Available online at www.sciencedirect.com
ScienceDirect
Stochastic Processes and their Applications 118 (2008) 368–388
stochastic processes and their applications
www.elsevier.com/locate/spa

# Hypoelliptic heat kernel inequalities on Lie groups

Tai Melcher*

*Department of Mathematics, University of Virginia, Charlottesville, VA 22903, United States*

Received 26 February 2007; accepted 21 April 2007
Available online 29 April 2007

**Abstract**

This paper discusses the existence of gradient estimates for the heat kernel of a second order hypoelliptic operator on a manifold. For elliptic operators, it is now standard that such estimates (satisfying certain conditions on coefficients) are equivalent to a lower bound on the Ricci tensor of the Riemannian metric. For hypoelliptic operators, the associated "Ricci curvature" takes on the value $-\infty$ at points of degeneracy of the semi-Riemannian metric. For this reason, the standard proofs for the elliptic theory fail in the hypoelliptic setting.

This paper presents recent results for hypoelliptic operators. Malliavin calculus methods transfer the problem to one of determining certain infinite dimensional estimates. Here, the underlying manifold is a Lie group, and the hypoelliptic operators are given by the sum of squares of left invariant vector fields. In particular, "$L^p$-type" gradient estimates hold for $p \in (1, \infty)$, and the $p = 2$ gradient estimate implies a Poincaré estimate in this context.


*MSC:* primary 22E30; secondary 60H07

*Keywords:* Heat kernels; Hypoellipticity; Malliavin calculus

## 1. Introduction

### 1.1. Background

Let $M$ be a manifold of dimension $d$, and let $\{X_i\}_{i=1}^k$ be a set of smooth vector fields on $M$ satisfying

$$T_m M = \operatorname{span}(\{X(m) : X \in \mathcal{L}\}), \quad \forall m \in M, \tag{HC}$$

* Tel.: +1 434 9822782; fax: +1 434 9823084.
*E-mail address:* melcher@virginia.edu.

0304-4149/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.spa.2007.04.012

where $\mathcal{L}$ is the Lie algebra of vector fields generated by the collection $\{X_i\}_{i=1}^k$. This assumption is the *Hörmander condition*, and the collection $\{X_i\}_{i=1}^k$ is a *Hörmander set*. Under this assumption, by a celebrated theorem of Hörmander, the operator

$$L = \sum_{i=1}^{k} X_i^2 \tag{1.1}$$

is hypoelliptic. Recall that a subelliptic operator $L$ is said to be *hypoelliptic* if $Lu \in C^\infty(\Omega)$ implies that $u \in C^\infty(\Omega)$, for all distributions $u \in C^\infty(\Omega)'$ on any open set $\Omega \subset_o M$.

**Notation 1.1.** Let $C_c^\infty(M)$ denote the set of smooth functions on $M$ with compact support, and let $C_b^\infty(M)$ denote the set of smooth, bounded functions on $M$. When $M = \mathbb{R}^n$, let $C_p^\infty(\mathbb{R}^n)$ denote those functions $f \in C^\infty(\mathbb{R}^n)$ such that $f$ and all of its partial derivatives have at most polynomial growth.

Let $\nabla = (X_1, \ldots, X_k)$. This paper continues the work begun in [10], considering $L^p$-type gradient inequalities of the form

$$|\nabla e^{tL/2} f|^p \leq K_p(t) e^{tL/2} |\nabla f|^p, \quad p \in [1, \infty), \tag{1.2}$$

for $f \in C_c^\infty(M)$ and $t > 0$. For $p = 1$, (1.2) is equivalent to a one parameter family of log Sobolev estimates for the heat kernel; for $p = 2$, (1.2) is equivalent to a one parameter family of Poincaré estimates. The former has implications for the hypercontractivity of an associated semigroup; see [13,14].

When $L$ is an elliptic operator, a lower bound on the Ricci curvature is equivalent to the estimate (1.2) holding with some coefficients $K_p > 0$ such that $K_p(0) = 1$ and $\dot{K}_p(0)$ exists. In particular, in the elliptic setting, (1.2) holds with exponential coefficients $K_p(t) = e^{pkt}$, where $-2k$ is the lower bound on the Ricci curvature; see for example [2–4]. However, an operator $L$ of the form (1.1) need not be elliptic. The principle symbol of $L$ at $\xi \in T_m^* M$ is given by $\sigma_L(\xi) = \sum_{i=1}^k [\xi(X_i)]^2$. By definition, the operator $L$ is degenerate at points $m \in M$ where there exists $0 \neq \xi \in T_m^* M$ such that $\sigma_L(\xi) = 0$. At points of degeneracy of $L$, the Ricci tensor is not well defined and should be interpreted to take the value $-\infty$ in some directions. Thus, there exists no lower bound on the Ricci curvature in this case. Nevertheless, it is reasonable to ask if inequalities of the form (1.2) might still hold, perhaps with some discontinuity in the coefficients $K_p$ near $t = 0$. In particular, under what conditions do functions $K_p(t) < \infty$ exist such that (1.2) is satisfied for all $f \in C_c^\infty(M)$ and $t > 0$?

The paper [10] addressed the special case of the real three-dimensional Heisenberg Lie group, and the estimate (1.2) was proved to hold for all $p > 1$ with a constant coefficient $K_p(t) \equiv K_p$, yielding a Poincaré estimate in this case. Using analytic methods in [20], Li was able to prove (1.2) on the Heisenberg group for $p = 1$, yielding the log Sobolev estimate. Here in this paper, the case is addressed where the manifold $M$ is a general Lie group and the vector fields $\{X_i\}_{i=1}^k$ are invariant under left translation.

Related results appear in Kusuoka and Stroock [19], Picard [26], and Auscher, Coulhon, Duong, and Hofmann [1]. Also, [1,6] include some potential applications of the result proven here.

### 1.2. Statement of results

Let $G$ be a $d$-dimensional Lie group with Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$ and identity element $e$. Let $L_g$ denote left translation by an element $g \in G$, and let $R_g$ denote right translation. Suppose

$\{X_i\}_{i=1}^k \subset \mathfrak{g}$ is a linearly independent Lie generating set; that is, there exists some $m \in \mathbb{N}$ such that

$$\begin{aligned}&\text{span}\{X_i, [X_{i_1}, X_{i_2}], [X_{i_1}, [X_{i_2}, X_{i_3}]], \ldots, [X_{i_1}, [\ldots, [X_{i_{m-1}}, X_{i_m}]\cdots]] :\\ &\quad i, i_r \in \{1, \ldots, k\}, r \in \{1, \ldots, m\}\} = \mathfrak{g}.\end{aligned} \tag{1.3}$$

**Notation 1.2.** Let $\Sigma = \Sigma_0 := \{X_1, \ldots, X_k\}$, and $\Sigma_r$ be defined inductively by

$$\Sigma_r := \{[X_i, V] : V \in \Sigma_{r-1}, i = 1, \ldots, k\},$$

for all $r \in \mathbb{N}$. Since $\{X_i\}_{i=1}^k$ is a Lie generating set, there is a finite $m$ such that

$$\text{span}(\cup_{r=0}^m \Sigma_r) = \mathfrak{g}.$$

Let $\mathfrak{g}_0 := \text{span}(\Sigma_0)$, and let $\{Y_j\}_{j=1}^{d-k} \subset \cup_{r=1}^m \Sigma_r$ be a basis of $\mathfrak{g}/\mathfrak{g}_0$. Define an inner product $\langle\cdot,\cdot\rangle$ on $\mathfrak{g}$ by making $\{X_i\}_{i=1}^k \cup \{Y_j\}_{j=1}^{d-k}$ an orthonormal set. Note then that $\{X_i\}_{i=1}^k$ is an orthonormal basis of $\mathfrak{g}_0$. Extend $\langle\cdot,\cdot\rangle$ to a right invariant metric on $G$ by defining $\langle\cdot,\cdot\rangle_g : T_gG \times T_gG \to \mathbb{R}$ as

$$\langle v, w\rangle_g := \langle R_{g^{-1}*}v, R_{g^{-1}*}w\rangle, \quad \text{for all } v, w \in T_gG.$$

The $g$ subscript will be suppressed when there is no chance of confusion.

**Notation 1.3.** Given an element $X \in \mathfrak{g}$, let $\tilde{X}$ denote the left invariant vector field on $G$ such that $\tilde{X}(e) = X$, where $e$ is the identity of $G$. Recall that the $\tilde{X}$ left invariant means that the vector field commutes with left translation in the following way:

$$\tilde{X}(f \circ L_g) = (\tilde{X}f) \circ L_g,$$

for all $f \in C^1(G)$. Similarly, let $\hat{X}$ denote the right invariant vector field associated to $X$.

**Definition 1.4.** The left invariant *gradient* on $G$ is the operator on $C^1(G)$ given by

$$\nabla := (\tilde{X}_1, \ldots, \tilde{X}_k).$$

The *subLaplacian* on $G$ is the second-order operator acting on $C^2(G)$, given by

$$L := \sum_{i=1}^k \tilde{X}_i^2.$$

**Remark 1.5.** Since $\{X_i\}_{i=1}^k$ is a Lie generating set, $\{\tilde{X}_i\}_{i=1}^k$ satisfies the Hörmander condition (HC) and Hörmander's theorem [15] implies that $L$ is hypoelliptic.

Let $L^2(G)$ denote the space of square integrable functions on $G$ with respect to right invariant Haar measure. Then $L$ is a densely defined, symmetric operator on $L^2(G)$ and the symmetric bilinear form associated to $L$ is given by $\mathcal{E}^0(f_1, f_2) := (-Lf_1, f_2)_{L^2(G)}$. Note that $\mathcal{E}^0$ is positive, and so $\mathcal{E}^0$ is closable. The minimal closure $\mathcal{E}$ is associated to a self-adjoint operator $\bar{L}$, which is an extension of $L$ called the Friedrichs extension of $L$.

**Definition 1.6.** Let $P_t$ denote the *heat semigroup* $e^{t\bar{L}/2}$, where $\bar{L}$ is the Friedrichs extension of $L|_{C_c^\infty(G)}$ to $L^2(G, dg)$ with $dg$ the right Haar measure on $G$. By the left invariance of $L$ and the satisfaction of the Hörmander condition, $P_t$ admits a left convolution kernel $p_t$ such that

$$P_t f(h) = f * p_t(h) = \int_G f(hg) p_t(g) dg,$$

for all $f \in C_c^\infty(G)$. The function $p_t$ is called the *heat kernel* of $G$.

