

# AN EXTENSION OF HÖRMANDER'S THEOREM FOR INFINITELY DEGENERATE SECOND-ORDER OPERATORS\*

DENIS R. BELL<sup>†</sup> AND SALAH-ELDIN A. MOHAMMED<sup>‡</sup>

ABSTRACT. We establish the hypoellipticity of a large class of highly degenerate second order differential operators of Hörmander type. The hypotheses of our theorem allow Hörmander's general Lie algebra condition to fail on a collection of hypersurfaces. The proof of the theorem is probabilistic in nature. It is based on the Malliavin calculus and requires new sharp estimates for diffusion processes in Euclidean space.

## 1. Introduction.

Let  $X_0, \dots, X_n$  denote a collection of smooth vector fields defined on an open subset  $D$  of  $\mathbf{R}^d$ , and  $c : D \rightarrow \mathbf{R}$  a smooth function. Consider the second order differential operator

$$L := \frac{1}{2} \sum_{i=1}^n X_i^2 + X_0 + c. \quad (1.1)$$

Let  $Lie(X_0, \dots, X_n)$  be the Lie algebra generated by the vector fields  $X_0, \dots, X_n$ . According to the theorem of Hörmander ([H], Theorem 1.1),  $L$  is hypoelliptic on  $D$  if the vector space  $Lie(X_0, \dots, X_n)(x)$  has dimension  $d$  at every  $x \in D$ . Hörmander's condition characterizes hypoellipticity for operators of the form (1.1) with analytic coefficients. However this is not the case if the vector fields  $X_0, \dots, X_n$  defining  $L$  are allowed to be smooth non-analytic. A striking illustration of the non-necessity of the Hörmander condition in the smooth *non-analytic* case is provided by a result of Kusuoka and Stroock, who have made a complete study of hypoellipticity for the class of differential operators on  $\mathbf{R}^3$  of the form

$$L_\sigma := \frac{\partial^2}{\partial x_1^2} + \sigma^2(x_1) \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}. \quad (1.2)$$

---

\* *Duke Mathematical Journal*, Vol.78, No. 3, (1995), 453-475. .

<sup>†</sup>The research of this author is supported in part by NSF Grant DMS-9121406.

<sup>‡</sup>The research of this author is supported in part by NSF Grants DMS-8907857 and DMS-9206785.

Here  $\sigma$  is assumed to be a  $C^\infty$  real-valued even function, non-decreasing on  $[0, \infty)$ , which vanishes (only) at zero. It is shown in ([K-S], Theorem 8.41) that  $L_\sigma$  is *hypoelliptic on  $\mathbf{R}^3$  if and only if  $\sigma$  satisfies the condition  $\lim_{s \rightarrow 0^+} s \log \sigma(s) = 0$* . In particular, the operator  $L_\sigma$  corresponding to  $\sigma(s) = \exp(-|s|^p)$  is hypoelliptic if  $p$  lies in the range  $(-1, 0)$ ; however any such operator fails to satisfy Hörmander's condition on the hyperplane  $x_1 = 0$ .

Let  $L$  be the operator defined in (1.1). The purpose of this paper is to establish a criterion for hypoellipticity sharper than that of Hörmander, in the case where  $L$  has smooth non-analytic coefficients. Our main theorem (Theorem 1.0) asserts the hypoellipticity of the operator  $L$  on  $D$  under hypotheses that allow Hörmander's general condition to *fail at an exponential rate* on a collection of surfaces in  $D$ . The precise sense in which Hörmander's condition fails is described in the hypotheses of Theorem 1.0 below. Our second main result gives a parallel criterion for hypoellipticity of the parabolic operator  $L + \frac{\partial}{\partial t}$  on  $\mathbf{R} \times D$  in the case when Hörmander's restricted condition *fails at an exponential rate* on a collection of surfaces in  $D$ .

In what follows, each vector field on  $D$  is considered as a column vector with respect to a fixed (canonical) basis for the vector space of all smooth vector fields on  $D$ . For any positive integer  $m$ , let  $E^{(m)}$  denote a matrix whose columns consist of

$$X_0; \cdots; X_n;$$

together with all vector fields of the form

$$[X_{i_1}, X_{i_2}]_{i_1, i_2=0}^n; \cdots; [X_{i_1}, [X_{i_2}, [X_{i_3}, \cdots, [X_{i_{m-1}}, X_{i_m}]] \cdots]]_{i_1, i_2, \dots, i_m=0}^n,$$

arranged in a specified order. The symbol  $[\cdot, \cdot]$  denotes the Lie bracket operation on vector fields. Denote by  $X^{(m)}$  the matrix obtained by deleting the column  $X_0$  from  $E^{(m)}$ . For any  $x \in D$  and  $m \geq 1$ , define  $\mu^{(m)}(x)$  and  $\lambda^{(m)}(x)$  to be the smallest eigenvalues of the matrices  $E^{(m)}(x) E^{(m)*}(x)$  and  $X^{(m)}(x) X^{(m)*}(x)$ , respectively. Note that  $\mu^{(m)}(x)$  and

$\lambda^{(m)}(x)$  are independent of the choice of the basis in the space of vector fields, and are also independent of the specific ordering of the columns referred to above.

Observe that  $\mu^{(m)}(x) > 0$  for some  $m \geq 1$  if and only if Hörmander's general condition holds for the operator  $L$  at  $x \in D$ . Similarly the condition  $\lambda^{(m)}(x) > 0$  for some  $m \geq 1$  is equivalent to Hörmander's general condition for the parabolic operator  $L + \frac{\partial}{\partial t}$  at  $(t, x) \in \mathbf{R} \times D$  for some  $t \in \mathbf{R}$  (and hence for all  $t \in \mathbf{R}$ ). We will say that  $x \in D$  is an *elliptic Hörmander point* for the operator  $L$  if there is an integer  $m \geq 1$  such that  $\mu^{(m)}(x) > 0$ . Similarly  $x \in D$  is said to be a *parabolic Hörmander point* for the operator  $L$  if there is an integer  $m \geq 1$  such that  $\lambda^{(m)}(x) > 0$ . The set of all elliptic Hörmander points in  $D$  is denoted by  $H_e$ , and the set of all parabolic Hörmander points is denoted by  $H$ . Note that the sets  $H_e$  and  $H$  are open in  $D$ . The points in the closed sets  $H_e^c, H^c$  will be called *elliptic non-Hörmander points* and *parabolic non-Hörmander points*, respectively. It is easy to see that  $H \subseteq H_e$  and  $H_e^c \subseteq H^c$ . When  $L$  is formally self-adjoint, it follows from Fefferman and Phong ([F-P]) that  $L$  is *not subelliptic* on  $H_e^c$ .

We are now ready to state our first main theorem.

**Theorem 1.0.**

*Suppose that the elliptic non-Hörmander set  $H_e^c$  is contained in a  $C^2$  submanifold  $M$  of  $D$ , of codimension 1. Assume that at every point in  $H_e^c$ , at least one of the vector fields  $X_1, \dots, X_n$  is transversal to  $M$ . Assume further that for every  $x \in H_e^c$ , there exists an integer  $m \geq 1$ , an open neighborhood  $U$  of  $x$ , and an exponent  $p \in (-1, 0)$  such that  $\mu^{(m)}(y) \geq \exp\{-[\rho(y, M)]^p\}$  for all  $y \in U$ , where  $\rho(y, M)$  is the Euclidean distance of  $y$  from  $M$ .*

*Then  $L$  is hypoelliptic on  $D$ .*

In Section 3, we will show that Theorem 1.0 is a consequence of the following theorem which deals with parabolic hypoellipticity when Hörmander's restricted condition is allowed to fail on a family of hypersurfaces in  $D$ .

**Theorem 1.1.**

Suppose that the parabolic non-Hörmander set  $H^c$  is contained in a  $C^2$  submanifold  $N$  of  $D$ , of codimension 1. Assume that at every point in  $H^c$ , at least one of the vector fields  $X_1, \dots, X_n$  is transversal to  $N$ . Assume further that for every  $x \in H^c$ , there exists an integer  $m \geq 1$ , an open neighborhood  $U_1$  of  $x$ , and an exponent  $p \in (-1, 0)$  such that  $\lambda^{(m)}(y) \geq \exp\{-[\rho(y, N)]^p\}$  for all  $y \in U_1$ .

Then the parabolic operator  $L + \frac{\partial}{\partial t}$  is hypoelliptic on  $\mathbf{R} \times D$ .

We make the following observations concerning the nature of the hypotheses in the above theorems.

- (i) The result of Kusuoka and Stroock cited above implies that the differential operator

$$L' := \frac{\partial^2}{\partial x_1^2} + \exp\left\{-\frac{1}{|x_1|}\right\} \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2} \quad (1.3)$$

is *not hypoelliptic* on  $\mathbf{R}^3$ . This example shows that the lower bound  $-1$  on  $p$  in Theorems 1.0 and 1.1 is *optimal*.

