Ergodicity of Markov Semigroups with Hörmander type generators in Infinite Dimensions *

Federica Dragoni*, Vasilis Kontis‡, Bogusław Zegarliński‡

* University of Padova

‡Imperial College London

Abstract: *We develop an effective strategy for proving strong ergodicity of (nonsymmetric) Markov semigroups associated to Hörmander type generators when the underlying configuration space is infinite dimensional.*

Contents

1 Introduction

After an initial development of a strategy for proving the log-Sobolev inequality for infinite dimensional Hörmander type generators $\mathcal{L}$ symmetric in $L_2(\mu)$ defined

*Supported by EPSRC EP/D05379X/1

with a suitable nonproduct measure μ ([22], [18], [20], [19]), one can envisage an extension of the established strategy (see e.g. [25]) for proving strong pointwise ergodicity for the corresponding Markov semigroups $P_t \equiv e^{t\mathcal{L}}$, (or in case of the compact spaces even in the uniform norm as in [14] and references therein). Still to obtain a fully fledged theory, which could include for example configuration spaces given by general noncompact nilpotent Lie groups other than Heisenberg type groups, one needs to conquer a (finite dimensional) problem of sub-Laplacian bounds (of the corresponding control distance). Unfortunately this is a VP-hard problem which will likely stay with us for more than quite a while. The other motivation for our work comes also from a desire to get a strategy for studying Markov semigroups of the above mentioned type which are not symmetric with respect to some a priori given reference measure in cases where the underlying configuration space is infinite dimensional and noncompact. In finite dimensions an interesting analysis in the L_2 framework with respect to a reference measure in particular involving the long time behaviour was provided in [24]. In a number of recent works an interesting progress has been made in understanding the sub-gradient bounds on finite dimensional sub-Riemaniann manifolds provided by compact and noncompact Lie groups. Many of the related works (as e.g. [9], [10], [23], [16] see also references therein) are heavily based on complicated stochastic analysis methods with sharp results obtained for Heisenberg type groups. Another insight and complementary understanding were achieved via a more analytic route one can find in [3] and [21] ([18]). In particular such bounds involving the length of the sub-gradient offer a nice way of getting smoothing and spectral properties as well as other interesting features coming from related entropy bounds for the heat kernel. In [4] an analog of the Orstein-Uhlenbeck processes was proposed and studied with the drift term provided by the logarithmic derivative of heat kernels on groups with some general theory involving L_2 subgradient bounds and a related Poincaré inequality. In [5] some stochastic analysis (in a Hilbert space along ideas [8]) is studied for certain infinite dimensional models of financial mathematics. The analysis there concentrates however on hypoellipticity aspects. For some other directions involving a hypoellipticity theme in infinite dimensions see e.g. also [6], [15] and references therein.
In this paper we construct and study Markov semigroups on infinite dimensional spaces provided for example as an infinite product of noncompact Lie groups (as e.g. nilpotent free groups), and formulate an effective condition for their exponential ergodicity in supremum norm. Our main tool is provided by a complete gradient bound, where the square of the gradient (or subgradient) is replaced by similar objects but with a family of fields which is closed with respect to taking commutators with the fields appearing in the definition of our Markov generator. We assume that our theory is furnished with some natural dilation operator

which when included in the generator with sufficiently large coefficient assures the exponential dumping. The use of a complete gradient, while it may not provide us with smoothing information, it proves to be very effective when the long time behaviour is concerned, giving also some extra information about the equilibrium measure. In a finite dimensional setup it provides an alternative view to [4]. On the other hand in a general situation when working in infinite dimensions we have no a priori reference measure and so no natural L_2 approach can be used.
The organisation of the paper is as follows. In section 2 we present the general framework with a number of simple examples, presenting a general idea in finite dimensions. In section 3 we construct a Markov semigroup in an infinite dimensional setup proving a strong approximation property (or as it is sometimes called a finite speed of propagation of information). This approximation is later used together with square of the (complete) gradient bounds to obtain the exponential decay to equilibrium in the supremum norm for a large class of initial configurations. Finally we conclude with a Poincaré type inequality with complete gradient form which allows us via general arguments to obtain exponential moments estimates for suitable (generalised) Lipschitz variables.

2 Finite-dimensional case

Consider smooth vector fields $X_1, \ldots, X_M$ on $\mathbb{R}^N$, satisfying the Hörmander condition with step $K > 1$. For $n \geq N$, by $(Z_k)_{k=1}^n$ we denote an adapted family of fields, containing a basis for the related sub-Riemannian geometry. So $Z_k = X_k$, for $k = 1, \ldots, M$, while the remaining $Z_{M+1}, \ldots, Z_n$ are ordered commutators of length between 2 and K.
For $m \leq M$, we consider the following operator:

$$\begin{aligned} \mathcal{L} &:= L + L_G + L_\alpha \\ L &:= \sum_{i=1}^m X_i^2 - \beta D \\ L_G &:= \sum_{i,j=1}^m G_{i,j}(x) X_i X_j \\ L_\alpha &:= \sum_{i=1}^m \alpha_i(x) X_i \end{aligned} \tag{1}$$

where $\beta \in (0, \infty)$ is a constant, D is a first order dilations generator satisfying

$$e^{sD} Z_k e^{-sD} = e^{s\lambda_k} Z_k \text{ and } [Z_k, D] = \lambda_k Z_k \text{ for some } \lambda_k > 0, \tag{2}$$

for $k = 1, .., n$, and $\alpha(x) = (\alpha_1(x), \dots, \alpha_m(x))$ is a smooth function, while $G(x) = \big(G_{i,j}(x)\big)_{i,j=1}^m$ is an $m \times m$-matrix, satisfying, for any $x \in \mathbb{R}^N$,

$$G^*(x) + I > 0, \quad \text{with } G^*_{ij}(x) \equiv \frac{1}{2}\left(G_{ij}(x) + G_{ji}(x)\right), \tag{3}$$

where I is the $m \times m$-identity-matrix.
Let us introduce the following condition on the geometry of the vector fields:

$$\exists \;\; c_{kjl} \in \mathbb{R} \;\text{ such that }\; [Z_k, X_j] = \sum_{l=1}^n c_{kjl} Z_l, \tag{4}$$

for any $k = 1, \dots, n$ and $j = 1, \dots, m$.

Remark 2.1. *Note that condition (4) is stronger than the Hörmander condition which implies a similar expression, but in general with non-constant coefficients* c_{kjl}*.*

Example 2.2. Here we give some examples of sub-Riemannian geometries which fit in the above framework and where condition (4) holds:

(1) The Heisenberg group: $X_1 = (1, 0, -\frac{y}{2})^T$ and $X_2 = (0, 1, \frac{x}{2})^T$ on $(x, y, z) \in \mathbb{R}^3$. In this case $Z_1 = X_1$, $Z_2 = X_2$ and $Z_3 = Z := [X_1, X_2] = (0, 0, 1)^T$. The family $\{Z_1, Z_2, Z_3\}$ forms a basis for the Lie algebra (here $n = N$). One can calculate that $c_{kjl} = 0$ for any $(k, j, l) \neq (1, 2, 3) \vee (2, 1, 3)$ while $c_{123} = 1$ and $c_{213} = -1$.

(2) The Grušin plane: $X_1 = (1, 0)^T$ and $X_2 = (0, x)^T$ on $(x, y) \in \mathbb{R}^2$. In this case $Z_1 = X_1$, $Z_2 = X_2$ and $Z_3 = Z := [X_1, X_2] = (0, 1)^T$ and $c_{kjl} = 0$ for any $(k, j, l) \neq (1, 2, 3) \vee (2, 1, 3)$ while $c_{123} = 1$ and $c_{213} = -1$. The family $\{Z_1, Z_2, Z_3\}$ contains a basis for the Lie algebra, given by $\{Z_1, Z_3\}$.

(3) The Martinet distribution: $X_1 = (1, 0, -y^2)^T$ and $X_2 = (0, 1, 0)^T$ on $(x, y, z) \in \mathbb{R}^3$. In this case $Z_1 = X_1$, $Z_2 = X_2$, $Z_3 = Z := [X_1, X_2] = (0, 0, 2y)^T$ and $Z_4 = [Z, X_2] = (0, 0, -2)^T$. Then $c_{kjl} = 0$ for any $(k, j, l) \neq (1, 2, 3) \vee (2, 1, 3) \vee (3, 2, 4)$ while $c_{123} = c_{324} = 1$ and $c_{213} = -1$.

Note that the last example (Martinet distribution) is a step 3 distribution while all the others are step 2, and it is the easiest sub-Riemannian geometry where normal geodesics occur.

For smooth functions f, we define

$$\Gamma(f) := \sum_{k=1}^n |Z_k f|^2, \tag{5}$$

which we call the complete gradient form (as opposed, for example, to the sub-gradient of a Lie group). Note here that in general it may be convenient later to include more fields Z_k than it would be necessary just to span the tangent space at any given point. The corresponding quadratic form is given by

$$\Gamma(f,g) = \sum_{k=1}^{n} \big(Z_k f\big)\big(Z_k g\big).$$

2.1 Associated Stochastic Differential Equation

Here we want to write the Stochastic Differential Equation having the operator $\mathcal{L}$ as generator. The SDE has the general form

$$d\xi(t) = \mu(\xi(t))dt + A(\xi(t)) \circ dW(t),$$

where $\mu(\xi(t)) \in \mathbb{R}^n$ is the so called drift part while $A(\xi(t))$ is a $N \times m$ matrix, W is an m-dimensional Brownian motion and by $\circ$ we mean the Stratonovich differential.
It is known that, given a second-order differential operator, it is possible to find a SDE having such an operator as generator, whenever the second-order part can be expressed as trace. In general the first-order part of the operator is related to the drift-part (i.e. the deterministic part of the SDE) while the second-order part is related to the stochastic part of the equation. In particular, the stochastic part has to be written as a Stratonovich differential whenever there is an explicit dependence on the space. We also recall that the Stratonovich differential can be always written in Itô formulation as follows:

$$A(\xi(t)) \circ dW(t) = A(\xi(t))dW(t) + \sum_{i=1}^{m} \nabla_{A^i} A^i(\xi(t))\, dt, \tag{6}$$

where A^i are rows of the matrix A and $\nabla_{A^i} A^j$ is the derivative of the vector field A^j along the vector field A^i, for any i, j. Our operator can be written as

$$\mathcal{L} = \left(\sum_{i=1}^{m} X_i^2 + \sum_{i,j=1}^{m} G_{i\,j}(x) X_i X_j \right) + \left(\sum_{i=1}^{m} \alpha_i(x) X_i - \beta D \right) =: \mathcal{L}_{\textbf{II-order}} + \mathcal{L}_{\textbf{I-order}}.$$

Note that to write the associated SDE, we do not need any assumption on D while we need to assume condition (3).

