

*PERTURBATION RESULTS CONCERNING GAUSSIAN ESTIMATES
AND HYPOELLIPTICITY FOR LEFT-INVARIANT LAPLACIANS ON
COMPACT GROUPS*

BY

QI HOU (Beijing) and LAURENT SALOFF-COSTE (Ithaca, NY)

Abstract. We study left-invariant Laplacians on compact connected groups that are form-comparable perturbations of bi-invariant Laplacians. Our results show that Gaussian bounds for derivatives of heat kernels enjoyed by certain bi-invariant Laplacians hold for their form-comparable perturbations. We further show that the parabolic operators associated with such left-invariant Laplacians, in particular, with the bi-invariant Laplacians, are hypoelliptic in various senses.

1. Introduction. Let (G, ν) be a compact connected metrizable group where ν is the normalized Haar measure. The main interest of the results developed here is for infinite-dimensional groups G , such as infinite products of compact Lie groups. Symmetric Gaussian convolution semigroups of measures on G exhibit a rich variety of behaviors [1, 2, 3]. In [4], Bendikov and Saloff-Coste showed that (1) when they exist, continuous density functions of symmetric central Gaussian convolution semigroups $(\mu_t)_{t>0}$ of measures on G admit spatial derivatives of all orders in certain directions, and (2) the derivatives satisfy certain Gaussian upper bounds if the density function satisfies $\lim_{t \rightarrow 0} t \log \mu_t(e) = 0$ (referred to as the (CK*) condition). The same authors studied various function spaces on G with norms involving different combinations of derivatives, and examined hypoellipticity properties related to the generators of such Gaussian semigroups. See [6, 7]. In these works, the semigroup being central, or equivalently, its generator being bi-invariant, plays an essential role.

In the present paper we consider symmetric (noncentral) convolution semigroups on G (referred to as perturbed semigroups) that are comparable to symmetric central Gaussian semigroups in the sense that their associated Dirichlet forms are comparable.

2020 *Mathematics Subject Classification*: Primary 22E66; Secondary 31C25, 60G51, 60B15.

Key words and phrases: hypoellipticity, Gaussian convolution semigroup, heat equation.

Received 25 October 2021; revised 4 August 2022.

Published online 8 November 2022.

More precisely, any generator $-\Delta$ of a symmetric central Gaussian semigroup $(\mu_t^\Delta)_{t>0}$ on G is a bi-invariant operator of the form $-\Delta = \sum_{i=1}^\infty X_i^2$, where $\{X_i\}_i$ forms a basis of left-invariant vector fields in some proper sense (i.e. a projective basis). In such a projective basis, the generator $-L$ of any symmetric Gaussian semigroup $(\mu_t^L)_{t>0}$ is of the form $-L = \sum_{i,j} a_{ij} X_i X_j$ (see [10, 11]). This operator L is only left-invariant in general. If further the following relation holds between the Dirichlet forms associated with Δ, L :

$$c\mathcal{E}_\Delta \leq \mathcal{E}_L \leq C\mathcal{E}_\Delta,$$

for some $c, C > 0$, we call L a *form-comparable perturbation* of Δ . In this paper we only consider those L that are form-comparable perturbations of bi-invariant Laplacians Δ .

We show that when the semigroup μ_t^Δ admits a continuous density function, denoted again by μ_t^Δ , so does the semigroup μ_t^L , and the density functions μ_t^Δ and μ_t^L both belong to certain smooth function spaces associated with Δ and L (Theorem 4.1). We then show that when μ_t^Δ further satisfies the (CK*) condition, (1) both density functions and their derivatives satisfy certain Gaussian estimates, and (2) both $\partial_t + \Delta$ and $\partial_t + L$ are hypoelliptic in various senses (Proposition 4.2 and Theorem 3.5). As in [7], our proof for the hypoellipticity properties follows the general heat kernel/semigroup approach by Kusuoka and Stroock [16], with some additional ideas to overcome the difficulties brought by dropping the bi-invariance assumption. Note that in the bi-invariant operator case, the hypoellipticity of $\partial_t + \Delta$ is new even though the hypoellipticity of Δ has been studied in [7, 9]. For a thorough discussion of hypoellipticity in this context, see [9].

We organize this paper as follows. In Section 2 we briefly introduce the setting and fix notations. In Section 3 we present the main theorem on hypoellipticity, Theorem 3.5, and give an example. Sections 4 and 5 are devoted to the proof of the hypoellipticity theorem, where other theorems regarding heat kernels are presented first and they constitute an important part of the proof. Some details are postponed to the Appendix.

2. Setting and notation. This section contains a minimal introduction to the setting of our study. For more details, see [3, 5, 8] and the references therein. Write $\mathbb{N} := \{0, 1, \dots\}$.

2.1. Basic structures. In this entire paper, let G be a compact connected metrizable group with identity element e . The group G can be viewed as the projective limit of a sequence of Lie groups [14, 15],

$$G = \varprojlim_{\alpha \in \mathbb{N}} G_\alpha.$$

Here the index set $\aleph$ is finite or countable; $G_\alpha = G/K_\alpha$ where $\{K_\alpha\}$ is a decreasing sequence of compact normal subgroups with $\bigcap_{\alpha \in \aleph} K_\alpha = \{e\}$; for $\alpha \leq \beta$, the map $\pi_{\alpha,\beta} : G_\alpha \rightarrow G_\beta$ is the projection map. Denote the projection map $G \rightarrow G_\alpha = G/K_\alpha$ by π_α .

The Bruhat test function space $\mathcal{B}(G)$ is defined as

$$\mathcal{B}(G) := \{f : G \rightarrow \mathbb{R} : f = \phi \circ \pi_\alpha \text{ for some } \alpha \in \aleph, \phi \in C^\infty(G_\alpha)\}.$$

Here $C^\infty(G_\alpha)$ denotes the set of all smooth functions on G_α ; $\mathcal{B}(G)$ is independent of the choice of $\{K_\alpha\}_{\alpha \in \aleph}$ [12, 7]. The Bruhat test functions are generalizations of cylindric functions on the infinite torus $\mathbb{T}^\infty$ (i.e. smooth functions that depend on only finitely many coordinates). Let $\mathcal{B}'(G)$ be the dual space of $\mathcal{B}(G)$ with the strong dual topology; elements of $\mathcal{B}'(G)$ are called *Bruhat distributions*. In Section 2.4 we introduce more smooth function spaces on G and their dual spaces as distribution spaces.

Lastly, we recall the following definitions of convolutions.

- The convolution of any two Borel measures μ_1, μ_2 on G , $\mu_1 * \mu_2$, is the measure defined by

$$\mu_1 * \mu_2(f) = \int_{G \times G} f(xy) d\mu_1(x) d\mu_2(y), \quad \forall f \in C(G).$$

- The convolution of any two functions $f, g \in C(G)$ is the function given by

$$f * g(x) = \int_G f(xy^{-1})g(y) d\nu(y) = \int_G f(y)g(y^{-1}x) d\nu(y).$$

- The convolution of a function $f \in C(G)$ and a Borel measure μ is defined as

$$\mu * f(x) = \int_G f(y^{-1}x) d\mu(y), \quad f * \mu(x) = \int_G f(xy^{-1}) d\mu(y).$$

- The convolution of a function $f \in \mathcal{B}(G)$ and a distribution $U \in \mathcal{B}'(G)$ is defined by

$$(f * U)(\phi) = U(\check{f} * \phi), \quad (U * f)(\phi) = U(\phi * \check{f}), \quad \forall \phi \in \mathcal{B}(G).$$

2.2. Gaussian semigroups and generators. Recall that a family $(\mu_t)_{t>0}$ of probability measures on G is called a *symmetric Gaussian convolution semigroup*, if it satisfies the following properties:

- (i) (semigroup property) $\mu_t * \mu_s = \mu_{t+s}$ for any $t, s > 0$;
- (ii) (weakly continuous) $\mu_t \rightarrow \delta_e$ weakly as $t \rightarrow 0$;
- (iii) (Gaussian) $t^{-1}\mu_t(V^c) \rightarrow 0$ as $t \rightarrow 0$ for any neighborhood V of the identity $e \in G$;
- (iv) (symmetric) $\check{\mu}_t = \mu_t$ for any $t > 0$; here $\check{\mu}_t$ is defined by $\check{\mu}_t(V) = \mu_t(V^{-1})$ for any Borel subset $V \subset G$.

For $t > 0$, set

$$H_t f(x) := f * \check{\mu}_t(x) = \int_G f(xy) d\mu_t(y), \quad \forall f \in C(G),$$

and then extend it to $L^2(G)$. This formula defines a semigroup of self-adjoint operators on $L^2(G)$. Let $-L$ be the (L^2) -infinitesimal generator of $(H_t)_{t>0}$,

$$-L f := \lim_{t \rightarrow 0} \frac{H_t f - f}{t},$$

with domain $\mathcal{D}(L)$ equal to the space of functions in $L^2(G)$ for which this limit exists in $L^2(G)$.

2.3. Projective bases and sub-Laplacians. The Lie algebra of G (denoted by $\mathfrak{g}$) is defined to be the projective limit of the Lie algebras of G_α (denoted by $\mathfrak{g}_\alpha$), with projection maps $d\pi_{\alpha,\beta}$ for $\alpha, \beta \in \mathbb{N}$, $\alpha \leq \beta$. A family $\{X_i\}_{i \in \mathcal{I}}$ of elements of $\mathfrak{g}$ is called a *projective basis* of $\mathfrak{g}$, if for each $\alpha \in \mathbb{N}$, there is a finite subset $\mathcal{I}_\alpha \subset \mathcal{I}$ such that $d\pi_\alpha(X_i) = 0$ for all $i \notin \mathcal{I}_\alpha$, and $\{d\pi_\alpha(X_i)\}_{i \in \mathcal{I}_\alpha}$ is a basis of the Lie algebra $\mathfrak{g}_\alpha$. The Lie algebra $\mathfrak{g}$ admits many projective bases [11, 15]. For future use, let

$$\mathbb{R}^{(\mathcal{I})} := \{\xi = (\xi_i)_{i \in \mathcal{I}} : \text{all but finitely many entries are zero}\}.$$

For a fixed projective basis $X = \{X_i\}_{i \in \mathcal{I}}$, to any infinite real symmetric nonnegative coefficient matrix $A = (a_{ij})_{\mathcal{I} \times \mathcal{I}}$, we associate the second-order left-invariant differential operator (acting on $\mathcal{B}(G)$)

$$(2.1) \quad L_A := - \sum_{i,j \in \mathcal{I}} a_{ij} X_i X_j.$$

Here A being real symmetric nonnegative means that $a_{ij} = a_{ji}$ are real numbers, and $\sum a_{ij} \xi_i \xi_j \geq 0$ for all $\xi = (\xi_i)_{i \in \mathcal{I}} \in \mathbb{R}^{(\mathcal{I})}$. We refer to these differential operators as *sub-Laplacians*. The matrix A can be degenerate. When A is positive definite, we call L_A a *Laplacian*.

There is a one-to-one correspondence between symmetric Guassian convolution semigroups $(\mu_t)_{t>0}$ and infinite real symmetric nonnegative matrices A [7, 10]. This correspondence is given by the fact that the infinitesimal generator of $(\mu_t)_{t>0}$ (and of the corresponding semigroup $(H_t)_{t>0}$) is of the form $-L_A$ for some matrix A . More precisely, this generator, computed on Bruhat test functions, has the form

$$\lim_{t \rightarrow 0} \frac{H_t f - f}{t} = -L_A f = \sum_{i,j \in \mathcal{I}} a_{ij} X_i X_j f.$$

Conversely, given a sub-Laplacian L , there is a semigroup that we denote by $(H_t^L)_{t>0}$, whose generator agrees with $-L$ in the above sense. The symmetric Gaussian semigroup corresponding to $(H_t^L)_{t>0}$ is denoted by $(\mu_t^L)_{t>0}$. The semigroup $(H_t^L)_{t>0}$ is Markov and commutes with left translations. See

e.g. [7] for more details. Let $(\mathcal{E}_L, \mathcal{F}_L)$ denote the Dirichlet form associated with $(H_t^L)_{t>0}$ with domain $\mathcal{F}_L$.