- (ii) Oleinik and Radekevich ([O-R], Theorem 2.5.3) have shown that if the elliptic non-Hörmander set  $H_e^c$  of  $L$  is compact and  $L$  satisfies the transversality condition of Theorem 1.0 at all points of  $H_e^c$ , then  $L$  is hypoelliptic. Example (1.3) shows that if the compactness assumption on  $H_e^c$  is dropped, then a further hypothesis such as the exponential degeneracy condition in Theorem 1.0 is required which controls the rate at which the Hörmander condition fails as one approaches  $H_e^c$ .
- (iii) In a series of papers, Y. Morimoto has obtained hypoellipticity results for operators with exponential degeneracies on hyperplanes (see, e.g., [Mo]). Morimoto's techniques are purely analytic and rely on the theory of pseudo differential operators. A special case of Theorem 1.0, allowing a degeneracy of exponential order at an isolated point, has been obtained by P. Malliavin ([M.2], Theorem 4.1).

- (iv) The following alternative form of Theorem 1.1 shows that the transversality condition can be weakened, provided the range of  $p$  is restricted appropriately. In the statement of Theorem 1.2 below, we will denote by  $X_i\phi$  the action of the vector field  $X_i$  on a given  $C^\infty$  real-valued function  $\phi : D \rightarrow \mathbf{R}$ , for  $0 \leq i \leq n+1$ . We have no reason to believe that the lower bound  $-\frac{2}{(18)^r}$  on the exponent  $p$  is optimal.

**Theorem 1.2.**

Let  $X_{n+1}$  denote the action of the operator  $L-c$  on smooth functions on  $D$ . Suppose that for every  $x \in D$  there exists an integer  $m \geq 1$  such that exactly one of the following two conditions holds:

(a)  $\lambda^{(m)}(x) > 0$ .

(b) There exists an integer  $r \geq 1$ , an open neighborhood  $U \subseteq D$  of  $x$ , a  $C^\infty$  function  $\phi : U \rightarrow \mathbf{R}$ , and an exponent  $p \in (-\frac{2}{(18)^r}, 0)$  such that

(i)  $\phi(x) = 0$ , and there exist  $1 \leq i_1, i_2, \dots, i_r \leq n+1$  such that

$$X_{i_1} X_{i_2} \cdots X_{i_r} \phi(x) \neq 0. \tag{1.4}$$

(ii)  $\lambda^{(m)}(y) \geq \exp(-|\phi(y)|^p)$ , for all  $y \in U$ .

Then the parabolic operator  $L + \frac{\partial}{\partial t}$  is hypoelliptic on  $\mathbf{R} \times D$ .

- (v) Without loss of generality, we may (and, from now on, will) assume that the maps  $X_0, \dots, X_n, c$  in (1.1) are smooth, defined on *the whole of*  $\mathbf{R}^d$  and have compact support. This follows from a simple argument using a partition of unity and the fact that hypoellipticity of  $L + \frac{\partial}{\partial t}$  is a local property.

We prove Theorems 1.1 and 1.2 by probabilistic methods. Completing a program initiated by Malliavin ([M.1] and [M.2]), Kusuoka and Stroock [K-S] developed the stochastic analysis required to give a probabilistic proof of Hörmander's theorem. The theory in [K-S] appears to break down when applied to *exponentially degenerate* diffusion processes.

Thus new ideas are required in order to prove the results described here. The essential new feature of this work is an analysis of the interaction between the diffusion process  $\xi$  associated with the operator  $L$  and the hypersurface  $N$  containing  $H^c$ . This analysis breaks down into two stages:

- (a) The distance between the diffusion  $\xi(t)$  and the hypersurface  $N$  is measured by the Itô process  $\phi(\xi(t))$ , where  $\phi$  is a local parametrization of  $N$ . We first isolate a class of small random intervals which we call *exponentially positive*. (See Section 3.) We then introduce Wiener functionals that correspond to higher-order time-moments of  $\phi(\xi(t))$  over exponentially positive random intervals. We derive new probabilistic lower bounds on the  $L^q$ -norms of the above Wiener functionals for arbitrary  $q \geq 1$ . These lower bounds are *sharp* in  $q$  in an asymptotic sense, as  $q \rightarrow \infty$ .
- (b) We then study the rate at which the above probabilistic lower bounds are degraded under the action of an exponentially degenerate diffusion covariance. This leads to sharp estimates which allow us to establish the hypoellipticity of  $L + \frac{\partial}{\partial t}$ .

Although our analysis employs the Malliavin calculus, the above estimates are obtained using only basic stochastic analytic tools, i.e. Itô's formula, Girsanov's theorem, the time change theorem for stochastic integrals, and elementary properties of the Wiener process. Thus our arguments provide insight into the relationship between the critical exponent  $-1$  occurring in Theorems 1.0, 1.1 and the stochastic structure underlying the hypoellipticity problem.

The paper is laid out as follows. In Section 2, we provide background material and establish notations. In Section 3, we give a proof of Theorem 1.1 based on two results, Lemmas 3.2 and 3.3, which constitute steps (a) and (b) above. We then show how Theorem 1.0 follows from Theorem 1.1. The modifications to the proof of Theorem 1.1 required to prove Theorem 1.2 are discussed at the end of Section 3. Finally, Section 4 contains the proofs of the lemmas assumed in Section 3.

An announcement of Theorem 1.1 has appeared in [B-M].

## 2. Differential operators and diffusion processes.

The purpose of this section is to establish the probabilistic framework upon which our proof of Theorem 1.1 will be based. In particular, we introduce Wiener space and summarize the elements of the Malliavin calculus, which plays a major role in our analysis of the regularity of the operator  $L$  in (1.1). The central result of the Malliavin calculus is an integration by parts formula for stochastic maps defined on the Wiener space. This formula yields the important result stated below as Theorem 2.1 ([S]). One of the authors ([B], Chapter 4) has given an elementary proof of Malliavin's integration by parts formula based on classical finite dimensional calculus. Further information on the Malliavin calculus can be found in ([B], [I-W], [S]) and the references therein.

Let  $\Omega$  denote the space of all continuous paths  $\omega : \mathbf{R}^+ \rightarrow \mathbf{R}^n$ ,  $\omega(0) = 0$ , endowed with the topology of uniform convergence on compact sets. Let  $P$  denote the standard Wiener measure on  $\Omega$ ,  $\mathcal{F}$  the Borel  $\sigma$ -field of  $\Omega$  completed under  $P$ , and  $(\mathcal{F}_t)_{t \geq 0}$  the filtration consisting of the sub- $\sigma$ -fields  $\mathcal{F}_t$  generated by the evaluations  $\{\omega \rightarrow \omega_u, 0 \leq u \leq t\}$  and including all sets of  $P$ -measure zero in  $\mathcal{F}$ . For each  $q \geq 1$  denote by  $L^q := L^q(\Omega, \mathbf{R}; P)$ , the Banach space of all  $q$ -integrable Wiener functionals  $f : \Omega \rightarrow \mathbf{R}$  with the norm  $\|f\|_q := (E|f|^q)^{1/q}$ . Throughout the paper we will denote the Euclidean norm on  $\mathbf{R}^d$  by  $|\cdot|$ . The corresponding norm on the space  $\mathbf{R}^{d \times d}$  of  $d \times d$  matrices will be denoted by  $\|\cdot\|$ .

According to a fundamental theorem of Wiener, the Hilbert space  $L^2$  decomposes as the direct sum

$$L^2 = \bigoplus_{k=0}^{\infty} Z_k,$$

where, for each  $k \geq 0$ ,  $Z_k$  is the space of  $k^{\text{th}}$  order homogeneous chaos. Define  $\mathcal{D}$  to be the following dense subspace of  $L^2$

$$\mathcal{D} := \left\{ f : f \in L^2, \sum_{k=1}^{\infty} k^2 \|\Pi_k f\|_2^2 < \infty \right\},$$

where  $\Pi_k : L^2 \rightarrow Z_k$  is the orthogonal projection map.

Let  $\mathcal{L} : \mathcal{D} \rightarrow \mathbf{R}$  denote the *number operator*

$$\mathcal{L}(f) := - \sum_{k=1}^{\infty} \left(\frac{k}{2}\right) \Pi_k f, \quad f \in \mathcal{D}$$

and  $\langle \cdot, \cdot \rangle : \mathcal{D} \times \mathcal{D} \rightarrow L^2(\Omega, \mathbf{R})$  an associated bilinear form, defined as the unique closed symmetric bilinear extension of the following form (which is denoted by the same symbol):

$$\langle f, g \rangle := \mathcal{L}(fg) - f\mathcal{L}(g) - g\mathcal{L}(f), \quad f, g \in \mathcal{D} : fg \in \mathcal{D}.$$

For any set  $\mathcal{G} \subset \mathcal{D}$ , define  $M(\mathcal{G})$  to be the union of  $\mathcal{G}$  and the set obtained by applying  $\mathcal{L}$  to all elements of  $\mathcal{G}$ , and  $\langle \cdot, \cdot \rangle$  to all pairs of elements of  $\mathcal{G}$ .

**Theorem 2.1.** (*Malliavin-Stroock*)

Suppose that  $\Theta = (\Theta_1, \Theta_2, \dots, \Theta_d) \in \mathcal{D}^d$  satisfies the following hypotheses

(i)  $\mathcal{G}_0 := \{\Theta_1, \Theta_2, \dots, \Theta_d\} \subset \mathcal{D}$ , and the sets defined inductively by  $\mathcal{G}_n := M(\mathcal{G}_{n-1})$ ,  $n \geq 1$  are all subsets of  $\mathcal{D}$ , and for every  $n, q \geq 1, \mathcal{G}_n \subset L^q$ .