The operator $P_t$ is a symmetric Markov semigroup. By Remark 1.5, $L$ is a hypoelliptic operator, and so $p_t$ is a smooth density on $G$. In the sequel, let $L$ denote its own Friedrichs extension. For the standard semigroup theory used here, see for example [7].

**Notation 1.7.** Let $K_p(t)$ be the best function such that

$$|\nabla P_t f|^p \le K_p(t) P_t |\nabla f|^p, \quad p \in [1, \infty), \tag{$I_p$}$$

for all $f \in C_c^\infty(G)$ and $t > 0$.

**Theorem 1.8.** *For all $p \in (1, \infty)$, $K_p(t) < \infty$ for all $t > 0$. If $G$ is a nilpotent Lie group, then there exists a constant $K_p < \infty$ such that $K_p(t) \le K_p$ for all $t > 0$.*

This theorem was established in [10] in the case of the real three-dimensional Heisenberg Lie group. The method of proof in this case is analogous. The heat kernel $p_t(g)dg$ may be realized as the distribution in $t$ of the Cartan rolling map on $G$, the process $\xi$ satisfying the Stratonovich stochastic differential equation

$$\mathrm{d}\xi_t = \sum_{i=1}^{k} \tilde{X}_i(\xi_t) \circ \mathrm{d}b_t^i, \quad \text{with } \xi_0 = e,$$

where $b^1, \dots, b^k$ are $k$ independent real-valued Brownian motions. Thus, for all $f \in C_c^\infty(G)$,

$$P_t f(e) = \mathbb{E}[f(\xi_t)].$$

Sections 2.2 and 2.3 discuss properties of $\xi$. This representation of $P_t$ transforms the finite dimensional problem to a problem on a Wiener space. Section 2.4 describes a standard "lifting" procedure which constructs vector fields $\mathbf{X}_i$ on a Wiener space from the vector fields $\tilde{X}_i$ via the map $\xi$. Then Malliavin's probabilistic techniques on proving hypoellipticity give componentwise bounds of $P_t(\tilde{X}_i f)(e) = \mathbb{E}[(\tilde{X}_i f)(\xi_t)] = \mathbb{E}[\mathbf{X}_i(f(\xi_t))]$. Section 2.1 reviews some calculus on Wiener space necessary for this argument.

Section 3 contains the proof of Theorem 1.8. Results from Section 2 show that for a Lie group $G$, $K_p(t) < \infty$ for all $t > 0$; however, this method does not give any estimates on the behavior of $K_p$ with respect to $t$. In a generalization of the Heisenberg scaling argument in [10], Section 3.2 addresses the special case of nilpotent and stratified groups. When $G$ is stratified, dilation arguments imply that the coefficients $K_p$ are independent of the $t$ parameter. When $G$ is nilpotent, covering $G$ with a stratified group shows that there is a constant $K_p$ such that $K_p(t) \le K_p$ for all $t > 0$, and this completes the proof of Theorem 1.8. This implies the following Poincaré estimate for the heat kernel measure in this context.

**Theorem 1.9.** *Suppose $G$ is a nilpotent Lie group with identity element $e$. Then*

$$P_t f^2(e) - (P_t f)^2(e) \le K_2 t P_t |\nabla f|^2(e),$$

*for all $f \in C_c^\infty(G)$ and $t > 0$, where $K_2$ is the constant in Theorem 1.8 for $p = 2$.*

Note that this theorem gives an improvement in the elliptic case with negative curvature, giving linear coefficients where the estimate was previously known only with coefficients of exponential growth. This is stated explicitly in Corollary 3.16. It could be conjectured that this is true for every Riemannian manifold; that is, for any Riemannian manifold equipped with a Laplace Beltrami operator, Poincaré estimates for the associated heat kernel hold with linear coefficients.

## 2. Wiener calculus over $G$

### 2.1. *Review of calculus on Wiener space*

This section contains a brief introduction to basic Wiener space definitions and notions of differentiability. For a more complete exposition, consult [8,17,25] and the references contained therein.

Let $(\mathscr{W}(\mathbb{R}^k), \mathcal{F}, \mu)$ denote a classical $k$-dimensional Wiener space. That is, $\mathscr{W} = \mathscr{W}(\mathbb{R}^k)$ is the Banach space of continuous paths $\omega : [0, 1] \to \mathbb{R}^k$ such that $\omega_0 = 0$, equipped with the supremum norm

$$\|\omega\| = \max_{t\in[0,1]} |\omega_t|,$$

$\mu$ is the standard Wiener measure, and $\mathcal{F}$ is the completion of the Borel $\sigma$-field on $\mathscr{W}$ with respect to $\mu$. By definition of $\mu$, the process

$$b_t(\omega) = (b_t^1(\omega), \ldots, b_t^k(\omega)) = \omega_t$$

is an $\mathbb{R}^k$ Brownian motion. For those $\omega \in \mathscr{W}$ which are absolutely continuous, let

$$E(\omega) := \int_0^1 |\dot{\omega}_s|^2 \mathrm{d}s$$

denote the energy of $\omega$. The Cameron–Martin space is the Hilbert space of finite energy paths,

$$\mathscr{H} = \mathscr{H}(\mathbb{R}^k) := \{\omega \in \mathscr{W}(\mathbb{R}^k) : \omega \text{ is absolutely continuous and } E(\omega) < \infty\},$$

equipped with the inner product

$$(h, k)_{\mathscr{H}} := \int_0^1 \dot{h}_s \cdot \dot{k}_s \mathrm{d}s, \quad \text{for all } h, k \in \mathscr{H}.$$

More generally, for any finite dimensional vector space $V$ equipped with an inner product, let $\mathscr{W}(V)$ denote path space on $V$, and $\mathscr{H}(V)$ denote the set of Cameron–Martin paths, where the definitions are completely analogous, replacing the inner products and norms where necessary.

**Definition 2.1.** Denote by $\mathcal{S}$ the class of *smooth cylinder functionals*, random variables $F : \mathscr{W} \to \mathbb{R}$ such that

$$F(\omega) = f(\omega_{t_1}, \ldots, \omega_{t_n}), \tag{2.1}$$

for some $n \geq 1$, $0 < t_1 < \cdots < t_n \leq 1$, and function $f \in C_p^\infty((\mathbb{R}^k)^n)$ (see Notation 1.1). For $E$ to be a real separable Hilbert space, let $\mathcal{S}_E$ be the set of $E$-valued smooth cylinder functions $F : \mathscr{W} \to E$ of the form

$$F = \sum_{j=1}^m F_j e_j, \tag{2.2}$$

for some $m \geq 1$, $e_j \in E$, and $F_j \in \mathcal{S}$.

**Definition 2.2.** Fix $h \in \mathscr{H}$. The directional derivative of a smooth cylinder functional $F \in \mathcal{S}$ of the form (2.1) along $h$ is given by

$$\partial_h F(\omega) := \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_0 F(\omega + \epsilon h) = \sum_{i=1}^{n} \nabla^i f(\omega_{t_1}, \ldots, \omega_{t_n}) \cdot h_{t_i},$$

where $\nabla^i f$ is the gradient of $f$ with respect to the $i$th variable.

The following integration by parts result is standard; see for example Theorem 8.2.2 of Hsu [16].

**Proposition 2.3.** *Let $F, G \in \mathcal{S}$ and $h \in \mathscr{H}$. Then*

$$(\partial_h F, G)_{\mathscr{H}} = (F, \partial_h^* G)_{\mathscr{H}},$$

*where* $\partial_h^* = -\partial_h + \int_0^1 \dot{h}_s \cdot \mathrm{d}b_s$.

**Definition 2.4.** The gradient of a smooth cylinder functional $F \in \mathcal{S}$ is the random process $D_t F$ taking values in $\mathscr{H}$ such that $(DF, h)_{\mathscr{H}} = \partial_h F$. It may be determined that, for $F$ of the form (2.1),

$$D_t F = \sum_{i=1}^{n} \nabla^i f(\omega_{t_1}, \ldots, \omega_{t_n})(t_i \wedge t),$$

where $s \wedge t = \min\{s, t\}$. For $F \in \mathcal{S}_E$ of the form (2.2), define the derivative $D_t F$ to be the random process taking values in $\mathscr{H} \otimes E$ given by

$$D_t F := \sum_{j=1}^{m} D_t F_j \otimes e_j.$$

Iterations of the derivative for smooth functionals $F \in \mathcal{S}$ are given by

$$D^k_{t_1,\ldots,t_k} F = D_{t_1} \cdots D_{t_k} F \in \mathscr{H}^{\otimes k},$$

for $k \in \mathbb{N}$. For $F \in \mathcal{S}_E$,

$$D^k F = \sum_{j=1}^{m} D^k F_j \otimes e_j,$$

and these are measurable functions are defined almost everywhere on $[0, 1]^k \times \mathscr{W}$. The operator $D$ on $\mathcal{S}_E$ is closable, and there exist closed extensions $D^k$ to $L^p(\mathscr{W}, \mathscr{H}^{\otimes k} \otimes E)$; see, for example [25], Theorem 8.28 of [16], or Theorem 8.5 of [17]. Denote the closure of the derivative operator also by $D$ and the domain of $D^k$ in $L^p([0, 1]^k \times \mathscr{W})$ by $\mathcal{D}^{k,p}$, which is the completion of the family of smooth Wiener functionals $\mathcal{S}$ with respect to the seminorm $\|\cdot\|_{k,p,E}$ on $\mathcal{S}_E$ given by

$$\|F\|_{k,p,E} := \left(\sum_{j=0}^{k} \mathbb{E}(\|D^j F\|^p_{\mathscr{H}^{\otimes j} \otimes E})\right)^{1/p},$$

for any $p \geq 1$. Let

$$\mathcal{D}^{k,\infty}(E) := \bigcap_{p>1} \mathcal{D}^{k,p}(E) \quad \text{and} \quad \mathcal{D}^{\infty}(E) := \bigcap_{p>1} \bigcap_{k \geq 1} \mathcal{D}^{k,p}(E).$$

When $E = \mathbb{R}$, write $\mathcal{D}^{k,p}(\mathbb{R}) = \mathcal{D}^{k,p}$, $\mathcal{D}^{k,\infty}(\mathbb{R}) = \mathcal{D}^{k,\infty}$, and $\mathcal{D}^{\infty}(\mathbb{R}) = \mathcal{D}^{\infty}$.