2.4. Smooth function spaces and distribution spaces. Let $I \subset \mathbb{R}$ be an open interval. Corresponding to $\mathcal{B}(G)$ defined in the first subsection, define the Bruhat test function space on $I \times G$, $\mathcal{B}(I \times G)$, as

$$\mathcal{B}(I \times G) := \{f : I \times G \rightarrow \mathbb{R} : f = \phi \circ \tilde{\pi}_\alpha \text{ for some } \alpha \in \aleph, \phi \in C_c^\infty(I \times G_\alpha)\}.$$

Here $\tilde{\pi}_\alpha : I \times G \rightarrow I \times G_\alpha$, $(t, x) \mapsto (t, \pi_\alpha(x))$, is the projection map, and $C_c^\infty(I \times G_\alpha)$ is the set of smooth functions on $I \times G_\alpha$ with compact support.

In the following we review some norms and seminorms and consider the function spaces as completions of $\mathcal{B}(G)$ with respect to these (semi)norms. See [6] for detailed discussions of these function spaces. Along the way we define the corresponding spaces on $I \times G$; they serve as generalizations of the classical test function space $\mathcal{D}(\mathbb{R}^n)$ to spaces of test functions in $I \times G$.

REMARK 2.1. We only define the “smooth” function spaces in the sense that the completions are with respect to (semi)norms of all orders. Spaces as completions with respect to (semi)norms of orders up to some $k \in \mathbb{N}$ can be correspondingly defined, and denoted accordingly with the “ ∞ ” superscripts replaced by k .

Spaces $\mathcal{C}_X^\infty(G)$ and $\mathcal{C}_X^\infty(I \times G)$. Given any projective basis $\{X_i\}_{i \in \mathcal{I}}$ and multi-index $l = (l_1, \dots, l_k) \in \mathcal{I}^k$ where $k \in \mathbb{N}$, for any $f \in \mathcal{B}(G)$, denote

$$X^l f(x) = X_{l_1} \cdots X_{l_k} f(x) =: D_x^k f(X_{l_1}, \dots, X_{l_k}).$$

For $k = 0$, this is taken as f itself. Let $\mathcal{C}_X^\infty(G)$ be the completion of $\mathcal{B}(G)$ with respect to the seminorms

$$\left\{ \|X^l f\|_{L^\infty(G)} = \sup_G |X^l f| : l \in \mathcal{I}^k, k \in \mathbb{N} \right\}.$$

Correspondingly, for any open interval $J \Subset I$, let $\mathcal{C}_{0,X}^\infty(J \times G)$ be the completion of $\mathcal{B}(J \times G)$ with respect to the seminorms

$$\left\{ \|\partial_t^m X^l f\|_{L^\infty(J \times G)} = \sup_{J \times G} |\partial_t^m X^l f| : l \in \mathcal{I}^k, m, k \in \mathbb{N} \right\}.$$

Here the subscript “0” refers to the vanishing of the function and all its derivatives on the boundary ∂J . To clarify notations, if $f = \phi \circ \tilde{\pi}_\alpha$ is as in the definition of $\mathcal{B}(J \times G)$, then

$$\partial_t^m X_{l_1} \cdots X_{l_k} f(t, x) = (\partial_t^m d\pi_\alpha(X_{l_1}) \cdots d\pi_\alpha(X_{l_k}) \phi)(t, \pi_\alpha(x)).$$

The spaces are equipped with the topology defined by the seminorms above. Note that for any two open intervals $I_1 \Subset I_2$, functions in $\mathcal{C}_{0,X}^\infty(I_1 \times G)$ can be naturally extended by 0 to $I_2 \times G$, thus $\mathcal{C}_{0,X}^\infty(I_1 \times G) \subset \mathcal{C}_{0,X}^\infty(I_2 \times G)$.

Finally, define $\mathcal{C}_X^\infty(I \times G)$ as the inductive limit of $\{\mathcal{C}_{0,X}^\infty(J \times G)\}_{J \in I}$.

Spaces $\mathcal{S}_X^\infty(G)$ and $\mathcal{S}_X^\infty(I \times G)$. For any $k \in \mathbb{N}$, $f \in \mathcal{B}(G)$, define a function $|D^k f|_X : G \rightarrow [0, +\infty]$ by

$$\begin{aligned} |D^k f|_X(x) &= |D_x^k f|_X \\ &:= \left(\sum_{l \in \mathcal{I}^k} |D_x^k f(X_{l_1}, \dots, X_{l_k})|^2 \right)^{1/2} = \left(\sum_{l \in \mathcal{I}^k} |X^l f(x)|^2 \right)^{1/2}. \end{aligned}$$

When $k = 1$, $|D_x^1 f|_X^2 = \sum_{i \in \mathcal{I}} |X_i f(x)|^2$ is the square of the gradient of f , often denoted by $\Gamma(f, f)$ in the literature (see e.g. [6]). Let $\mathcal{S}_X^\infty(G)$ be the completion of $\mathcal{B}(G)$ with respect to the norms $\{S_X^k\}_{k \in \mathbb{N}}$ defined as

$$S_X^k(f) := \sup_{0 \leq m \leq k} \| |D^m f|_X \|_{L^\infty(G)} = \sup_{0 \leq m \leq k} \sup_{x \in G} |D_x^m f|_X.$$

REMARK 2.2. It is shown in [6] that for any two projective bases $\{X_i\}_{i \in \mathcal{I}}$ and $\{Z_i\}_{i \in \mathcal{I}}$ such that $\sum X_i^2 = \sum Z_i^2$, for all $x \in G$ and $k \in \mathbb{N}$, $|D_x^k f|_X = |D_x^k f|_Z$ and $S_X^k(f) = S_Z^k(f)$. In other words, the space $\mathcal{S}_X^\infty(G)$ depends on the projective basis $\{X_i\}_{i \in \mathcal{I}}$ through its associated sum of squares operator

$$L = - \sum_{i \in \mathcal{I}} X_i^2.$$

As $\{X_i\}_{i \in \mathcal{I}}$ is taken as a projective basis, L is what we called a Laplacian earlier. Below, we write $S_L^k(f) = S_X^k(f)$, $|D_x^k f|_L = |D_x^k f|_X$, $\mathcal{S}_L^\infty(G) = \mathcal{S}_X^\infty(G)$.

For any projective basis $X = \{X_i\}_{i \in \mathcal{I}}$ and $L = - \sum X_i^2$, for any open interval $J \Subset I$, the space $\mathcal{S}_{0,X}^\infty(J \times G) = \mathcal{S}_{0,L}^\infty(J \times G)$ is defined as the completion of $\mathcal{B}(J \times G)$ with respect to the (semi)norms

$$\left\{ S_L^{k,p}(J \times G, f) := \sup_{0 \leq a \leq p} \sup_{t \in J} S_L^k(\partial_t^a f) = \sup_{\substack{0 \leq a \leq p \\ 0 \leq m \leq k}} \sup_{(t,x) \in J \times G} |D_x^m(\partial_t^a f)|_L \right\}_{k,p \in \mathbb{N}}.$$

Here writing $J \times G$ in $S_L^{k,p}(J \times G, f)$ is to clarify the set on which the supremum is taken. The space $\mathcal{S}_L^\infty(I \times G)$ is defined as the inductive limit of the spaces $\{\mathcal{S}_{0,L}^\infty(J \times G)\}_{J \in I}$.

Spaces $\mathcal{T}_L^\infty(G)$ and $\mathcal{T}_L^\infty(I \times G)$. It is not known if $\mathcal{S}_L^\infty(G)$ is contained in the $C(G)$ -domain of L [6, p. 178]. There, to address this problem, the authors introduced the space $\mathcal{T}_L^\infty(G)$ which we now recall.

Fix any projective basis $X = \{X_i\}_{i \in \mathcal{I}}$ and let $L = - \sum_{i \in \mathcal{I}} X_i^2$. For any $k \in \mathbb{N}$ and $\lambda = (\lambda_0, \lambda_1, \dots, \lambda_k) \in \mathbb{N}^{k+1}$, and any $f \in \mathcal{B}(G)$, let

$$|D_x^{k,\lambda} f|_X = |D_x^{k,\lambda} f|_L := \left(\sum_{l \in \mathcal{I}^k} |L^{\lambda_0} X_{l_1} L^{\lambda_1} X_{l_2} L^{\lambda_2} \dots X_{l_k} L^{\lambda_k} f(x)|^2 \right)^{1/2}.$$

The subscripts can be written as X or L because the quantity depends only on the sum of squares associated with the projective basis X ; see also

Remark 2.2. Now we take the supremum over k and λ to build a norm. We first introduce the following two abbreviations.

- For any $l \in \mathcal{I}^k$ and $\lambda \in \mathbb{N}^{k+1}$, set

$$P_L^{l,\lambda} f := L^{\lambda_0} X_{l_1} L^{\lambda_1} X_{l_2} L^{\lambda_2} \cdots X_{l_k} L^{\lambda_k} f.$$

For any bi-invariant Laplacian Δ , $P_\Delta^{l,\lambda} f = X_{l_1} X_{l_2} \cdots X_{l_k} \Delta^{\lambda_0 + \cdots + \lambda_k} f$.

- For any $k, m \in \mathbb{N}$, let $A(k, m)$ denote the set of all possible $(k+1)$ -tuples of integers that sum up to m , i.e.,

$$A(k, m) := \left\{ \lambda = (\lambda_0, \lambda_1, \dots, \lambda_k) : \lambda_i \in \mathbb{N}, \sum_{i=0}^k \lambda_i = m \right\}.$$

Using these notations, for any $N \in \mathbb{N}$, define

$$M_X^N(f) = M_L^N(f) :=$$

$$\sup_{x \in G} \sup_{\substack{k, m \in \mathbb{N} \\ k+2m \leq N}} \sup_{\lambda \in A(k, m)} |D_x^{k,\lambda} f|_L = \sup_{x \in G} \sup_{\substack{k, m \in \mathbb{N} \\ k+2m \leq N}} \sup_{\lambda \in A(k, m)} \left(\sum_{l \in \mathcal{I}^k} |P_L^{l,\lambda} f(x)|^2 \right)^{1/2}.$$

Define $\mathcal{T}_L^\infty(G)$ as the completion of $\mathcal{B}(G)$ with respect to the norms $\{M_L^N\}_{N \in \mathbb{N}}$. For any open interval $J \Subset I$, define $\mathcal{T}_{0,L}^\infty(J \times G)$ as the completion of $\mathcal{B}(J \times G)$ with respect to the set of (semi)norms

$$\begin{aligned} \left\{ M_L^{N,p}(J \times G, f) := \sup_{0 \leq a \leq p} \sup_{t \in J} M_L^N(\partial_t^a f) \right. \\ \left. = \sup_{\substack{a, k, m \in \mathbb{N} \\ 0 \leq a \leq p \\ k+2m \leq N}} \sup_{(t, x) \in J \times G} \sup_{\lambda \in A(k, m)} \left(\sum_{l \in \mathcal{I}^k} |P_L^{l,\lambda} \partial_t^a f(t, x)|^2 \right)^{1/2} \right\}, \end{aligned}$$

where $N, p \in \mathbb{N}$. As before, the space $\mathcal{T}_L^\infty(I \times G)$ is defined as the inductive limit of $\{\mathcal{T}_{0,L}^\infty(J \times G)\}_{J \Subset I}$.