(ii) The (Malliavin) covariance matrix  $C$  defined by  $C := [\langle \Theta_i, \Theta_j \rangle]_{i,j=1}^d \in GL(d)$  a.s. and

$$\det C^{-1} \in L^q,$$

for all  $q \geq 1$ . Then the measure  $P \circ \Theta^{-1}$  is absolutely continuous with respect to Lebesgue measure on  $\mathbf{R}^d$  and has a  $C^\infty$  density.

Let  $W := (W_1, \dots, W_n) : \mathbf{R}^+ \times \Omega \rightarrow \mathbf{R}^n$  denote standard Brownian motion defined on the complete filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$  by  $W(t, \omega) = \omega(t)$ ,  $t \in \mathbf{R}^+$ ,  $\omega \in \Omega$ . Consider the diffusion process

$$\left. \begin{aligned} d\xi^x(t) &= \sum_{i=1}^n X_i(\xi^x(t)) \circ dW_i(t) + X_0(\xi^x(t)) dt, & t > 0, \\ \xi^x(0) &= x, \end{aligned} \right\} \quad (2.1)$$



where the  $\circ dW_i$  denote *Stratonovich* differentials. Then  $\xi^x$  is a time-homogeneous Markov process.

Define a family of finite Borel measures  $\nu(x, t, \cdot)$  by

$$\nu(x, t, B) := \int_{\xi^x(t)^{-1}(B)} \exp\left\{\int_0^t c(\xi^x(u)) du\right\} dP$$

for  $t \geq 0$ ,  $x \in \mathbf{R}^d$ , and every Borel set  $B \subseteq \mathbf{R}^d$ . The associated semigroup  $(P_t)_{t \geq 0}$  on the space  $\mathcal{C}_b(\mathbf{R}^d, \mathbf{R})$  of all bounded continuous functions  $\phi : \mathbf{R}^d \rightarrow \mathbf{R}$ , is defined by

$$P_t(\phi)(x) := \int_{\mathbf{R}^d} \phi(y) \nu(x, t, dy), \quad (2.2)$$

for  $t \geq 0$ ,  $x \in \mathbf{R}^d$ , and  $\phi \in \mathcal{C}_b(\mathbf{R}^d, \mathbf{R})$ . Then for any bounded smooth initial function  $\phi : \mathbf{R}^d \rightarrow \mathbf{R}$ , the function  $u(t, x) := P_t(\phi)(x)$  is  $C^\infty$  on  $\mathbf{R}^+ \times \mathbf{R}^d$ , and is the unique solution of the Cauchy initial-value problem

$$\left. \begin{aligned} \frac{\partial u}{\partial t} &= Lu, & t > 0 \\ u(0, \cdot) &= \phi. \end{aligned} \right\} \quad (2.3)$$

For all  $t \geq 0$ , the map

$$\begin{aligned} \xi^x(t, \cdot) &: \Omega \rightarrow \mathbf{R}^d \\ \omega &\mapsto \xi^x(t, \omega) \end{aligned}$$

satisfies hypothesis (i) of Theorem 2.1. The corresponding covariance matrix  $C(t, x)$  (defined in Theorem 2.1(ii)) is given by

$$C(t, x) = Y^x(t) \int_0^t Z^x(s) X(\xi^x(s)) X(\xi^x(s))^* Z^x(s)^* ds [Y^x(t)]^*. \quad (2.4)$$

Here  $X$  is the  $d \times n$  matrix-valued function  $[X_1, \dots, X_n]$ , where  $X_1, \dots, X_n$  are the vector fields in (1.1),  $Y^x(t)$  is the a.s. (non-singular)  $d \times d$  matrix-valued process obtained by differentiating the stochastic flow  $x \rightarrow \xi^x(t, \omega)$  on  $\mathbf{R}^d$  with respect to  $x$  for a.a.  $\omega \in \Omega$ .

The inverse process  $Y^x(t)^{-1}$  is denoted by  $Z^x(t)$ . It can be shown that the processes  $Y^x$  and  $Z^x$  satisfy the stochastic integral equations

$$Y^x(t) = I + \sum_{i=1}^n \int_0^t DX_i(\xi^x(s))Y^x(s) \circ dW_i(s) + \int_0^t DX_0(\xi^x(s))Y^x(s) ds, \quad (2.5)$$

and

$$Z^x(t) = I - \sum_{i=1}^n \int_0^t Z^x(s)DX_i(\xi^x(s)) \circ dW_i(s) - \int_0^t Z^x(s)DX_0(\xi^x(s)) ds. \quad (2.6)$$

See ([K-S], p.3-4), ([B], p.75). We will apply Theorem 2.1 with  $\theta := \xi^x(t)$ . This yields the following result:

**Theorem 2.2.**

*Define  $\Delta(t, x) := \det C(t, x)$ , where  $C(t, x)$  is as in (2.4). Suppose that  $C(t, x) \in GL(d)$  a.s. and  $\Delta(t, x)^{-1} \in L^q$  for all  $q \geq 1$ ,  $t > 0$  and  $x \in \mathbf{R}^d$ . Then the transition probabilities  $p(t, x, dy) := P(\xi^x(t) \in dy)$  admit a family of densities  $p(t, x, y)$  with respect to Lebesgue measure on  $\mathbf{R}^d$ . Furthermore, if the map*

$$\mathbf{R}^+ \times \mathbf{R}^d \rightarrow L^q(\Omega, \mathbf{R})$$

$$(t, x) \mapsto \Delta(t, x)^{-1}$$

*is locally bounded for every  $q \geq 1$ , then the densities  $p(t, x, y)$  can be chosen to be  $C^\infty$  in  $(t, x, y)$ . In particular, the semigroup  $(P_t)_{t \geq 0}$  defined in (2.2) is smoothing, i.e. for all  $t > 0$ ,  $P_t$  takes continuous functions to smooth functions, and the function  $u(t, x) := P_t(\phi)(x)$  defines a smooth solution to the Cauchy problem (2.3) for any continuous initial function  $\phi : \mathbf{R}^d \rightarrow \mathbf{R}$  with compact support.*

Theorem 2.2 establishes a key relationship between invertibility of the covariance matrices  $C(t, x)$  and the regularity of the differential operator  $L$ . This relationship can be further developed to yield the following criterion for hypoellipticity of the parabolic operator  $L + \frac{\partial}{\partial t}$ , due to Kusuoka and Stroock. See the proof of Theorem (8.6) and also

Theorem (8.13) in ([K-S], pp. 49-52). Recall that a differential operator  $A$  on  $\mathbf{R} \times D$  is *hypoelliptic* on  $\mathbf{R} \times D$  if for every distribution  $\phi$  on  $\mathbf{R} \times D$ ,  $A\phi$  is  $C^\infty$  implies that  $\phi$  is a  $C^\infty$  function.

**Theorem 2.3.**

Let  $D$  be an open set in  $\mathbf{R}^d$ . Suppose that for every  $q \geq 1$  and every  $x$  in  $D$ , there exists a neighborhood  $V \subseteq D$  of  $x$  such that

$$\lim_{t \rightarrow 0^+} t \log \left\{ \sup_{y \in V} \|\Delta(t, y)^{-1}\|_q \right\} = 0. \tag{2.7}$$

Then the differential operator  $L + \frac{\partial}{\partial t}$  is hypoelliptic on  $\mathbf{R} \times D$ .

**3. Proofs of Theorems.**

*Proof of Theorem 1.1.*

Throughout this section,  $T > 0$  will denote a fixed time. We first introduce a term that will be used extensively in the sequel.

**Definition.**

A (non-negative) random variable  $X$  is *exponentially positive* if there exist positive constants  $c_1$  and  $c_2$  (which we will refer to as the *characteristics* of  $X$ ) such that

$$P(X < \epsilon) < \exp(-c_1 \epsilon^{-1})$$

for all  $\epsilon \in (0, c_2)$ .

We will also make frequent use of the following well-known result, ([I-W], Lemma 10.5, p. 398).

**Lemma 3.1.**

Let  $y : [0, T] \times \Omega \rightarrow \mathbf{R}^d$  be an Itô process of the form

$$dy(t) = \sum_{i=1}^n a_i(t) dW_i(t) + b(t) dt, \quad 0 \leq t \leq T, \quad (3.1)$$

where  $a_1, \dots, a_n, b : [0, T] \times \Omega \rightarrow \mathbf{R}^d$  are measurable  $(\mathcal{F}_t)_{0 \leq t \leq T}$ -adapted processes, all bounded a.s. by a deterministic constant  $c_3$ . Let  $r > 0$  and define

$$\tau := \inf\{s > 0 : |y(s) - y(0)| = r\} \wedge T. \quad (3.2)$$

Then  $\tau$  is an exponentially positive  $(\mathcal{F}_t)_{0 \leq t \leq T}$ -stopping time, and the characteristics of  $\tau$  depend only on  $r, c_3, n$  and  $d$ .

We now state the two key lemmas upon which our argument turns. The first yields sharp probabilistic lower bounds when applied to diffusion processes with at least one non-zero initial time diffusion coefficient.