We denote by $\sigma(x)$ the $n\times m$ matrix whose rows are the vector fields $X_1(x),\dots,X_m(x)$. We first write the drift part which comes from the first-order part of the operator, that is, for any smooth function f,

$$\mathcal{L}_{\textbf{I-order}}\, f = \sum_{i=1}^{m} \alpha_i(x) X_i(f) - \beta D = \alpha^T(x)\sigma^T(x)\nabla f - \beta D^T \nabla f.$$

Note that $\sigma^T(x)\nabla f =: D_{\mathcal{X}} f$ is the horizontal gradient of f (or to be more precise is the coordinate-vector of the horizontal gradient $\mathcal{X} f$ written in the basis of the vector fields $X_1,\dots,X_m$). The drift part for the associated SDE is:

$$\mu(\xi(t)) = \alpha(\xi(t))\sigma(\xi(t)) - \beta D(\xi(t)).$$

Now we want to write explicitly the stochastic part of the equation. Let us first assume that G is symmetric (i.e. $G = G^*$) and introduce

$$B(x) := \sqrt{I + G(x)},$$

where I is the $m\times m$-identity-matrix. Note that $B = (B_{ij})_{i,j=1}^m$ is well-defined since $I + G$ is symmetric and we have assumed condition (3). We are going to show that

$$\mathcal{L}_{\textbf{II-order}} := \sum_{i=1}^{m} X_i^2 + \sum_{i,j=1}^{m} G_{i\,j}(x) X_i X_j$$

is the generator of

$$d\xi(t) = \sigma(\xi(t))B(\xi(t)) \,\circ dW(t) = \left(\sum_{i,j=1}^{m} B_{j\,i}(\xi(t))X_j^l(\xi(t)) \circ dW_i\right)_{l=1}^{N} = \sum_{i=1}^{m} Y_i \circ dW_i,$$

with W_i the standard Brownian motion and denoting $Y_i := \sum_{j=1}^m B_{j\,i} X_j$. It is known (see e.g. [12], [11]) that the generator of the stochastic equation $d\xi_i = Y_i \circ dW_i$ is given by $\sum_i Y_i^2$, so we rest just to calculate it:

$$\sum_i Y_i^2 = \sum_i \left(\sum_j B_{ij} X_j\right)^2 = \sum_i \sum_{lm} B_{il} X_l B_{im} X_m = \sum_{lm}\sum_i B_{il} B_{il} X_l X_m$$
$$= \sum_{lm} (B\, B^t)_{lm} X_l X_m.$$

Using the fact that G is symmetric, we get

$$\sum_i Y_i^2 = \sum_{lm} (B^2)_{lm} X_l X_m = \sum_{lm} (I+G)_{lm} X_l X_m = \sum_l X_l^2 + \sum_{lm} G_{lm} X_l X_m$$
$$= \mathcal{L}_{\textbf{II-order}}.$$

Therefore, under the assumption that G is symmetric (and so is B), the associated SDE is

$$d\xi(t) = \big(\alpha(\xi(t))\sigma(\xi(t)) - \beta D(\xi(t))\big)dt + \sigma(\xi(t))\big(I + G(\xi(t))\big) \circ dW(t). \tag{7}$$

Let us now see what happens when the matrix G is not symmetric. Note that

$$\sum_{i,j} G_{ij} X_i X_j = \sum_{i,j} \left(\frac{G_{ij} + G_{ji}}{2}\right) X_i X_j + \sum_{i,j} \left(\frac{G_{ij} - G_{ji}}{2}\right) X_i X_j$$
$$\equiv \sum_{i,j} G^*_{ij} X_i X_j + \sum_{i,j} G^{aSym}_{ij} X_i X_j.$$

Since for G^{aSym}, the antisymmetric part of G, we have

$$\sum_{i,j} G^{aSym}_{ij} X_i X_j = \frac{1}{2} \sum_{i,j} G^{aSym}_{ij} [X_i, X_j],$$

therefore the antisymmetric part of G gives an extra first order part (i.e. an extra term in the drift part) depending on the commutators. Thus, under assumption (3), the SDE associated to the operator $\mathcal{L}$ is

$$d\xi(t) = \left\{\alpha(\xi(t))\sigma(\xi(t)) - \beta D(\xi(t)) + \frac{1}{2} \sum_{i,j=1}^{m} G^{aSym}_{ij} \big[X_i(\xi(t)), X_j(\xi(t))\big]\right\} dt$$
$$+ \sigma(\xi(t))\big(I + G^*(\xi(t))\big) \circ dW(t). \tag{8}$$

Remark 2.3. *Without assumption* (3) *the II-order part of the operator cannot be written as a trace and therefore is not a generator of a stochastic process. The same condition will arise in order to find an exponential decay for the semigroup associated to the operator.*

2.2 Existence of a limit measure

Let $(P_t)_{t\geq 0}$ denote the semigroup generated by $\mathcal{L}$, where $\mathcal{L}$ is given by (1). We show that one can extract a subsequence $(P_{t_k})_{k=1}^{\infty}$ which converges weakly to a probability measure on $\mathbb{R}^N$. Here and in the sequel we use the notation $d(x) = d(x, 0)$, where d is a metric on $\mathbb{R}^N$.

Lemma 2.4. *Let ρ be a smooth function such that $\rho(x) = 0$ for $d(x) < 1$ and $\rho \to \infty$ as $d(x) \to \infty$. Assume*

1. $\sum_{i=1}^{m} X_i^2 \rho + \sum_{i,j=1}^{m} G_{ij} X_i X_j \rho \leq C_1$

2. $\sum_{i=1}^m |X_i\rho|^2 \leq C_2$

3. $P_t\rho \leq c_1 P_t D\rho + c_2$

for some constants $C_1, C_2, c_1, c_2 > 0$. *Then there exists a constant* $K \in (0,\infty)$ *such that* $P_t\rho \leq K$ *for all* $t > 0$.

Proof. We have

$$\begin{aligned}\partial_t P_t\rho &= P_t\mathcal{L}\rho \\ &= P_t\sum_{i=1}^m X_i^2\rho - \beta P_t D\rho + P_t\sum_{i=1}^m \alpha_i X_i\rho + P_t\sum_{i,j=1}^m G_{ij}X_iX_j\rho \\ &\leq C_1 - \frac{\beta}{c_1}P_t\rho - \frac{c_2\beta}{c_1} + \max_i \|\alpha_i\|_\infty\sqrt{n}C_2\end{aligned}$$

using our assumptions. Integrating this inequality we get

$$P_t\rho \leq e^{-\eta t}\rho + \frac{\beta}{c_1\eta}\left(1 - e^{-\eta t}\right)$$

with $\eta = C_1 - \frac{c_2\beta}{c_1} + \max_i \|\alpha_i\|_\infty\sqrt{n}C_2$, which is bounded for all $t > 0$. □

Remark 2.5. *The first two assumptions of Lemma 2.4 can be relaxed to*

1. $\sum_{i=1}^m X_i^2\rho + \sum_{i,j=1}^m G_{ij}X_iX_j\rho \leq C_1\rho + \tilde{C}_1$ *and*

2. $\sum_{i=1}^m |X_i\rho|^2 \leq C_2\rho + \tilde{C}_2$

respectively, for some constants $C_1, \tilde{C}_1, C_2, \tilde{C}_2 > 0$, *at the expense of having to take* β *large enough to ensure that the coefficient of* $P_t\rho$ *in the proof is negative.*

The function ρ can be thought of as a cut-off of an appropriate distance function.

Example 2.6. *We illustrate this in case of a Lie group of Heisenberg type* $\mathbb{G} = \left(\mathbb{R}^{m+l}, \circ, \delta_\lambda\right)$, *with left-invariant vector fields* $X_1, \ldots, X_m$. *Such a group is naturally equipped with dilations* $\delta_\lambda(x,t) = (\lambda x, \lambda^2 t)$, *where* $(x,t) \in \mathbb{R}^m \times \mathbb{R}^l$, *which form a 1-parameter family of homomorphisms. Here, one may define the following smooth homogeneous gauge (also known as the Folland-Kaplan gauge, see e.g. [7])*

$$N(x,t) = \left(|x|^4 + 16|t|^2\right)^{\frac{1}{4}}, \tag{9}$$

where $|\cdot|$ *denotes the Euclidean norm. A computation then shows that the sub-gradient and the sub-Laplacian of this gauge function read*

$$\sum_{i=1}^m |X_iN|^2 = \frac{|x|^2}{N^2} \tag{10}$$

and
$$\sum_{i=1}^{m} X_i^2 N = 3\frac{|x|^2}{N^3} \tag{11}$$
respectively, while the dilation operator is the generator of $(\delta_\lambda)_{\lambda>0}$ given by
$$\begin{aligned} D &= \partial_\lambda \mid_{\lambda=1} \delta_\lambda(x,t) \\ &= \sum_{i=1}^{m} x_i \partial_{x_i} + 2\sum_{i=1}^{n} t_i \partial_{t_i}. \end{aligned}$$
Since $\partial_{x_i} N = N^{-3}|x|^2 x_i$ and $\partial_{t_i} N = 8N^{-3} t_i$, we have
$$DN = N^{-3}\left(|x|^4 + 16|t|^2\right) = N$$
and therefore $P_t N = P_t DN$. Moreover, if we introduce a cut-off function $g : \mathbb{R} \to \mathbb{R}$ such that $g(x) = 0$ on $[0,1]$, $g(x) = x$ for $x \geq 2$ and g is continuous and smooth on $(1,2)$, then the function $\hat{N}(x,t) = g(N(x,t))$ is such that $\sum_{i=1}^m X_i^2 \hat{N} + |X_i \hat{N}|^2$ is bounded, since $|x| \leq N(x,t)$. Hence, in this case, a function satisfying the assumptions of Lemma 2.4 exists.

Similarly one can construct suitable ρ for other (noncompact) homogeneous Lie groups using a smooth (outside the origin) homogeneous norm (of [17], [7]).

Theorem 2.7. *There exists a sequence $\{t_k\}_{k=1}^\infty \subset \mathbb{R}$ and a probability measure ν on $\mathbb{R}^n$ such that for all bounded and Lipschitz f*
$$P_{t_k} f \to \int f d\nu$$
as $k \to \infty$.

Proof. For $L > 0$, we define sets $\Upsilon_L = \{\rho \leq L\}$ for which we have, by Markov's inequality and Lemma 2.4,
$$P_t(\Upsilon_L) \geq 1 - \frac{K}{L},$$
for some constant $K > 0$. Therefore $(P_t)_{t>0}$ represents a tight family of measures on $\mathbb{R}^N$ and we deduce from Prokhorov's Theorem that there exists a convergent subsequence $P_{t_k} \to \lim_{k\to\infty} P_{t_k} =: \nu$ in the weak sense. $\square$

2.3 Complete Gradient Bounds

We start by proving the following bound for the semigroup P_t and the complete gradient Γ defined in (5).

Theorem 2.8. *Let $\mathcal{L}$ be the operator defined in (1), under the assumptions (2) and (3) and let P_t be the semigroup associated to $\mathcal{L}$. Let us also assume that G_{ij} is constant and $\|Z_k\alpha_i\|_\infty < \infty$ for all $k = 1 \dots n$ and $i, j = 1 \dots m$. If (4) holds, then there exists $\kappa \in \mathbb{R}$ such that*

$$\Gamma(P_t f) \leq e^{-\kappa t} P_t \Gamma(f). \tag{12}$$

Moreover, there exists $b_0 \in (0, \infty)$ such that for all $\beta > b_0$ we have $\kappa \in (0, \infty)$.