**Definition 2.5.** Let $D^*$ denote the $L^2(\mu)$-adjoint of the derivative operator $D$, which has its domain in $L^2(\mathscr{W} \times [0, 1], \mathscr{H})$ consisting of functions $G$ such that

$$|\mathbb{E}[(DF, G)_{\mathscr{H}}]| \leq C \|F\|_{L^2(\mu)},$$

for all $F \in \mathcal{D}^{1,2}$, where $C$ is a constant depending on $G$. For those functions $G$ in the domain of $D^*$, $D^*G$ is the element of $L^2(\mu)$ such that

$$\mathbb{E}[FD^*G] = \mathbb{E}[(DF, G)_{\mathscr{H}}].$$

It is known that $D$ is a continuous operator from $\mathcal{D}^\infty$ to $\mathcal{D}^\infty(\mathscr{H})$, and similarly, $D^*$ is continuous from $\mathcal{D}^\infty(\mathscr{H})$ to $\mathcal{D}^\infty$; see for example Theorem V-8.1 and its corollary in [17].

Malliavin [21,22] introduced the notion of derivatives of Wiener functionals and applied it to the regularity of probability laws induced by the solutions to stochastic differential equations at fixed times. The notion of Sobolev spaces of Wiener functionals was first introduced by Shigekawa [28] and Stroock [29,30].

### 2.2. *Rolling map*

Now, let $G$ be a Lie group with identity $e$ and Lie algebra $\mathrm{Lie}(G) = \mathfrak{g}$, and suppose $\{X_i\}_{i=1}^k \subset \mathfrak{g}$ is a linearly independent Lie generating set, in the sense of Eq. (1.3). Recall that $\{X_i\}_{i=1}^k$ is an orthonormal basis of the subspace $\mathfrak{g}_0 = \mathrm{span}(\{X_i\}_{i=1}^k)$ with respect to the inner product defined on $\mathfrak{g}$.

**Notation 2.6.** Let $\mathrm{Ad} : G \to \mathrm{End}(\mathfrak{g})$ denote the adjoint representation of $G$ with differential $\mathrm{ad} := \mathrm{d}(\mathrm{Ad}) : \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$. That is, $\mathrm{Ad}(g) = \mathrm{Ad}_g = L_{g*}R_{g^{-1}*}$, for all $g \in G$, and $\mathrm{ad}(X) = \mathrm{ad}_X = [X, \cdot]$, for all $X \in \mathfrak{g}$. For any function $\varphi \in C^1(G)$, define $\hat{\nabla}\varphi, \tilde{\nabla}\varphi : G \to \mathfrak{g}$ such that, for any $g \in G$ and $X \in \mathfrak{g}$,

$$\langle \hat{\nabla}\varphi(g), X\rangle := \langle \mathrm{d}\varphi(g), R_{g*}X\rangle = (\hat{X}\varphi)(g)$$
$$\langle \tilde{\nabla}\varphi(g), X\rangle := \langle \mathrm{d}\varphi(g), L_{g*}X\rangle = (\tilde{X}\varphi)(g).$$

The sequel will use the following facts:

$$\begin{aligned}\langle \hat{\nabla}\varphi(g), X\rangle &= \langle \mathrm{d}\varphi(g), L_{g*}L_{g^{-1}*}R_{g*}X\rangle \\ &= \langle \mathrm{d}\varphi(g), L_{g*}\mathrm{Ad}_{g^{-1}}X\rangle = \langle \tilde{\nabla}\varphi(g), \mathrm{Ad}_{g^{-1}}X\rangle \end{aligned} \tag{2.3}$$

and similarly

$$\langle \tilde{\nabla}\varphi(g), X\rangle = \langle \hat{\nabla}\varphi(g), \mathrm{Ad}_g X\rangle. \tag{2.4}$$

Now suppose $\{b_t^i\}_{i=1}^k$ are $k$ independent real-valued Brownian motions. Then

$$\vec{b}_t := X_i b_t^i := \sum_{i=1}^k X_i b_t^i$$

is a $(\mathfrak{g}_0, \langle\cdot,\cdot\rangle)$ Brownian motion. In the sequel, the convention of summing over repeated upper and lower indices will be observed. Let $\xi : [0, 1] \times \mathscr{W} \to G$ denote the solution to the Stratonovich stochastic differential equation

$$\mathrm{d}\xi_t = \xi_t \circ \mathrm{d}\vec{b}_t := L_{\xi_t *} \circ \mathrm{d}\vec{b}_t = L_{\xi_t *}X_i \circ \mathrm{d}b_t^i = \tilde{X}_i(\xi_t) \circ \mathrm{d}b_t^i, \quad \text{with } \xi_0 = e. \tag{2.5}$$

The solution $\xi$ exists by the standard theory; see, for example, Theorem V-1.1 of [17]. Additionally, Remark V-10.3 of [17] implies that $P_t = e^{tL/2}$, with $L = \sum_{i=1}^k \tilde{X}_i^2$, is the associated Markov diffusion semigroup to $\xi$, where $P_t$ is as defined in Definition 1.6; that is, $\nu_t := (\xi_t)_*\mu = p_t(g)dg$ is the density of the transition probability of the diffusion process $\xi_t$, where $dg$ denotes right Haar measure, and

$$(P_t f)(e) = \mathbb{E}[f(\xi_t)], \tag{2.6}$$

for any $f \in C_c^\infty(G)$, where the right hand side is the expectation conditioned on $\xi_0 = e$.

The following theorem is proved in [23].

**Theorem 2.7.** *For any $f \in C_c^\infty(G)$, $f(\xi_t) \in \mathcal{D}^\infty$ for all $t \in [0, 1]$. In particular, $D[f(\xi_t)] \in \mathscr{H} \otimes \mathbb{R}^k$ and*

$$(D[f(\xi_t)])^i = \left\langle \hat{\nabla} f(\xi_t), \int_0^{\cdot\wedge t} \mathrm{Ad}_{\xi_\tau} X_i \mathrm{d}\tau \right\rangle, \tag{2.7}$$

*for $i = 1, \ldots, k$, componentwise in $\mathscr{H}$, and, for any $h \in \mathscr{H}$,*

$$\partial_h f(\xi_t) = \left\langle \hat{\nabla} f(\xi_t), \int_0^t \mathrm{Ad}_{\xi_s} X_i \dot{h}_s^i \mathrm{d}s \right\rangle = \left\langle \mathrm{d}f(\xi_t), R_{\xi_t *} \int_0^t \mathrm{Ad}_{\xi_s} X_i \dot{h}_s^i \mathrm{d}s \right\rangle. \tag{2.8}$$

**Notation 2.8.** Let $\bigcap_{p>1} L^p(\mu) =: L^{\infty-}(\mu)$.

## 2.3. Covariance matrix

The Malliavin covariance matrix of $\xi$ is the matrix $\sigma_t(\omega) := \xi_t'(\omega)\xi_t'(\omega)^* : T_{\xi_t(\omega)}G \to T_{\xi_t(\omega)}G$, where $\xi_t'(\omega) : \mathscr{H} \to T_{\xi_t(\omega)}G$ is the Frechet derivative given by

$$\xi_t'(\omega)h := \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_0 \xi_t(\omega + \epsilon h),$$

for all $h \in \mathscr{H}$, and its adjoint $\xi_t'(\omega)^* : T_{\xi_t(\omega)} \to \mathscr{H}$ is computed relative to the Cameron–Martin inner product on $\mathscr{H}$ and the chosen metric on $G$. Note that Eq. (2.8) implies that

$$\xi_t'(\omega)h = R_{\xi_t *} \int_0^t \mathrm{Ad}_{\xi_s} X_i \dot{h}_s^i \mathrm{d}s \tag{2.9}$$

**Notation 2.9.** In the following, let $\mathrm{Ad}_{\xi_t}^\dagger$ denote the adjoint of $\mathrm{Ad}_{\xi_t}$ as an operator on $\mathfrak{g}$, and let $P : \mathfrak{g} \to \mathfrak{g}_0$ be orthogonal projection onto the subspace $\mathfrak{g}_0$.

**Theorem 2.10.** *The Malliavin covariance matrix of $\xi$ is*

$$\sigma_t := \xi_t'(\omega)\xi_t'(\omega)^* = R_{\xi_t *} \left( \int_0^t \mathrm{Ad}_{\xi_s} P \mathrm{Ad}_{\xi_s}^\dagger \mathrm{d}s \right) R_{\xi_t *}^{\mathrm{tr}}. \tag{2.10}$$

*Let $\bar{\sigma}_t = \int_0^t \mathrm{Ad}_{\xi_s} P \mathrm{Ad}_{\xi_s}^\dagger \mathrm{d}s$, and $\Delta_t := \det \bar{\sigma}_t$. Then $\Delta_t > 0$ a.e., and so $\bar{\sigma}_t$ is invertible a.e. for $t > 0$. Moreover,*

$$\Delta_t^{-1} \in L^{\infty-}(\mu).$$

**Proof.** To determine $\sigma_t = \xi_t'(\omega)\xi_t'(\omega)^*$, first compute $\xi_t'(\omega)^* : T_{\xi_t(\omega)}G \to \mathscr{H}$, the adjoint in $\xi_t'(\omega)$ with respect to the Cameron–Martin inner product and the right invariant metric on $TG$. By Eq. (2.9), for any $X \in \mathfrak{g}$,

$$\begin{aligned}(\xi_t'(\omega)^*(R_{\xi_t *}X), h)_{\mathscr{H}} &= \langle R_{\xi_t *}X, \xi_t'(\omega)h\rangle \\ &= \left\langle R_{\xi_t *}X, R_{\xi_t *}\int_0^t \mathrm{Ad}_{\xi_s} X_i \dot{h}_s^i \mathrm{d}s\right\rangle \\ &= \left\langle X, \int_0^t \mathrm{Ad}_{\xi_s} X_i \dot{h}_s^i \mathrm{d}s\right\rangle = \int_0^t \left\langle \mathrm{Ad}_{\xi_s}^{\dagger} X, X_i\right\rangle \dot{h}_s^i \mathrm{d}s,\end{aligned}$$

where the penultimate equality follows from the right invariance of the metric on $G$. It then follows that

$$\frac{\mathrm{d}}{\mathrm{d}s}[\xi_t'(\omega)^*(R_{\xi_t *}X)]_s^i = 1_{s\le t}\langle \mathrm{Ad}_{\xi_s}^{\dagger} X, X_i\rangle, \tag{2.11}$$

componentwise in $\mathscr{H}$. Combining Eqs. (2.9) and (2.11),

$$\begin{aligned}\xi_t'(\omega)\xi_t'(\omega)^*(R_{\xi_t *}X) &= R_{\xi_t *}\int_0^t \mathrm{Ad}_{\xi_s} X_i \frac{\mathrm{d}}{\mathrm{d}s}[\xi_t'(\omega)^*(R_{\xi_t *}X)]_s^i \mathrm{d}s \\ &= \sum_{i=1}^k R_{\xi_t *}\int_0^t \mathrm{Ad}_{\xi_s} X_i \langle \mathrm{Ad}_{\xi_s}^{\dagger} X, X_i\rangle \mathrm{d}s \\ &= R_{\xi_t *}\int_0^t \mathrm{Ad}_{\xi_s} P \mathrm{Ad}_{\xi_s}^{\dagger} X \mathrm{d}s,\end{aligned}$$

and Eq. (2.10) follows.