**Lemma 3.2.**

Let  $y : [0, T] \times \Omega \rightarrow \mathbf{R}^d$  be the Itô process in Lemma 3.1. Suppose that  $\tau \leq T$  is an exponentially positive  $(\mathcal{F}_t)_{0 \leq t \leq T}$ -stopping time such that at least one diffusion coefficient  $a_i$  satisfies the condition: a.s.,  $|a_i(s)| \geq \delta$ , for all  $0 \leq s \leq \tau$ , for some deterministic  $\delta > 0$ . Then for every  $m \geq 2$ , there exist positive constants  $c_4, c_5$  and  $T_0$  such that for all  $t \in (0, T_0)$  and  $\epsilon \in (0, c_4 t^{m+1})$ , the following holds

$$P\left(\int_0^{t \wedge \tau} |y(u)|^m du < \epsilon\right) < \exp\left\{-c_5 \epsilon^{-\frac{1}{m+1}}\right\}. \quad (3.3)$$

The constants  $c_4$  and  $c_5$  can be chosen to depend only on  $m, c_3, \delta$ , and the characteristics of  $\tau$ . The constant  $T_0$  depends only on the characteristics of  $\tau$ .

The following result describes precisely how the estimate (3.3) transforms under composition of the integrand with a function that vanishes at zero, at (an appropriate) exponential rate.

**Lemma 3.3.**

Let  $\tau$  be an exponentially positive  $(\mathcal{F}_t)_{0 \leq t \leq T}$ -stopping time and let  $p \in (-1, 0)$ .

Suppose  $y$  is an Itô process as in Lemma 3.1. Suppose further that  $y$  and  $\tau$  satisfy an estimate of the form (3.3) for some  $m > -\frac{p}{p+1}$ . Then there exist positive constants  $T_1$ ,  $c_6$ ,  $c_7$  and  $q > 1$  such that for all  $t \in (0, T_1)$  and all  $\epsilon < \exp\{-c_6 t^{-\frac{1}{q}}\}$ , the following holds

$$P\left(\int_0^{t \wedge \tau} \exp(-|y(u)|^p) du < \epsilon\right) < \exp\{-c_7 |\log \epsilon|^q\}. \quad (3.4)$$

Furthermore, the constants  $T_1$ ,  $c_6$ ,  $c_7$  and  $q$  are completely determined by  $c_3$  in Lemma 3.1,  $c_4$ ,  $c_5$ , and  $m$  in (3.3),  $p$ , and the characteristics of  $\tau$ .

Our goal is to verify the hypothesis of Theorem 2.3. Firstly, note that (2.7) can be reduced to a more basic estimate by means of the following observation.

Let  $\Delta(t, x)$  be as defined in Theorem 2.2. Then for every  $q \geq 1$  and every bounded set  $V \subset \mathbf{R}^d$  there exists a positive constant  $c_8$  such that for all  $t \in (0, T)$  and  $x \in V$ , we have

$$\|\Delta(t, x)^{-1}\|_{2q}^{2q} \leq c_8 \left\{ 1 + \sum_{j=1}^{\infty} P\left(Q(t, x) < j^{-\frac{1}{2dq}}\right) \right\}, \quad (3.5)$$

where

$$Q(t, x) := \inf \left\{ \sum_{i=1}^n \int_0^t \langle Z^x(u) X_i(\xi^x(u)), h \rangle^2 du : h \in \mathbf{R}^d, |h| = 1 \right\}. \quad (3.6)$$

Secondly, we state the following lemma which contains a local version of the hypotheses underlying Theorem 1.1.

**Lemma 3.4.**

Suppose that the conditions of Theorem 1.1 are satisfied. Then for every  $x \in D$  there exists an integer  $m \geq 1$  such that exactly one of the following two conditions holds:

- (a)  $\lambda^{(m)}(x) > 0$ .

(b) There exists an open neighborhood  $U \subseteq D$  of  $x$ , a  $C^2$  function  $\phi : U \rightarrow \mathbf{R}$ , and an exponent  $p \in (-1, 0)$  such that

- (i)  $\phi(x) = 0$  and  $\nabla \phi(x) \cdot X_i(x) \neq 0$ , for at least one  $i = 1, \dots, n$ ,
- (ii)  $\lambda^{(m)}(y) \geq \exp(-|\phi(y)|^p)$ , for all  $y \in U$ .

We now assume the hypotheses and notations of Theorem 1.1. Without loss of generality we take  $D = \mathbf{R}^d$ . Suppose that  $m$  is an integer for which the conclusion of Lemma 3.4 holds for the point  $x_0 \in \mathbf{R}^d$ . Assume that  $t \in (0, T)$  and  $x$  lies in a fixed bounded neighborhood  $W$  of  $x_0$ . Define

$$\tau_1 := \inf \left\{ s > 0 : |\xi^x(s) - x| \vee \|Z^x(s) - I\| = \frac{1}{2} \right\} \wedge T. \quad (3.7)$$

By Lemma 3.1,  $\tau_1$  is an exponentially positive stopping time with characteristics independent of  $x \in W$ .

Let  $S^d := \{h \in \mathbf{R}^d : |h| = 1\}$  denote the unit sphere in  $\mathbf{R}^d$ . Suppose  $h \in S^d$ . Then

$$P \left( \sum_{i=1}^n \int_0^t \langle Z^x(u) X_i(\xi^x(u)), h \rangle^2 du < \epsilon \right) \leq P(A \cap E) + P(A \cap E^c),$$

where

$$A := \left( \sum_{i=1}^n \int_0^{t \wedge \tau_1} \langle Z^x(u) X_i(\xi^x(u)), h \rangle^2 du < \epsilon \right)$$

and

$$E := \left( \sum_{i=1}^n \int_0^{t \wedge \tau_1} \left[ \sum_{j=1}^n \langle Z^x(u) [X_i, X_j](\xi^x(u)), h \rangle^2 + \langle Z^x(u) \left\{ [X_i, X_0] + \frac{1}{2} \sum_{j,k=1}^n [X_i, [X_j, X_k]] \right\}(\xi^x(u)), h \rangle^2 \right] du < \epsilon^\alpha \right),$$

with  $\alpha = \frac{1}{18}$ . By the Kusuoka-Stroock-Norris lemma ([N], [B], Lemma 6.5; cf. [K-S], Theorem A.24), there exist positive constants  $c_9$  and  $c_{10}$  such that

$$P(A \cap E^c) \leq c_9 \exp(-c_{10} \epsilon^{-\alpha}). \quad (3.8)$$

The constants  $c_9$  and  $c_{10}$  are independent of  $h \in S^d$ . Note that  $E \subseteq F \cap G$ , where

$$F := \left( \sum_{i,j=1}^n \int_0^{t \wedge \tau_1} \langle Z^x(u)[X_i, X_j](\xi^x(u)), h \rangle^2 du < \epsilon^\alpha \right)$$

$$G := \left( \sum_{i=1}^n \int_0^{t \wedge \tau_1} \langle Z^x(u) \left\{ [X_i, X_0] + \frac{1}{2} \sum_{j,k=1}^n [X_i, [X_j, X_k]] \right\} (\xi^x(u)), h \rangle^2 du < \epsilon^\alpha \right).$$

Thus

$$P \left( \sum_{i=1}^n \int_0^t \langle Z^x(u) X_i(\xi^x(u)), h \rangle^2 du < \epsilon \right) \leq c_9 \exp(-c_{10} \epsilon^{-\alpha}) + P(A \cap F \cap G). \quad (3.9)$$

Applying a similar argument to  $P(A \cap F \cap G)$  gives

$$P(A \cap F \cap G) \leq c_{11} \exp(-c_{12} \epsilon^{-\alpha^2}) + P(A \cap F \cap G \cap H) \quad (3.10)$$

where

$$H := \left( \sum_{i,j,k=1}^n \int_0^{t \wedge \tau_1} \langle Z^x(u) [X_i, [X_j, X_k]](\xi^x(u)), h \rangle^2 du < \epsilon^{\alpha^2} \right).$$

It is easy to check that

$$G \cap H \subseteq \left( \sum_{i=1}^n \int_0^{t \wedge \tau_1} \langle Z^x(u) [X_i, X_0](\xi^x(u)), h \rangle^2 du < \epsilon^{r_1} \right)$$

for some  $r_1 \in (0, 1)$  and sufficiently small  $\epsilon > 0$ .

Thus

$$A \cap F \cap G \cap H \subseteq \left( \int_0^{t \wedge \tau_1} \left\{ \langle \sum_{i=1}^n (Z^x(u) X_i(\xi^x(u)), h \rangle^2 \right. \right. \\ \left. \left. + \sum_{i,j=0}^n \langle Z^x(u) [X_i, X_j](\xi^x(u)), h \rangle^2 \right\} du < \epsilon^{r_2} \right)$$

for some  $r_2 \in (0, 1)$  and sufficiently small  $\epsilon > 0$ .