Proof. The proof follows the Bakry-Emery type strategy (see e.g. [1], [2]) with suitable modifications required by our setup. Let us set $f_s := P_s f$. Note that it is sufficient to prove that

$$\frac{d}{ds} P_{t-s}\Gamma(f_s) \leq -\kappa \ P_{t-s}\Gamma(f_s), \tag{13}$$

which gives estimate (12) after integration over $s \in [0, t]$.
To prove (13), we remark that

$$\frac{d}{ds} P_{t-s}\Gamma(f_s) = P_{t-s}\big(-\mathcal{L}\Gamma(f_s) + 2\Gamma(f_s, \mathcal{L}f_s)\big),$$

since $\Gamma(f, g)$ is defined as a bilinear form.
Using the explicit expressions for $\Gamma(f)$ and $\mathcal{L}$, the previous relation becomes:

$$\begin{aligned} \frac{d}{ds} P_{t-s}\Gamma(f_s) &= P_{t-s}\sum_k \Big(-\mathcal{L}|Z_k f_s|^2 + 2(Z_k f_s)(Z_k \mathcal{L} f_s)\Big) \\ &= P_{t-s}\sum_k \Big(-\mathcal{L}|Z_k f_s|^2 + 2(Z_k f_s)(\mathcal{L} Z_k f_s) + 2(Z_k f_s)\big([Z_k, \mathcal{L}] f_s\big)\Big). \end{aligned}$$

We note that

$$\begin{aligned} I_k :=-\mathcal{L}|Z_k f_s|^2 + 2(Z_k f_s)(\mathcal{L} Z_k f_s) &= -2\sum_j |X_j Z_k f_s|^2 - 2\sum_{i,j} G_{ij}(X_i Z_k f_s)(X_j Z_k f_s) \\ &= -2\sum_{i,j}\big(G_{ij} + \delta_{ij}\big)(X_i Z_k f_s)(X_j Z_k f_s), \end{aligned} \tag{14}$$

where $\delta_{ij} = 0$, for $i \neq j$ and $\delta_{ii} = 1$, for any $i, j = 1, \dots, m$.
The third term is more difficult to estimate since it depends on the commutators. For this purpose, we need to use assumption (4). Let us set

$$J_k := 2(Z_k f_s)\big([Z_k, \mathcal{L}] f_s\big),$$

and calculate the commutators $[Z_k, \mathcal{L}]f_s$, i.e.

$$[Z_k, \mathcal{L}]f_s = [Z_k, L]f_s + [Z_k, L_G]f_s + [Z_k, L_\alpha]f_s.$$

Recalling the definition of $\mathcal{L}$ and noting that $[Z, Y^2]f = [Z, Y]Yf + Y[Z, Y]f$, using assumption (2), for the operator L we get

$$\begin{aligned}[Z_k, L]f_s &= \sum_j [Z_k, X_j^2]f_s - \beta\lambda_k(Z_k f_s) \\ &= \sum_j \left\{[Z_k, X_j](X_j f_s) + X_j[Z_k, X_j]f_s\right\} - \beta\lambda_k(Z_k f_s).\end{aligned}$$

Using condition (4), we obtain

$$[Z_k, L]f_s = \sum_{j,l} c_{kjl}\left\{(Z_l X_j f_s) + (X_j Z_l f_s)\right\} - \beta\lambda_k(Z_k f_s).$$

We are going to use the negative term I_k, in order to control mixed terms like $(Z_k f_s)(X_j Z_k f_s)$. To this end we have to rewrite the term $(Z_l X_j f_s)$ in a more suitable form, using once more assumption (4), i.e.

$$\begin{aligned}\sum_{j,l} c_{kjl}\left\{(Z_l X_j f_s) + (X_j Z_l f_s)\right\} &= \sum_{j,l} c_{kjl}\left\{2(X_j Z_l f_s) + [Z_l, X_j]f_s\right\} \\ &= 2\sum_{j,l} c_{kjl}(X_j Z_l f_s) + \sum_{j,l,n} c_{kjl}c_{ljn}(Z_n f_s),\end{aligned}$$

Hence, summing up we get

$$\begin{aligned}\sum_k J_k =& -2\beta\sum_k \lambda_k|Z_k f_s|^2 + 4\sum_{k,j,l} c_{kjl}(Z_k f_s)(X_j Z_l f_s) && (15)\\ &+ 2\sum_{k,j,l,n} c_{kjl}c_{ljn}(Z_k f_s)(Z_n f_s) + 2\sum_k (Z_k f_s)\left\{[Z_k, L_G]f_s + [Z_k, L_\alpha]f_s\right\}. && (16)\end{aligned}$$

We can now similarly estimate the remaining terms. Note that

$$[Z, XY]f = [Z, X]Yf + X[Z, Y]f\,,$$

and thus

$$\begin{aligned}[Z_k, L_G]f_s &= \sum_{i,j} G_{ij}\left\{[Z_k, X_i](X_j f_s) + X_i[Z_k, X_j]f_s\right\} \\ &= \sum_{i,j,l}(G_{ij} + G_{ji})c_{kil}(X_j Z_l f_s) + \sum_{i,j,l,n} G_{ij}c_{kil}c_{ljn}(Z_n f_s).\end{aligned}$$

Moreover,

$$
\begin{aligned}
[Z_k, L_\alpha] f_s &= \sum_i (Z_k \alpha_i) X_i f_s + \sum_i \alpha_i [Z_k, X_i] f_s \\
&= \sum_i (Z_k \alpha_i) X_i f_s + \sum_{i,l} \alpha_i c_{kil} (Z_l f_s).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
2 \sum_k (Z_k f_s)[Z_k, L_\alpha] f_s &= 2 \sum_{k,i} (Z_k f_s)(Z_k \alpha_i)(X_i f_s) + 2 \sum_{k,i,l} \alpha_i c_{kil} (Z_k f_s)(Z_l f_s) \\
&\leq \sum_{k,i} \|Z_k \alpha_i\|_\infty \left(|Z_k f_s|^2 + |X_i f_s|^2 \right) + 2 \sum_{k,i,l} \alpha_i c_{kil} (Z_k f_s)(Z_l f_s) \\
&\leq \max_k \sum_i \|Z_k \alpha_i\|_\infty \Gamma(f_s) + \max_i \sum_k \|Z_k \alpha_i\|_\infty \Gamma(f_s) \\
&\quad + 2 \sum_{k,i,l} \alpha_i c_{kil} (Z_k f_s)(Z_l f_s),
\end{aligned}
$$

where we used Young's inequality to estimate the first term. Combining the above, (15) becomes

$$
\begin{aligned}
\sum_k J_k =& - 2\beta \sum_k \lambda_k |Z_k f_s|^2 + 4 \sum_{k,j,l} c_{kjl} (Z_k f_s)(X_j Z_l f_s) + \\
&+ 2 \sum_{k,j,l,n} c_{kjl} c_{ljn} (Z_k f_s)(Z_n f_s) + 2 \sum_{k,i,j,l} (G_{ij} + G_{ji}) c_{kil} (Z_k f_s)(X_j Z_l f_s) \\
&+ 2 \sum_{k,i,j,l,n} G_{ij} c_{kil} c_{ljn} (Z_k f_s)(Z_n f_s) + 2 \sum_{k,i,l} \alpha_i c_{kil} (Z_k f_s)(Z_l f_s) + \eta \Gamma(f_s),
\end{aligned}
\tag{17}
$$

with $\eta = \max_k \sum_i \|Z_k \alpha_i\|_\infty + \max_i \sum_k \|Z_k \alpha_i\|_\infty$. We start by estimating the terms in (17) where $(X_j Z_l f_s)$ does not appear. Let us set $\lambda_* := \min_k \lambda_k > 0$, then $-2\beta \sum_k \lambda_k |Z_k f_s|^2 \leq -2\beta \lambda_* \Gamma(f_s)$ (recalling that $\beta > 0$). The other terms can be treated similarly. Recalling that α_i, G_{ij} and c_{kjl} are in general non positive, we get:

$$
\begin{aligned}
2\sum_{k,j,l,n} c_{kjl}c_{ljn}(Z_kf_s)(Z_nf_s) &+ 2\sum_{k,i,j,l,n} G_{ij}c_{kil}c_{ljn}(Z_kf_s)(Z_nf_s)\\
&= 2\sum_{k,i,j,l,n}\big(G_{ij}+\delta_{ij}\big)c_{kil}c_{ljn}(Z_kf_s)(Z_nf_s)\\
&\le \sum_{k,i,j,l,n}\big|G_{ij}+\delta_{ij}\big||c_{kil}||c_{ljn}|\big\{|Z_nf_s|^2+|Z_kf_s|^2\big\}\\
&\le \sup_k\sum_{n,i,j,l}\big|G_{ij}+\delta_{ij}\big|\big\{|c_{kil}||c_{ljn}|+|c_{nil}||c_{ljk}|\big\}\Gamma(f_s) =: C_1\Gamma(f_s)
\end{aligned}
$$

and

$$
2\sum_{k,i,l}\alpha_i c_{kil}(Z_kf_s)(Z_lf_s) \le \sup_k\sum_{i,l}|\alpha_i|\big(|c_{kil}|+|c_{lik}|\big)\Gamma(f_s) =: C_2\Gamma(f_s).
$$

By (14) and (17), we have the following estimate:

$$
\begin{aligned}
\frac{d}{ds}P_{t-s}\Gamma(f_s) &= \sum_k (I_k+J_k)\\
&\le \big(-2\beta\lambda_* + C_1 + C_2 + \eta\big)\Gamma(f_s) + 4\sum_{k,j,l} c_{kjl}(Z_kf_s)(X_jZ_lf_s)\\
&\quad + 2\sum_{k,i,j,l}(G_{ij}+G_{ji})c_{kil}(Z_kf_s)(X_jZ_lf_s)\\
&\quad - 2\sum_{k,i,j}\big(G_{ij}+\delta_{ij}\big)(X_iZ_kf_s)(X_jZ_kf_s).
\end{aligned}
$$

The idea is to use Young's inequality to estimate the remaining terms with a non-positive part depending on $(X_iZ_kf_s)$ and a positive part depending just on $\Gamma(f_s)$. Let us recall that $D_{\mathcal{X}}Z_kf_s = (X_1Z_kf_s,\dots,X_1Z_kf_s)\in\mathbb{R}^m$ is the horizontal gradient of Z_kf_s, then

$$
\sum_k I_k = -2\sum_{k,i,j}\big(G_{ij}+\delta_{ij}\big)(X_iZ_kf_s)(X_jZ_kf_s) = -2\sum_k\Big\langle\big(G^*+I\big)D_{\mathcal{X}}Z_kf_s, D_{\mathcal{X}}Z_kf_s\Big\rangle,
$$

since $\langle Aa,a\rangle = 0$, for any $a\in\mathbb{R}^m$, whenever A is an antisymmetric matrix (recall $G^*_{ij} = \frac{G_{ij}+G_{ji}}{2}$).