The proof that $\Delta_t > 0$ and $\Delta_t^{-1} \in L^{\infty-}(\mu)$ is by now standard, and relies on satisfaction of the Hörmander bracket condition, $\mathrm{Lie}(\{X_i\}_{i=1}^k) = \mathfrak{g}$; for example, a simple adaptation of the proof of Theorem 8.6 in Driver [8] will work. ∎

**Remark 2.11.** By the general theory, Theorem 2.10 implies $\nu_t = \mathrm{Law}(\xi_t)$ is a smooth measure; see for example Remark V-10.3 of [17].

### 2.4. *Lifted vector fields and $L^2$-adjoints*

Throughout this section, $t \in [0, 1]$ will be fixed.

**Definition 2.12.** Given $X \in \mathfrak{g}$, let $\tilde{X}$ be the associated left invariant vector field on $G$. Define the "*lifted vector field*" $\mathbf{X}$ of $\tilde{X}$ as

$$\mathbf{X} = \mathbf{X}^t := \xi_t'(\omega)^*[\xi_t'(\omega)\xi_t'(\omega)^*]^{-1}\tilde{X}(\xi_t) = \xi_t'(\omega)^*\sigma_t^{-1}\tilde{X}(\xi_t) \in \mathscr{H}, \tag{2.12}$$

acting on functions $F \in \mathcal{D}^{1,2}$ by

$$\mathbf{X}F = (DF, \mathbf{X})_{\mathscr{H}}.$$

**Proposition 2.13.** *For any $X \in \mathfrak{g}$, $\mathbf{X} \in \mathcal{D}^{\infty}(\mathscr{H})$, and*

$$\mathbf{X}[f(\xi_t)] = (\tilde{X}f)(\xi_t),$$

*for any $f \in C^{\infty}(G)$.*

**Proof.** Combining Eqs. (2.10) and (2.11) gives

$$\frac{\mathrm{d}}{\mathrm{d}s}\mathbf{X}^i_s = 1_{s\le t}\left\langle \mathrm{Ad}^\dagger_{\xi_s}\left(\int_0^t \mathrm{Ad}_{\xi_r} P\,\mathrm{Ad}^\dagger_{\xi_r}\mathrm{d}r\right)^{-1}\mathrm{Ad}_{\xi_t}X, X_i\right\rangle.$$

Thus, rewrite Eq. (2.12) explicitly as

$$\begin{aligned}\mathbf{X}^i &= \int_0^{\cdot\wedge t}\left\langle \mathrm{Ad}^\dagger_{\xi_s}\left(\int_0^t \mathrm{Ad}_{\xi_r} P\mathrm{Ad}^\dagger_{\xi_r}\mathrm{d}r\right)^{-1}\mathrm{Ad}_{\xi_t}X, X_i\right\rangle \mathrm{d}s\\ &= \left\langle\left(\int_0^{\cdot\wedge t}\mathrm{Ad}^\dagger_{\xi_s}\mathrm{d}s\right)\left(\int_0^t \mathrm{Ad}_{\xi_s}P\mathrm{Ad}^\dagger_{\xi_s}\mathrm{d}s\right)^{-1}\mathrm{Ad}_{\xi_t}X, X_i\right\rangle. \qquad (2.13)\end{aligned}$$

A standard argument shows that $\overline{W} = \int_0^\cdot \mathrm{Ad}_{\xi_s}\mathrm{d}s \in \mathcal{D}^\infty(\mathscr{H}(\mathrm{End}(\mathfrak{g})))$; see for example Proposition 5 of [24]. Note that $W_t^\dagger = \mathrm{Ad}^\dagger_{\xi_t} : \mathscr{W} \to \mathrm{End}(\mathfrak{g})$ satisfies the differential equation

$$\mathrm{d}W_t^\dagger = \mathrm{ad}^\dagger_{X_i} W_t^\dagger \circ \mathrm{d}b_t^i, \quad \text{with } W_0^\dagger = I,$$

which is linear with smooth coefficients. Similarly, one may show that

$$\overline{W}^\dagger := \int_0^\cdot \mathrm{Ad}^\dagger_{\xi_s}\mathrm{d}s \in \mathcal{D}^\infty(\mathscr{H}(\mathrm{End}(\mathfrak{g}))).$$

Also, Theorem 2.10 implies that

$$\bar\sigma_t^{-1} = \left(\int_0^t \mathrm{Ad}_{\xi_s}P\mathrm{Ad}^\dagger_{\xi_s}\mathrm{d}s\right)^{-1}$$

exists and is in $L^{\infty-}(\mu)$ componentwise. Thus, Eq. (2.13) implies that $\mathbf{X} \in \mathcal{D}^\infty(\mathscr{H})$.

For $f \in C^\infty(G)$ and $(h^1, \ldots, h^k) \in \mathscr{H}$, by Eq. (2.8),

$$\partial_h[f(\xi_t)] = (D[f(\xi_t)], h)_{\mathscr{H}} = \left\langle \hat\nabla f(\xi_t), \int_0^t \mathrm{Ad}_{\xi_s}X_i\dot h^i_s\mathrm{d}s\right\rangle,$$

and so

$$\begin{aligned}\mathbf{X}[f(\xi_t)] &= (D[f(\xi_t)], \mathbf{X})_{\mathscr{H}}\\ &= \left\langle\hat\nabla f(\xi_t), \int_0^t \mathrm{Ad}_{\xi_s}X_i\left\langle \mathrm{Ad}^\dagger_{\xi_s}\left(\int_0^t \mathrm{Ad}_{\xi_r}P\mathrm{Ad}^\dagger_{\xi_r}\mathrm{d}r\right)^{-1}\mathrm{Ad}_{\xi_t}X, X_i\right\rangle\mathrm{d}s\right\rangle\\ &= \left\langle\hat\nabla f(\xi_t), \int_0^t \mathrm{Ad}_{\xi_s}P\mathrm{Ad}^\dagger_{\xi_s}\left(\int_0^t \mathrm{Ad}_{\xi_r}P\mathrm{Ad}^\dagger_{\xi_r}\mathrm{d}r\right)^{-1}\mathrm{Ad}_{\xi_t}X\mathrm{d}s\right\rangle\\ &= \langle\hat\nabla f(\xi_t), \mathrm{Ad}_{\xi_t}X\rangle = \langle\tilde\nabla f(\xi_t), \mathrm{Ad}_{\xi_t^{-1}}\mathrm{Ad}_{\xi_t}X\rangle\\ &= (\tilde X f)(\xi_t),\end{aligned}$$

where the penultimate equality used Eq. (2.3). ■

**Definition 2.14.** For a vector field $\mathbf{X}$ acting on functions of $\mathscr{W}$, denote the adjoint of $\mathbf{X}$ in the $L^2(\mu)$ inner product by $\mathbf{X}^*$, which has domain in $L^2(\mu)$ consisting of functions $G$ such that for all $F \in \mathcal{D}^{1,2}$,

$$\mathbb{E}[(\mathbf{X}F)G] \le c\|F\|_{L^2(\mu)}$$

for some constant $c$. For functions $G$ in the domain of $\mathbf{X}^*$,

$$\mathbb{E}[F(\mathbf{X}^* G)] = \mathbb{E}[(\mathbf{X}F)G]$$

for all $F \in \mathcal{D}^{1,2}$.

Note that for any lifted vector field $\mathbf{X}$ acting on function $F \in \mathcal{D}^{1,2}$ as defined in Definition 2.12,

$$\mathbb{E}[\mathbf{X}F] = \mathbb{E}[(DF, \mathbf{X})_{\mathscr{H}}] = \mathbb{E}[F D^* \mathbf{X}].$$

Thus, $\mathbf{X}^* = \mathbf{X}^* 1 = D^* \mathbf{X}$ a.s. Recall that $D^*$ is a continuous operator from $\mathcal{D}^\infty(\mathscr{H})$ into $\mathcal{D}^\infty$; see for example Theorem V-8.1 and its corollary in [17]. Thus, for $\mathbf{X}$ a vector field on $W$ as defined in Eq. (2.12), Proposition 2.13 implies that $D^* \mathbf{X} \in \mathcal{D}^\infty$. This proves the following proposition.

**Proposition 2.15.** *Let $\tilde{X}$ be a left invariant vector field on $G$. Then, for the vector field on $\mathscr{W}$ defined by*

$$\mathbf{X} = \xi_t'(\omega)^* [\xi_t'(\omega) \xi_t'(\omega)^*]^{-1} \tilde{X}(\xi_t(\omega)),$$

*$\mathbf{X}^* \in \mathcal{D}^\infty$, where $\mathbf{X}^*$ is the $L^2(\mu)$-adjoint of $\mathbf{X}$.*

## 3. Lie group inequalities

Again, let $G$ be a Lie group with identity $e$ and a Lie algebra $\mathrm{Lie}(G) = \mathfrak{g}$, and suppose $\{X_i\}_{i=1}^k \subset \mathfrak{g}$ is a Hörmander set, in the sense of Eq. (1.3). The gradient $\nabla = (\tilde{X}_1, \ldots, \tilde{X}_k)$ and the subLaplacian $L = \sum_{i=1}^k \tilde{X}_i^2$ are operators on smooth functions of $G$ with compact support. Let $L$ also denote the self-adjoint extension of the subLaplacian, and $P_t = \mathrm{e}^{tL/2}$ be the heat semigroup as in Definition 1.6.