Combining this with (3.9) and (3.10), one obtains

$$P \left( \sum_{i=1}^n \int_0^t \langle Z^x(u) X_i(\xi^x(u)), h \rangle^2 du < \epsilon \right) \\ \leq c_9 \exp(-c_{10} \epsilon^{-\alpha}) + c_{11} \exp(-c_{12} \epsilon^{-\alpha^2}) \\ + P \left( \int_0^{t \wedge \tau_1} \left\{ \sum_{i=1}^n \langle Z^x(u) X_i(\xi^x(u)), h \rangle^2 + \sum_{i,j=0}^n \langle Z^x(u) [X_i, X_j](\xi^x(u)), h \rangle^2 \right\} du < \epsilon^{r_2} \right). \quad (3.11)$$

Iterating the argument used to derive (3.11) proves the following:

For each  $m \geq 1$ , there exist positive constants  $c_{13}$  and  $c_{14}$  and exponents  $r_3$  and  $r_4 \in (0, 1)$ , all independent of  $h \in S^d$ , such that for all  $t \in (0, T)$ ,  $x \in W$ , and  $\epsilon \in (0, c_{14})$ , one has

$$\begin{aligned} & P\left(\sum_{i=1}^n \int_0^t \langle Z^x(u) X_i(\xi^x(u)), h \rangle^2 du < \epsilon\right) \\ & \leq \exp(-c_{14}\epsilon^{-r_3}) + P\left(\sum_{j=1}^N \int_0^{t \wedge \tau_1} \langle Z^x(u) K_j(\xi^x(u)), h \rangle^2 du < \epsilon^{r_4}\right). \end{aligned} \quad (3.12)$$

Here the vector fields  $K_1, \dots, K_N$  are the columns of the matrix function  $X^{(m)}$ . Applying a straightforward compactness argument (cf. [B], Lemma 6.8) to (3.12), one obtains

$$\begin{aligned} P(Q(t, x) < \epsilon) & \leq \exp(-c_{15}\epsilon^{-r_3}) \\ & + c_{16}\epsilon^{-d} \sup\left\{P\left(\sum_{j=1}^N \int_0^{t \wedge \tau_1} \langle Z^x(u) K_j(\xi^x(u)), h \rangle^2 du < c_{17}\epsilon^{r_4}\right) : |h| = 1\right\} \end{aligned} \quad (3.13)$$

for  $\epsilon \in (0, c_{18})$  and positive constants  $c_{15}, c_{16}, c_{17}, c_{18}$ .

The argument used to derive the estimate (3.13) above is due to Kusuoka and Stroock ([K-S]).

Since (3.8) implies  $\|Z^x(u) - I\| \leq \frac{1}{2}$ , for all  $0 \leq u \leq \tau_1$ , it is easy to deduce from (3.13) that

$$P(Q(t, x) < \epsilon) \leq \exp(-c_{15}\epsilon^{-r_3}) + c_{16}\epsilon^{-d} P\left(\int_0^{t \wedge \tau_1} \lambda^{(m)}(\xi^x(u)) du < c'_{18}\epsilon^{r_4}\right). \quad (3.14)$$

We now consider each of the two cases (a) and (b) delineated in the conclusion of Lemma 3.4. Suppose first that (a) holds at  $x_0$  for some  $m \geq 1$ . Then by continuity of  $\lambda^{(m)}$  there exist  $\rho > 0$  and  $\delta > 0$  such that

$$\lambda^{(m)}(y) \geq \delta \quad (3.15)$$



for all  $y \in B_\rho(x_0)$ , where  $B_\rho(x_0)$  denotes the open ball in  $\mathbf{R}^d$  with center  $x_0$  and radius  $\rho$ . Let  $V := B_{\rho/2}(x_0)$ , assume  $x \in V$ , and let  $\tau_2$  denote the first exit time of  $\xi^x$  from  $V$ . Then (3.14) and (3.15) imply

$$P(Q(t, x) < \epsilon) \leq \exp(-c_{15}\epsilon^{-r_3}) + c_{16}\epsilon^{-d}P\left(\tau_1 \wedge \tau_2 \wedge t < \frac{c_{18}\epsilon^{r_4}}{\delta}\right) \quad (3.16)$$

$$\leq c_{19} \exp(-c_{20}\epsilon^{-c_{21}r_5}) \quad (3.17)$$

provided  $t > \frac{c_{18}\epsilon^{r_4}}{\delta}$ , where  $r_5 := r_3 \wedge r_4$  and  $c_{19}$ ,  $c_{20}$  and  $c_{21}$  are positive constants, independent of  $(t, x) \in (0, T) \times V$ . Substituting (3.17) into (3.5) yields, for every  $q \geq 1$ , the following inequality

$$\|\Delta(t, x)^{-1}\|_{2q}^{2q} \leq c_8 \left\{ \left( \frac{\delta t}{c_{18}} \right)^{-\frac{2dq}{r_4}} + A(t) \right\},$$

where

$$\begin{aligned} A(t) &:= 1 + \sum_{j=k}^{\infty} c_{19} \exp(-c_{20}j^{r_6}), \\ &\leq 1 + \sum_{j=1}^{\infty} c_{19} \exp(-c_{20}j^{r_6}) < \infty, \end{aligned}$$

$r_6 := \frac{c_{21}r_5}{2dq} > 0$ , and  $k := \left\lceil \left( \frac{\delta t}{c_{18}} \right)^{-\frac{2dq}{r_4}} \right\rceil$  is the integer part of  $\left( \frac{\delta t}{c_{18}} \right)^{-\frac{2dq}{r_4}}$ . We conclude that  $\|\Delta(t, x)^{-1}\|_{2q}$  grows no faster than a power of  $t$  as  $t \downarrow 0$ , uniformly with respect to  $x \in V$ . Hence (2.7) is satisfied.

We now turn to the case where (b) of Lemma 3.4 holds at the point  $x_0$ . By the transversality condition b(i) we may choose  $\rho > 0$  small enough to ensure that  $B_\rho(x_0) \subset U$  and such that

$$|\nabla\phi(x) \cdot X_i(x)| \geq \frac{1}{2} |\nabla\phi(x_0) \cdot X_i(x_0)| > 0$$

for some  $1 \leq i \leq n$  and every  $x \in B_\rho(x_0)$ . Let  $V := B_{\rho/2}(x_0)$ . Assume  $x \in V$  and let  $\tau_3$  denote the first exit time of  $\xi^x$  from  $B_{\rho/2}(x)$ . In view of Lemma 3.4 (b)(ii), (3.14) implies

$$P(Q(t, x) < \epsilon) \leq \exp(-c_{15}\epsilon^{-r_3}) + c_{16}\epsilon^{-d}P\left(\int_0^{t \wedge \tau_1 \wedge \tau_3} \exp(-|\eta^x(u)|^p) du < c_{18}\epsilon^{r_4}\right) \quad (3.18)$$

where  $\eta^x(t)$  denotes the process  $\phi(\xi^x(t))$ ,  $t \leq \tau_3$ . Applying Itô's lemma to compute  $\eta^x(t)$  gives

$$d\eta^x(t) = \sum_{i=1}^n \nabla \phi(\xi^x(t)) \cdot X_i(\xi^x(t)) dW_i(t) + (L - c)\phi(\xi^x(t)) dt. \quad (3.19)$$

Lemma 3.4 (b)(i), Lemma 3.1, and (3.19) imply that the process  $y := \eta^x$  and the stopping time  $\tau := \tau_1 \wedge \tau_2$  satisfy the hypotheses of Lemma 3.2. Hence (3.3) is satisfied for *every*  $m > 1$  with  $\tau = \tau_1 \wedge \tau_3$  and  $y = \eta^x$ . Thus, by Lemma 3.3 there exist positive constants  $c_6, c_7, T_1$  and  $q' > 1$ , all independent of  $x \in V$ , such that for all  $t \in (0, T_1)$  and  $\epsilon < \exp(-c_6 t^{-\frac{1}{q'}})$

$$P\left(\int_0^{t \wedge \tau_1 \wedge \tau_3} \exp(-|\eta^x(u)|^p) du < \epsilon\right) < \exp\{-c_7 |\log \epsilon|^{q'}\}. \quad (3.20)$$

Substituting this into (3.18) gives

$$P(Q(t, x) < \epsilon) \leq \exp(-c_{15} \epsilon^{-r_3}) + c_{16} \epsilon^{-d} \exp(-c_7 |\log \epsilon^{r_4}|^{q'}) \quad (3.21)$$

for  $t \in (0, T_1)$  and  $\epsilon < \exp(-c_6 t^{-\frac{1}{q'}})$ . Combining (3.21) with (3.5), we arrive at

$$\|\Delta(t, x)^{-1}\|_{2q}^{2q} \leq c_8 \{\exp(2dq c_6 t^{-\frac{1}{q'}}) + c_{22}\}, \quad 0 < t < T_1 \quad (3.22)$$

where

$$c_{22} := 1 + \sum_{j=1}^{\infty} \left\{ \exp\left(-c_{15} j^{\frac{r_3}{2dq}}\right) + c_{16} j^{1/2q} \exp(-c_7 |\log j^{-\frac{r_4}{2dq}}|^{q'}) \right\} < \infty.$$

Note that the constants  $c_6, c_8$  and  $c_{22}$  can all be chosen to be independent of  $x \in V$ . The right hand side of (3.22) *explodes exponentially fast* as  $t \downarrow 0$ . However, since  $q' > 1$  we conclude that (2.7) holds also for this case, and the proof of Theorem 1.1 is complete.  $\square$