Distribution spaces $\mathcal{T}'_L(G)$ and $\mathcal{T}'_L(I \times G)$. Let $\mathcal{T}'_L(G)$ and $\mathcal{T}'_L(I \times G)$ be the dual spaces of $\mathcal{T}_L^\infty(G)$ and $\mathcal{T}_L^\infty(I \times G)$, respectively, equipped with the strong dual topology. Recall that it means that for any $U \in \mathcal{T}'_L(I \times G)$ and any precompact open subset $J \Subset I$, there exist some $C(J) > 0$ and $N(J), p(J) \in \mathbb{N}$ (written as N, p below) such that for any $f \in \mathcal{T}_{0,L}^\infty(J \times G)$,

$$|U(f)| \leq C(J) M_L^{N,p}(f).$$

REMARK 2.3. In this paper, we deal with distributional solutions U of the heat equation $(\partial_t + L)U = F$. Because $\mathcal{C}_X^\infty$ and $\mathcal{S}_L^\infty$ may not be contained in the domain of L , to make sense of LU we do not consider their dual spaces.

To summarize, the above spaces on $I \times G$ satisfy

$$\begin{aligned} \mathcal{B}(I \times G) &\subset \mathcal{T}_L^\infty(I \times G) \subset \mathcal{S}_L^\infty(I \times G) \\ &\subset \mathcal{C}_X^\infty(I \times G) \subset \mathcal{T}'_L(I \times G) \subset \mathcal{B}'_L(I \times G); \end{aligned}$$

the same inclusions hold for the corresponding spaces on G .

By the previously recalled definitions of convolutions, the convolution of any two functions $f, g \in C(\mathbb{R} \times G)$ (at least one with compact support) is the function given by

$$(2.2) \quad f * g(s, x) = \int_{\mathbb{R}} \int_G f(t, z) g(s - t, z^{-1}x) dt d\nu(z).$$

Notation. To emphasize the difference between convolutions on G and on $I \times G$, in the rest of the paper we write $\star$ instead of $*$ for convolutions involving both time and space. For example, we write $f \star g$ for (2.2).

We also use the following convolution. Let f be any function with the property that $f, \check{f} \in \mathcal{T}_L^\infty(\mathbb{R} \times G)$, where $\check{f}(s, x) = f(-s, x^{-1})$. Let U be any distribution in $\mathcal{T}'_L(I \times G)$ with compact support in $I \times G$. We may define their convolutions as

$$(f \star U)(\phi) = U(\check{f} \star \phi), \quad (U \star f)(\phi) = U(\phi \star \check{f}), \quad \forall \phi \in \mathcal{T}_L^\infty(I \times G).$$

These convolutions are well-defined, because (1) both $\check{f} \star \phi$ and $\phi \star \check{f}$ belong to $\mathcal{T}_L(\mathbb{R} \times G)$; (2) as U has compact support in $I \times G$, there is some function $\eta \in \mathcal{T}_L(I \times G)$ with compact support such that $U = \eta U$.

2.5. Bi-invariant Laplacians and their perturbations. By definition, a symmetric Gaussian semigroup $(\mu_t)_{t>0}$ is called *central* if for all $t > 0$, $\mu_t(a^{-1}Ba) = \mu_t(B)$ for any Borel set $B \subset G$ and any $a \in G$. Central semigroups commute with any other factor in convolutions. Their generators are bi-invariant Laplacians. In this paper we use the symbol Δ to denote a bi-invariant Laplacian, and L for a general left-invariant sub-Laplacian that may or may not be bi-invariant. As a bi-invariant operator, Δ satisfies

$$\Delta Zf = Z\Delta f$$

for any left-invariant vector field $Z \in \mathfrak{g}$ and any smooth function $f \in \mathcal{B}(G)$.

Let L, P be two sub-Laplacians. We say that L, P are *form-comparable perturbations* of each other if their corresponding Dirichlet forms $(\mathcal{E}_L, \mathcal{F}_L)$ and $(\mathcal{E}_P, \mathcal{F}_P)$ are comparable. That is, there exist $c, C > 0$ such that for any $f \in \mathcal{B}(G)$,

$$(2.3) \quad c\mathcal{E}_P(f, f) \leq \mathcal{E}_L(f, f) \leq C\mathcal{E}_P(f, f).$$

Note that (2.3) implies that $\mathcal{F}_L = \mathcal{F}_P$. In the present paper we study perturbations of bi-invariant Laplacians; these perturbations have to be Laplacians themselves.

Given a bi-invariant Laplacian Δ and a left-invariant Laplacian L , we can mix them up to define the function/distribution spaces $\mathcal{T}_{\Delta,L}^\infty(I \times G)$, $\mathcal{T}'_{\Delta,L}(I \times G)$ as follows. Suppose $\Delta = -\sum_{i \in \mathcal{I}} X_i^2$. The space $\mathcal{T}_{\Delta,L}^\infty(I \times G)$ is the inductive limit of $\{\mathcal{T}_{0,\Delta,L}^\infty(J \times G)\}_{J \in I}$, where $\mathcal{T}_{0,\Delta,L}^\infty(J \times G)$ for each open interval $J \in I$ is defined as the completion of $\mathcal{B}(J \times G)$ with respect to the (semi)norms

$$(2.4) \quad M_{\Delta,L}^{N,p}(J \times G, f) := \sup_{\substack{0 \leq a \leq p \\ 2b+k+2m \leq N \\ (t,x) \in J \times G}} \sup_{\substack{\lambda \in \Lambda(k,m) \\ \lambda = (\lambda_0, \dots, \lambda_k)}} \left(\sum_{l \in \mathcal{I}^k} |L^{\lambda_0} X_{l_1} L^{\lambda_1} \dots X_{l_k} L^{\lambda_k} \Delta^b \partial_t^a f(t, x)|^2 \right)^{1/2}.$$

$\mathcal{T}'_{\Delta,L}(I \times G)$ is the strong topological dual of $\mathcal{T}_{\Delta,L}^\infty(I \times G)$.

REMARK 2.4. Suppose $L = -\sum_{i \in \mathcal{I}} Y_i^2$ for another projective basis $\{Y_i\}_{i \in \mathcal{I}}$. In the Appendix, we show that the norms defined using $\{Y_i\}_{i \in \mathcal{I}}$ instead of $\{X_i\}_{i \in \mathcal{I}}$ or using both $\{X_i\}_{i \in \mathcal{I}}$ and $\{Y_i\}_{i \in \mathcal{I}}$ are equivalent to the norms (2.4).

2.6. Properties of the spaces. We record here some properties of the above function spaces that are particularly useful in the present paper. See [6] for more details. The first three items hold for function spaces associated with general left-invariant sub-Laplacians $L = -\sum Y_i^2$.

- (i) All the function spaces on G introduced above ($\mathcal{B}(G)$, $\mathcal{C}_Y^k(G)$, $\mathcal{S}_L^k(G)$, $\mathcal{T}_L^k(G)$, $k \in \mathbb{N} \cup \{\infty\}$) are algebras for pointwise multiplication. See [6, Sections 2, 3]. Based on this fact it is straightforward to check that the corresponding function spaces on $I \times G$ are algebras for pointwise multiplication as well.
- (ii) Let $\mathfrak{S}(G)$ be any of the function spaces $\mathcal{B}(G)$, $\mathcal{C}_Y^k(G)$, $\mathcal{S}_L^k(G)$, and $\mathcal{T}_L^k(G)$. For any Borel measure μ , for any function $f \in \mathfrak{S}(G)$, we have $\mu * f \in \mathfrak{S}(G)$, and the map (convolution with μ on the left)

$$\mu * : \mathfrak{S}(G) \rightarrow \mathfrak{S}(G), \quad f \mapsto \mu * f,$$

is continuous. Let $\|\cdot\|$ denote any norm involved in the definitions of these spaces. Then $\|\mu * f\| \leq |\mu| \|f\|$ where $|\mu|$ is the total mass of μ . The same statement holds for $\mu \star : \mathfrak{S}(I \times G) \rightarrow \mathfrak{S}(I \times G)$, if the measure μ is on $I \times G$.

- (iii) For any sequence $\phi_n \in L^1(G)$ with $\phi_n \rightarrow \delta_e$ weakly as $n \rightarrow \infty$, for any $f \in \mathfrak{S}(G)$ as in the previous item, the sequence $f_n := \phi_n * f$ converges to f in $\mathfrak{S}(G)$. The same statement holds for any sequence $\phi_n \in L^1((0, 1) \times G)$ with $\phi_n \rightarrow \delta_{(0,e)}$ weakly and any $f \in \mathfrak{S}(I \times G)$.

When $\mathfrak{S} = \mathcal{B}(G)$ or $\mathcal{B}(I \times G)$, conclusions of (ii) and (iii) hold too for convolutions on the right. When the space $\mathfrak{S}$ is any of the other three types,

however, due to the fact that the differential operators are left-invariant (they act on the rightmost function in a convolution), the order of the convolutions in (ii) and (iii) is crucial. For instance, even if the approximation of identity $\{\phi_n\}_{n \in \mathbb{N}}$ in (iii) consists of very nice functions that are in $\mathcal{B}(G)$ or $\mathcal{B}(\mathbb{R} \times G)$, for an arbitrary function f in $\mathcal{T}_L^k(G)$ or $\mathcal{T}_L^{k,p}(I \times G)$ it is not clear in general if $f * \phi_n$ converges to f in the $\mathcal{T}$ -space. On the other hand, using a trick mentioned in [6, 7] regarding the use of certain right-invariant vector fields, one can make some statements for convolutions on the right. We list a few relevant results here. As above, the results are true for spaces of each finite order and on G or $I \times G$. For simplicity we only write the superscript ∞ and do not specify the underlying space.

- (iv) For a bi-invariant Laplacian (denoted by Δ), conclusions of (ii) and (iii) hold for convolutions on the right when $\mathfrak{S} = \mathcal{S}_\Delta^\infty$ or $\mathcal{T}_\Delta^\infty$.
- (v) For a left-invariant Laplacian L that is form-comparable to a bi-invariant Laplacian Δ , by Lemma 6.1 in the Appendix, $\mathcal{S}_L^\infty = \mathcal{S}_\Delta^\infty$. Hence (ii) and (iii) hold for convolutions on the right when $\mathfrak{S} = \mathcal{S}_L^\infty$.

Finally, we record a result for $\mathcal{C}$ -type spaces. For any left-invariant vector field Z , let $\check{Z}$ be the right-invariant vector field on G such that $\check{Z}(e) = Z(e)$. For any projective basis $Y = \{Y_i\}_{i \in \mathcal{I}}$ of left-invariant vector fields, in [7] the authors considered a right-invariant version of the $\mathcal{C}_Y^k$ space ($k \in \mathbb{N}$), denoted by $\mathcal{RC}_Y^k$, which is analogously defined as the completion of the Bruhat space with respect to the seminorms $\|\check{Y}^l f\|_\infty$. A special case is when Y is a so-called special projective basis, which roughly speaking requires that each Y_i is a finite linear combination of $\{\check{Y}_j\}_{j \in \mathcal{I}}$: $Y_i = \sum_{j \in J(i)} a_{ij} \check{Y}_j$, where each a_{ij} is in $\mathcal{B}(G)$ and $J(i)$ is a finite index set, and vice versa for each $\check{Y}_i$. When Y is a special projective basis, $\mathcal{C}_Y^k = \mathcal{RC}_Y^k$. For a bi-invariant Laplacian, Δ , there exists a special projective basis $X = \{X_i\}_{i \in \mathcal{I}}$ such that $\Delta = -\sum X_i^2$. See [7, Section 4.2]. Hence the next item follows.

- (vi) For a bi-invariant Laplacian, Δ , and any special projective basis, $X = \{X_i\}_{i \in \mathcal{I}}$, such that $\Delta = -\sum X_i^2$, (ii) and (iii) hold for convolutions on the right when $\mathfrak{S} = \mathcal{C}_X^\infty = \mathcal{RC}_X^\infty$.

3. Statement of the main results. As preparations for the main results, we first review some definitions mentioned in the Introduction.