Analogously,

$$
\begin{aligned}
4\sum_{k,j,l} c_{kjl}(Z_kf_s)(X_jZ_lf_s) &+ 2\sum_{k,i,j,l}(G_{ij}+G_{ji})\,c_{kil}(Z_kf_s)(X_jZ_lf_s)\\
&= 4\sum_{k,i,j,l}(\delta_{ij}+G^*_{ij})\,c_{kil}(Z_kf_s)(X_jZ_lf_s) =: I'.
\end{aligned}
$$

For the sake of simplicity, let us denote $a_{jl} := \sum_{k,i}(\delta_{ij} + G^*_{ij})c_{kil}(Z_k f_s)$ and $b_{jl} := (X_j Z_l f_s)$. Young's inequality tells that $a_{jl} b_{jl} \leq \frac{\varepsilon a_{jl}^2}{2} + \frac{b_{jl}^2}{2\varepsilon}$, for any $\varepsilon > 0$. Thus

$$I' \leq 2 \sum_{jl} \left(\varepsilon \left(\sum_{k,i} (\delta_{ij} + G^*_{ij})\, c_{kil}(Z_k f_s) \right)^2 + \frac{|X_j Z_l f_s|^2}{\varepsilon} \right).$$

We can estimate the first part as follows:

$$\begin{aligned}
&2\varepsilon \sum_{jl} 2 \left(\sum_{k,i} (\delta_{ij} + G^*_{ij})\, c_{kil}(Z_k f_s) \right)^2 \\
&\quad = 2\varepsilon \sum_{j,l} \left(\sum_{k,n} \left[\sum_i \left(\delta_{ij} + G^*_{ij}\right) c_{kil}(Z_n f_s) \right] \left[\sum_i \left(\delta_{ij} + G^*_{ij}\right) c_{nil}(Z_k f_s) \right] \right) \\
&\quad \leq \varepsilon \sum_{j,l,k,n} \left\{ \left[\sum_i \left(\delta_{ij} + G^*_{ij}\right) c_{kil} \right]^2 |Z_n f_s|^2 + \left[\sum_i \left(\delta_{ij} + G^*_{ij}\right) c_{nil} \right]^2 |Z_k f_s|^2 \right\} \\
&\qquad\qquad = 2\varepsilon \sum_{k,l,j} \left[\sum_i \left(\delta_{ij} + G^*_{ij}\right) c_{kil} \right]^2 \Gamma(f_s) =: \varepsilon\, C_3 \Gamma(f_s),
\end{aligned}$$

while

$$2 \sum_{jl} \frac{|X_j Z_l f_s|^2}{\varepsilon} = 2 \sum_k \left\langle \frac{1}{\varepsilon} D_{\mathcal{X}} Z_k f_s, D_{\mathcal{X}} Z_k f_s \right\rangle.$$

Therefore we can conclude

$$\begin{aligned}
&\frac{d}{ds} P_{t-s} \Gamma(f_s) = \sum_k (I_k + J_k) \\
&\leq \big(-2\beta\lambda_* + C_1 + C_2 + \varepsilon\, C_3 \big) \Gamma(f_s) - 2 \sum_k \left\langle \left(G^* + I - \frac{I}{\varepsilon} \right) D_{\mathcal{X}} Z_k f_s, D_{\mathcal{X}} Z_k f_s \right\rangle.
\end{aligned}$$

By assumption (3), there exists $\delta > 0$ such that

$$G^* + I \geq \delta I. \tag{18}$$

Choosing $\varepsilon = \frac{1}{\delta}$, where δ is (the biggest number) such that (18) holds, the inner product in the above estimate is non-negative. Therefore

$$\frac{d}{ds} P_{t-s} \Gamma(f_s) \leq \left(-2\beta\lambda_* + C_1 + C_2 + \eta + \frac{1}{\delta}\, C_3 \right) \Gamma(f_s),$$

which gives the theorem with

$$\begin{aligned}\kappa :=& 2\beta\lambda_* - C_1 - C_2 - \eta - \frac{1}{\delta}\, C_3 \\ =& 2\beta \min_k \lambda_k - \sup_k \sum_{n,i,j,l} \big|G_{ij} + \delta_{ij}\big|\big\{|c_{kil}||c_{ljn}| + |c_{nil}||c_{ljk}|\big\} \\ & - \sup_k \sum_{i,l} |\alpha_i|\big(|c_{kil}| + |c_{lik}|\big) - \max_k \sum_i \|Z_k\alpha_i\|_\infty - \max_i \sum_k \|Z_k\alpha_i\|_\infty \\ & - \frac{2}{\delta} \sum_{k,l,j} \left[\sum_i \big(\delta_{ij} + G^*_{ij}\big) c_{kil}\right]^2 . \end{aligned} \tag{19}$$

Finally, by choosing $\beta > \frac{1}{2\lambda_*}\left(C_1 + C_2 + \eta + \frac{1}{\delta}\, C_3\right)$ we can ensure that $\kappa > 0$. $\square$

Remark 2.9. *More generally, for a non-constant matrix $G = G(x)$ the theorem continues to hold under the additional assumption that the quantities $\|Z_k G_{ij}\|_\infty$ are bounded for all $k = 1 \dots n$ and $i, j = 1 \dots m$.*

Remark 2.10 (Case $G = 0$)**.** *Whenever $G = 0$, we can choose $\delta = 1$ in the constant κ. Then*

$$\begin{aligned}\kappa =& 2\beta \min_k \lambda_k - \max_k \sum_i \|Z_k\alpha_i\|_\infty - \max_i \sum_k \|Z_k\alpha_i\|_\infty \\ & - \sup_k \left\{ \sum_{n,j,l} \big\{|c_{kjl}||c_{ljn}| + |c_{njl}||c_{ljk}|\big\} - \sum_{i,l} |\alpha_i|\big(|c_{kil}| + |c_{lik}|\big) \right\} - 2\sum_{k,l,j} c^2_{kjl}.\end{aligned}$$

Remark 2.11 (Optimal constant in the case $G = 0$)**.** *The constant found is a priori not always optimal. In fact, to deduce constant C_1 and C_3 we used two different estimates: $2\mu\nu ab \leq \mu\nu(a^2 + b^2)$ for the first constant, $2\mu\nu ab \leq \mu^2 a^2 + \nu^2 b^2$ for the second one.*
It is possible to give examples when the first estimate is optimal (i.e. if $a = b$ but $\mu \neq \nu$ e.g. $\mu = \nu^{-1}$) and examples when the second one is the optimal one (i.e. $\mu a = \nu b$ but $a \neq b$ e.g. $a = b^{-1}$). Therefore we could use both estimates, finding two different constants:

$$\begin{aligned} C'_1 &= 2\sum_{kjl} c^2_{kjl} = C_2, \\ C'_2 &= 2\sup_n \sum_{kjl} c_{kjl} c_{njl}. \end{aligned}$$

It is clear that $C_2 < C'_2$ (taking $k = n$ in C'_2) while C_1 is often smaller than C'_1 because in general very few c_{kjl} are different from 0. Therefore in C_1 many terms *vanish, but it is not a always true. Therefore in the case $G_{ij} = 0$, the optimal constant is given by $\overline{\kappa} := 2\beta\lambda_* - \overline{C}_1 - C_2 - C_3 - \eta$, with $\overline{C}_1 := \min\{C_1, C_1'\}$. In the same way, one could write the optimal constant in the case $G_{ij} \neq 0$.*

Example 2.12. (1) *In the Heisenberg group, one can consider the operator $D := xX_1 + yX_2 + 2zZ = x\partial_x + y\partial_y + 2z\partial_z$ which satisfies assumption (2) with $\lambda_1 = 1$, $\lambda_2 = 1$ and $\lambda_3 = 2$; therefore $\lambda_* := \min_k \lambda_k = 1$. In the simplest case where $\alpha_i \equiv G_{ij} \equiv 0$, using Remark 2.10 and recalling that $c_{123} = 1 = -c_{213}$ and $c_{kjl} = 0$ otherwise, we see that $\kappa = 2\beta - 4$ so that $\kappa > 0$ for any $\beta > 2$.*
(2) *In the Grušin plane the dilation operator is given by $D := x\partial_x + 2y\partial_y$. Assumption (2) is satisfied with $\lambda_1 = \lambda_2 = 1$ and $\lambda_3 = 2$ and thus $\lambda_* = 1$ as in the Heisenberg group.*
(3) *The dilation operator for the Martinet distribution is $D := x\partial_x + y\partial_y + 3z\partial_z$ so $\lambda_1 = \lambda_2 = 1$, while $\lambda_3 = 2$ and $\lambda_4 = 3$. Hence $\lambda_* = 1$.*

If we do not assume (4), then by the Hörmander condition we know that

$$\exists \ c_{kjl} \in C^\infty(\mathbb{R}^n) \text{ such that } [Z_k, X_j] = \sum_{l=1}^n c_{kjl}(x) Z_l. \tag{20}$$

Looking at the above calculations, we get and extra term in $\sum_k J_k$ where the horizontal derivatives of the coefficients appear, more precisely

$$2\sum_{k,j,l}(X_j c_{kjl})(Z_l f_s)(Z_k f_s) \le \sup_k \sum_{jl} \left\{|X_j c_{kjl}(x)| + |X_j c_{ljk}(x)|\right\}\Gamma(f_s) =: C_4(x)\Gamma(f_s).$$

This implies that:

$$\frac{d}{ds}P_{t-s}\Gamma(f_s) \le \left(-2\beta\lambda_* + C_1(x) + C_2(x) + \eta + \frac{1}{\delta}\, C_3(x) + C_4(x)\right)\Gamma(f_s),$$

Therefore Theorem 2.13 holds in a stronger form since the exponential in the estimate depends now on time and space.

Theorem 2.13. *Let $X_1, \ldots, X_m$ be smooth vector fields satisfying the Hörmander condition, $\mathcal{L}$ the operator defined in (1), satisfying the assumptions (2) and (3), and let P_t be the associated semigroup. Let us also assume that G_{ij} is constant and $\|Z_k\alpha_i\|_\infty < \infty$ for all $k = 1 \ldots n$ and $i, j = 1 \ldots m$. Then there exists a smooth function $\kappa(x)$ such that*

$$\Gamma(P_t f) \le e^{-\kappa(x)t} P_t \Gamma(f). \tag{21}$$

Moreover, under the additional assumption that the functions $c_{kjl}(x)$ and their horizontal derivatives $X_j c_{kjl}(x)$ are bounded in $x \in \mathbb{R}^n$, there exists $\widehat{b_0} \in (0, \infty)$ such that for all $\beta > \widehat{b_0}$ we have $\kappa > 0$.

The following result provides an extension of Theorem 2.8 to an l_q-gradient bound.

Theorem 2.14. *Let $q > 1$. Under the assumptions of Theorem 2.8, there exists a constant $\kappa' \in \mathbb{R}$ such that*

$$\Gamma(P_t f)^{\frac{q}{2}} \leq e^{-\kappa' t} P_t \Gamma(f)^{\frac{q}{2}}.$$

Moreover, there exists $b_0' > 0$ such that for $\beta > b_0'$ we have $\kappa' > 0$.