*Proof of Theorem 1.0.*

We employ a technique used by [K-S]. The idea is to imbed the operator  $L$  of Theorem 1.0 in an another operator  $\tilde{L}$  defined on a  $(d+1)$ -dimensional domain and satisfying

all the conditions of Theorem 1.1. The operator  $\tilde{L}$  would then be parabolic hypoelliptic, and this will then imply that  $L$  is hypoelliptic on  $D$ . Assume that the operator  $L$  satisfies all the hypotheses of Theorem 1.0. Pick a smooth non-negative real-valued function  $\rho$  on  $(0, 1)$  such that both  $\rho(s)$  and its derivative  $\rho'(s)$  are bounded away from zero for all  $s \in (0, 1)$ . Define the operator  $\tilde{L}$  on the domain  $D \times (0, 1)$  by  $\tilde{L} := \rho(s)L + \frac{1}{2} \frac{\partial^2}{\partial s^2}$ . It is easy to see that  $\tilde{L}$  has the Hörmander form

$$\tilde{L} := \frac{1}{2} \sum_{i=1}^{n+1} \tilde{X}_i^2 + \tilde{X}_0 + \tilde{c}.$$

where  $\tilde{X}_0(x, s) := \rho(s)X_0(x)$ ,  $\tilde{X}_i(x, s) := \rho(s)^{1/2} X_i(x)$ ,  $1 \leq i \leq n$ ,  $\tilde{X}_{n+1}(x, s) := \frac{\partial}{\partial s}$ ,  $\tilde{c}(x, s) := \rho(s)c(x)$  for  $(x, s) \in D \times (0, 1)$ . For each positive integer  $m$ , let  $\tilde{X}^{(m)}$  be the matrix with columns  $\tilde{X}_1, \dots, \tilde{X}_{n+1}$  together with all iterated Lie brackets formed from  $\tilde{X}_0, \dots, \tilde{X}_{n+1}$  with no more than  $m$  iterations. Denote by  $\tilde{\lambda}^{(m)}(x, s)$  the smallest eigenvalue of  $\tilde{X}^{(m)}(x, s) \tilde{X}^{(m)*}(x, s)$ , for  $(x, s) \in D \times (0, 1)$ . Since  $\rho(s)$  and  $\rho'(s)$  are bounded away from zero for all  $s \in (0, 1)$ , it is easy to see that there are positive constants  $\delta_m$  such that  $\tilde{\lambda}^{(m)}(x, s) \geq \delta_m \mu^{(m)}(x)$  for all  $(x, s) \in D \times (0, 1)$ , and all  $m \geq 1$ . Let the parabolic non-Hörmander set of the operator  $\tilde{L}$  be  $\tilde{H}^c \subset D \times (0, 1)$ . Then it follows from the last inequality that  $\tilde{H}^c \subseteq H_e^c \times (0, 1)$ . Since  $H_e^c$  is contained in a  $C^2$  hypersurface  $M$  in  $D$ , then  $\tilde{H}^c$  is contained in the  $C^2$  hypersurface  $M \times (0, 1)$  in  $D \times (0, 1)$ . By the hypotheses of Theorem 1.0, at least one the vector fields  $X_1, \dots, X_n$  is transversal to  $M$  at every point of  $H_e^c$ . Therefore one of the vectors fields  $\tilde{X}_1, \dots, \tilde{X}_{n+1}$  is transversal to  $M \times (0, 1)$  at every point of  $\tilde{H}^c$ . The exponential degeneracy condition of Theorem 1.0 easily implies that  $\tilde{\lambda}^{(m)}$  satisfies the corresponding degeneracy condition of Theorem 1.1 on the set  $\tilde{H}^c$ . Hence we conclude from Theorem 1.1 that the operator  $\tilde{L} + \frac{\partial}{\partial t}$  is hypoelliptic on  $\mathbf{R} \times D \times (0, 1)$ . Therefore  $\tilde{L}$  is hypoelliptic on  $D \times (0, 1)$ , and consequently  $L$  is hypoelliptic on  $D$ .  $\square$

*Proof of Theorem 1.2.*

The proof of Theorem 1.2 requires the following modified form of Lemma 3.2.

**Lemma 3.5.**

Suppose that (1.4) holds, and let  $\eta^x$  denote the Itô process in (3.19). Let  $\tau$  be an exponentially positive  $(\mathcal{F}_t)_{0 \leq t \leq T}$ -stopping time. Then there exist positive constants  $c'_1, c'_2$ , and  $c'_3$ , depending only on the characteristics of  $\tau$ , such that for all  $t \in (0, c'_1)$  and  $\epsilon \in (0, c'_2 t^{(18)^r})$ , one has

$$P\left(\int_0^{t \wedge \tau} |\eta^x(u)|^2 du < \epsilon\right) < \exp(-c'_3 \epsilon^{-\frac{1}{(18)^r}}). \quad (3.23)$$

The argument used to prove Lemma 3.3 now shows that (3.4) is satisfied with  $y := \eta^x$ , provided  $p$  lies in the restricted range  $(-\frac{2}{(18)^r}, 0)$ . Once this is established, the proof of Theorem 1.2 follows along the same lines as that of Theorem 1.1.  $\square$

**4. Proofs of lemmas.**

A proof of Lemma 3.1 can be found in ([I-W], Lemma 10.5, p. 398). Our proof of Lemma 3.2 builds upon two preliminary results, which we first state and prove.

**Proposition 4.1.**

Suppose  $m \geq 2$  and  $a > 0$ . Let  $B : [0, \infty) \times \Omega \rightarrow \mathbf{R}$  be a one-dimensional Brownian motion. Then there exists a positive constant  $c_{23}$  such that

$$P\left(\int_0^a |B(u)|^m du < \epsilon\right) \leq \sqrt{2} \exp\left(-c_{23} a^{1+\frac{2}{m}} \epsilon^{-\frac{2}{m}}\right)$$

for every  $\epsilon > 0$ . The constant  $c_{23}$  may be chosen to be  $2^{-7}$ .

*Proof.*

The result is known to hold for  $m = 2$ , with  $c_{23} = 2^{-7}$  (cf. [I-W], Lemma V.10.6, p. 399).

For  $m > 2$  we apply Hölder's inequality and use the result for  $m = 2$  to obtain

$$\begin{aligned} P\left(\int_0^a |B(u)|^m du < \epsilon\right) &\leq P\left(\int_0^a |B(u)|^2 du < a^{1-\frac{2}{m}} \epsilon^{\frac{2}{m}}\right) \\ &\leq \sqrt{2} \exp(-c_{23} a^{1+\frac{2}{m}} \epsilon^{-\frac{2}{m}}) \end{aligned}$$

for every  $\epsilon > 0$ . This proves the proposition.  $\square$

**Proposition 4.2.**

Assume the notation and hypotheses of Lemma 3.2. Then for every  $m \geq 2$  and  $q > 0$ , there exist positive constants  $c_{24}$  and  $c_{25}$  such that for all  $\epsilon > 0$ ,

$$P\left(\int_0^\tau |y(u)|^m du < \epsilon, \tau \geq \epsilon^q\right) < c_{24} \exp\left(-c_{25} \epsilon^{q + \frac{2(q-1)}{m}}\right).$$

The constants  $c_{24}$  and  $c_{25}$  depend only on  $c_3$  (the bound for the drift and diffusion coefficients of  $y$ ),  $\delta$ ,  $m$ , and  $q$ . In particular, they are independent of  $y(0)$  and the characteristics of  $\tau$ .

*Proof.*

Expressing (3.1) in coordinates, it is sufficient to treat the case  $d = 1$ ,  $n \geq 1$ . In this case, the process  $y$  may be written in the form

$$y(t) = B(\tau_4(t)) + \int_0^t b(u) du, \quad 0 \leq t \leq T,$$

where

$$\tau_4(t) := \int_0^t |a(u)|^2 du, \quad 0 \leq t \leq T,$$

and  $B : [0, \infty) \times \Omega \rightarrow \mathbf{R}$  is a one-dimensional  $(\mathcal{F}_{\tau_4^{-1}(t)} : 0 \leq t \leq T)$ -adapted Brownian motion started at  $B(0) = y(0)$ . By assumption, we may take  $|a(u)| \geq \delta > 0$  for  $0 \leq u \leq \tau$ . Hence  $\tau \geq \epsilon^q$  implies  $\tau_4(\tau) \geq \delta^2 \epsilon^q$ . Furthermore, the function  $\tau_4(t)$  is strictly increasing on  $(0, \tau)$  and changes of the time variable yield

$$\begin{aligned} \int_0^\tau |y(u)|^m du &\geq c_{26}^{-1} \int_0^{\tau_4(\tau)} |y(\tau_4^{-1}(s))|^m ds \\ &= c_{26}^{-1} \int_0^{\tau_4(\tau)} \left| B(s) + \int_0^s \frac{b(\tau_4^{-1}(u))}{|a(\tau_4^{-1}(u))|^2} du \right|^m ds, \end{aligned}$$

where  $c_{26} := nc_3^2$ . Thus

$$\begin{aligned} &P\left(\int_0^\tau |y(u)|^m du < \epsilon, \tau \geq \epsilon^q\right) \\ &\leq P\left(\int_0^{\delta^2 \epsilon^q} \left| B(s) + \int_0^s \frac{b(\tau_4^{-1}(u))}{|a(\tau_4^{-1}(u))|^2} du \right|^m ds < c_{26} \epsilon, \tau_4(\tau) \geq \delta^2 \epsilon^q\right). \end{aligned} \quad (4.1)$$