The following lemmas were proved in [10] in the context of the Heisenberg Lie group (Lemmas 2.3, 2.4). The proofs are identical in the general Lie group case.

**Lemma 3.1.** *By the left invariance of $\nabla$ and $P_t$, the inequality $(I_p)$ holds for all $g \in G$, $f \in C_c^\infty(G)$, and $t > 0$, if and only if,*

$$|\nabla P_t f|^p(e) \le K_p(t) P_t |\nabla f|^p(e),$$

*for all $f \in C_c^\infty(G)$ and $t > 0$, where $e \in G$ is the identity element.*

**Lemma 3.2.** *For $X \in \mathfrak{g}$,*

$$\tilde{X} P_t f(e) = P_t \hat{X} f(e)$$

*for all $f \in C_c^\infty(G)$. More generally,*

$$\hat{X} P_t f = P_t \hat{X} f,$$

*from which the previous equation follows, since $\hat{X} = \tilde{X}$ at $e$.*

(The proof of Lemma 3.2 is actually easier than its analogue Lemma 2.4 in [10], since working with functions with compact support – versus functions with polynomial growth – requires only the invariance of the Haar measure to justify passing the derivative through the integral.)

### 3.1. $L^p$-type gradient estimate ($p > 1$)

**Notation 3.3.** For each $r \in \{0, 1, \dots, m\}$, let $\Lambda^r = \Lambda^{k,r}$ be the set of multi-indices $\alpha = (\alpha_0, \alpha_1, \dots, \alpha_r) \in \{1, \dots, k\}^{r+1}$. For any $\alpha \in \Lambda^r$, define

$$\alpha' := (\alpha_1, \dots, \alpha_r) \quad \text{and}$$
$$\overline{\alpha} := (\alpha_r, \dots, \alpha_0) = \alpha \text{ reversed.}$$

Define the order of $\alpha$ by $|\alpha| := r + 1$. Let

$$X_\alpha = [X_{\alpha_r}, [\cdots, [X_{\alpha_1}, X_{\alpha_0}] \cdots]] = \mathrm{ad}_{X_{\alpha_r}} \cdots \mathrm{ad}_{X_{\alpha_1}} X_{\alpha_0} \quad \text{and}$$
$$X^\alpha = X_{\alpha_r} \cdots X_{\alpha_0}.$$

When $r = 0$ and $|\alpha| = 1$, that is, $\alpha = (\alpha_0)$, then $X^\alpha = X_{\alpha_0} = X_\alpha$. For each $\alpha \in \Lambda^r$, there exist $\epsilon_{\beta,\alpha} \in \mathbb{Z}$ such that

$$X_\alpha = \sum_{\beta \in \Lambda^r} \epsilon_{\beta,\alpha} X^\beta.$$

**Proposition 3.4.** *For any $X \in \mathfrak{g}$, $\hat{X}$ may be written as*

$$\hat{X} = \sum_{r=0}^{m} \sum_{\alpha \in \Lambda^r} c_\alpha \tilde{X}^\alpha, \tag{3.1}$$

*with $c_\alpha : G \to \mathbb{R}$ (some of these are $0$) such that $c_\alpha(\xi_t) \in \mathcal{D}^\infty$, for all $t \in [0, 1]$.*

**Proof.** Recall from Notation 1.2 that

$$\Sigma_r = \{[X_{i_1}, [\dots, [X_{i_{r-1}}, X_{i_r}] \cdots]] : i_1, \dots, i_r \in \{1, \dots, k\}\}$$
$$= \{X_\alpha : \alpha \in \Lambda^r\},$$

for $r = 0, \dots, m$. Recall also from Notation 1.2 that $\{X_i, Y_j : i \in \{1, \dots, k\}, j \in \{1, \dots, d-k\}\}$ of $\mathfrak{g}$ is an orthonormal basis, where $d = \dim(G)$ and, for each $j \in \{1, \dots, d-k\}$, $Y_j$ is some commutator $X_{\alpha(j)} \in \Sigma_{r(j)}$ for some $\alpha(j) \in \Lambda^{r(j)}$, $r(j) \in \{1, \dots, m\}$. Thus, for any $g \in G$ and $X \in \mathfrak{g}$,

$$\begin{aligned}
\hat{X}(g) &= R_{g*}X = L_{g*}L_{g^{-1}*}R_{g*}X = L_{g*}\mathrm{Ad}_{g^{-1}}X \\
&= L_{g*}\left(\sum_{i=1}^{k} \langle \mathrm{Ad}_{g^{-1}}X, X_i \rangle X_i + \sum_{j=1}^{d-k} \langle \mathrm{Ad}_{g^{-1}}X, Y_j \rangle Y_j\right) \\
&= L_{g*}\left(\sum_{i=1}^{k} \langle \mathrm{Ad}_{g^{-1}}X, X_i \rangle X_i + \sum_{j=1}^{d-k} \sum_{\alpha \in \Lambda^{r(j)}} \epsilon_{\alpha,\alpha(j)} \langle \mathrm{Ad}_{g^{-1}}X, Y_j \rangle X^\alpha\right) \\
&= \sum_{i=1}^{k} \langle \mathrm{Ad}_{g^{-1}}X, X_i \rangle \tilde{X}_i(g) + \sum_{j=1}^{d-k} \sum_{\alpha \in \Lambda^{r(j)}} \epsilon_{\alpha,\alpha(j)} \langle \mathrm{Ad}_{g^{-1}}X, Y_j \rangle \tilde{X}^\alpha(g)
\end{aligned}$$

where $\epsilon_{\alpha,\alpha(j)} \in \mathbb{Z}$. So

$$\hat{X}(g) = \sum_{r=0}^{m} \sum_{\alpha \in \Lambda^r} c_\alpha \tilde{X}^\alpha(g),$$

where

$$c_\alpha(g) = \begin{cases} \langle \mathrm{Ad}_{g^{-1}} X, X_i \rangle & \text{when } r = 0 \text{ and } \alpha = (i) \\ \epsilon \langle \mathrm{Ad}_{g^{-1}} X, Y_j \rangle, \epsilon \in \mathbb{Z} & \text{when } r \in \{1, \dots, m\}. \end{cases}$$

Note that $\mathrm{Ad}_{\xi_t}$ satisfies the Stratonovich stochastic differential equation

$$\mathrm{dAd}_\xi = \mathrm{Ad}_\xi \circ \mathrm{ad}_{db} = \mathrm{Ad}_\xi \mathrm{ad}_{X_i} \circ \mathrm{d}b^i, \quad \text{with } \mathrm{Ad}_{\xi_0} = I.$$

By differentiating the identity $\mathrm{Ad}_{\xi_t} \mathrm{Ad}_{\xi_t}^{-1} = I$, one may verify that $\mathrm{Ad}_{\xi_t}^{-1} = \mathrm{Ad}_{\xi_t^{-1}}$ satisfies

$$\mathrm{dAd}_{\xi^{-1}} = - \circ \mathrm{ad}_{db} \mathrm{Ad}_{\xi^{-1}} = -\mathrm{ad}_{X_i} \mathrm{Ad}_{\xi^{-1}} \circ \mathrm{d}b^i, \quad \text{with } \mathrm{Ad}_{\xi_0^{-1}} = I$$

a linear differential equation with smooth coefficients. Then by Theorem V-10.1 of Ikeda and Watanabe [17], $\mathrm{Ad}_{\xi_t^{-1}} \in \mathcal{D}^\infty(\mathrm{End}(\mathfrak{g}))$ componentwise with respect to some basis.

The function $u : \mathrm{End}(\mathfrak{g}) \to \mathbb{R}$ given by $u(W) = \langle WX, Y \rangle$ is a smooth function for any fixed $X, Y \in \mathfrak{g}$. Thus, $u(\mathrm{Ad}_{\xi_t^{-1}}) \in \mathcal{D}^\infty$ for all $t \in [0, 1]$. Since $c_\alpha(\xi_t) = \epsilon u(\mathrm{Ad}_{\xi_t^{-1}})$, with $Y = X_i$ or $Y_j$, this implies that $c_\alpha(\xi_t) \in \mathcal{D}^\infty$, for all $\alpha \in \Lambda^r$. ■

**Theorem 3.5.** *For all $p \in (1, \infty)$, $K_p(t) < \infty$, where $K_p(t)$ are the functions defined in Notation 1.7.*

**Proof.** Lemma 3.1 implies that the inequality $(I_p)$ is translation invariant on groups. Thus, it suffices to determine a finite coefficient $K_p(t)$ such that the inequality holds at the identity.

Note that for any $X \in \mathfrak{g}$, Lemma 3.2 and Eq. (3.1) imply that

$$|\tilde{X} P_t f|^2(e) = |\hat{X} P_t f|^2(e) = |P_t \hat{X} f|^2(e) \leq C \sum_{r=0}^{m} \sum_{\alpha \in \Lambda^r} |P_t c_\alpha \tilde{X}^\alpha f|^2(e),$$

for a constant $C = C(k, m)$. Eq. (2.6) implies that, for any $f \in C_c^\infty(G)$, $P_t f(e) = \mathbb{E}[f(\xi_t)]$, where $\xi$ is the solution to the Stratonovich Eq. (2.5). Thus, for any $\alpha \in \Lambda^r$,

$$\begin{aligned} |P_t c_\alpha \tilde{X}^\alpha f|(e) &\leq \mathbb{E}|c_\alpha(\xi_t)(\tilde{X}^\alpha f)(\xi_t)| = \mathbb{E}|c_\alpha(\xi_t)\mathbf{X}^{\alpha'}[(\tilde{X}_{\alpha_0} f)(\xi_t)]| \\ &= \mathbb{E}\left|\left(\mathbf{X}^{\overline{\alpha'}}\right)^* [c_\alpha(\xi_t)](\tilde{X}_{\alpha_0} f)(\xi_t)\right| \\ &\leq \left(\mathbb{E}\left|\left(\mathbf{X}^{\overline{\alpha'}}\right)^* [c_\alpha(\xi_t)]\right|^q\right)^{1/q} (\mathbb{E}|(\tilde{X}_{\alpha_0} f)(\xi_t)|^p)^{1/p} \\ &= \left(\mathbb{E}\left|\left(\mathbf{X}^{\overline{\alpha'}}\right)^* [c_\alpha(\xi_t)]\right|^q\right)^{1/q} (P_t|\tilde{X}_{\alpha_0} f|^p(e))^{1/p} \\ &\leq \left(\mathbb{E}\left|\left(\mathbf{X}^{\overline{\alpha'}}\right)^* [c_\alpha(\xi_t)]\right|^q\right)^{1/q} (P_t|\nabla f|^p(e))^{1/p}, \end{aligned}$$

by Hölder's inequality, where $q$ is the conjugate exponent to $p$, $\mathbf{X}^\alpha$ is the lifted vector field on $W$ of the vector field $\tilde{X}^\alpha$, as defined in Eq. (2.12), and $(\mathbf{X}^\alpha)^* = \mathbf{X}_{\alpha_r}^* \cdots \mathbf{X}_{\alpha_0}^*$ (so $(\mathbf{X}^{\overline{\alpha'}})^* = \mathbf{X}_{\alpha_1}^* \cdots \mathbf{X}_{\alpha_r}^*$). Propositions 2.15 and 3.4 imply that $(\mathbf{X}^{\overline{\alpha'}})^*[c_\alpha(\xi_t)] \in L^{\infty-}(\mu)$, for all $\alpha \in \Lambda^r$.