Proof. As before, we follow the strategy outlined in [1]. For simplicity we treat the case $L_G = L_\alpha = 0$. The proof of the general case follows from the proof of Theorem 2.8 and arguments similar to the ones below. We aim to show that

$$\frac{d}{ds} P_{t-s} \Gamma(f_s)^{\frac{q}{2}} \leq -\kappa' P_{t-s} \Gamma(f)^{\frac{q}{2}}$$

with $f_s = P_s f$ as above. To this end, we note that

$$\begin{aligned}\frac{d}{ds} P_{t-s} \Gamma(f_s)^{\frac{q}{2}} &= P_{t-s} \left(-\mathcal{L}\Gamma(f_s)^{\frac{q}{2}} + q\Gamma(f_s)^{\frac{q}{2}-1} \Gamma(f_s, \mathcal{L} f_s) \right) \\ &= P_{t-s} \left(-\frac{q}{2} \frac{\mathcal{L}\Gamma(f_s)}{\Gamma(f_s)^{1-\frac{q}{2}}} + \frac{q}{2}\left(1 - \frac{q}{2}\right) \frac{\bar{\Gamma}(\Gamma(f_s))}{\Gamma(f_s)^{2-\frac{q}{2}}} + q \frac{\Gamma(f_s, \mathcal{L} f_s)}{\Gamma(f_s)^{1-\frac{q}{2}}} \right)\end{aligned}$$

where we made use of the diffusion property for the generator $\mathcal{L}$ and set $2\bar{\Gamma}(f) := \mathcal{L}(f^2) - 2f\mathcal{L}f = \sum_{i=1}^m |X_i f|^2$. In what follows, the variables i, k in the sums run over the ranges $\{1, \dots m\}$ and $\{1, \dots, n\}$ respectively. For the first term, we have

$$\begin{aligned}\mathcal{L}\Gamma(f_s) &= \sum_{i,k} X_i^2 (Z_k f_s)^2 - \beta \sum_k D(Z_k f_s)^2 \\ &= 2\sum_{i,k} |X_i Z_k f_s|^2 + 2\sum_k (Z_k f_s)(\mathcal{L} Z_k f_s).\end{aligned}$$

On the other hand, the second term can be estimated as follows:

$$\begin{aligned}\bar{\Gamma}(\Gamma(f_s)) &= \sum_i \left(X_i \left(\sum_k (Z_k f_s)^2 \right) \right)^2 = 4 \sum_i \left(\sum_k (Z_k f_s)(X_i Z_k f_s) \right)^2 \\ &\leq 4 \left(\sum_k |Z_k f_s|^2 \right) \left(\sum_{i,k} |X_i Z_k f_s|^2 \right) = 4\Gamma(f_s) \left(\sum_{i,k} |X_i Z_k f_s|^2 \right).\end{aligned}$$

Finally

$$\begin{aligned}\Gamma(f_s, \mathcal{L} f_s) &= \sum_{i,k} (Z_k f_s)(Z_k X_i^2 f_s) - \beta \sum_k (Z_k f_s)(Z_k D f_s) \\ &= \sum_k (Z_k f_s)(\mathcal{L} Z_k f_s) + \sum_k (Z_k f_s)([Z_k, \mathcal{L}] f_s),\end{aligned}$$

Combining the above, we obtain

$$\begin{aligned}
&\frac{d}{ds}P_{t-s}\Gamma(f_s)^{\frac{q}{2}} \leq \\
&\quad P_{t-s}\left(-q\frac{\left(\sum_{i,k}|X_iZ_kf_s|^2+\sum_{i,k}(Z_kf_s)(\mathcal{L}Z_kf_s)\right)}{\Gamma(f_s)^{1-\frac{q}{2}}}\right)\\
&\quad + P_{t-s}\left(q\frac{\left(\sum_{i,k}(Z_kf_s)(Z_k\mathcal{L}f_s)\right)}{\Gamma(f_s)^{1-\frac{q}{2}}}\right) + P_{t-s}\left(q(2-q)\frac{\sum_{i,k}|X_iZ_kf_s|^2}{\Gamma(f_s)^{1-\frac{q}{2}}}\right)\\
&= P_{t-s}\left(\frac{q\left(\sum_{i,k}((Z_kf_s)[Z_k,X_i^2]f_s+(1-q)|X_iZ_kf_s|^2)-\beta\sum_k(Z_kf_s)[Z_k,D]f_s\right)}{\Gamma(f_s)^{1-\frac{q}{2}}}\right)\\
&\leq P_{t-s}\left(\frac{q}{\Gamma(f_s)^{1-\frac{q}{2}}}\sum_{i,k}(Z_kf_s)[Z_k,X_i^2]f_s-\beta\lambda_*q|\nabla f_s|^q-\frac{q(q-1)}{\Gamma(f_s)^{1-\frac{q}{2}}}\sum_{i,k}|X_iZ_kf_s|^2\right),
\end{aligned}$$

where we made use of assumption (2) and $\lambda_* = \inf_k \lambda_k$. The fact that $q > 1$ is crucial, since this makes the coefficient in front of the last term in the above expression non-zero, hence allowing us to use it to control the mixed derivatives coming from the first term as follows. Observe that using assumption (4), we can write

$$\begin{aligned}
\sum_{i,k}(Z_kf_s)([Z_k,X_i^2]f_s) &= \sum_{i,k}(Z_kf_s)\left([Z_k,X_i](X_if_s)+X_i[Z_k,X_i]f_s\right)\\
&= \sum_{i,k}(Z_kf_s)\left(\sum_l c_{kil}\left(Z_lX_if_s+X_iZ_lf_s\right)\right)\\
&= \sum_{i,k}(Z_kf_s)\left(\sum_l c_{kil}(2X_iZ_lf_s+[Z_l,X_i]f_s)\right)\\
&= \sum_{i,k}(Z_kf_s)\left(\sum_l 2c_{kil}X_iZ_lf_s+\sum_{l,m}c_{kil}c_{lim}Z_mf_s\right)\\
&\leq 2\sum_{i,k,l}c_{kil}(Z_kf_s)(X_iZ_lf_s)\\
&\qquad + \sum_{i,k,l,m}|c_{kil}c_{lim}|\left(\frac{|Z_kf_s|^2+|Z_mf_s|^2}{2}\right),
\end{aligned}$$

where in the last step we used the Cauchy-Schwarz inequality. Thus

$$
\begin{aligned}
&\frac{d}{ds}P_{t-s}\Gamma(f_s)^{\frac{q}{2}} \leq \\
&\quad P_{t-s}\left(\frac{q}{\Gamma(f_s)^{1-\frac{q}{2}}}\sum_{i,k,l,m}\left(2c_{kil}(Z_kf_s)(X_iZ_lf_s)+|c_{kil}c_{lim}|\left(\frac{|Z_kf_s|^2+|Z_mf_s|^2}{2}\right)\right)\right) \\
&\quad - P_{t-s}\left(\beta\lambda_*q\Gamma(f_s)^{\frac{q}{2}}+\frac{q(q-1)}{\Gamma(f_s)^{1-\frac{q}{2}}}\sum_{i,k}|X_iZ_kf_s|^2\right) \\
&\leq P_{t-s}\left(\frac{q}{\Gamma(f_s)^{1-\frac{q}{2}}}\sum_{i,k,l}2c_{ikl}(Z_kf_s)(X_iZ_lf_s)-\frac{q(q-1)\sum_{i,k}|X_iZ_kf_s|^2}{\Gamma(f_s)^{1-\frac{q}{2}}}+C_1\Gamma(f_s)^{\frac{q}{2}}\right)
\end{aligned}
$$

where $C_1 = q\left(\frac{1}{2}\max_k\sum_{i,l,m}|c_{kil}c_{lim}|+\frac{1}{2}\max_m\sum_{i,l,k}|c_{kil}c_{lim}|-\beta\lambda_*\right)$. Using the inequality

$$
\frac{q}{\Gamma(f_s)^{1-\frac{q}{2}}}\sum_{i,k,l}2c_{ikl}(Z_kf_s)(X_iZ_lf_s) \leq C_2\Gamma(f_s)^{\frac{q}{2}}+\frac{q(q-1)}{\Gamma(f_s)^{1-\frac{q}{2}}}\sum_{i,k}|X_iZ_kf_s|^2.
$$

with $C_2 = \frac{q}{q-1}\left(\max_k\sum_{i,l}c_{ikl}^2\right)$, we get

$$
\frac{d}{ds}P_{t-s}\Gamma(f_s)^{\frac{q}{2}} \leq -\kappa' P_{t-s}\Gamma(f_s)^{\frac{q}{2}}
$$

with $\kappa' = -C_1 - C_2$ which is strictly positive for

$$
\beta > \frac{1}{\lambda_*}\left(\frac{1}{2}\max_k\sum_{i,l,m}|c_{kil}c_{lim}|+\frac{1}{2}\max_m\sum_{i,l,k}|c_{kil}c_{lim}|+\frac{1}{q-1}\left(\max_k\sum_{i,l}c_{ikl}^2\right)\right).
$$

Integration of the above differential inequality ends the proof. □

Summarising, the key idea of our estimates is contained in the assumption of completeness of the set of fields $\{Z_k\}$ in the sense that their commutators with the fields appearing in the generator do not give essentially new fields. In case of free nilpotent Lie groups ([13]) they can be chosen simply by taking all the fields generated by the fields defining the generator. In some cases (as for example groups of rank 2, when the fields have linear coefficients) our procedure will work with the usual square of the gradient form. Then our method applied to the complete gradient $\{Z_k\}$ provides some other useful information on monotonicity of derivatives.