DEFINITION 3.1 (Property (CK*)). Let $(\mu_t)_{t>0}$ be a symmetric Gaussian convolution semigroup. We say that $(\mu_t)_{t>0}$ satisfies *Property (CK*)* if for any $t > 0$, μ_t admits a continuous density $\mu_t(\cdot)$ with respect to the Haar measure ν , and the density function satisfies

$$(3.1) \quad \lim_{t \rightarrow 0^+} t \log \mu_t(e) = 0.$$

REMARK 3.2. Let Δ be a bi-invariant Laplacian satisfying Property (CK*). Let L be a left-invariant Laplacian that is form-comparable to Δ . Then there are some $\beta, \gamma > 0$ such that $\mathcal{E}_{\beta\Delta} \leq \mathcal{E}_L \leq \mathcal{E}_{\gamma\Delta}$. Denote the two operators' associated (symmetric) semigroups as μ_t^Δ, μ_t^L , respectively. Then $\mu_t^L = \mu_t^{L-\beta\Delta} * \mu_t^{\beta\Delta}$, indicating that μ_t^L admits a continuous density because $\mu_t^{\beta\Delta}$ does. Moreover, because μ_t^L is symmetric and $\|\mu_t^L\|_{L^\infty(G)} = \mu_t^L(e)$,

$$\mu_t^L(e) = \mu_t^{L-\beta\Delta} * \mu_t^{\beta\Delta}(e) \leq \mu_t^{\beta\Delta}(e).$$

Hence μ_t^L also satisfies (CK*).

DEFINITION 3.3 (Hypoellipticity). Let L be a left-invariant sub-Laplacian on G and $I \subset \mathbb{R}$ be an open interval. Let $\mathfrak{A}$ be a space of distributions in time and space. Let $\mathfrak{S}$ be a space of continuous functions on $I \times G$. The associated parabolic operator $\partial_t + L$ is said to be $\mathfrak{A}$ - $\mathfrak{S}$ -hypoelliptic if for any $U \in \mathfrak{A}$ and $F \in \mathcal{B}'(I \times G)$ such that

$$(\partial_t + L)U = F \quad \text{in } \mathcal{B}'(I \times G),$$

and for any open subset $J \times \Omega \subset I \times G$ such that

$$\forall \varphi \in \mathcal{B}_c(J \times \Omega), \quad \varphi F \in \mathfrak{S},$$

U satisfies

$$\forall \varphi \in \mathcal{B}_c(J \times \Omega), \quad \varphi U \in \mathfrak{S}.$$

Here $\mathcal{B}_c(J \times \Omega)$ consists of functions in $\mathcal{B}(I \times G)$ with compact support in $J \times \Omega$. For example, $\mathfrak{A}$ can be $\mathcal{T}'_{\Delta,L}(I \times G)$, and $\mathfrak{S}$ can be $C(I \times G)$.

REMARK 3.4. Observe that for any $N, p \in \mathbb{N} \cup \{\infty\}$, functions in $\mathcal{S}_L^{N,p}(I \times G)$ have compact supports in $I \times G$, whereas functions in $C^p(I \rightarrow \mathcal{S}_L^N(G))$ need not. Nevertheless, by definition, being $\mathcal{T}'_L(I \times G)$ - $\mathcal{S}_L^{N,p}(I \times G)$ -hypoelliptic is equivalent to being $\mathcal{T}'_L(I \times G)$ - $C^p(I \rightarrow \mathcal{S}_L^N(G))$ -hypoelliptic. The same is true for $\mathcal{C}$ - and $\mathcal{T}$ -type spaces.

The main results of this paper regarding hypoellipticity are summarized in the following theorem.

THEOREM 3.5. Let (G, ν) be a compact connected metrizable group with normalized Haar measure ν , let $I \subset \mathbb{R}$ be an open interval. Let $(\mu_t^\Delta)_{t>0}$ be a symmetric central Gaussian convolution semigroup on G with generator $-\Delta$ and write $\Delta = -\sum X_i^2$ for some special projective basis $X = \{X_i\}_{i \in \mathcal{I}}$. Suppose $(\mu_t^\Delta)_{t>0}$ satisfies Property (CK*). Let L be any form-comparable perturbation of Δ in the sense of (2.3). Then the parabolic operator $\partial_t + L$ is $\mathcal{T}'_{\Delta,L}$ - $\mathfrak{S}$ -hypoelliptic. Here $\mathfrak{S}$ can be $\mathcal{T}_\Delta^{N,p}$, $\mathcal{S}_L^{N,p} = \mathcal{S}_\Delta^{N,p}$, $\mathcal{C}_X^{N,p}$, where $N, p \in \mathbb{N} \cup \{\infty\}$; all spaces are on $I \times G$.

REMARK 3.6. Together with Remark 3.4, Theorem 3.5 implies that L is $\mathcal{T}'_{\Delta,L}(I \times G)$ - $C(I \times G)$ -hypoelliptic, and more generally, L is $\mathcal{T}'_{\Delta,L}(I \times G)$ - $C^p(I \rightarrow \mathcal{C}_X^N(G))$ -hypoelliptic for $N, p \in \mathbb{N} \cup \{\infty\}$.

REMARK 3.7. Note that in Theorem 3.5, the only $\mathcal{T}$ -type space we consider in hypoellipticity is the $\mathcal{T}_\Delta$ space; our proof does not cover the $\mathcal{T}_L$ and $\mathcal{T}_{\Delta,L}$ cases.

Recall that given $\Delta = -\sum X_i^2$, all other sub-Laplacians on G are of the form $L_A = -\sum a_{ij} X_i X_j$ for some matrix $A = (a_{ij})_{i,j \in \mathcal{I}}$. The next example provides a simple well-known condition on the matrix A such that L_A is a form-comparable perturbation of Δ in the sense of (2.3).

EXAMPLE 3.8. For any number $0 < \epsilon < 1$, we say that a matrix $A = (a_{ij})_{i,j \in \mathcal{I}}$ is ϵ -diagonally dominant if for any $i \in \mathcal{I}$,

$$\epsilon |a_{ii}| > \sum_{j \neq i} |a_{ij}|.$$

Let (G, ν) be a compact connected group and $\Delta = -\sum_{i \in \mathcal{I}} X_i^2$ be a bi-invariant Laplacian on G as before. Let $L_A := -\sum_{i,j \in \mathcal{I}} a_{ij} X_i X_j$ be a Laplacian on G where $A = (a_{ij})_{i,j \in \mathcal{I}}$ is an ϵ -diagonally dominant matrix for some $0 < \epsilon < 1$. Then L_A is a form-comparable perturbation of Δ . Theorem 3.5 applies if $(\mu_t^\Delta)_{t>0}$ satisfies Property (CK*).

4. Proof of Theorem 3.5: properties of the perturbed heat kernel. In this section we prove that certain regularity properties of the heat kernel are preserved by form-comparable perturbations of bi-invariant Laplacians. These results are useful in the proof of Theorem 3.5 and are interesting in themselves. For simplicity we write $\mathcal{T}_{\Delta,L}$ for $\mathcal{T}_{\Delta,L}^\infty$.

THEOREM 4.1. *Let G be a compact connected metrizable group. Let $(\mu_t^\Delta)_{t>0}$ be a symmetric central Gaussian semigroup on G with generator $-\Delta$. Assume that $(\mu_t^\Delta)_{t>0}$ admits a continuous density function, denoted again by μ_t^Δ . Then for any left-invariant Laplacian L that is form-comparable to Δ , the corresponding semigroup $(\mu_t^L)_{t>0}$ admits a continuous density function (denoted again by μ_t^L), and the density functions μ_t^Δ and μ_t^L both belong to $C^\infty((0, \infty) \rightarrow \mathcal{T}_{\Delta,L}(G))$.*

Note that by Theorem 4.1, for any bump function $b(t)$, say in $C_c^\infty((0, 1))$, the product functions $b(t)\mu_t^\Delta$ and $b(t)\mu_t^L$ belong to $\mathcal{T}_{\Delta,L}((0, 1) \times G)$.

Proof of Theorem 4.1. We first show that $\mu_t^\Delta \in C^\infty((0, \infty) \rightarrow \mathcal{T}_{\Delta,L}(G))$. For convenience, we sometimes write the density function $\mu_t^\Delta(x)$ as $\mu^\Delta(t, x)$. Let $\{X_i\}_{i \in \mathcal{I}}$, $\{Y_i\}_{i \in \mathcal{I}}$ be two projective bases such that $\Delta = -\sum X_i^2$, $L = -\sum Y_i^2$. By [4], $\mu_t^\Delta \in C^\infty((0, \infty) \rightarrow \mathcal{T}_\Delta(G))$. Because the seminorms of $\mathcal{T}_{\Delta,L}$ can be given using either $\{X_i\}_{i \in \mathcal{I}}$ or $\{Y_i\}_{i \in \mathcal{I}}$ (see Remark 2.4), for simplicity

of notations we use $\{Y_i\}_{i \in \mathcal{I}}$ here, and because $\partial_t \mu_t^\Delta = -\Delta \mu_t^\Delta$ in $C(G)$, to show that $\mu_t^\Delta \in C^\infty((0, \infty) \rightarrow \mathcal{T}_{\Delta, L}(G))$, it suffices to show that for any $N, p \in \mathbb{N}$ and any $t > 0$,

$$M_L^N(\partial_t^p \mu_t^\Delta) = \sup_{\substack{k, m \in \mathbb{N} \\ k+2m \leq N}} \sup_{x \in G} \sup_{\lambda \in \Lambda(k, m)} \left(\sum_{l \in \mathcal{I}^k} |P_L^{l, \lambda} \partial_t^p \mu_t^\Delta(x)|^2 \right)^{1/2} < \infty.$$

We now compute $P_L^{l, \lambda} \mu_t^\Delta(x) = L^{\lambda_0} Y_{l_1} L^{\lambda_1} \dots Y_{l_k} L^{\lambda_k} \mu_t^\Delta(x)$. First, let $\epsilon := \frac{1}{2k+1}$. For any fixed α with $0 < \alpha < \epsilon C^{-1} = \frac{1}{(2k+1)C}$ (C as in $\mathcal{E}_L \leq C\mathcal{E}_\Delta$), we have $\mathcal{E}_{\epsilon\Delta - \alpha L} \geq \delta \mathcal{E}_\Delta$ for some small $\delta > 0$. Thus as Δ commutes with L , μ_t^Δ can be decomposed as

$$\mu_t^\Delta = \mu_t^{(1-\epsilon)\Delta} * \mu_t^{\epsilon\Delta - \alpha L} * \mu_t^{\alpha L}.$$

Because μ_t^Δ commutes with any function in convolution,

$$\begin{aligned} Y_{l_k} L^{\lambda_k} \mu_t^\Delta &= Y_{l_k} (\mu_t^{(1-\epsilon)\Delta} * \mu_t^{\epsilon\Delta - \alpha L} * L^{\lambda_k} \mu_t^{\alpha L}) \\ &= Y_{l_k} (\mu_t^{(1-2\epsilon)\Delta} * \mu_t^{\epsilon\Delta - \alpha L} * L^{\lambda_k} \mu_t^{\alpha L} * \mu_t^{\epsilon\Delta}) \\ &= \mu_t^{(1-2\epsilon)\Delta} * \mu_t^{\epsilon\Delta - \alpha L} * L^{\lambda_k} \mu_t^{\alpha L} * Y_{l_k} \mu_t^{\epsilon\Delta}. \end{aligned}$$

So

$$\begin{aligned} P_L^{l, \lambda} \mu_t^\Delta(x) &= L^{\lambda_0} Y_{l_1} L^{\lambda_1} \dots Y_{l_k} L^{\lambda_k} \mu_t^\Delta(x) \\ &= L^{\lambda_0} Y_{l_1} L^{\lambda_1} \dots Y_{l_{k-1}} L^{\lambda_{k-1}} (\mu_t^{(1-2\epsilon)\Delta} * \mu_t^{\epsilon\Delta - \alpha L} * L^{\lambda_k} \mu_t^{\alpha L} * Y_{l_k} \mu_t^{\epsilon\Delta})(x) \\ &= \mu_t^{\epsilon\Delta - \alpha L} * L^{\lambda_k} \mu_t^{\alpha L} * Y_{l_k} \mu_t^{\epsilon\Delta} * (L^{\lambda_0} Y_{l_1} L^{\lambda_1} \dots Y_{l_{k-1}} L^{\lambda_{k-1}} \mu_t^{(1-2\epsilon)\Delta})(x). \end{aligned}$$