So in particular, using the above with $X = X_i$ gives

$$
\begin{aligned}
|\nabla P_t f|^p(e) &= \left(\sum_{i=1}^{k} |\tilde{X}_i P_t f|^2(e)\right)^{p/2} \\
&\leq C \left[\sum_{i=1}^{k}\sum_{r=0}^{m}\sum_{\alpha\in\Lambda^r}\left(\mathbb{E}\left|\left(\mathbf{X}^{\overline{\alpha'}}\right)^{*}[c_{i,\alpha}(\xi_t)]\right|^q\right)^{p/q}\right] P_t|\nabla f|^p(e),
\end{aligned}
$$

where $C = C(k, m, p)$ and $q = \frac{p}{p-1}$. Thus, the inequality ($I_p$) holds, with

$$
C_p(t) = C(k, m, p)\sum_{i=1}^{k}\sum_{r=0}^{m}\sum_{\alpha\in\Lambda^r}\left(\mathbb{E}\left|\left(\mathbf{X}^{\overline{\alpha'}}\right)^{*}[c_{i,\alpha}(\xi_t)]\right|^q\right)^{p/q}. \tag{3.2}
$$

Therefore, $K_p(t) \leq C_p(t) < \infty$ for all $t > 0$ and $p \in (1, \infty)$. ■

It is important to note that, in this general Lie group case, there is currently no good control over the behavior of the functions $C_p$ in Eq. (3.2) with respect to $t$. In fact, from certain scaling arguments, it is expected that $C_p(t) \to \infty$ as $t \to 0$; see for example [5,18]. However, these coefficients are almost certainly not optimal.

To explore cases where the behavior of these coefficients is more understood, it will become useful to extend the set of test functions considered. The following proposition relaxes the condition of compact support to boundedness with bounded first order derivatives.

**Proposition 3.6.** *For all $p \in (1, \infty)$,*

$$|\nabla P_t f|^p \leq K_p(t) P_t|\nabla f|^p$$

*for all $f \in C_b^\infty(G)$ with bounded derivatives of first order and $t > 0$.*

**Proof.** Let $f \in C_b^\infty(G)$ with bounded first order derivatives, and let $\varphi_m \in C_c^\infty(G, [0, 1])$ be a sequence of functions such that $\varphi_m \uparrow 1$, $\varphi_m(g) = 1$ when $|g| \leq m$ (for some norm on $G$), and $\sup_m \sup_{g\in G} |\tilde{X}\varphi_m| < \infty$ for all $X \in \mathfrak{g}$; see Lemma 3.6 of [9]. Then $f_m = \varphi_m f \in C_c^\infty(G)$, and so there exists an optimal function $K_p(t) < \infty$ such that

$$|\nabla P_t f_m|^p \leq K_p(t) P_t|\nabla f_m|^p$$

for all $t > 0$. For any $X \in \mathfrak{g}$,

$$
\begin{aligned}
\lim_{m\to\infty} |\tilde{X} f_m - \tilde{X} f| &= \lim_{m\to\infty} |(\tilde{X}\varphi_m) f + \varphi_m \tilde{X} f - \tilde{X} f| \\
&\leq \lim_{m\to\infty} |\tilde{X}\varphi_m||f| + |\varphi_m - 1||\tilde{X} f| = 0
\end{aligned}
$$

which implies that $|\nabla f_m| \to |\nabla f|$ boundedly. Thus, by the dominated convergence theorem,

$$\lim_{m\to\infty} P_t|\nabla f_m|^p = P_t|\nabla f|^p.$$

Similarly,

$$
\begin{aligned}
\lim_{m\to\infty} |\tilde{X} P_t f_m - \tilde{X} P_t f| &= \lim_{m\to\infty} |P_t \hat{X} f_m - P_t \hat{X} f| \\
&\leq \lim_{m\to\infty} P_t|\hat{X} f_m - \hat{X} f| \\
&\leq \lim_{m\to\infty} P_t(|\hat{X}\varphi_m||f|) + P_t(|\varphi_m - 1||\hat{X} f|) = 0
\end{aligned}
$$

by dominated convergence, and hence

$$\lim_{m\to\infty} |\nabla P_t f_m| = |\nabla P_t f|.$$

Thus,

$$|\nabla P_t f|^p = \lim_{m\to\infty} |\nabla P_t f_m|^p \le K_p(t) \lim_{m\to\infty} P_t |\nabla f_m|^p = K_p(t) P_t |\nabla f|^p. \quad \blacksquare$$

### 3.2. Poincaré inequality

The following result is a direct corollary to Theorem 3.5. The proof is completely analogous to the proof of Theorem 4.2 in [10] in the Heisenberg Lie group context.

**Theorem 3.7** (*Poincaré Inequality*). *Let $K_2(t)$ be the best function for which $(I_p)$ holds for $p = 2$, and let $p_t(g)dg$ be the hypoelliptic heat kernel. Then*

$$\int_G f^2(g) p_t(g) dg - \left(\int_G f(g) p_t(g) dg\right)^2 \le \Lambda(t) \int_G |\nabla f|^2(g) p_t(g) dg, \tag{3.3}$$

*for all $f \in C_c^\infty(G)$ and $t > 0$, where*

$$\Lambda(t) = \int_0^t K_2(s) \mathrm{d}s.$$

**Proof.** Let $F_t(g) = (P_t f)(g)$. Then

$$\frac{\mathrm{d}}{\mathrm{d}s} P_{t-s} F_s^2 = P_{t-s}\left(-\frac{1}{2} L F_s^2 + F_s L F_s\right) = -P_{t-s} |\nabla F_s|^2.$$

Integrating this equation on $s$ gives us

$$\begin{aligned} P_t f^2 - (P_t f)^2 &= \int_0^t P_{t-s} |\nabla F_s|^2 \mathrm{d}s \\ &= \int_0^t P_{t-s} |\nabla P_s f|^2 \mathrm{d}s \\ &\le \int_0^t K_2(s) P_{t-s} P_s |\nabla f|^2 \mathrm{d}s = \left(\int_0^t K_2(s) \mathrm{d}s\right) \cdot P_t |\nabla f|^2 \end{aligned}$$

where the inequality follows from Theorem 3.5. Evaluating the above at $e \in G$ gives the desired result. $\blacksquare$

This theorem is less useful in the general Lie group case, because nothing is known about the integrability of $K_p(t)$. However, the next two sections show that when $G$ is a nilpotent Lie group, $K_p(t)$ is a bounded function for all $p \in (1, \infty)$. In particular, when $p = 2$, this implies that the Poincaré inequality holds with $\Lambda(t) < \infty$, for all $t > 0$.

#### 3.2.1. Stratified nilpotent Lie groups

**Definition 3.8.** A Lie algebra $\mathfrak{g}$ is said to be *nilpotent* if $\mathrm{ad}_X$ is a nilpotent endomorphism of $\mathfrak{g}$ for all $X \in \mathfrak{g}$, that is, if there exists $m \in \mathbb{N}$ such that

$$\mathrm{ad}_{Y_1} \cdots \mathrm{ad}_{Y_{m-1}} Y_m = [Y_1, [\ldots, [Y_{m-1}, Y_m] \cdots]] = 0,$$

for any $Y_1, \ldots, Y_m \in \mathfrak{g}$. If $m$ is the smallest number for which the above equality holds, $\mathfrak{g}$ is nilpotent of step $m$. A Lie group $G$ is nilpotent if $\mathfrak{g} = \mathrm{Lie}(G)$ is a nilpotent Lie algebra.

**Definition 3.9.** A family of *dilations* on a Lie algebra $\mathfrak{g}$ is a family of algebra automorphisms $\{\Phi_r\}_{r>0}$ on $\mathfrak{g}$ of the form $\Phi_r = \exp(W \log r)$, where $W$ is a diagonalizable linear operator on $\mathfrak{g}$ with positive eigenvalues.

**Definition 3.10.** A *stratified* group $G$ is a simply connected nilpotent group for which there exists a subset of the Lie algebra $V_1 \subset \mathfrak{g}$, such that $\mathfrak{g} = \oplus_{j=1}^m V_j$ with $V_{j+1} = [V_1, V_j]$, for $j = 1, \ldots, m-1$, and $V_{m+1} = [V_1, V_m] = \{0\}$.