3 Extension to infinite dimensions

Let $\mathbf{\Lambda} \Subset \mathbb{Z}^d$ be a finite subset of the d-dimensional lattice. For each $k \in \mathbb{Z}^d$, we consider isomorphic copies of the vector fields $Z_1, \ldots, Z_n$ denoted by $Z_{k,1}, \ldots, Z_{k,n}$ (and similarly the isomorphic copies of $X_i = Z_i$, $i = 1, \ldots, m$). As in the finite dimensional case, let

$$\mathcal{L}_k^{(1)} = \sum_{i=1}^{m} X_{k,i}^2 - \beta D_k + \sum_{i,i'=1}^{m} G_{ii'} X_{k,i} X_{k,i'},$$
$$\mathcal{L}_k^{(2)} = \sum_{i=1}^{m} \alpha_{k,i} X_{k,i},$$

where $G = (G_{ij})$ is a constant matrix satisfying (3) and $\alpha_{k,i} = \alpha_{k,i}(\omega),\ \omega \in (\mathbb{R}^N)^{\mathbb{Z}^d}$. We will work under the additional assumption that the range of interaction is finite. In other words $\alpha_{k,i}$ will depend only on coordinates around k, i.e. $Z_{k,i}\alpha_{j,r} = 0$ whenever $|k - j| > R$. Here, the distance of two points on the lattice is defined as $|k-j| = \sum_{l=1}^{d} |k_l - j_l|$, where $k = (k_1, \ldots k_l)$ and $j = (j_1, \ldots j_l)$. For $k \in \mathbb{Z}^d$, $\Lambda \subset \mathbb{Z}^d$, we set $dist(k, \Lambda) \equiv \inf\{|k - l|) : l \in \Lambda\}$. In addition, we will assume that the quantities $\|Z_{k,r}\alpha_{j,i}\|_\infty$ are uniformly bounded in $k, j \in \mathbb{Z}^d$. By $\Lambda(f)$ we will denote a localisation set for a function f, meaning that f depends only on coordinates indexed by points in $\Lambda(f)$.
We consider a Markov semigroup $P_t^{\mathbf{\Lambda}}$, defined via its generator

$$\mathcal{L}_{\mathbf{\Lambda}} = \sum_{k \in \mathbb{Z}^d} \mathcal{L}_k^{(1)} + \sum_{k \in \mathbf{\Lambda}} \mathcal{L}_k^{(2)}$$

and we define $\Gamma_k = \sum_{r=1}^{n} |Z_{k,r} f|^2, \Gamma_{\mathbf{\Lambda}} = \sum_{k \in \mathbf{\Lambda}} \Gamma_k$ and $\Gamma = \sum_{k \in \mathbb{Z}^d} \Gamma_k$. This definition is motivated by the fact that the generators $\mathcal{L}_{\mathbf{\Lambda}}$ approximate, as $\mathbf{\Lambda} \uparrow \mathbb{Z}^d$, the infinite dimensional generator

$$\mathcal{L} = \sum_{k \in \mathbb{Z}^d} \left(\mathcal{L}_k^{(1)} + \mathcal{L}_k^{(2)} \right) \equiv \sum_{k \in \mathbb{Z}^d} \mathcal{L}_k.$$

This construction, the details of which are presented below, allows us to approximate the infinite dimensional semigroup $(e^{t\mathcal{L}})_{t \geq 0}$ by Markov semigroups, which are easier to study.

3.1 Strong Approximation Property

Given a finite set $\mathbf{\Lambda} \subset \mathbb{Z}^d$, for a cylinder function f such that $\Lambda(f) \subset \mathbf{\Lambda}$ we introduce $f_s = P_s^{\mathbf{\Lambda}} f$ and start similarly as before by considering

$$\begin{aligned}
\partial_s P_{t-s}^{\mathbf{\Lambda}} \Gamma_k(f_s) &= P_{t-s}^{\mathbf{\Lambda}} \left(-\mathcal{L}_{\mathbf{\Lambda}} \Gamma_k(f_s) + 2\Gamma_k(f_s, \mathcal{L}_{\mathbf{\Lambda}} f_s)\right) \\
&= P_{t-s}^{\mathbf{\Lambda}} \sum_{r=1}^{n} \left(-\mathcal{L}_{\mathbf{\Lambda}} |Z_{k,r}(f_s)|^2 + 2(Z_{k,r} f_s)\mathcal{L}_{\mathbf{\Lambda}} Z_{k,r} f_s + 2(Z_{k,r} f_s)[Z_{k,r}, \mathcal{L}_{\mathbf{\Lambda}}] f_s\right) \\
&= P_{t-s}^{\mathbf{\Lambda}} \sum_{r=1}^{n} \left(-2\bar{\Gamma}(Z_{k,r} f_s) + 2(Z_{k,r} f_s)[Z_{k,r}, \mathcal{L}_{\mathbf{\Lambda}}] f_s\right)
\end{aligned}$$

where $\bar{\Gamma}(f) = \sum_{j \in \mathbb{Z}^d} \left(\sum_{i,i'=1}^{m} (G_{ii'} + \delta_{ii'})(X_{j,i} f)(X_{j,i'} f)\right)$. Next, we notice that

$$\begin{aligned}
[Z_{k,r}, \mathcal{L}_{\mathbf{\Lambda}}] &= \sum_{j \in \mathbb{Z}^d} \sum_{i=1}^{m} \left[Z_{k,r}, X_{j,i}^2 - \beta D_j\right] + \sum_{j \in \mathbb{Z}^d} \sum_{i,i'=1}^{m} [Z_{k,r}, G_{ii'} X_{j,i} X_{j,i'}] \\
&\quad + \sum_{j \in \mathbf{\Lambda}} \sum_{i=1}^{m} [Z_{k,r}, \alpha_{j,i} X_{j,i}] \\
&= \sum_{i=1}^{m} \left([Z_{k,r}, X_{k,i}^2 - \beta D_k] + \alpha_{k,i}[Z_{k,r} X_{k,i}]\right) \\
&\quad + \sum_{i,i'=1}^{m} [Z_{k,r}, G_{ii'} X_{k,i} X_{k,i'}] + \sum_{j \in \mathbf{\Lambda}} \sum_{i=1}^{m} (Z_{k,r} \alpha_{j,i}) X_{j,i},
\end{aligned}$$

because $Z_{k,r}$ and $X_{j,i}$ commute when $j \neq k$ for all r, i. Combining the above, we arrive at

$$\begin{aligned}
\partial_s P_{t-s}^{\mathbf{\Lambda}} \Gamma_k(f_s) &= 2P_{t-s}^{\mathbf{\Lambda}} \sum_{r=1}^{n} \Bigg(-\bar{\Gamma}(Z_{k,r} f_s) + \sum_{i=1}^{m} (Z_{k,r} f_s)[Z_{k,r}, X_{k,i}^2 - \beta D_k] f_s \\
&\quad + \sum_{i,i'=1}^{m} (Z_{k,r} f_s)[Z_{k,r}, G_{ii'} X_{k,i} X_{k,i'}] f_s + \sum_{i=1}^{m} \alpha_{k,i} (Z_{k,r} f_s)[Z_{k,r} X_{k,i}] f_s \Bigg) \\
&\quad + 2P_{t-s}^{\mathbf{\Lambda}} \sum_{j \in \mathbf{\Lambda}} \sum_{i=1}^{m} \sum_{r=1}^{n} (Z_{k,r} f_s)(Z_{k,r} \alpha_{j,i})(X_{j,i} f_s) \\
&\leq 2(-\beta \lambda_* + \tilde{C}) P_{t-s}^{\mathbf{\Lambda}} \Gamma_k(f_s) + 2P_{t-s}^{\mathbf{\Lambda}} \sum_{j \in \mathbf{\Lambda}} \sum_{i=1}^{m} \sum_{r=1}^{n} (Z_{k,r} f_s)(Z_{k,r} \alpha_{j,i})(X_{j,i} f_s),
\end{aligned}$$

with some constant $\tilde{C}$ dependent only on the structure constants c_{kjl}, G_{ij} and $||\alpha_{k,i}||$. For the last sum, we use Young's inequality to get

$$
\begin{aligned}
2\sum_{j\in\mathbf{\Lambda}}\sum_{i=1}^{m}\sum_{r=1}^{n}&(Z_{k,r}f_s)(Z_{k,r}\alpha_{j,i})(X_{j,i}f_s) \leq \\
&\sum_{i=1}^{m}\sum_{r=1}^{n}|Z_{k,r}\alpha_{k,i}|\left(|Z_{k,r}f_s|^2+|X_{k,i}f_s|^2\right) \\
&+\sum_{j\in\mathbf{\Lambda},j\neq k}\sum_{i=1}^{m}\sum_{r=1}^{n}|Z_{k,r}\alpha_{j,i}|\left(|Z_{k,r}f_s|^2+|X_{j,i}f_s|^2\right) \\
&\leq A_k\Gamma_k(f_s)+\sum_{j\in\mathbf{\Lambda},j\neq k}M_{k,j}\Gamma_j(f_s)
\end{aligned}
$$

with

$$
\begin{aligned}
A_k &= \max_{r=1,\dots,n}\sum_{i=1}^{m}\|Z_{k,r}\alpha_{k,i}\|_\infty+\max_{i=1,\dots,m}\sum_{r=1}^{n}\|Z_{k,r}\alpha_{k,i}\|_\infty \\
&\quad+\sum_{j\in\mathbf{\Lambda},j\neq k}\max_{r=1,\dots,n}\sum_{i=1}^{m}\|Z_{k,r}\alpha_{j,i}\|_\infty, \\
M_{k,j} &= \max_{i=1,\dots,m}\left(\sum_{r=1}^{n}\|Z_{k,r}\alpha_{j,i}\|_\infty\right),
\end{aligned}
$$

which are finite quantities by our assumptions. We therefore arrive at

$$
\partial_s P^{\mathbf{\Lambda}}_{t-s}\Gamma_k(f_s)\leq-\bar{\kappa}P^{\mathbf{\Lambda}}_{t-s}\Gamma_k(f_s)+\sum_{j\in\mathbf{\Lambda},j\neq k}M_{k,j}P^{\mathbf{\Lambda}}_{t-s}\Gamma_j(f_s) \tag{22}
$$

with

$$
C\equiv\tilde{C}+\sup_{k\in\mathbb{Z}^d}A_k \qquad \text{and} \qquad \bar{\kappa}\equiv 2(\beta\lambda_*-C).
$$

Solving this differential inequality, we obtain the following bound:

Lemma 3.1. *There exists constants $\bar{\kappa}\in\mathbb{R}$ and $M_{k,j}\in(0,\infty)$, $M_{k,j}\equiv 0$ for $|j-k|>R$, such that for any $\mathbf{\Lambda}\subset\mathbb{Z}^d$ and any smooth cylinder function f with $\Lambda(f)\subset\mathbf{\Lambda}$, we have*

$$
\Gamma_k(P^{\mathbf{\Lambda}}_t f)\leq e^{-\bar{\kappa}t}P^{\mathbf{\Lambda}}_t\Gamma_k(f)+\sum_{j\in\mathbf{\Lambda},j\neq k}M_{k,j}\int_0^t ds\, e^{-\bar{\kappa}(t-s)}\,P^{\mathbf{\Lambda}}_{t-s}\Gamma_j(P^{\mathbf{\Lambda}}_s f). \tag{23}
$$

Remark 3.2. *One can use this lemma to get gradient bounds in l_q, $q\geq 1$ norms for vectors Γ_k, $k\in\mathbb{Z}^d$.*

Remark 3.3. *For a matrix $\hat{G} = ((\hat{G}^{kk'}_{ii'})^m_{i,i'=1})_{k,k'\in\mathbb{Z}^d}$ satisfying $\hat{G}^* + I \geq 0$ and $\sum_{k,k'\in\mathbb{Z}^d}\sum_{i,i'=1}^m |\hat{G}^{kk'}_{ii'}| < \infty$, it is possible to repeat the above argument for the generator given by*

$$\hat{\mathcal{L}}_{\mathbf{\Lambda}} = \mathcal{L}_{\mathbf{\Lambda}} + \sum_{k,k'\in\mathbb{Z}^d}\sum_{i,i'=1}^m \hat{G}^{kk'}_{ii'} X_{k,i} X_{k',i'}.$$