Repeating this decomposition process k times, we get

$$(4.1) \quad \begin{aligned} P_L^{l, \lambda} \mu_t^\Delta(x) &= (\mu_t^{\epsilon\Delta - \alpha L} * L^{\lambda_k} \mu_t^{\alpha L} * Y_{l_k} \mu_t^{\epsilon\Delta}) * \dots \\ &\quad * (\mu_t^{\epsilon\Delta - \alpha L} * L^{\lambda_1} \mu_t^{\alpha L} * Y_{l_1} \mu_t^{\epsilon\Delta}) * \mu_t^{\epsilon\Delta - \alpha L} * L^{\lambda_0} \mu_t^{\alpha L}(x). \end{aligned}$$

To estimate

$$\left(\sum_{l \in \mathcal{I}^k} |P_L^{l, \lambda} \mu_t^\Delta(x)|^2 \right)^{1/2} = \|(P_L^{l, \lambda} \mu_t^\Delta(x))\|_{l^2},$$

i.e., the l^2 norm of the $\mathcal{I}^k$ -indexed vector $(P_L^{l, \lambda} \mu_t^\Delta(x))$, we use the following bound: for any l^2 vector functions $\mathbf{u}(x) = (u_i(x))_i$, $\mathbf{v}(x) = (v_j(x))_j$, where $u_i, v_j \in C(G)$ for each i, j ,

$$(4.2) \quad \|\mathbf{u} * \mathbf{v}(x)\|_{l^2} \leq (\|\mathbf{u}\|_{l^2} * \|\mathbf{v}\|_{l^2})(x).$$

To see this, we apply the Minkowski inequality twice:

$$\begin{aligned} \|\mathbf{u} * \mathbf{v}(x)\|_{l^2}^2 &= \sum_i \left(\sum_j |u_i * v_j(x)|^2 \right) \\ &= \sum_i \|u_i * \mathbf{v}(x)\|_{l^2}^2 \leq \sum_i (|u_i| * \|\mathbf{v}\|_{l^2}(x))^2 \leq (\|\mathbf{u}\|_{l^2} * \|\mathbf{v}\|_{l^2}(x))^2. \end{aligned}$$

Repeatedly applying (4.2), we get

$$(4.3) \quad \begin{aligned} & \| (P_L^{\lambda} \mu_t^{\Delta}(x)) \|_{l^2} \\ & \leq \mu_t^{\epsilon\Delta - \alpha L} * |L^{\lambda_k} \mu_t^{\alpha L}| * \|(Y_{l_k} \mu_t^{\epsilon\Delta})_{l_k}\|_{l^2} * \dots * \mu_t^{\epsilon\Delta - \alpha L} * |L^{\lambda_1} \mu_t^{\alpha L}| \\ & \quad * \|(Y_{l_1} \mu_t^{\epsilon\Delta})_{l_1}\|_{l^2} * \mu_t^{\epsilon\Delta - \alpha L} * |L^{\lambda_0} \mu_t^{\alpha L}|(x), \end{aligned}$$

where each component is a continuous function on $(0, \infty) \times G$. The same is true if we replace $\mu_t^{\epsilon\Delta - \alpha L}$, $\mu_t^{\alpha L}$, $\mu_t^{\epsilon\Delta}$ in (4.3) by any of their time derivatives. Hence $M_L^N(\partial_t^p \mu_t^{\Delta}) = M_L^N(G, \partial_t^p \mu_t^{\Delta}) < \infty$ for any $N, p \in \mathbb{N}$ and any $t > 0$, and μ_t^{Δ} belongs to $C^\infty((0, \infty) \rightarrow \mathcal{T}_{\Delta, L}(G))$.

Finally, by decomposing the semigroup $(\mu_t^L)_{t>0}$ into convolutions similarly to the above we find that $(\mu_t^L)_{t>0}$ admits a continuous density function which further belongs to $C^\infty((0, \infty) \rightarrow \mathcal{T}_{\Delta, L}(G))$. For example, for any $k, n \in \mathbb{N}$, similar to (4.1), we can fix some $0 < \beta < c/2$ (c as in $c\mathcal{E}_\Delta \leq \mathcal{E}_L$) and decompose as follows:

$$(4.4) \quad \begin{aligned} & L^{\lambda_0} Y_{l_1} L^{\lambda_1} \dots Y_{l_k} L^{\lambda_k} \Delta^n \mu_t^L \\ & = L^{\lambda_0} Y_{l_1} L^{\lambda_1} \dots Y_{l_k} (\Delta^n \mu_t^{\beta\Delta} * \mu_t^{\frac{1}{2}L - \beta\Delta} * L^{\lambda_k} \mu_t^{\frac{1}{2}L}) \\ & = \Delta^n \mu_t^{\frac{1}{2}\beta\Delta} * \mu_t^{\frac{1}{2}L - \beta\Delta} * L^{\lambda_k} \mu_t^{\frac{1}{2}L} * L^{\lambda_0} Y_{l_1} L^{\lambda_1} \dots Y_{l_{k-1}} L^{\lambda_{k-1}} \mu_t^{\frac{1}{2}\beta\Delta}. \end{aligned}$$

The last term is then in the form of (4.1). We can thus bound the terms $(\sum_{l \in \mathcal{I}^k} |P_L^{l, \lambda} \Delta^n \partial_t^p \mu_t^L(x)|^2)^{1/2}$ as was done for μ_t^{Δ} . ■

Next we show that μ_t^L satisfies similar off-diagonal Gaussian estimates to those for μ_t^{Δ} ; the latter are obtained in [4].

PROPOSITION 4.2. *Let G be a compact connected metrizable group. Let $(\mu_t^{\Delta})_{t>0}$ be a symmetric central Gaussian semigroup on G with generator $-\Delta$. Suppose $(\mu_t^{\Delta})_{t>0}$ satisfies Property (CK*). Let L be any form-comparable perturbation of Δ . Then for any compact set $K \subset G$ with $e \notin K$, and any $T > 0$, $N \in \mathbb{N}$, $\sigma \geq 0$, $A, \alpha > 0$,*

$$(4.5) \quad \sup_{0 < t < T} \frac{e^{AM_L(\alpha t)}}{t^\sigma} M_{\Delta, L}^N(K, \mu_t^L) < +\infty.$$

Here $M_L(t) := \log \mu_t^L(e)$.

Proof. As in (4.4), (4.3) of the previous theorem, the terms

$$\left(\sum_{l \in \mathcal{I}^k} |P_L^{l, \lambda} \Delta^n \mu_t^L(x)|^2 \right)^{1/2} = \|(P_L^{\lambda} \Delta^n \mu_t^L(x))\|_{l^2}$$

are bounded above by convolutions of terms of the form $|\Delta^n \mu_t^{\Delta}|$, $|D_x^1 \Delta^n \mu_t^{\Delta}|_L$ (equivalent to $|D_x^1 \Delta^n \mu_t^{\Delta}|_\Delta$), $|L^{\lambda}| \mu_t^L(x)$, and $\mu_t^{\theta L - \epsilon \Delta}(x)$, where $\theta, \epsilon \in \mathbb{R}$ are such that $\mathcal{E}_{\theta L - \epsilon \Delta}$ is comparable with $\mathcal{E}_\Delta$. The terms μ_t^{Δ} , μ_t^L , $\mu_t^{\theta L - \epsilon \Delta}$ satisfy the following estimates.

- (i) By [4, Section 4.2], for any compact set $K \subset G$ with $e \notin K$, and any $T > 0$, $N \in \mathbb{N}$, $\sigma \geq 0$, $A, \alpha > 0$,

$$(4.6) \quad \sup_{0 < t < T} \frac{e^{AM_\Delta(\alpha t)}}{t^\sigma} M_\Delta^N(K, \mu_t^\Delta) < +\infty,$$

where $M_\Delta(t) := \log \mu_t^\Delta(e)$.

- (ii) For the term μ_t^L , by Remark 3.2, μ_t^L satisfies Property (CK*) and $\mu_t^L(e) \leq \mu_t^{\beta\Delta}(e)$ for some $\beta > 0$. Then

$$M_L(t) = \log \mu_t^L(e) \leq \log \mu_t^{\beta\Delta}(e) = o(1/t).$$

By [2, Section 3] (see also [4, Theorem 4.8]), for any $\psi \in \mathcal{B}(G)$ with $|D_x^1 \psi|_L \leq 1$, $|L\psi| \leq 1$,

$$\mu_t^L(x) \leq \exp \left\{ M_L(t) - \frac{c_L(\psi(x) - \psi(e))^2}{t} \right\}$$

for some $c_L > 0$. It follows that for any compact set K with $e \notin K$,

$$\lim_{t \rightarrow 0} \sup_{x \in K} \mu_t^L(x) = 0.$$

Further, by [13, Theorem 4], for any $a \in \mathbb{N}$,

$$(4.7) \quad \lim_{t \rightarrow 0} \sup_{x \in K} |\partial_t^a \mu_t^L(x)| = \lim_{t \rightarrow 0} \sup_{x \in K} |L^a \mu_t^L(x)| = 0.$$

- (iii) Applying the same arguments to $\mu_t^{\theta L - \epsilon \Delta}$ as to μ_t^L (because $\mathcal{E}_{\theta L - \epsilon \Delta}$ is comparable to $\mathcal{E}_\Delta$) gives

$$(4.8) \quad \lim_{t \rightarrow 0} \sup_{x \in K} |\mu_t^{\theta L - \epsilon \Delta}(x)| = 0.$$

As discussed above, each $t^{-\sigma} e^{AM_L(\alpha t)} (\sum_{l \in \mathcal{I}^k} |P_L^{l, \lambda} \Delta^n \mu_t^L(x)|^2)^{1/2}$ is bounded above by a convolution of terms all of which satisfy the off-diagonal bound as described in Lemma 4.3 below (see (4.6)–(4.8)). The desired bound (4.5) thus follows from Lemma 4.3. ■

LEMMA 4.3. *If two families of functions $(u_t)_{t>0}$ and $(v_t)_{t>0}$ on G both satisfy the following off-diagonal bound: for any compact set K with $e \notin K$,*

$$\lim_{t \rightarrow 0} \sup_{x \in K} u_t(x) = \lim_{t \rightarrow 0} \sup_{x \in K} v_t(x) = 0,$$

*then $u_t * v_t$ also satisfies this off-diagonal bound.*

Proof. Let V_1 be an open neighborhood of e with $\overline{V_1} \cap K = \emptyset$. For any $x \in K$, let U_x be a small open neighborhood of x that is away from V_1 . Let $x(U_x^c)^{-1} = \{xy^{-1} \mid y \in U_x^c\}$; then $x(U_x^c)^{-1}$ is away from e , i.e. $x(U_x^c)^{-1} \subset (U_e)^c$ for some open neighborhood U_e of e . Let $V_x \Subset U_x$ be a smaller open neighborhood of x . We may pick U_e such that for any $\tilde{x} \in V_x$, $\tilde{x}(U_x^c)^{-1} \subset (U_e)^c$.

