamma(x_1)} m(y) p_t(y,z)\,dy \geq \epsilon \int_{B_\gamma(x_1)} p_t(y,z)\,dy.$$

To bound $p_t(y,z)$ from below, there are two cases. First suppose that $\mathcal{D}(x) = \mathbf{R}^d$. Then by definition of $\mathcal{D}(x)$, we have that $\mathcal{C}^o$ is d-dimensional, and hence $\mathcal{D}(y) = \mathbf{R}^d$ for all $y \in \mathbf{R}^d$. Theorem 2.5 (a) implies that for any $y \in B_\gamma(x_1)$, $z \in \mathcal{D}(x)$ there exists $t > 0$ such that $p_t(x,z) > 0$. Since the transition density is a continuous function in all of its arguments, there exists an open neighborood U of (t,x,z) in $(0,\infty) \times B_\gamma(x_1) \times \mathbf{R}^d$ such that $p_s(x',z') \geq c > 0$ for $(s,x',z') \in U$. In particular, for almost every y in an open ball centered at z

$$m(y) \geq \epsilon c > 0.$$

Since m is continuous it follows that $m(z) \geq \epsilon c > 0$. For the second case, suppose that $\mathcal{D}(x) \neq \mathbf{R}^d$. In particular, this implies that $\mathcal{C}^o$ has dimension $l < d$ and

$x \notin \mathcal{D}(x)$. Take $z \in \mathcal{D}(x)$ and decrease $\delta > 0$ so that for every $y \in B_\delta(x)$, $z \in \mathcal{D}(y)$. Following now in the same way as in the previous case we finish the proof of the result.

□

Appendix

Here we prove Lemma 4.5. We recall that this result is the slight modification of the criterion for positivity of the density given by Ben-Arous Léandre [2] which was applied without proof in Section 4. Such an extension is needed in this paper since the drift vector field X_0 was not assumed to be globally Lipschitzian and its derivatives were not assumed to be globally bounded.

The proof of Lemma 4.5 is almost identical to (and in some parts simpler than) the proof of Proposition 4.2.2 of [12]. The basic difference needed to remove these assumptions on X_0 is that we need to compare the stopped process $x_{t\wedge\tau_n}$ with another process $x_t^{(n)}$ such that $x_t^{(n)}$ solves an SDE whose coefficients satisfy the required Lipschitzian and boundedness conditions and

$$x_{t\wedge\tau_n} = x_{t\wedge\tau_n}^{(n)} \quad \text{for all} \quad t \geq 0.$$

This localization procedure is relatively standard but we include the details for completeness.

To do such a comparison, for any integer $n \geq 1$ let $X_0^{(n)}$ be a C^∞ vector field on $\mathbf{R}^d$ satisfying

$$X_0^{(n)}(x) = \begin{cases} X_0(x) & \text{for } |x| \leq n \\ 0 & \text{for } |x| \geq n+1 \end{cases}.$$

For $x \in \mathbf{R}^d$, $n \in \mathbf{N}$, $t > 0$ and $H = (H^j) \in C([0,t] : \mathbf{R}^r)$ let $\Phi_t^{x,n}(H)$ denote the solution of the equation

$$\Phi_t^{x,n}(H) = x + \int_0^t X_0^{(n)}(\Phi_s^{x,n}(H))\, ds + \sum_{j=1}^r X_j H_t^j.$$

Let $J_{s,t}^{x,n} = J_{s,t}^{x,n}(H)$ denote the $d \times d$ matrix-valued solution of the equation

$$J_{s,t}^{x,n} = \mathrm{Id}_{d\times d} + \int_s^t DX_0^{(n)}(\Phi_u^{x,n}(H)) J_{s,u}^{x,n}\, du$$

and $M_t^{x,n}(H)$ denote the matrix

$$(M_t^{x,n}(H))_{lm} = \sum_{j=1}^r \int_0^t (J_{s,t}^{x,n}(H)X_j)^l (J_{s,t}^{x,n}(H)X_j)^m\, ds.$$

Proof of Lemma 4.5. As in [12], our goal is to use Malliavin calculus to bound $p_t^n(x,z)$ from below by a quantity which is positive if the covariance matrix $M_t^{x,n}(H)$ is invertible. For brevity of notation during this proof, we will write the functional $\Phi_t^{x,n}(\,\cdot\,)$ simply as $\Phi(\cdot)$. Let $H_\cdot = \int_0^\cdot h_u\, du$, $h \in L^2([0,\infty) : \mathbf{R}^r)$ be as in the statement of the lemma and let $k_l(s)$ denote the lth row of the matrix $k_{lj}(s) = (J_{s,t}^{x,n}(H)X_j)_l$. For $y \in \mathbf{R}^d$, let

$$(T_y W)(t) = W(t) + \sum_{l=1}^d y_l \int_0^t k_l(s)\, ds \quad \text{and} \quad g(y,W) = \Phi(T_y W) - \Phi(W)$$