Proposition 3.4 (Finite speed of propagation of information)**.** *Let f be a smooth function and assume $\Lambda(f) \subset \mathbf{\Lambda} \Subset \mathbb{Z}^d$. For $k \notin \Lambda(f)$, we have*

$$\|\Gamma_k(P^{\mathbf{\Lambda}}_t f)\|_\infty \leq e^{N_k(\log C - \log N_k + 2 + \log t) + Ct} \sum_{j\in\mathbb{Z}^d} \|\Gamma_j f\|_\infty,$$

where $N_k = \left[\frac{dist(k,\Lambda(f))}{R}\right]$ and $C > 0$ is a constant. Hence, for any $\sigma > 0$ there exists $\tau > 1$ such that if $N_k \geq \tau t$

$$\|\Gamma_k(P^{\mathbf{\Lambda}}_t f)\|_\infty \leq e^{-\sigma t - \sigma N_k} \sum_{j\in\mathbb{Z}^d} \|\Gamma_j f\|_\infty.$$

Proof. We argue similarly as in [14] (see also references given there). From Lemma 3.1, we have

$$||\Gamma_k(P^{\mathbf{\Lambda}}_t f)||_\infty \leq e^{-\bar{\kappa}t}||\Gamma_k(f)||_\infty + \sum_{j\in\mathbf{\Lambda}, j\neq k} M_{k,j} \int_0^t ds\, e^{-\bar{\kappa}(t-s)}\, ||\Gamma_j(P^{\mathbf{\Lambda}}_s f)||_\infty$$

with $M_{kj} \equiv 0$, $|j-k| \geq R$ and $\kappa \in (0,\infty)$. This implies

$$\begin{aligned}\|\Gamma_k(P_t f)\|_\infty &\leq \|\Gamma_k f\|_\infty + \sum_{j\in\mathbf{\Lambda}} M_{kj} \int_0^t \|\Gamma_j(f_s)\|_\infty ds \\ &= \sum_{j\in\mathbf{\Lambda}} M_{kj} \int_0^t \|\Gamma_j(f_s)\|_\infty ds,\end{aligned}$$

since $k \notin \Lambda(f)$. We may iterate the above to get

$$\|\Gamma_k(P_t f)\|_\infty \leq C^{N_k} \frac{t^{N_k}}{N_k!} e^{Ct} \sum_{j\in\mathbb{Z}^d} \|\Gamma_j f\|_\infty,$$

with some constant $C > 0$ and $N_k = \left[\frac{dist(k,\Lambda(f))}{R}\right]$. Since $N_k! > e^{N_k \log N_k - 2N_k}$,

$$\|\Gamma_k(P_t f)\|_\infty \leq e^{N_k(\log C - \log N_k + 2 + \log t) + Ct} \sum_{j\in\mathbb{Z}^d} \|\Gamma_j f\|_\infty.$$

Now, given $\sigma > 0$, we may choose $\tau \geq 1$ large enough so that $\log \frac{C}{\tau} + 2 + \frac{C}{\tau} \leq -2\sigma$. If $N_k \geq \tau t$, we then get

$$\begin{aligned} N_k \left(\log C - \log N_k + 2 + \log t\right) + Ct &\leq N_k \left(\log \frac{C}{\tau} + 2 + \frac{C}{\tau}\right) \\ &\leq N_k(-2\sigma) \leq -\sigma N_k - \sigma t, \end{aligned}$$

as required. $\square$

Theorem 3.5. *For any $t > 0$ and any continuous function f the following limit exists in the uniform norm*

$$\lim_{\mathbf{\Lambda} \uparrow \mathbb{Z}^d} P_t^{\mathbf{\Lambda}} f =: P_t f$$

and defines a Markov semigroup.

Proof. It is sufficient to prove the existence of the limit for smooth cylinder functions. To this end pick $\mathbf{\Lambda}_1, \mathbf{\Lambda}_2 \Subset \mathbb{Z}^d$ such that $\Lambda(f) \subset \mathbf{\Lambda}_1 \subset \mathbf{\Lambda}_2$ and choose a sequence $\mathbf{\Lambda}^{(n)}$ such that $\mathbf{\Lambda}^{(0)} = \mathbf{\Lambda}_1$, $\mathbf{\Lambda}^{(\mathcal{N})} = \mathbf{\Lambda}_2$ and $\mathbf{\Lambda}^{(n+1)} \setminus \mathbf{\Lambda}^{(n)} = \{j_n\}$ is a singleton for any $n = 0, \ldots, \mathcal{N} - 1$. Using the Fundamental Theorem of Calculus and the fact that $P_t^{\mathbf{\Lambda}^{(n)}}$ is contractive, we have

$$\begin{aligned} \left\| P_t^{\mathbf{\Lambda}_2} f - P_t^{\mathbf{\Lambda}_1} f \right\|_\infty &\leq \sum_{n=0}^{\mathcal{N}-1} \left\| P_t^{\mathbf{\Lambda}^{(n)}} f - P_t^{\mathbf{\Lambda}^{(n+1)}} f \right\|_\infty \\ &\leq \sum_{n=0}^{\mathcal{N}-1} \int_0^t \left\| P_{t-s}^{\mathbf{\Lambda}^{(n)}} (\mathcal{L}_{\mathbf{\Lambda}^{(n)}} - \mathcal{L}_{\mathbf{\Lambda}^{(n+1)}}) P_s^{\mathbf{\Lambda}^{(n+1)}} f \right\|_\infty ds \\ &\leq \sum_{n=0}^{\mathcal{N}-1} \int_0^t \left\| (\mathcal{L}_{\mathbf{\Lambda}^{(n)}} - \mathcal{L}_{\mathbf{\Lambda}^{(n+1)}}) P_s^{\mathbf{\Lambda}^{(n+1)}} f \right\|_\infty ds \\ &\leq \sum_{n=0}^{\mathcal{N}-1} \int_0^t \left\| \sum_{i=1}^m \alpha_{j_n,i} X_{j_n,i} P_s^{\mathbf{\Lambda}^{(n+1)}} f \right\|_\infty ds \\ &\leq \sum_{n=0}^{\mathcal{N}-1} \int_0^t \left\| \sum_{i=1}^m \alpha_{j_n,i}^2 \right\|_\infty^{\frac{1}{2}} \left\| \sum_{i=1}^m \left| X_{j_n,i} P_s^{\mathbf{\Lambda}^{(n+1)}} f \right|^2 \right\|_\infty^{\frac{1}{2}} ds. \\ &\leq m \sum_{n=0}^{\mathcal{N}-1} \int_0^t \|\alpha_{j_n,i}\|_\infty \sqrt{\|\Gamma_{j_n}(P_s^{\mathbf{\Lambda}^{(n+1)}} f)\|_\infty} ds \end{aligned}$$

Let $\sigma > 0$. Since $j_n \notin \mathbf{\Lambda}_1$, we can apply Lemma 3.4 to conclude that

$$\left\|P_t^{\mathbf{\Lambda}_2} f - P_t^{\mathbf{\Lambda}_1} f\right\|_\infty \leq m \sum_{n=0}^{\mathcal{N}-1} \|\alpha_{j_n,i}\|_\infty \int_0^t e^{-\sigma s - \sigma N_{j_n}} \sqrt{\sum_{j\in\mathbb{Z}^d} \|\Gamma_j f\|_\infty} ds$$

$$\leq m\mathcal{N} e^{-\sigma \overline{N}} \, \frac{1-e^{-\sigma t}}{\sigma} \max_{n=1,\ldots,\mathcal{N}-1} \|\alpha_{j_n,i}\|_\infty \sqrt{\sum_{j\in\mathbb{Z}^d} \|\Gamma_j f\|_\infty} ds \quad (24)$$

provided that $N_{j_n} = \left[\frac{dist(j_n,\Lambda(f))}{R}\right] \geq \overline{N} \geq \tau t$ for some $\tau > 1$ large enough, where $\overline{N} = \left[\frac{dist(\mathbf{\Lambda}_1,\Lambda(f))}{R}\right]$. We have therefore established that if $(\mathbf{\Lambda}_n \Subset \mathbb{Z}^d)_{n=0}^\infty$ is a sequence such that $\mathbf{\Lambda}_n \uparrow \mathbb{Z}^d$ as $n \to \infty$, then $(P_t^{\mathbf{\Lambda}_n} f)_{n=0}^\infty$ is a Cauchy sequence. $\square$

3.2 Existence of a limit measure

Let $\zeta \in \mathbb{N}$ be such that $\sum_{k\in\mathbb{Z}^d}(1+|k|)^{-\zeta} < \infty$. For $\mathscr{K} \in \mathbb{N}$, define sets

$$\Omega_{\mathscr{K}} = \left\{\omega \in (\mathbb{R}^N)^{\mathbb{Z}^d} : \sum_{k\in\mathbb{Z}^d} (1+|k|)^{-\zeta} d(\omega_k) < \mathscr{K}\right\}$$

and let

$$\Omega := \cup_{\mathscr{K}\in\mathbb{N}}(\Omega_{\mathscr{K}}) = \left\{\omega \in (\mathbb{R}^N)^{\mathbb{Z}^d} : \sum_{k\in\mathbb{Z}^d} (1+|k|)^{-\zeta} d(\omega_k) < \infty\right\}.$$

For $j \in \mathbb{Z}^d$ and $\omega \in (\mathbb{R}^N)^{\mathbb{Z}^d}$ we consider the (semi-)distance $d_j(\omega) \equiv d(\omega_j)$ (recall that we write $d(x,0) = d(x)$ where d is a metric on $\mathbb{R}^N$). The corresponding cut-off ρ_j then satisfies, (similarly as in Lemma 2.4),

$$P_t^{\mathbf{\Lambda}} \sum_{j\in\mathbf{\Lambda}} \rho_j \leq K_{\mathbf{\Lambda}}$$

for some constant $K_{\mathbf{\Lambda}} > 0$ and all $t > 0$. If we define $\Upsilon_L^{\mathbf{\Lambda}} = \{\sum_{j\in\mathbf{\Lambda}} \rho_j \leq L\}$, arguing as in Section 2.2, we can extract a convergent subsequence $P_{t_k}^{\mathbf{\Lambda}}$ such that for all bounded continuous f and $\omega \in \Omega$ we have $P_{t_k}^{\mathbf{\Lambda}} f(\omega) \to \nu_{\mathbf{\Lambda},\omega}(f)$.

3.3 Ergodicity of the semigroup

By Section 3.2 and Theorem 3.5 we have that for $\omega \in \Omega$ there exists a measure ν_ω such that

$$P_{t_k}^{\mathbf{\Lambda}} f(\omega) \to \nu_\omega(f)$$

as $k \to \infty$ and $\mathbf{\Lambda} \uparrow \mathbb{Z}^d$. Moreover, by Markov's inequality, for all $\omega \in \Omega$

$$\nu_\omega(\Omega_{\mathscr{K}}) \geq 1 - \frac{1}{\mathscr{K}} \sup_{k \in \mathbb{Z}^d} \left(\int d(x_k) \nu_\omega(dx) \right) \sum_{k \in \mathbb{Z}^d} (1 + |k|)^{-\zeta}$$

and thus $\nu_\omega(\Omega) = 1$. We will show the following result.