For a general exposition on nilpotent Lie groups and dilations, see [11,12] and the references contained therein. If $G$ is a stratified Lie group, a natural family of dilations may be defined on $\mathfrak{g}$ by setting $\Phi_r(X) = r^j X$, for all $X \in V_j$. The generator $W$ of this dilation acts on parts of the vector space decomposition by $WV_j = jV_j$, for each $j = 1, \ldots, m$. The automorphism $\Phi_r$ induces a group dilation $\phi_r$ via the exponential maps, $\phi_r = \exp \circ \Phi_r \circ \exp^{-1}$. Since $G$ is a simply connected nilpotent group, the exponential map is in fact a global diffeomorphism on $\mathfrak{g}$, and $\exp^{-1}$ exists everywhere on $G$; see for example Theorem 3.6.2 of Varadarajan [31]. Then for each $X \in V_1$,

$$\begin{aligned}\tilde{X}(f \circ \phi_r)(g) &= \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_0 (f \circ \phi_r)(g e^{\epsilon X}) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_0 f(\phi_r(g)\phi_r(e^{\epsilon X})) \\ &= \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_0 f(\phi_r(g) e^{r\epsilon X}) = \left.\frac{\mathrm{d}}{\mathrm{d}\epsilon}\right|_0 r f(\phi_r(g) e^{\epsilon X}) = r(\tilde{X} f \circ \phi_r)(g),\end{aligned} \tag{3.4}$$

for all $f \in C^1(G)$, where the second equality used the fact that $\phi_r$ is a homomorphism. Let $\{X_i\}_{i=1}^k \subset V_1$ be a basis of $V_1$, and consider the operators $\nabla = (\tilde{X}_1, \ldots, \tilde{X}_k)$ and $L = \sum_{i=1}^k \tilde{X}_i^2$. Eq. (3.4) implies that

$$\nabla(f \circ \phi_r) = r(\nabla f) \circ \phi_r, \tag{3.5}$$

and thus the following proposition is obtained:

**Proposition 3.11.** *Let $L$ denote the self-adjoint extension of $\sum_{i=1}^k \tilde{X}_i^2$, and $P_t = \mathrm{e}^{tL/2}$ be as in Definition 1.6. Then*

$$L(f \circ \phi_r) = r^2 (Lf) \circ \phi_r$$

*and*

$$P_t(f \circ \phi_r) = \mathrm{e}^{tL/2}(f \circ \phi_r) = (\mathrm{e}^{r^2 tL/2} f) \circ \phi_r = (P_{r^2 t} f) \circ \phi_r, \tag{3.6}$$

*for any $f \in C_c^\infty(G)$.*

**Proof.** Let $\mathcal{E}^0(f, h) := \sum_{i=1}^k (\tilde{X}_i f, \tilde{X}_i h)_{L^2(G)}$ be a Dirichlet form associated to $L$. Recall from Section 1 that $\mathcal{E}^0$ has a closed extension $\mathcal{E}$. By definition,

$$f_1 \in C_c^\infty(G) \quad \text{and} \quad Lf_1 = h \Longleftrightarrow \mathcal{E}(f_1, f_2) = (h, f_2), \quad \forall f_2 \in \mathrm{Dom}(\mathcal{E}).$$

Now note that

$$\begin{aligned}\mathcal{E}^0(f\circ\phi_r, f\circ\phi_r) &= \sum_{i=1}^k \int_G \langle \tilde{X}_i(f\circ\phi_r)\rangle^2(g)dg\\ &= \sum_{i=1}^k r^2 \int |(\tilde{X}_i f)\circ\phi_r|^2(g)dg\\ &= \sum_{i=1}^k r^2 \int |\tilde{X}_i f|^2(g) J(r^{-1})dg = r^2 J(r^{-1})\mathcal{E}^0(f,f),\end{aligned}$$

where $J(r)$ is the Jacobian of the transformation $\phi_r$,

$$J(r) = \prod_{j=1}^m (r^j)^{d_j}$$

with $d_j = \dim(V_j)$. Thus, $J(r^{-1}) = J(r)^{-1}$. So $f \in \mathrm{Dom}(\mathcal{E})$ implies that $f\circ\phi_r \in \mathrm{Dom}(\mathcal{E})$, and, in general, $\mathcal{E}(f\circ\phi_r, h\circ\phi_r) = r^2 J(r^{-1})\mathcal{E}(f,h)$, for $f, h \in \mathrm{Dom}(\mathcal{E})$. Replacing $h$ here by $h\circ\phi_{r^{-1}}$ gives

$$\begin{aligned}\mathcal{E}(f\circ\phi_r, h) &= r^2 J(r^{-1})\mathcal{E}(f, h\circ\phi_{r^{-1}})\\ &= r^2 J(r^{-1})(Lf, h\circ\phi_{r^{-1}})_{L^2(G)}\\ &= r^2 J(r^{-1})J(r)(Lf\circ\phi_r, h)_{L^2(G)} = r^2(Lf\circ\phi_r, h)_{L^2(G)},\end{aligned}$$

which implies that if $f \in \mathrm{Dom}(L)$, then $f\circ\phi_r \in \mathrm{Dom}(L)$ and $L(f\circ\phi_r) = r^2 Lf\circ\phi_r$.

Now, for $r > 0$, let $U_r : L^2(G) \to L^2(G)$ be the unitary operator given by $U_r f = \frac{1}{\sqrt{J(r^{-1})}} f\circ\phi_r$. Then

$$LU_r = r^2 U_r L = U_r(r^2 L)$$

as operators, and thus $U_r^{-1} L U_r = r^2 L$. Then

$$U_r^{-1}\mathrm{e}^{tL/2}U_r = \mathrm{e}^{tU_r^{-1}LU_r/2} = \mathrm{e}^{r^2 tL/2},$$

from which it follows that

$$r^2\mathrm{e}^{tL/2}(f\circ\phi_r) = \mathrm{e}^{tL/2}U_r f = U_r \mathrm{e}^{r^2 tL/2} f = r^2(\mathrm{e}^{r^2 tL/2} f)\circ\phi_r. \quad \blacksquare$$

This gives the following proposition.

**Proposition 3.12.** *Suppose $G$ is a stratified Lie group with vector space decomposition $\oplus_{j=1}^m V_j$. Let $\{X_i\}_{i=1}^k \subset V_1$, $\nabla$, and $L$ be as above, and let $p \in (1,\infty)$. If $K_p$ is the best constant such that*

$$|\nabla P_1 f|^p \le K_p P_1 |\nabla f|^p,$$

*for all $f \in C_c^\infty(G)$, then $K_p(t) = K_p$ for all $t > 0$, where $K_p(t)$ is the function defined in Notation 1.7.*

**Proof.** By Eqs. (3.5) and (3.6),

$$\begin{aligned}|\nabla P_t(f\circ\phi_{t^{-1/2}})|^p &= |\nabla[(P_1 f)\circ\phi_{t^{-1/2}}]|^p = |t^{-1/2}(\nabla P_1 f)\circ\phi_{t^{-1/2}}|^p\\ &\le K_p t^{-p/2}(P_1|\nabla f|^p)\circ\phi_{t^{-1/2}} = K_p t^{-p/2} P_t(|\nabla f|^p\circ\phi_{t^{-1/2}})\\ &= K_p P_t(|\nabla f\circ\phi_{t^{-1/2}}|^p).\end{aligned}$$

Replacing $f$ by $f \circ \phi_{t^{1/2}}$ in the above computation proves the assertion. Moreover, reversing the above argument shows that $|\nabla P_t f|^p \le K_p P_t |\nabla f|^p$ implies that $|\nabla P_1 f|^p \le K_p P_1 |\nabla f|^p$. ■

*3.2.2. Nilpotent Lie groups*

Now let $G$ be a general nilpotent Lie group. Because not all nilpotent Lie groups admit dilations, the functions $K_p(t)$ are not scale invariant in this context. However, covering $G$ with a group which has a family of dilations adapted to its structure, shows that there exists some constant $K_p < \infty$ for which $K_p(t) < K_p$ for all $t > 0$.

**Definition 3.13.** Let $\mathcal{L} = \mathcal{L}(k, m)$ be the *free nilpotent Lie algebra of step m with k generators* $\{e_i\}_{i=1}^k$. Then $\mathcal{L}$ is the unique (up to isomorphism) nilpotent Lie algebra of rank $m$ such that, for every nilpotent Lie algebra $\mathfrak{g}$ of rank $m$ and map $\tilde{\Pi} : \{e_1, \ldots, e_k\} \to \mathfrak{g}$, there exists a unique homomorphism $\Pi : \mathcal{L} \to \mathfrak{g}$ which extends $\tilde{\Pi}$. Let $\mathcal{N} = \mathcal{N}(k, m)$ be the *free nilpotent Lie group of rank m with k generators*, which is the simply connected group of $\mathcal{L}(k, m)$.

The Lie algebra $\mathcal{L}(k, m)$ admits a vector space decomposition by setting $V_1 = \text{span}\{e_1, \ldots, e_k\}$. Thus, $\mathcal{N}$ is a stratified Lie group with Hörmander set $\{e_i\}_{i=1}^k \subset \mathcal{L}$; for definitions and further details, see [33]. Let $\nabla_{\mathcal{L}} = (\tilde{e}_1, \ldots, \tilde{e}_k)$, $\mathscr{L} = \sum_{i=1}^k \tilde{e}_i^2$, and $\mathscr{P}_t = e^{t\mathscr{L}/2}$. Theorem 3.5 and Proposition 3.12 imply that, for all $p \in (1, \infty)$, there exist constants $K_p^{\mathcal{L}} < \infty$ such that

$$|\nabla_{\mathcal{L}} \mathscr{P}_t f|^p \le K_p^{\mathcal{L}} \mathscr{P}_t |\nabla_{\mathcal{L}} f|^p, \tag{3.7}$$

for all $f \in C_c^\infty(\mathcal{N})$ and $t > 0$.

**Proposition 3.14.** *Let $G$ be a nilpotent group of step $m$ with Hörmander set $\{X_i\}_{i=1}^k$. Then $K_p(t) \le K_p^{\mathcal{L}}$ for all $t > 0$, where $K_p(t)$ is the function defined in Notation 1.7.*

**Proof.** By definition of $\mathcal{L} = \mathcal{L}(k, m)$, there exists a unique Lie algebra homomorphism $\Pi : \mathcal{L} \to \mathfrak{g}$ such that $\Pi(e_i) = X_i$. Then $\Pi$ induces a group homomorphism $\pi : \mathcal{N} \to G$ via the exponential maps,

$$\pi = \exp_G \circ\, \Pi \circ \exp_{\mathcal{N}}^{-1}.$$

Again, because $\mathcal{N}$ is a simply connected nilpotent Lie group, the exponential map on $\mathcal{L}$ is a global diffeomorphism. Note that $\pi_* = \Pi$,

$$\begin{array}{ccc} \mathcal{L}(k,m) & \xrightarrow{\Pi} & \mathfrak{g} \\ {\scriptstyle \exp_{\mathcal{N}}}\downarrow & & \downarrow{\scriptstyle \exp_G} \\ \mathcal{N}(k,m) & \xrightarrow[\pi]{} & G \end{array}$$

and the vector fields $\tilde{X}_i$ and $\tilde{e}_i$ are $\pi$-related; that is,

$$\tilde{e}_\alpha (f \circ \pi) = (\tilde{X}_\alpha f) \circ \pi,$$

for any multi-index $\alpha \in \Lambda^r$ and $f \in C_c^\infty(G)$. Note that $f \circ \pi \in C_b^\infty(\mathcal{N})$ and has bounded first order derivatives. Thus, by Proposition 3.6,

$$|\nabla P_t f|^p(e) = |\nabla_{\mathcal{L}} \mathscr{P}_t (f \circ \pi)|^p(e_{\mathcal{N}}) \le K_p^{\mathcal{L}} \mathscr{P}_t |\nabla_{\mathcal{L}} (f \circ \pi)|^p(e_{\mathcal{N}}) = K_p^{\mathcal{L}} P_t |\nabla f|^p(e),$$

where $e_{\mathcal{N}}$ is the identity element of $\mathcal{N}$. Since $K_p(t)$ is the best constant for which

$$|\nabla P_t f|^p(e) \leq K_p(t) P_t |\nabla f|^p(e)$$

holds, the above implies that $K_p(t) \leq K_p^{\mathcal{L}}$ for all $t > 0$. ■

This method of lifting the vector fields to a free nilpotent Lie algebra was learned from [32, 33]. A generalization of this procedure may be found in [27].