We now define a (bounded) process  $h : [0, \infty) \times \Omega \rightarrow \mathbf{R}$  by

$$h(u) := \begin{cases} \frac{b(\tau_4^{-1}(u))}{|a(\tau_4^{-1}(u))|^2}, & u \in (0, \tau_4(\tau)) \\ \frac{b(\tau)}{|a(\tau)|^2}, & u \geq \tau_4(\tau). \end{cases}$$

and we denote by  $B'$  the process

$$B'(s) := B(s) + \int_0^s h(u) du, \quad 0 \leq s \leq \tau_4(T).$$

By the Girsanov theorem,  $B'$  is an  $(\mathcal{F}_{\tau_4^{-1}(t)} : 0 \leq t \leq T)$ -adapted Brownian motion on  $\Omega$  with respect to the measure

$$dP' := \left\{ \exp \left( - \int_0^{\tau_4(T)} h(u) dB(u) - \frac{1}{2} \int_0^{\tau_4(T)} h^2(u) du \right) \right\} dP.$$

Denote by  $\Omega_\epsilon$  the event

$$\Omega_\epsilon := \left( \int_0^{\delta^2 \epsilon^q} |B'(s)|^m ds < c_{26} \epsilon \right),$$

and by  $G$  the Girsanov density

$$G := \exp \left( - \int_0^{\tau_4(T)} h(u) dB(u) - \frac{1}{2} \int_0^{\tau_4(T)} h^2(u) du \right).$$

We now apply Hölder's inequality to (4.1) to obtain

$$\begin{aligned} P \left( \int_0^\tau |y(u)|^m du < \epsilon, \tau \geq \epsilon^q \right) &\leq P(\Omega_\epsilon) \\ &\leq \sqrt{E(G^{-2})P'(\Omega_\epsilon)}. \end{aligned}$$

By Proposition 4.1, we have

$$P'(\Omega_\epsilon) \leq \sqrt{2} \exp \left( -2c_{25} \epsilon^{q + \frac{2(q-1)}{m}} \right), \quad (4.2)$$

where  $c_{25} := \frac{1}{2} c_{23} c_{26}^{-\frac{2}{m}} \delta^{2(1 + \frac{2}{m})}$ . The boundedness of  $h$  and  $\tau_4(T)$  imply the existence of a constant  $c_{27}$ , depending only on the bounds of the foregoing quantities, such that

$$G^{-2} \leq c_{27} \exp \left( 2 \int_0^{\tau_4(T)} h(u) dB(u) - 2 \int_0^{\tau_4(T)} h^2(u) du \right). \quad (4.3)$$

The desired conclusion follows from (4.2) and (4.3), together with the fact that the exponential on the right hand side of (4.3) is a Girsanov density (note that it is obtained by replacing  $h$  by  $(-2h)$  in the relation defining  $G$ ), and therefore has expectation equal to

1.  $\square$

*Proof of Lemma 3.2.*

Note that for every  $q > 0$ , we may write

$$P\left(\int_0^{t \wedge \tau} |y(u)|^m du < \epsilon\right) \leq P_1 + P_2, \quad (4.4)$$

where

$$P_1 := P\left(\int_0^{t \wedge \tau} |y(u)|^m du < \epsilon, t \wedge \tau \geq \epsilon^q\right)$$

and

$$P_2 := P(t \wedge \tau < \epsilon^q).$$

By Proposition 4.2,

$$P_1 < c_{24} \exp\left(-c_{25} \epsilon^{q + \frac{2(q-1)}{m}}\right), \quad (4.5)$$

where  $c_{24}$  and  $c_{25}$  are independent of  $t$ . Now  $\tau$  is exponentially positive; so if  $T_0 > t > \epsilon^q$ , then

$$P_2 < \exp(-c_{27} \epsilon^{-q}), \quad (4.6)$$

where  $c_{27}$  and  $T_0$  denote the characteristics of  $\tau$ . Combining (4.4), (4.5) and (4.6), we obtain

$$P\left(\int_0^{t \wedge \tau} |y(u)|^m du < \epsilon\right) \leq c_{24} \exp\left(-c_{25} \epsilon^{q + \frac{2(q-1)}{m}}\right) + \exp(-c_{27} \epsilon^{-q})$$

for  $t \in (0, T_0)$  and  $0 < \epsilon < t^{\frac{1}{q}}$ . The lemma now follows by choosing as  $q$  the value for which the two exponents  $\left\{q + \frac{2(q-1)}{m}\right\}$  and  $-q$  coincide, namely  $\frac{1}{m+1}$ .  $\square$

*Proof of Lemma 3.3.*

Choose and fix  $m \geq \max\left\{-\frac{p}{1+p}, 2\right\}$  and set  $q := -\frac{m}{p(m+1)}$ ; so  $q > 1$ . Define a function  $\psi : [0, \infty) \rightarrow [0, 1)$  by

$$\psi(z) := \begin{cases} \exp(-z^{\frac{p}{m}}) & z > 0 \\ 0 & z = 0. \end{cases}$$

Note that

- (i)  $\psi$  is strictly increasing;
- (ii)  $\psi$  is convex in an interval  $(0, c_{28})$ , for some positive constant  $c_{28}$ .

Furthermore

$$P\left(\int_0^{t \wedge \tau} \exp(-|y(u)|^p) du < \epsilon\right) = P\left(\int_0^{t \wedge \tau} \psi(|y(u)|^m) du < \epsilon\right).$$

We break the proof of the lemma into two cases. Firstly, suppose that  $|y(0)| \geq c_{29}$  ( $:= (\frac{1}{2}c_{28})^{\frac{1}{m}}$ ). Let  $\tau_5 := \inf\{s > 0 : |y(s) - y(0)| = \frac{1}{2}c_{29}\} \wedge \tau$ . Then there exists a positive constant  $c_{30}$ , determined by  $m$ ,  $p$ , and  $\delta$ , such that

$$P\left(\int_0^{t \wedge \tau} \exp(-|y(u)|^p) du < \epsilon\right) \leq P(t \wedge \tau_5 \leq c_{30}\epsilon). \quad (4.7)$$

Applying Lemma 3.1 to the right hand side of (4.7), we deduce the existence of positive constants  $c_{31}$  and  $c_{32}$ , such that if  $t \in (0, c_{31})$  and  $0 < \epsilon < \frac{t}{c_{30}}$ , then

$$P\left(\int_0^{t \wedge \tau} \exp(-|y(u)|^p) du < \epsilon\right) \leq \exp\left(-\frac{c_{32}}{\epsilon}\right).$$

The constants  $c_{30}, c_{31}, c_{32}$  depend only on  $m, p, c_3$ , and the characteristics of  $\tau$ . Thus the conclusion of the lemma holds in this case.

Alternatively, suppose that  $|y(0)| < c_{29}$ . We now set

$$\tau_6 := \inf\{s > 0 : |y(s)|^m = c_{28}\} \wedge \tau.$$

Jensen's inequality yields

$$\begin{aligned} P\left(\int_0^{t \wedge \tau} \psi(|y(u)|^m) du < \epsilon\right) &\leq P\left(\int_0^{t \wedge \tau_6} |y(u)|^m du \leq (t \wedge \tau_6) \psi^{-1}\left(\frac{\epsilon}{t \wedge \tau_6}\right)\right) \\ &\leq P_1 + P_2 \end{aligned}$$

where

$$P_1 := P\left(t \wedge \tau_6 \leq \frac{\epsilon}{c_{28}}\right)$$



and

$$P_2 := P\left(\int_0^{t \wedge \tau_6} |y(u)|^m du \leq (t \wedge \tau_6) \psi^{-1}\left(\frac{\epsilon}{t \wedge \tau_6}\right), \quad t \wedge \tau_6 > \frac{\epsilon}{c_{28}}\right).$$

Note that  $P_1$  is of the same form as the probability on the right hand side of (4.7), and hence satisfies a similar estimate.

We now consider  $P_2$ . An elementary argument shows that the convexity of  $\psi$  in the interval  $(0, c_{28})$  implies that the function  $\theta(u) := u\psi^{-1}\left(\frac{\epsilon}{u}\right)$  is increasing for  $u > \frac{\epsilon}{c_{28}}$ . In particular, if  $t \wedge \tau_6 > \frac{\epsilon}{c_{28}}$  then

$$(t \wedge \tau_6) \psi^{-1}\left(\frac{\epsilon}{t \wedge \tau_6}\right) \leq T \psi^{-1}\left(\frac{\epsilon}{T}\right)$$

where  $T$  is any upper bound for  $t$ . This implies

$$P_2 \leq P\left(\int_0^{t \wedge \tau_6} |y(u)|^m du \leq T \psi^{-1}\left(\frac{\epsilon}{T}\right)\right). \quad (4.8)$$

We now apply Lemma 3.2 to estimate the right hand side of (4.8). Thus

$$\begin{aligned} P_2 &< \exp\left(-c_5 \left\{T \psi^{-1}\left(\frac{\epsilon}{T}\right)\right\}^{-\frac{1}{m+1}}\right) \\ &\leq \exp(-c_7 |\log \epsilon|^q), \end{aligned} \quad (4.9)$$

for all  $0 < t < c_{34}$  and  $\epsilon < \exp\left(-c_{33} t^{-\frac{1}{q}}\right)$ , where  $c_7$ ,  $c_{33}$ , and  $c_{34}$  are positive constants exhibiting the appropriate dependence. Clearly, (4.9) gives an estimate for  $P_2$  of the required form, and the proof of the lemma is complete.  $\square$

*Proof of Lemma 3.4.*

It is sufficient to show that the conditions of Theorem 1.1 imply that conclusions (b)(i),(ii) of Lemma 3.4 hold for every  $x \in H^c$ .