For any such $\tilde{x}$ (i.e., $\tilde{x} \in V_x$),

$$\begin{aligned}
 (4.9) \quad u_t * v_t(\tilde{x}) &= \int_G u_t(\tilde{x}y^{-1})v_t(y) d\nu(y) \\
 &= \int_{y \in (U_x)^c} u_t(\tilde{x}y^{-1})v_t(y) d\nu(y) + \int_{y \in U_x} u_t(\tilde{x}y^{-1})v_t(y) d\nu(y) \\
 &\leq \sup_{a \in (U_e)^c} |u_t(a)| \cdot \|v_t\|_\infty \cdot \nu(G) + \|u_t\|_\infty \cdot \sup_{y \in U_x} |v_t(y)| \cdot \nu(G).
 \end{aligned}$$

Because K is compact, using some finite collection $\{V_{x_n}\}_{1 \leq n \leq N_0}$ to cover K , with each $x_n \in K$ and V_{x_n} chosen in the above way, the estimate (4.9) then implies that

$$\lim_{t \rightarrow 0} \sup_{x \in K} |u_t * v_t(x)| = 0. \blacksquare$$

REMARK 4.4. As a consequence, for any $0 < a < b < \infty$, any $A, \alpha, T > 0$, $\sigma \geq 0$, $N, p \in \mathbb{N}$, and any compact set K with $e \notin K$,

$$\sup_{0 < \tau < T} \frac{\exp\{AM_L(\alpha\tau)\}}{\tau^\sigma} M_{\Delta, L}^{N, p}((a\tau, b\tau) \times K, \mu^L) < \infty.$$

5. Proof of Theorem 3.5: hypoellipticity. We now use results in the previous section to prove hypoellipticity properties of the parabolic operator $\partial_t + L$. By definition of $\mathcal{T}'_{\Delta, L}$ - $\mathfrak{S}$ -hypoellipticity, for any $U \in \mathcal{T}'_{\Delta, L}(I \times G)$ and $F \in \mathcal{B}'(I \times G)$ such that $(\partial_t + L)U = F$, for any open subset $I' \times \Omega' \subset I \times G$ such that $\psi F \in \mathfrak{S}$ for any $\psi \in \mathcal{B}_c(I' \times \Omega')$, we need to show that $\psi U \in \mathfrak{S}$ for any $\psi \in \mathcal{B}_c(I' \times \Omega')$.

Recall that in Theorem 3.5, $\mathfrak{S}$ can be $\mathcal{T}_\Delta^{N, p}$, $\mathcal{S}_L^{N, p}$, $\mathcal{C}_X^{N, p}$, where all spaces are on $I \times G$, and $X = \{X_i\}_{i \in \mathcal{I}}$ is any special projective basis such that $\Delta = -\sum X_i^2$. We prove the theorem for the case

$$\mathfrak{S} = \mathcal{T}_\Delta^\infty(I \times G) =: \mathcal{T}_\Delta(I \times G);$$

the proofs for the other cases are very similar.

Fix any $\psi \in \mathcal{B}_c(I' \times \Omega')$. To show that $\psi U \in \mathcal{T}_\Delta(I \times G)$, we construct an approximation sequence in $\mathcal{T}_\Delta(I \times G)$. Let $I_0 \times \Omega_0, I_1 \times \Omega_1, I_2 \times \Omega_2$ be open sets such that

$$\text{supp } \psi \subset I_0 \times \Omega_0 \Subset I_1 \times \Omega_1 \Subset I_2 \times \Omega_2 \Subset I' \times \Omega'.$$

Pick some $\eta \in \mathcal{B}(I \times G)$ with $\eta \equiv 1$ on $I_1 \times \Omega_1$ and $\text{supp } \eta \subset I_2 \times \Omega_2$. Fix some bump function $\rho \in C_c^\infty((1, 2))$ that satisfies $\rho \geq 0$ and $\int_{\mathbb{R}} \rho(t) dt = 1$. For any $\tau > 0$, let $\rho_\tau(t) := (1/\tau)\rho(t/\tau)$. Then ρ_τ is supported in $(\tau, 2\tau)$. Let c_0 be a positive number to be determined in Lemma 5.1. For any $(\alpha, \tau) \in$

$[0, 1] \times [0, c_0] \setminus \{(0, 0)\}$, define

$$\tilde{U}_{\alpha, \tau} := \begin{cases} (\rho_\alpha \mu^\Delta) \star (\eta U) \star (\rho_\tau \mu^L) & \text{when } (\alpha, \tau) \in (0, 1] \times (0, c_0], \\ \tilde{U}_{0, \tau} = (\eta U) \star (\rho_\tau \mu^L) & \text{when } \tau > 0, \alpha = 0, \\ \tilde{U}_{\alpha, 0} = (\rho_\alpha \mu^\Delta) \star (\eta U) & \text{when } \alpha > 0, \tau = 0. \end{cases}$$

Recall that we use $\star$ to emphasize that the convolution is in time and space (see (2.2)). The two-parameter sequence $\{\psi \tilde{U}_{\alpha, \tau}\}_{\alpha, \tau}$ is our approximation sequence. The following lemma shows that $\psi \tilde{U}_{\alpha, \tau} \in \mathcal{T}_\Delta(I \times G)$ for all $(\alpha, \tau) \in [0, 1] \times [0, c_0] \setminus \{(0, 0)\}$.

LEMMA 5.1. *For any $W \in \mathcal{T}'_{\Delta, L}(I \times G)$ with compact support in $I \times G$, there exists some $0 < c_0 < 1$ such that for any fixed $0 < \tau \leq c_0$,*

$$W_\tau := W \star (\rho_\tau \mu^L) \in \mathcal{T}_\Delta(I \times G).$$

Proof. First note that $\rho_\tau \mu^L \in \mathcal{T}_\Delta((0, 2) \times G)$. The convolution W_τ can be interpreted as a continuous function

$$\begin{aligned} W_\tau(s, x) &= W((t, y) \mapsto \rho_\tau(s - t) \mu_{s-t}^L(x^{-1}y)) \\ &= W((t, y) \mapsto \rho_\tau(s - t) \mu_{s-t}^L(y^{-1}x)). \end{aligned}$$

So the function (shift of the time by $\tau/2$ for μ^L)

$$w_\tau : \mathbb{R} \times G \rightarrow \mathbb{R}, \quad (s, x) \mapsto W((t, y) \mapsto \rho_\tau(s - t) \mu_{s-t-\tau/2}^L(y^{-1}x)),$$

is continuous. In fact, $w_\tau \in C_c^\infty(\mathbb{R} \rightarrow C(G))$. As $\mu_{s-t}^L = \mu_{s-t-\tau/2}^L * \mu_{\tau/2}^L$,

$$\begin{aligned} W_\tau(s, x) &= W((t, y) \mapsto \rho_\tau(s - t) \int_G \mu_{s-t-\tau/2}^L(y^{-1}xz^{-1}) \mu_{\tau/2}^L(z) d\nu(z)) \\ &= \int_G W((t, y) \mapsto \rho_\tau(s - t) \mu_{s-t-\tau/2}^L(y^{-1}xz^{-1})) \mu_{\tau/2}^L(z) d\nu(z) \\ &= (w_\tau(s, \cdot) * \mu_{\tau/2}^L)(x). \end{aligned}$$

Because $\mu_{\tau/2}^L \in \mathcal{T}_\Delta(G)$, the convolution W_τ is in $\mathcal{T}_\Delta(\mathbb{R} \times G)$. When $\tau > 0$ is small enough, $W_\tau \in \mathcal{T}_\Delta(I \times G)$. ■

REMARK 5.2. The same proof shows that $W \star (\rho_\alpha \mu^\Delta) = (\rho_\alpha \mu^\Delta) \star W \in \mathcal{T}_\Delta(\mathbb{R} \times G)$. The above method further shows that $W \star (\rho_\alpha \mu^L)$ and $W \star (\rho_\alpha \mu^\Delta)$ belong to $\mathcal{T}_{\Delta, L}(\mathbb{R} \times G)$, but this fact is not needed in the proof below.

REMARK 5.3. It is not clear if $W \star \rho_\tau \mu^L$ converges to W in $\mathcal{T}'_{\Delta, L}(I \times G)$ (which is convergence in a very weak sense), because for any $\phi \in \mathcal{T}_{\Delta, L}(I \times G)$,

$$(W \star \rho_\tau \mu^L)(\phi) = W(\phi \star (\check{\rho}_\tau \check{\mu}^L)),$$

and it is not clear if $\phi \star (\check{\rho}_\tau \check{\mu}^L)$ converges to ϕ in $\mathcal{T}_{\Delta, L}(I \times G)$. See Section 2.6. For this reason, we use the two-parameter approximation sequence $\{\psi \tilde{U}_{\alpha, \tau}\}$ to approximate $\psi \eta U = \psi U$.

We now take the following steps to prove the convergence of the approximation sequence $\{\psi\tilde{U}_{\alpha,\tau}\}_{\alpha,\tau}$.

STEP 1. For any fixed $\tau \in (0, c_0]$, think of $\psi\tilde{U}_{\alpha,\tau}$ as a $\mathcal{T}_\Delta(I \times G)$ -valued function in α . Then

$$\psi\tilde{U}_{\alpha,\tau} \in C([0, 1] \rightarrow \mathcal{T}_\Delta(I \times G)).$$

To prove this, it suffices to check that $\tilde{U}_{\alpha,\tau}$ is continuous at $\alpha = 0$. This is true because $\rho_\alpha \mu^\Delta \star ((\eta U) \star \rho_\tau \mu^L) \rightarrow (\eta U) \star \rho_\tau \mu^L$ as $\alpha \rightarrow 0$ in $\mathcal{T}_\Delta(I \times G)$.

STEP 2. As $\tau \rightarrow 0$, the sequence $\{\psi\tilde{U}_{\alpha,\tau}\}_{\tau>0}$ converges uniformly to some function G_α in $C([0, 1] \rightarrow \mathcal{T}_\Delta(I \times G))$, because for any $N, p \in \mathbb{N}$,

$$\sup_{0 \leq \alpha \leq 1} \sup_{0 < \tau \leq c_0} M_\Delta^{N,p}(I_0 \times \Omega_0, \partial_\tau \tilde{U}_{\alpha,\tau}) \leq \sup_{0 < \tau \leq c_0} M_\Delta^{N,p}(I_0 \times \Omega_0, \partial_\tau \tilde{U}_{0,\tau}) < \infty.$$

We prove this fact in Proposition 5.4 below.

STEP 3. For $0 < \alpha \leq 1$, $G_\alpha = \psi\tilde{U}_{\alpha,0}$, because $(\rho_\alpha \mu^\Delta \star (\eta U)) \star \rho_\tau \mu^L \rightarrow \rho_\alpha \mu^\Delta \star (\eta U)$ in $C(I \times G)$ as $\tau \rightarrow 0$.

STEP 4. Because μ_t^Δ commutes with any function in convolution, we have $\tilde{U}_{\alpha,0} \rightarrow \eta U$ in $\mathcal{T}'_{\Delta,L}(I \times G)$ as $\alpha \rightarrow 0$. Hence $G_0 = \psi\eta U = \psi U$, and $\psi U \in \mathcal{T}_\Delta(I \times G)$.

To complete the proof of Theorem 3.5, it remains to verify Step 2, which we address in the next proposition.

PROPOSITION 5.4. *Under the hypotheses of Theorem 3.5, using the notations introduced at the beginning of this section,*

$$\sup_{0 < \tau \leq c_0} M_\Delta^{N,p}(I_0 \times \Omega_0, \partial_\tau \tilde{U}_{0,\tau}) < \infty.$$

Proof. For short we write $\tilde{U}_\tau$ for $\tilde{U}_{0,\tau}$. By computation, $\partial_\tau \rho_\tau(s) = -\partial_s \bar{\rho}_\tau(s)$ where $\bar{\rho}_\tau(s) := (s/\tau^2)\rho(s/\tau)$, and

$$\begin{aligned} \partial_\tau(\tilde{U}_\tau(s, x)) &= U((t, y) \mapsto -\eta(t, y)\partial_s \bar{\rho}_\tau(s-t)\mu_{s-t}^L(x^{-1}y)) \\ &= U(-\eta\partial_s(\bar{\rho}_\tau(s-\cdot)\mathcal{L}_{x^{-1}}\mu_{s-\cdot}^L)) + U(\eta\bar{\rho}_\tau(s-\cdot)\partial_s \mathcal{L}_{x^{-1}}\mu_{s-\cdot}^L). \end{aligned}$$

Here $\mathcal{L}$ represents left-translation, i.e., for any $x \in G$ and any function f , $\mathcal{L}_x f(y) = f(xy)$. In the first term, rewrite the function inside U as

$$\begin{aligned} -\eta(t, y)\partial_s(\bar{\rho}_\tau(s-t)\mathcal{L}_{x^{-1}}\mu_{s-t}^L(y)) \\ = \partial_t(\eta(t, y)\bar{\rho}_\tau(s-t)\mathcal{L}_{x^{-1}}\mu_{s-t}^L(y)) - \partial_t\eta(t, y) \cdot \bar{\rho}_\tau(s-t)\mathcal{L}_{x^{-1}}\mu_{s-t}^L(y). \end{aligned}$$

For the second term, use $\partial_s \mathcal{L}_{x^{-1}}\mu_{s-t}^L(y) = -\mathcal{L}_{x^{-1}}L\mu_{s-t}^L(y) = -L\mathcal{L}_{x^{-1}}\mu_{s-t}^L(y)$.