**Remark 3.15.** Note that the above argument is independent of the minimality of the Hörmander set $\{X_i\}_{i=1}^k$. So suppose that the collection $\{X_i\}_{i=1}^k$ spans the Lie algebra $\mathfrak{g}$. Since $G$ is a nilpotent Lie group (and thus unimodular,) it is then well known that the operator $L = \sum_{i=1}^k \tilde{X}_i^2$ is in fact the Laplace–Beltrami operator on the Riemannian manifold $(G, \langle\cdot,\cdot\rangle)$. Then it is well known that the inequality $(I_p)$ holds with exponential coefficients:

$$|\nabla P_t f|^p \leq \mathrm{e}^{pkt} P_t |\nabla f|^p,$$

where $-2k$ is a lower bound on the Ricci curvature; see for example Theorem 1.1 in [10]. Proposition 3.14 improves this result by implying that there exists a $K_p < \infty$ independent of $t$ such that

$$|\nabla P_t f|^p \leq K_p P_t |\nabla f|^p,$$

for all $f \in C_p^\infty(G)$ and $t > 0$. This implies the following corollary.

**Corollary 3.16.** *Let $G$ be a nilpotent Lie group of step $m$, and $\{X_i\}_{i=1}^k \subset \mathfrak{g}$ such that $\{X_i\}_{i=1}^k$ spans the Lie algebra $\mathfrak{g}$. Then, for $K_p(t)$ as in Notation 1.7,*

$$K_p(t) \leq \min\{K_p^{\mathcal{L}}, \mathrm{e}^{pkt}\},$$

*where $K_p^{\mathcal{L}}$ is the best constant so that $(I_p)$ holds on $\mathcal{L}(k, m)$ and $-2k$ is a lower bound on the Ricci curvature associated to the Riemannian metric determined by $L = \sum_{i=1}^k \tilde{X}_i^2$.*

This also gives the following Poincaré inequality for nilpotent Lie groups.

**Corollary 3.17.** *Suppose $G$ is a nilpotent Lie group, and let $K_2$ be a finite constant for which the inequality $(I_p)$ holds for $p = 2$. Then the inequality (3.3) holds with $\Lambda(t) = K_2 t$, for all $t > 0$.*

## Acknowledgements

I thank Bruce Driver for suggesting this problem and for many valuable discussions throughout the preparation of this work. The author was supported in part by NSF Grants 99-71036 and DMS 0202939.

## References

[1] Pascal Auscher, Thierry Coulhon, Xuan Thinh Duong, Steve Hofmann, Riesz transform on manifolds and heat kernel regularity, Ann. Sci. École Norm. Sup. (4) 37 (6) (2004) 911–957. MR MR2119242 (2005k:58043).

[2] Dominique Bakry, Ricci curvature and dimension for diffusion semigroups, in: Stochastic Processes and their Applications in Mathematics and Physics, Bielefeld, 1985, in: Math. Appl., vol. 61, Kluwer Acad. Publ., Dordrecht, 1990, pp. 21–31. MR 92e:58231.

[3] Dominique Bakry, Michel Émery, Hypercontractivité de semi-groupes de diffusion, C. R. Acad. Sci. Paris Sér. I Math. 299 (15) (1984) 775–778. MR 86f:60097.

[4] Dominique Bakry, Michel Émery, Diffusions hypercontractives, in: Séminaire de probabilités, XIX, 1983/84, in: Lecture Notes in Math., vol. 1123, Springer, Berlin, 1985, pp. 177–206. MR 88j:60131.

[5] Denis R. Bell, Salah Eldin A. Mohammed, The Malliavin calculus and stochastic delay equations, J. Funct. Anal. 99 (1) (1991) 75–99. MR 92k:60124.

[6] Thierry Coulhon, Xuan Thinh Duong, Riesz transform and related inequalities on noncompact Riemannian manifolds, Comm. Pure Appl. Math. 56 (12) (2003) 1728–1751. MR 2 001 444.

[7] Edward Brian Davies, One-parameter Semigroups, in: London Mathematical Society Monographs, vol. 15, Academic Press Inc. [Harcourt Brace Jovanovich Publishers], London, 1980, MR 82i:47060.

[8] Bruce K. Driver, Curved Wiener space analysis, in: Real and Stochastic Analysis, in: Trends Math., Birkhäuser Boston, Boston, MA, 2004, pp. 43–198. MR2090752 (2005g:58066).

[9] Bruce K. Driver, Leonard Gross, Hilbert spaces of holomorphic functions on complex Lie groups, in: New Trends in Stochastic Analysis, Charingworth, 1994, World Sci. Publishing, River Edge, NJ, 1997, pp. 76–106. MR 2000h:46029.

[10] Bruce K. Driver, Tai Melcher, Hypoelliptic heat kernel inequalities on the Heisenberg group, J. Funct. Anal. 221 (2) (2005) 340–365.

[11] Gerald B. Folland, Elias M. Stein, Hardy Spaces on Homogeneous Groups, Princeton University Press, Princeton, NJ, 1982.

[12] Roe W. Goodman, Nilpotent Lie Groups: Structure and Applications to Analysis, in: Lecture Notes in Mathematics, vol. 562, Springer-Verlag, Berlin, 1976, MR 56 #537.

[13] Leonard Gross, Logarithmic Sobolev inequalities, Amer. J. Math. 97 (4) (1975) 1061–1083. MR 54 #8263.

[14] Leonard Gross, Logarithmic Sobolev inequalities and contractivity properties of semigroups, in: Dirichlet Forms, Varenna, 1992, in: Lecture Notes in Math., vol. 1563, Springer, Berlin, 1993, pp. 54–88. MR 95h:47061.

[15] Lars Hörmander, Hypoelliptic second order differential equations, Acta Math. 119 (1967) 147–171. MR 36 #5526.

[16] Elton P. Hsu, Stochastic Analysis on Manifolds, in: Graduate Studies in Mathematics, vol. 38, American Mathematical Society, Providence, RI, 2002, MR 2003c:58026.

[17] Nobuyuki Ikeda, Shinzo Watanabe, Stochastic Differential Equations and Diffusion Processes, second ed., in: North-Holland Mathematical Library, vol. 24, North-Holland Publishing Co., Amsterdam, 1989, MR 90m:60069.

[18] Shigeo Kusuoka, Daniel Stroock, Applications of the Malliavin calculus. II, J. Fac. Sci. Univ. Tokyo Sect. IA Math. 32 (1) (1985) 1–76. MR 86k:60100b.

[19] Shigeo Kusuoka, Daniel Stroock, Applications of the Malliavin calculus. III, J. Fac. Sci. Univ. Tokyo Sect. IA Math. 34 (2) (1987) 391–442. MR 89c:60093.

[20] Hong-Quan Li, Estimation optimale du gradient du semi-groupe de la chaleur sur le groupe de Heisenberg, J. Funct. Anal. 236 (2) (2006) 369–394. MR MR2240167.

[21] Paul Malliavin, $C^k$-hypoellipticity with degeneracy, in: Stochastic Analysis (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1978), Academic Press, New York, 1978, pp. 199–214. MR 80i:58045a.

[22] Paul Malliavin, $C^k$-hypoellipticity with degeneracy. II, in: Stochastic Analysis (Proc. Internat. Conf., Northwestern Univ., Evanston, Ill., 1978), Academic Press, New York, 1978, pp. 327–340. MR 80i:58045b.

[23] Tai Melcher, Malliavin calculus for Lie group-valued Wiener functions, Preprint, http://faculty.virginia.edu/melcher, 2007.

[24] Tai Melcher, Some convergence arguments for matrix group-valued SDE solutions, http://faculty.virginia.edu/melcher, 2007.

[25] David Nualart, The Malliavin calculus and related topics, in: Probability and its Applications, Springer-Verlag, New York, 1995, MR 96k:60130.

[26] Jean Picard, Gradient estimates for some diffusion semigroups, Probab. Theory Related Fields 122 (4) (2002) 593–612. MR 2003d:58056.

[27] Linda Preiss Rothschild, E.M. Stein, Hypoelliptic differential operators and nilpotent groups, Acta Math. 137 (3–4) (1976) 247–320. MR 55 #9171.

[28] Ichiro Shigekawa, Derivatives of Wiener functionals and absolute continuity of induced measures, J. Math. Kyoto Univ. 20 (2) (1980) 263–289. MR 83g:60051.

[29] Daniel W. Stroock, The Malliavin calculus and its application to second order parabolic differential equations. I, Math. Syst. Theory 14 (1) (1981) 25–65. MR 84d:60092a.

[30] Daniel W. Stroock, The Malliavin calculus and its application to second order parabolic differential equations. II, Math. Syst. Theory 14 (2) (1981) 141–171. MR 84d:60092b.

[31] V.S. Varadarajan, Lie groups, Lie algebras, and their Representations, in: Graduate Texts in Mathematics, vol. 102, Springer-Verlag, New York, 1984, Reprint of the 1974 edition MR 85e:22001.

[32] N.Th. Varopoulos, Analysis on nilpotent groups, J. Funct. Anal. 66 (3) (1986) 406–431. MR 88h:22014.
[33] N.Th. Varopoulos, L. Saloff-Coste, T. Coulhon, Analysis and Geometry on Groups, in: Cambridge Tracts in Mathematics, vol. 100, Cambridge University Press, Cambridge, 1992, MR 95f:43008.