where $W(t) = (W^1(t), \dots, W^r(t))$ denotes the standard r-dimensional Wiener process on $(\Omega, \mathscr{F}, \mathbf{P})$. For $\beta > 1$, define cutoff functions $\mathcal{K}_\beta, \alpha_\beta \in C(\mathbf{R} : [0,1])$ by

$$\mathcal{K}_\beta(x) = \begin{cases} 0 & \text{if } |x| \geq \beta \\ 1 & \text{if } |x| \leq \beta - 1 \end{cases} \quad \text{and} \quad \alpha_\beta(x) = \begin{cases} 0 & \text{if } |x| \leq \frac{1}{\beta} \\ 1 & \text{if } |x| \geq \frac{2}{\beta} \end{cases},$$

and set

$$\mathcal{H}_\beta = \mathcal{K}_\beta(\|g(\cdot, W)\|_{C^2(B_1(0):\mathbf{R}^d)})\alpha_\beta(|\det \partial_j g^i(0)|).$$

Under our assumptions, one can check that (see [13], Example 1.2.1, Theorem 2.2.2 and surrounding text) $g(\cdot, W(\omega)) \in C^\infty(\mathbf{R}^d)$ for a.s. $\omega \in \Omega$.

Now let $f : \mathbf{R}^d \to [0, \infty)$ be bounded, measurable and $\rho : \mathbf{R}^r \to (0, \infty)$ be a measurable function satisfying $\int_{\mathbf{R}^r} \rho(y)\, dy = 1$. Observe that

$$\begin{aligned} \mathbf{E}_x f(x_{t\wedge\tau_n}) &= \int_{\mathbf{R}^r} \mathbf{E}_x f(x_{t\wedge\tau_n})\rho(y)\, dy \\ &= \int_{\mathbf{R}^r} \mathbf{E} f(\Phi(W))\mathbf{1}_{\{\|\Phi(W)\|_t \leq n\}}\rho(y)\, dy \end{aligned}$$

where

$$\{\|\Phi(W)\|_t \leq n\} = \left\{\omega \in \Omega : \sup_{s\in[0,t]} |\Phi_s^{x,n}(W(\omega))| \leq n\right\}.$$

Girsanov's theorem then gives

$$\begin{aligned} &\int_{\mathbf{R}^r} \mathbf{E} f(\Phi(W))\mathbf{1}_{\{\|\Phi(W)\|_t \leq n\}}\rho(y)\, dy \\ &= \int_{\mathbf{R}^r} \mathbf{E} f(\Phi(T_y W))\mathbf{1}_{\{\|\Phi(T_y W)\|_t \leq n\}} G(y)\rho(y)\, dy \end{aligned}$$

where $G(y) > 0$ is the Radon-Nikodym derivative in the Girsanov change of measure formula. Using this equality we see that for any $c_\beta > 0$

$$\begin{aligned} \mathbf{E}_x f(x_{t\wedge\tau_n}) &\geq \int_{\mathbf{R}^r} \mathbf{E} f(\Phi(T_y W))\mathbf{1}_{\{\|\Phi(T_y W)\|_t \leq n\}} G(y)\rho(y)\, dy \\ &\geq \mathbf{E}\mathcal{H}_\beta \int_{|y|\leq c_\beta} f(g(y) + \Phi(W))\mathbf{1}_{\{\|\Phi(T_y W)\|_t \leq n\}} G(y)\rho(y)\, dy \\ &\geq \mathbf{E}\mathcal{H}_\beta \mathbf{1}_{\{\sup_{|y|\leq c_\beta} \|\Phi(T_y W)\|_t \leq n\}} \int_{|y|\leq c_\beta} f(g(y) + \Phi(W)) G(y)\rho(y)\, dy. \end{aligned}$$

Let $A_\beta = \{\sup_{|y|\leq c_\beta} \|\Phi(T_y W)\|_t \leq n\}$. By Lemma 4.2.1 of [12], for any $\beta > 1$ there exist constants $c_\beta \in (0, \beta^{-1})$ and $\delta_\beta > 0$ such that any mapping $G : B_1(0) \to \mathbf{R}^d$ with $G(0) = 0$, $\|G\|_{C^2(B_1(0))} \leq \beta$ and $|\det \partial_j g^i(0)| \geq \frac{1}{\beta}$ is diffeomorphic from $B_{c_\beta}(0) \subset \mathbf{R}^d$ into a neighborhood of $B_{\delta_\beta}(0) \subset \mathbf{R}^d$. In particular, we find that after changing variables twice

$$\begin{aligned}\mathbf{E}_x f(x_{t\wedge\tau_n}) &\geq \mathbf{E}\mathcal{H}_\beta 1_{A_\beta}\int_{|y|\leq c_\beta} f(g(y)+\Phi(W))G(y)\rho(y)\,dy\\ &\geq \mathbf{E}\mathcal{H}_\beta 1_{A_\beta}\int_{|z|\leq\delta_\beta} f(z+\Phi(W))G(g^{-1}(z))\rho(g^{-1}(z))|\det\partial_j g^i(g^{-1}(z))|\,dz\\ &= \mathbf{E}\mathcal{H}_\beta 1_{A_\beta}\int_{|z-\Phi(W)|\leq\delta_\beta} f(z)G(g^{-1}(z-\Phi(W)))\\ &\qquad\times\rho(g^{-1}(z-\Phi(W)))|\det\partial_j g^i(g^{-1}(z-\Phi(W)))|\,dz.\end{aligned}$$