We get

$$\begin{aligned}
 \partial_\tau(\tilde{U}_\tau(s, x)) &= -((\partial_t + L)U)(\eta\bar{\rho}_\tau(s - \cdot)\mathcal{L}_{x^{-1}}\mu_{s-\cdot}^L) \\
 &\quad - U(\partial_1\eta \cdot \bar{\rho}_\tau(s - \cdot)\mathcal{L}_{x^{-1}}\mu_{s-\cdot}^L) + (\eta LU - L(\eta U))(\bar{\rho}_\tau(s - \cdot)\mathcal{L}_{x^{-1}}\mu_{s-\cdot}^L) \\
 &:= I_\tau(s, x) + II_\tau(s, x) + III_\tau(s, x).
 \end{aligned}$$

Here $\partial_1\eta(t, y) := \partial_t\eta(t, y)$, the partial derivative of η with respect to the first variable. We estimate each part separately.

Estimate of I_τ . For

$$I_\tau(s, x) = -((\partial_t + L)U)(\eta\bar{\rho}_\tau(s - \cdot)\mathcal{L}_{x^{-1}}\mu_{s-\cdot}^L) = -(\eta F) \star (\bar{\rho}_\tau\mu^L)(s, x),$$

to use the condition that $\eta F \in \mathcal{T}_\Delta(I \times G)$ with support in $I_2 \times \Omega_2$, we use a trick mentioned in [6, 7] which works as follows. Because Δ is bi-invariant, $\Delta = -\sum X_i^2 = -\sum \check{X}_i^2$. In [6] it is shown that (see Lemma 4.1 and the proof of Lemma 5.2 there)

$$\begin{aligned}
 M_\Delta^N(f) &= \sup_{x \in G} \sup_{\substack{k, m \in \mathbb{N} \\ k+2m \leq N}} \left(\sum_{l \in \mathcal{I}^k} |X_{l_1} \cdots X_{l_k} \Delta^m f(x)|^2 \right)^{1/2} \\
 &= \sup_{x \in G} \sup_{\substack{k, m \in \mathbb{N} \\ k+2m \leq N}} \left(\sum_{l \in \mathcal{I}^k} |\check{X}_{l_1} \cdots \check{X}_{l_k} \Delta^m f(x)|^2 \right)^{1/2}.
 \end{aligned}$$

The same is true for the $M_\Delta^{N,p}$ norms. Hence for any $0 < \tau \leq c_0$, using the “right-invariant” expression of $M_\Delta^{N,p}$, we get

$$M_\Delta^{N,p}(I_\tau) \leq M_\Delta^{N,p}(\eta F) \star \bar{\rho}_\tau\mu^L \leq 2M_\Delta^{N,p}(\eta F).$$

Estimate of II_τ . To estimate $M_\Delta^{N,p}(I_0 \times \Omega_0, II_\tau)$, as in the proof of Lemma 5.1, let

$$(5.1) \quad u_\tau(s, \omega) := U((t, y) \mapsto \partial_t\eta(t, y)\bar{\rho}_\tau(s - t)\mu_{s-t-\tau/2}^L(y^{-1}\omega)).$$

Then

$$II_\tau(s, x) = -u_\tau(s, \cdot) * \mu_{\tau/2}^L(x).$$

Note that $s \in I_0$, whereas for the function $\partial_t\eta(t, y)\bar{\rho}_\tau(s - t)$ to be nonzero, we need $t \in I_1^c$ and $\tau < s - t < 2\tau$. So

$$M_\Delta^{N,p}(I_0 \times \Omega_0, II_\tau) = 0 \quad \text{for } 0 < \tau < \tau_0,$$

where $\tau_0 := \min \{d(I_0, I_1^c)/2, c_0\}$. For $\tau_0 \leq \tau \leq c_0$, as $u_\tau \in C_c^\infty(\mathbb{R} \rightarrow C(G))$, by Minkowski's inequality,

$$\begin{aligned} & M_{\Delta}^{N,p}(I_0 \times \Omega_0, II_\tau) \\ & \leq \sup_{\substack{0 \leq a \leq p \\ k+2m \leq N}} \sup_{\lambda \in \Lambda(k,m)} \sup_{(s,x) \in I \times G} \left\{ |\partial_s^a u_\tau| * \left(\sum_{l \in \mathbb{Z}^k} |P_{\Delta}^{l,\lambda} \mu_{\tau/2}^L|^2 \right)^{1/2}(x) \right\}. \end{aligned}$$

Because $U \in \mathcal{T}'_{\Delta,L}(I \times G)$, there exist some $C > 0$ and $N', p' \in \mathbb{N}$ that depend on $N, p, U, \text{supp } \eta, \rho$, such that for any $0 \leq a \leq p$,

$$\|\partial_s^a u_\tau\|_{L^\infty(I \times G)} \leq C \frac{1}{\tau^{p'}} M_{\Delta,L}^{N',p'}(\eta) M_{\Delta,L}^{N',p'}((\tau/2, 3\tau/2) \times G, \mu^L).$$

Here the term $\tau^{-p'}$ is from taking derivatives of $\bar{\rho}_\tau$. Hence

$$\begin{aligned} & \sup_{0 < \tau \leq c_0} M_{\Delta}^{N,p}(I_0 \times \Omega_0, II_\tau) = \sup_{\tau_0 \leq \tau \leq c_0} M_{\Delta}^{N,p}(I_0 \times \Omega_0, II_\tau) \\ & \leq \frac{C\nu(G)}{\tau_0^{p'}} M_{\Delta,L}^{N',p'}(\eta) M_{\Delta,L}^{N',p'}\left(\left(\frac{\tau_0}{2}, \frac{3c_0}{2}\right) \times G, \mu^L\right) \sup_{\tau_0 \leq \tau \leq c_0} M_{\Delta}^N(\mu_{\tau/2}^L) < \infty. \end{aligned}$$

Estimate of III_τ . For the last part III_τ , let $\tilde{V} := \eta LU - L(\eta U)$. Then $\tilde{V}$ is supported away from $I_1 \times \Omega_1$, and

$$III_\tau(s, x) = \tilde{V}(\Phi \bar{\rho}_\tau(s - \cdot) \mathcal{L}_{x^{-1}} \mu_{s-\cdot}^L)$$

for some hollow-shaped function $\Phi \in \mathcal{B}(I \times G)$. More precisely, for some $J_1 \times \Theta_1$ and $J_2 \times \Theta_2$ satisfying

$$I_0 \times \Omega_0 \Subset J_1 \times \Theta_1 \Subset J_2 \times \Theta_2 \Subset I_1 \times \Omega_1,$$

Φ satisfies

$$\text{supp } \Phi \subset I' \times \Omega' \setminus J_1 \times \Theta_1, \quad \Phi \equiv 1 \text{ on } I_2 \times \Omega_2 \setminus J_2 \times \Theta_2.$$

As in the decomposition of II_τ , III_τ can be written as the convolution

$$III_\tau(s, x) = ((\Phi \tilde{V}) \star \bar{\rho}_\tau \mathcal{L}_{-\tau/2} \mu^L)(s, \cdot) * \mu_{\tau/2}^L(x),$$

where

$$((\Phi \tilde{V}) \star \bar{\rho}_\tau \mathcal{L}_{-\tau/2} \mu^L)(s, w) = \tilde{V}((t, y) \mapsto \Phi(t, y) \bar{\rho}_\tau(s - t) \mu_{s-t-\tau/2}^L(y^{-1}w)).$$

So by Minkowski's inequality,

$$\begin{aligned} & M_{\Delta}^{N,p}(I_0 \times \Omega_0, III_\tau) \\ & \leq \sup_{\substack{(s,x) \in I_0 \times \Omega_0 \\ \lambda \in \Lambda(k,m) \\ a,k,m}} \left\{ |\partial_s^a ((\Phi \tilde{V}) \star \bar{\rho}_\tau \mathcal{L}_{-\tau/2} \mu^L)(s, \cdot)| * \left(\sum_{l \in \mathbb{Z}^k} |P_{\Delta}^{l,\lambda} \mu_{\tau/2}^L|^2 \right)^{1/2}(x) \right\}, \end{aligned}$$

where the supremum is over $\{(a, k, m) : 0 \leq a \leq p, k + 2m \leq N\}$. Let Θ_0 be an open set satisfying $\Omega_0 \Subset \Theta_0 \Subset \Theta_1$. We split the convolution into two

parts,

$$\begin{aligned}
& |\partial_s^a((\Phi\tilde{V}) \star \bar{\rho}_\tau \mathcal{L}_{-\tau/2} \mu^L)(s, \cdot)| * \left(\sum_{l \in \mathcal{I}^k} |P_\Delta^{l,\lambda} \mu_{\tau/2}^L|^2 \right)^{1/2}(x) \\
&= \int_{\Theta_0} |\partial_s^a((\Phi\tilde{V}) \star \bar{\rho}_\tau \mathcal{L}_{-\tau/2} \mu^L)(s, y)| \left(\sum_{l \in \mathcal{I}^k} |(P_\Delta^{l,\lambda} \mu_{\tau/2}^L)(y^{-1}x)|^2 \right)^{1/2} d\nu(y) \\
&\quad + \int_{\Theta_0^c} |\partial_s^a((\Phi\tilde{V}) \star \bar{\rho}_\tau \mathcal{L}_{-\tau/2} \mu^L)(s, y)| \left(\sum_{l \in \mathcal{I}^k} |(P_\Delta^{l,\lambda} \mu_{\tau/2}^L)(y^{-1}x)|^2 \right)^{1/2} d\nu(y).
\end{aligned}$$

The first integral is bounded above by

$$\begin{aligned}
& \nu(G) \|(\Phi\tilde{V}) \star \bar{\rho}_\tau \mathcal{L}_{-\tau/2} \mu^L\|_{C^p(I_0 \rightarrow L^\infty(\Theta_0))} M_\Delta^N(\mu_{\tau/2}^L) \\
&\leq \nu(G) \frac{C_1}{\tau^{p_1}} M_{\Delta,L}^{N_1,p_1}((\tau/2, 3\tau/2) \times (\Theta_1^c)^{-1}\Theta_0, \mu^L) M_\Delta^N(\mu_{\tau/2}^L)
\end{aligned}$$

for some $C_1 > 0$ and $N_1, p_1 \in \mathbb{N}$ that depend on N, p, ρ, Φ , and the distribution $\tilde{V}$ (in other words, on U). Here for any $0 < \tau < 1$,

$$M_\Delta^N(\mu_{\tau/2}^L) \leq C'_1 \frac{e^{AM(\alpha\tau)}}{\tau^N}$$

for some constants $C'_1, \alpha, A > 0$, $M(s) := \log \mu_s^L(e)$. See the proof of Proposition 4.2, or [4, Theorem 4.8] and [13, Theorem 3].