Suppose all the conditions of Theorem 1.1 are satisfied. Fix any  $x \in H^c$ . Since  $N$  is a  $C^2$  hypersurface in  $\mathbf{R}^d$ , there is a  $C^2$  chart  $(V, \theta)$  centered at  $x$  such that  $\theta :=$

$(\theta_1, \theta_2) : V \rightarrow \mathbf{R}^{d-1} \times \mathbf{R}$  is a  $C^2$  diffeomorphism onto the open set  $\theta(V)$ ,  $\theta(x) = (0, 0)$ , and  $\theta(N \cap V) = (\mathbf{R}^{d-1} \times \{0\}) \cap \theta(V)$ . Furthermore the coordinate maps  $\theta_1 : V \rightarrow \mathbf{R}^{d-1}$  and  $\theta_2 : V \rightarrow \mathbf{R}$  are  $C^2$  with  $\nabla\theta_2(z) \neq 0$  for all  $z \in V$  and  $N \cap V = \theta_2^{-1}\{0\} \cap V$ . By the transversality hypothesis, there exists  $i = 1, \dots, n$  such that  $X_i(x) \notin T_x N$ , the tangent space to  $N$  at  $x$ . Now  $T_x N = [D\theta_2(x)]^{-1}\{0\}$ , where  $D\theta_2(x)$  is the Fréchet derivative of  $\theta_2$  at  $x$ . Thus  $\nabla\theta_2(x) \cdot X_i(x) \neq 0$ . Let  $U_1$  be the open neighborhood of  $x$  given in the statement of Theorem 1.1. Choose an open ball  $V_1 := B(x, \delta_1) \subseteq U_1 \cap V$  centered at  $x$  and with radius  $\delta_1 > 0$ . Define  $V_2 \subset V_1$  to be the ball  $B(x, \delta_1/2)$  of center  $x$  and radius  $\delta_1/2$ . Then a simple argument using the triangle inequality shows that  $\rho(y, N) = \rho(y, N \cap V_1)$  for all  $y \in V_2$ . Now let  $y \in V_2$  and  $z \in N \cap V_1$ . Then by the Lipschitz property of  $\theta$  there is a positive constant  $k$  such that

$$|y - z| > k|\theta(y) - \theta(z)| \geq k|\theta_2(y)|.$$

Therefore  $\rho(y, N) \geq k|\theta_2(y)|$  for all  $y \in V_2$ . Set  $U := V_2$  and  $\phi := k\theta_2|_U$ . If  $m \geq 1$  and  $p \in (-1, 0)$  are as in Theorem 1.1, it is easy to see that

$$\lambda^{(m)}(y) \geq \exp\{-[\rho(y, N)]^p\} \geq \exp\{-|\phi(y)|^p\},$$

for all  $y \in U$ . Therefore (b)(ii) is satisfied, and the proof of Lemma 3.4 is complete.  $\square$

*Proof of Lemma 3.5.*

Suppose  $f : \mathbf{R}^d \rightarrow \mathbf{R}$  is any  $C^\infty$  function. Recall that  $L - c$  has been relabeled as  $X_{n+1}$ . For each  $1 \leq j \leq n+1$ , one has

$$P\left(\int_0^{t \wedge \tau} |f(\xi^x(u))|^2 du < \epsilon\right) \leq P_1 + P_2$$

where

$$P_1 := P\left(\int_0^{t \wedge \tau} |X_j f(\xi^x(u))|^2 du < \epsilon^{\frac{1}{18}}\right)$$

and

$$P_2 := P\left(\int_0^{t \wedge \tau} |f(\xi^x(u))|^2 du < \epsilon, \sum_{i=1}^{n+1} \int_0^{t \wedge \tau} |X_i f(\xi^x(u))|^2 du \geq \epsilon^\alpha\right).$$

where  $\alpha = \frac{1}{18}$ .

Note that by Itô's lemma,

$$df(\xi^x(t)) = \sum_{i=1}^n X_i f(\xi^x(t)) dW_i(t) + (L - c)f(\xi^x(t)) dt.$$

The Kusuoka-Stroock-Norris lemma shows that, for sufficiently small  $\epsilon > 0$ , one has

$$P_2 \leq \exp(-c_{35}\epsilon^{-\alpha}),$$

where  $\alpha = \frac{1}{18}$ . Thus, for  $1 \leq j \leq n + 1$  and sufficiently small  $\epsilon$ ,

$$P\left(\int_0^{t \wedge \tau} |f(\xi^x(u))|^2 du < \epsilon\right) \leq \exp(-c_{35}\epsilon^{-\alpha}) + P\left(\int_0^{t \wedge \tau} |X_j f(\xi^x(u))|^2 du < \epsilon^\alpha\right).$$

We now iterate the above relation  $r$  times, starting with the smooth function  $\phi$  as the initial choice for  $f$ . This gives the inequality

$$\begin{aligned} P\left(\int_0^{t \wedge \tau} |\eta^x(u)|^2 du < \epsilon\right) &\leq \exp(-c_{35}\epsilon^{-\alpha}) + \exp(-c_{36}\epsilon^{-\alpha^2}) + \cdots + \exp(-c_{35+r}\epsilon^{-\alpha^r}) + \\ &+ P\left(\int_0^{t \wedge \tau} |X_{i_1} X_{i_2} \cdots X_{i_r} \phi(\xi^x(u))|^2 du < \epsilon^{\alpha^r}\right). \end{aligned} \quad (4.11)$$

Now  $X_{i_1} X_{i_2} \cdots X_{i_r} \phi$  is a continuous function and  $X_{i_1} X_{i_2} \cdots X_{i_r} \phi(x) \neq 0$ . Therefore by an argument similar to the one used to derive (3.13) from (3.10), we obtain positive constants  $c'_4, c'_5, c'_{37}$  depending only on the characteristics of  $\tau$ , such that

$$P\left(\int_0^{t \wedge \tau} |X_{i_1} X_{i_2} \cdots X_{i_r} \phi(\xi^x(u))|^2 du < \epsilon^{\alpha^r}\right) < \exp(-c'_{29}\epsilon^{-\alpha^r}). \quad (4.12)$$

for  $t \in (0, c'_4)$  and  $\epsilon \in (0, c'_5 t^{(18)^r})$ . Substituting (4.12) into (4.11) yields the desired conclusion. This completes the proof of Lemma 3.5.  $\square$

## References

- [B] Bell, D. R., *The Malliavin Calculus*, Pitman Monographs and Surveys in Pure and Applied Mathematics, Vol. 34, Longman, 1987.
- [B-M] Bell, D. R., and Mohammed, S.-E. A., Opérateurs paraboliques hypoelliptiques avec dégénérescences exponentielles, *C.R. Acad. Sci. Paris*, t. 317, Série I, (1993), 1059-1064.
- [F-P] Fefferman, C., and Phong, D., Subelliptic eigenvalue problems, *Conference on Harmonic Analysis*, in honor of A. Zygmund, Wadsworth, Belmont, California, 1981, 590-606.
- [H] Hörmander, L., Hypoelliptic second order differential equations, *Acta Math.* 119:3-4, 1967, 147-171.
- [I-W] Ikeda, N., and Watanabe, S., *Stochastic Differential Equations and Diffusion Processes*, Second Edition, North-Holland-Kodansha, 1989.
- [K-S] Kusuoka, S., and Stroock, D., Applications of the Malliavin calculus, Part II, *Journal of Faculty of Science, University of Tokyo*, Sec. 1A, Vol. 32, No. 1, 1985, 1-76.
- [M.1] Malliavin, P., Stochastic calculus of variations and hypoelliptic operators, *Proceedings of the International Conference on Stochastic Differential Equations*, Kyoto, Kinokuniya, 1976, 195-263..
- [M.2] Malliavin, P.,  $C^k$ -hypoellipticity with degeneracy, part II, *Stochastic Analysis*, A. Friedman and M. Pinsky (eds), 1978, 327-340.
- [Mo] Morimoto, Y., Hypoellipticity for infinitely degenerate elliptic operators, *Osaka J. Math.*, 24, 1987, 13-35.
- [N] Norris, J., Simplified Malliavin calculus, *Seminaire de Probabilites XX*, 1986, 101-130.
- [O-R] Oleinik, O. A., and Radekevich, E. V., *Second Order Equations with Non-negative Characteristic Form*, *Amer. Math. Soc.*, Providence, Rhode Island, and Plenum Press, 1973.
- [S] Stroock, D., The Malliavin calculus, a functional analytic approach, *J. Funct. Anal.* 44, 1981, 212-257.

Department of Mathematics  
University of North Florida,  
Jacksonville, Florida 32216.  
Email: dbell@unflvm.

Department of Mathematics,  
Southern Illinois University at Carbondale,  
Carbondale, Illinois 62901.  
Email: salah@c-math1.siu.edu