Therefore we deduce the following inequality

$$\begin{aligned}p_t(x,z) \geq &\mathbf{E}\mathcal{H}_\beta \mathbf{1}_{A_\beta}\mathbf{1}_{\{|z-\Phi(W)|\leq\delta_\beta\}}G(g^{-1}(z-\Phi(W))\\ &\times\rho(g^{-1}(z-\Phi(W)))|\det\partial_j g^i(g^{-1}(z-\Phi(W))|.\end{aligned}$$

By construction, if $\mathcal{H}_\beta\neq 0$ and $|z-\Phi(W)|\leq\delta_\beta$ then

$$G(g^{-1}(z-\Phi(W))\rho(g^{-1}(z-\Phi(W)))|\det\partial_j g^i(g^{-1}(z-\Phi(W))|>0.$$

Thus it remains to prove that $\beta>0$ can be chosen large enough so that the event

$$A_\beta\cap\Big\{|z-\Phi(W)|\,\leq\,\delta_\beta,\,|\det\partial_j g^i(0)|\,\geq\,2\beta^{-1},\,\|g(\cdot,W)\|_{C^2(B_1(0))}\,\leq\,\beta-1\Big\}$$

has positive probability. Note that this can be shown by following exactly the same line of reasoning starting in the last paragraph of p. 1777 of [10].

□

References

[1] Avanti Athreya, Tiffany Kolba, and Jonathan C. Mattingly. Propogating lyapunov functions to prove noise-induced stability. *arXiv: 1111.1755*, (1):1–41, 2011.

[2] G. Ben Arous and R. Léandre. Décroissance exponentielle du noyau de la chaleur sur la diagonale. II. *Probab. Theory Related Fields*, 90(3):377–402, 1991.

[3] Jeremiah Birrell, David P. Herzog, and Jan Wehr. The transition from ergodic to explosive behavior in a family of stochastic differential equations. *Stochastic Processes and their Applications*, 122(4):1519 – 1539, 2012.

[4] Lars Hörmander. Hypoelliptic second order differential equations. *Acta Math.*, 119:147–171, 1967.

[5] V. Jurdjevic and I. Kupka. Control systems on semisimple Lie groups and their homogeneous spaces. *Ann. Inst. Fourier (Grenoble)*, 31(4):vi, 151–179, 1981.

[6] V. Jurdjevic and I. Kupka. Polynomial control systems. *Math. Ann.*, 272(3):361–368, 1985.

[7] Velimir Jurdjevic. *Geometric control theory*, volume 52 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1997.

[8] S. Kusuoka and D. Stroock. Applications of the Malliavin calculus. II. *J. Fac. Sci. Univ. Tokyo Sect. IA Math.*, 32(1):1–76, 1985.

[9] J. C. Mattingly, A. M. Stuart, and D. J. Higham. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Process. Appl.*, 101(2):185–232, 2002.

[10] Jonathan C. Mattingly and Étienne Pardoux. Malliavin calculus for the stochastic 2D Navier-Stokes equation. *Comm. Pure Appl. Math.*, 59(12):1742–1790, 2006.

[11] James Norris. Simplified Malliavin calculus. In *Séminaire de Probabilités, XX, 1984/85*, volume 1204 of *Lecture Notes in Math.*, pages 101–130. Springer, Berlin, 1986.

[12] David Nualart. Analysis on Wiener space and anticipating stochastic calculus. In *Lectures on probability theory and statistics (Saint-Flour, 1995)*, volume 1690 of *Lecture Notes in Math.*, pages 123–227. Springer, Berlin, 1998.

[13] David Nualart. *The Malliavin calculus and related topics.* Probability and its Applications (New York). Springer-Verlag, Berlin, second edition, 2006.

[14] Marco Romito. Ergodicity of the finite dimensional approximation of the 3D Navier-Stokes equations forced by a degenerate noise. *J. Statist. Phys.*, 114(1-2):155–177, 2004.

[15] D. Stroock and S. R. S. Varadhan. On degenerate elliptic-parabolic operators of second order and their associated diffusions. *Comm. Pure Appl. Math.*, 25:651–713, 1972.

[16] Daniel W. Stroock and S. R. S. Varadhan. On the support of diffusion processes with applications to the strong maximum principle. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. III: Probability theory*, pages 333–359, Berkeley, Calif., 1972. Univ. California Press.