The second integral is bounded above by

$$\begin{aligned}
& \nu(G) \|(\Phi\tilde{V}) \star \bar{\rho}_\tau \mathcal{L}_{-\tau/2} \mu^L\|_{C^p(I_0 \rightarrow L^\infty(G))} M_\Delta^N((\Theta_0^c)^{-1}\Omega_0, \mu_{\tau/2}^L) \\
&\leq \nu(G) \frac{C_1}{\tau^{p_1}} M_{\Delta,L}^{N_1,p_1}((\tau/2, 3\tau/2) \times G, \mu^L) M_\Delta^N((\Theta_0^c)^{-1}\Omega_0, \mu_{\tau/2}^L).
\end{aligned}$$

Because $\overline{(\Theta_1^c)^{-1}\Theta_0}$ and $\overline{(\Theta_0^c)^{-1}\Omega_0}$ do not contain e , applying the off-diagonal Gaussian estimate of μ_t^L with $\tau/2 \leq t \leq 3\tau/2$ in Proposition 4.2 (see Remark 4.4) then shows that

$$\sup_{0 < \tau \leq c_0} M_\Delta^{N,p}(I_0 \times \Omega_0, III_\tau) < \infty.$$

Combining the estimates for I_τ , II_τ , III_τ , we conclude that

$$\sup_{0 < \tau \leq c_0} M_\Delta^{N,p}(I_0 \times \Omega_0, \partial_\tau \tilde{U}_\tau) < \infty. \blacksquare$$

6. Appendix. In this appendix we prove some equivalence relations between function space (semi)norms. The following lemma justifies the use of only $\{X_i\}_{i \in \mathcal{I}}$ in the definition of the $M_{\Delta,L}^N$ norms in (2.4).

LEMMA 6.1. *Let G be a compact connected metrizable group as before. Let L_1, L_2 be two form-comparable left-invariant sub-Laplacians on G . Let*

$\{X_i\}_{i \in \mathcal{I}}, \{Y_i\}_{i \in \mathcal{I}}$ be two projective bases such that

$$L_1 = - \sum_{i \in \mathcal{I}} X_i^2, \quad L_2 = - \sum_{i \in \mathcal{I}} Y_i^2.$$

Then for any $k, p \in \mathbb{N}$, the (semi)norms $S_{L_1}^{k,p}$ and $S_{L_2}^{k,p}$ are comparable.

Proof. Under the projective basis $\{X_i\}_{i \in \mathcal{I}}$, $L_2 = - \sum_{i,j \in \mathcal{I}} a_{ij} X_i X_j$ for some real symmetric nonnegative coefficient matrix $A = (a_{ij})_{i,j \in \mathcal{I}}$. Because L_1 and L_2 are form-comparable, we have $c\mathcal{E}_{L_1} \leq \mathcal{E}_{L_2} \leq C\mathcal{E}_{L_1}$ for some $c, C > 0$, which is equivalent to the condition

$$(6.1) \quad c \sum \xi_i^2 \leq \sum a_{ij} \xi_i \xi_j \leq C \sum \xi_i^2$$

for any $\xi = (\xi_i)_{i \in \mathcal{I}} \in \mathbb{R}^{(\mathcal{I})}$.

Because $\{X_i\}_{i \in \mathcal{I}}$ is a projective basis, there exists a matrix $(T_j^i)_{i,j \in \mathcal{I}}$ such that $Y_j = \sum_{i \in \mathcal{I}} T_j^i X_i$. Define a map $T : \mathbb{R}^{(\mathcal{I})} \rightarrow \mathbb{R}^{(\mathcal{I})}$ as

$$T(\xi) = T((\xi_i)_{i \in \mathcal{I}}) = (\eta_j)_{j \in \mathcal{I}}, \quad \text{where} \quad \eta_j = \sum_{i \in \mathcal{I}} T_j^i \xi_i.$$

Then $L = - \sum a_{ij} X_i X_j = - \sum Y_j^2$ implies that

$$(6.2) \quad \sum_{i,j \in \mathcal{I}} a_{ij} \xi_i \xi_j = \sum_{j \in \mathcal{I}} \left(\sum_{i \in \mathcal{I}} T_j^i \xi_i \right)^2 = \|T(\xi)\|_{l^2}^2,$$

and (6.1) is thus equivalent to

$$(6.3) \quad c \leq \|T\|_{l^2 \rightarrow l^2} \leq C.$$

To show the equivalence of the two (semi)norms $S_{L_1}^{k,p}$ and $S_{L_2}^{k,p}$ on $I \times G$ it suffices to show the equivalence of $S_{L_1}^k$ and $S_{L_2}^k$ on G . The proof is essentially a change of variable for the k -linear form

$$(X_{i_1}, \dots, X_{i_k}) \mapsto D_x^k f(X_{i_1}, \dots, X_{i_k}) = X_{i_1} \cdots X_{i_k} f(x).$$

More precisely, for any $f \in \mathcal{B}(G)$ and any $1 \leq r \leq k$,

$$\begin{aligned} & \left(\sum_{j_1, \dots, j_k} |Y_{j_1} \cdots Y_{j_k} f(x)|^2 \right)^{1/2} \\ &= \left(\sum_{j_1, \dots, j_{r-1}, j_{r+1}, \dots, j_k} \sum_{j_r} \left| Y_{j_1} \cdots \left(\sum_i T_{j_r}^i X_i \right) \cdots Y_{j_k} f(x) \right|^2 \right)^{1/2} \\ &= \left(\sum_{j_1, \dots, j_{r-1}, j_{r+1}, \dots, j_k} \sum_{j_r} \left| \sum_i T_{j_r}^i Y_{j_1} \cdots (X_i \cdots Y_{j_k} f(x)) \right|^2 \right)^{1/2} \\ &= \left(\sum_{j_1, \dots, j_{r-1}, j_{r+1}, \dots, j_k} \|T((Y_{j_1} \cdots Y_{j_{r-1}} X_i Y_{j_{r+1}} \cdots Y_{j_k} f(x)))_{i \in \mathcal{I}}\|_{l^2}^2 \right)^{1/2}. \end{aligned}$$

In the last line, $(Y_{j_1} \cdots Y_{j_{r-1}} X_i Y_{j_{r+1}} \cdots Y_{j_k} f(x))_{i \in \mathcal{I}}$ denotes the vector indexed by i with entries $Y_{j_1} \cdots Y_{j_{r-1}} X_i Y_{j_{r+1}} \cdots Y_{j_k} f(x)$ (the other indices $j_1, \dots, j_{r-1}, j_{r+1}, \dots, j_k$ are fixed). Hence by (6.3),

$$\begin{aligned} & \left(\sum_{j_1, \dots, j_k} |Y_{j_1} Y_{j_2} \cdots Y_{j_k} f(x)|^2 \right)^{1/2} \\ & \leq C \left(\sum_{j_1, \dots, j_{r-1}, j_{r+1}, \dots, j_k} \|(Y_{j_1} \cdots X_i \cdots Y_{j_k} f(x))_i\|_{\ell^2}^2 \right)^{1/2}. \end{aligned}$$

Repeating this step gives

$$\left(\sum_{j_1, \dots, j_k} |Y_{j_1} \cdots Y_{j_k} f(x)|^2 \right)^{1/2} \leq C^k \left(\sum_{i_1, \dots, i_k} |X_{i_1} \cdots X_{i_k} f(x)|^2 \right)^{1/2}.$$

Similarly we have

$$c^k \left(\sum_{i_1, \dots, i_k} |X_{i_1} \cdots X_{i_k} f(x)|^2 \right)^{1/2} \leq \left(\sum_{j_1, \dots, j_k} |Y_{j_1} \cdots Y_{j_k} f(x)|^2 \right)^{1/2}.$$

So the norms $S_{L_1}^k$ and $S_{L_2}^k$, and with time derivatives added in, the (semi)norms $S_{L_1}^{k,p}$ and $S_{L_2}^{k,p}$, are equivalent respectively. ■

When $L_1 = \Delta$ is bi-invariant and $L_2 = L$ is some left-invariant form-comparable perturbation of Δ , by repeating the proof of the above lemma with Δ and L inserted in the differential operator chain, we conclude that the norms defined by taking supremum of any of the following are equivalent:

- (i) $(\sum |L^{\lambda_0} X_{l_1} L^{\lambda_1} X_{l_2} L^{\lambda_2} \cdots X_{l_k} L^{\lambda_k} \Delta^b f|^2)^{1/2}$;
- (ii) $(\sum |L^{\lambda_0} Y_{l_1} L^{\lambda_1} Y_{l_2} L^{\lambda_2} \cdots Y_{l_k} L^{\lambda_k} \Delta^b f|^2)^{1/2}$;
- (iii) $(\sum |L^{\lambda_0} X_{l_1} Y_{l'_1} L^{\lambda_1} X_{l_2} Y_{l'_2} L^{\lambda_2} \cdots X_{l_k} Y_{l'_k} L^{\lambda_k} \Delta^b f|^2)^{1/2}$.

Acknowledgements. The authors are partially supported by NSF grant DMS 1707589.

REFERENCES

- [1] A. Bendikov and L. Saloff-Coste, *Elliptic diffusions on infinite products*, J. Reine Angew. Math. 493 (1997), 171–220.
- [2] A. Bendikov and L. Saloff-Coste, *On- and off-diagonal heat kernel behaviors on certain infinite dimensional local Dirichlet spaces*, Amer. J. Math. 122 (2000), 1205–1263.
- [3] A. Bendikov and L. Saloff-Coste, *Central Gaussian semigroups of measures with continuous density*, J. Funct. Anal. 186 (2001), 206–268.
- [4] A. Bendikov and L. Saloff-Coste, *Gaussian bounds for derivatives of central Gaussian semigroups on compact groups*, Trans. Amer. Math. Soc. 354 (2002), 1279–1298.
- [5] A. Bendikov and L. Saloff-Coste, *Central Gaussian convolution semigroups on compact groups: a survey*, Infin. Dimens. Anal. Quantum Probab. Related Topics 6 (2003), 629–659.

- [6] A. Bendikov and L. Saloff-Coste, *Spaces of smooth functions and distributions on infinite-dimensional compact groups*, J. Funct. Anal. 218 (2005), 168–218.
- [7] A. Bendikov and L. Saloff-Coste, *Hypoelliptic bi-invariant Laplacians on infinite dimensional compact groups*, Canad. J. Math. 58 (2006), 691–725.
- [8] A. Bendikov and L. Saloff-Coste, *Brownian motions on compact groups of infinite dimension*, in: Heat Kernels and Analysis on Manifolds, Graphs, and Metric Spaces (Paris, 2002), Contemp. Math. 338, Amer. Math. Soc., Providence, RI, 2003, 41–63.
- [9] A. Bendikov and L. Saloff-Coste, *On the hypoellipticity of sub-Laplacians on infinite dimensional compact groups*, Forum Math. 15 (2003), 135–163.
- [10] E. Born, *An explicit Lévy–Hinč'in formula for convolution semigroups on locally compact groups*, J. Theoret. Probab. 2 (1989), 325–342.
- [11] E. Born, *Projective Lie algebra bases of a locally compact group and uniform differentiability*, Math. Z. 200 (1989), 279–292.
- [12] F. Bruhat, *Distributions sur un groupe localement compact et applications à l'étude des représentations des groupes $\mathfrak{p}$ -adiques*, Bull. Soc. Math. France 89 (1961), 43–75.
- [13] E. B. Davies, *Non-Gaussian aspects of heat kernel behaviour*, J. London Math. Soc. (2) 55 (1997), 105–125.
- [14] V. M. Glushkov, *The structure of locally compact groups and Hilbert's fifth problem*, Amer. Math. Soc. Transl. Ser. (2) 15 (1960), 55–93.
- [15] E. R. van Kampen, *The structure of a compact connected group*, Amer. J. Math. 57 (1935), 301–308.
- [16] S. Kusuoka and D. Stroock, *Applications of the Malliavin calculus. II*, J. Fac. Sci. Univ. Tokyo Sect. IA Math. 32 (1985), 1–76.

Qi Hou
Yanqi Lake Beijing Institute of
Mathematical Sciences and Applications
Yanqihu, Huairou District
Beijing, China
E-mail: houqi@bimsa.cn
qh55@cornell.edu

Laurent Saloff-Coste
Department of Mathematics
Cornell University
567 Malott Hall
Ithaca, New York, United States
E-mail: lsc@math.cornell.edu