Theorem 3.6. *There exists $t_0 > 0$ such that for $t > t_0$, bounded smooth cylinder function f and any $\omega, \tilde{\omega} \in \Omega$,*

$$|P_t f(\omega) - P_t f(\tilde{\omega})| \leq \mathcal{C}(f, \omega, \tilde{\omega}) e^{-\varpi t},$$

where $\varpi > 0$ is a constant and $\mathcal{C}(f, \omega, \tilde{\omega})$ depends only on f, ω and $\tilde{\omega}$.

Proof. We choose $\mathbf{\Lambda} = \mathbf{\Lambda}(t)$ such that $\operatorname{diam}(\mathbf{\Lambda}) = \varkappa t$ for some $\varkappa > 0$ to be determined later, and order the elements of $\mathbf{\Lambda}$ lexicographically. For $\omega, \tilde{\omega} \in \Omega$ we can choose a suitable sequence $(\omega^k)_{k \in \mathbb{Z}^d}$ that interpolates between ω and $\tilde{\omega}$ and such that each element differs from the previous one only in single coordinate. Moreover for all $\mathbf{\Lambda} \Subset \mathbb{Z}^d$,

$$\begin{aligned} |P_t f(\omega) - P_t f(\tilde{\omega})| \leq & |P_t f(\omega) - P_t^{\mathbf{\Lambda}} f(\omega)| + |P_t^{\mathbf{\Lambda}} f(\omega) - P_t^{\mathbf{\Lambda}} f(\tilde{\omega})| \\ & + |P_t f(\tilde{\omega}) - P_t^{\mathbf{\Lambda}} f(\tilde{\omega})|. \end{aligned}$$

By the proof of Theorem 3.5 and the fact that $\operatorname{diam}(\mathbf{\Lambda}) = \varkappa t$, we can find $T > 0$ such that $t > T/\varkappa$ implies

$$\left| P_t f(\omega) - P_t^{\mathbf{\Lambda}} f(\omega) \right| + \left| P_t f(\tilde{\omega}) - P_t^{\mathbf{\Lambda}} f(\tilde{\omega}) \right| \leq \mathcal{C}_1(f, \omega, \tilde{\omega}) e^{-\theta t/2},$$

where $\theta \in (0, \infty)$ and $\mathcal{C}_1(f, \omega, \tilde{\omega})$ is a finite constant depending on the cylinder function f and configurations $\omega, \tilde{\omega}$. We also have

$$\left| P_t^{\mathbf{\Lambda}} f(\omega) - P_t^{\mathbf{\Lambda}} f(\tilde{\omega}) \right| \leq \sum_{k \in \mathbf{\Lambda}^R} \left| P_t^{\mathbf{\Lambda}} f(\omega^{k+1}) - P_t^{\mathbf{\Lambda}} f(\omega^k) \right|,$$

with $\mathbf{\Lambda}^R = \{k \in \mathbb{Z}^d : dist(k, \mathbf{\Lambda}) \leq R\}$ where R is the range of interaction. Let $\gamma : [0, t_k] \to \Omega$ be an admissible path connecting ω^k to ω^{k+1}, such that $\dot{\gamma}_s = 1$ (recall that ω^k and ω^{k+1} differ only in the k^{th} coordinate, so $t_k = d(\omega_k, \tilde{\omega}_k)$). The differential inequality (22) implies that

$$\partial_s P_{t-s}^{\mathbf{\Lambda}} \Gamma_k (P_s^{\mathbf{\Lambda}} f) \leq -(\bar{\kappa} - \max_{j \in \mathbb{Z}^d} M_{k,j}) P_{t-s}^{\mathbf{\Lambda}} \Gamma_{\mathbf{\Lambda}} (P_s^{\mathbf{\Lambda}} f)$$

(recall that $M_{k,j} \equiv 0$ when $|j - k| > R$), which after integration gives

$$\Gamma_k (P_t^{\mathbf{\Lambda}} f) \leq e^{-\varsigma t} P_t^{\mathbf{\Lambda}} (\Gamma_{\mathbf{\Lambda}} f)$$

with some $\varsigma \in \mathbb{R}$ which is positive for large β and can be made independent of k by our assumption that the quantities $\|Z_{k,r}\alpha_{j,i}\|_\infty$ are uniformly bounded in $k, j \in \mathbb{Z}^d$. This observation together with contractivity property of $P_t^{\mathbf{\Lambda}}$ imply

$$
\begin{aligned}
\sum_{k\in\mathbf{\Lambda}^R} \left|P_t^{\mathbf{\Lambda}} f(\omega^{k+1}) - P_t^{\mathbf{\Lambda}} f(\omega^k)\right| &\leq \sum_{k\in\mathbf{\Lambda}^R} \int_0^{t_k} \sqrt{\Gamma_k(P_t^{\mathbf{\Lambda}} f(\gamma_s))} ds \\
&\leq \sum_{k\in\mathbf{\Lambda}^R} (d(\omega_k) + d(\tilde{\omega}_k)) \left\|\Gamma_k(P_t^{\mathbf{\Lambda}} f)\right\|_\infty^{\frac{1}{2}} \leq \sum_{k\in\mathbf{\Lambda}^R} (d(\omega_k) + d(\tilde{\omega}_k)) e^{-\frac{\varsigma t}{2}} \left\|\Gamma_{\mathbf{\Lambda}}(f)\right\|_\infty^{\frac{1}{2}} \\
&\leq e^{-\frac{\varsigma t}{2}} \left(\sum_{k\in\mathbb{Z}^d} \|\Gamma_k(f)\|_\infty\right)^{\frac{1}{2}} \sum_{k\in\mathbf{\Lambda}^R} (d(\omega_k) + d(\tilde{\omega}_k)) \qquad (25) \\
&\leq e^{-\frac{\varsigma t}{2}} \left(\sum_{k\in\mathbb{Z}^d} \|\Gamma_k(f)\|_\infty\right)^{\frac{1}{2}} (C_\omega + C_{\tilde{\omega}})(1 + \varkappa t)^\zeta
\end{aligned}
$$

using that $|k| \leq \varkappa t$ since $k \in \mathbf{\Lambda}$, with $C_\omega \equiv \sum_{k\in\mathbb{R}^{\mathbb{Z}^d}} (1 + |k|)^{-\zeta} d(\omega_k)$ which is finite since $\omega \in \Omega$ (and similarly for $C_{\tilde{\omega}}$). Hence there exists a constant $\mathcal{C}_2(f, \omega, \tilde{\omega})$ such that

$$\left|P_t^{\mathbf{\Lambda}} f(\omega) - P_t^{\mathbf{\Lambda}} f(\tilde{\omega})\right| \leq \mathcal{C}_2(f, \omega, \tilde{\omega})(1 + \varkappa t)^\zeta e^{-\varsigma t/2}.$$

Combining the above we conclude that $t > T/\varkappa \equiv t_0$ implies

$$|P_t f(\omega) - P_t f(\tilde{\omega})| \leq \mathcal{C}(f, \omega, \tilde{\omega}) e^{-\varpi t},$$

for some constants $\varpi > 0$ and $\mathcal{C}(f, \omega, \tilde{\omega})$ depending only on the cylinder function f and configurations $\omega, \tilde{\omega}$. □

Remark 3.7. *We note that in fact the estimate (25) is sufficiently strong to include the configurations with exponential growth for which* $\sum_k e^{-\gamma|k|} d(\omega_k) < \infty$ *with any* $\gamma < \varsigma/2$ *which is much more than a set of measure one.*

3.4 Properties of the Invariant Measure

Recall the following representation of a covariance

$$P_t f^2 - (P_t f)^2 = 2 \int_0^t \partial_s P_s \bar{\Gamma}(P_{t-s} f) ds \qquad (26)$$

Since $\bar{\Gamma} \leq \Gamma$, if we have the following bound

$$\Gamma(P_\tau f) \leq e^{-\kappa\tau} P_\tau \Gamma(f)$$

then (26) implies the following result

Theorem 3.8. *Under the conditions on the generator $\mathcal{L}$ there exists $\beta_0 \in (0, \infty)$ such that for all $\beta > \beta$, any differentiable function f, at any $t > 0$*

$$P_t f^2 - (P_t f)^2 \leq \frac{2}{\kappa}(1 - e^{-\kappa t})\, P_t \Gamma(f). \tag{27}$$

Hence the unique P_t-invariant measure ν satisfies

$$\nu(f - \nu f)^2 \leq \frac{2}{\kappa}\, \nu \Gamma(f).$$

We mention that by abstract arguments, (see e.g. [14], Exercise 2.9, and references therein), the Poincaré type inequality (27) implies a uniform in $t > 0$ exponential bound

$$P_t e^{\delta f} < Const\, e^{\frac{\delta^2}{\kappa}\Gamma(f)} e^{\delta P_t f} \tag{28}$$

provided

$$\frac{\delta^2}{\kappa} ||\Gamma(f)||_\infty \leq 1. \tag{29}$$

Application of this property yields the following exponential bound result:

Corollary 3.9. *Under the condition of the Theorem 3.8 the invariant measure ν satisfies the following exponential bound*

$$\nu\left(e^{\delta f}\right) < Const\, e^{\frac{\delta^2}{\kappa}\Gamma(f)} e^{\delta \nu(f)}$$

for any function f satisfying (27) and for which νf is well defined.

Remark *An interesting question arises, which was also a part of motivation to our work, whether the measure ν can satisfy stronger coercive inequalities as for example Log-Sobolev inequality. The known strategy of [2] to obtain log-Sobolev requires bounds with Γ_1 and and unfortunately fails in cases of interest to us in this paper.*

Remark *Note that knowing a bit of regularity one can slightly optimize (26) as follows. First we use*

$$P_t f^2 - (P_t f)^2 \leq \int_0^{t-\varepsilon} ds 2 P_s \Gamma(P_{t-s} f) + \int_{t-\varepsilon}^{t} ds 2 P_s \Gamma(P_{t-s} f)$$

If one would have the following regularity estimate

$$\Gamma(P_\varepsilon f) \leq \bar{c}(\varepsilon) \Gamma_1(f)$$

then we get

$$\int_0^{t-\varepsilon} ds 2 P_s \Gamma(P_{t-s} f) = \int_0^{t-\varepsilon} ds 2 P_s \Gamma(P_{t-\varepsilon-s} P_\varepsilon f) \leq 2\bar{c}(\varepsilon) \int_0^{t-\varepsilon} ds e^{-\kappa(t-\varepsilon-s)} P_{t-\varepsilon} \Gamma_1(f)$$

$$\leq \frac{2}{\kappa}\bar{c}(\varepsilon)(1-e^{-\kappa(t-\varepsilon)})P_{t-\varepsilon}\Gamma_1(f) \leq \frac{2}{\kappa}\bar{c}(\varepsilon)P_{t-\varepsilon}\Gamma_1(f)$$

On the other hand we have

$$\int_{t-\varepsilon}^{t} ds 2P_s\Gamma(P_{t-s}f) \leq \frac{2}{\kappa}(1-e^{-\kappa\varepsilon})P_t\Gamma(f) \leq \varepsilon\frac{2}{\kappa}\varepsilon P_t\Gamma(f)$$

Hence, with $\gamma(\varepsilon)$ given as an inverse function of $c(\varepsilon) \equiv \frac{2}{\kappa}\bar{c}(\frac{\kappa}{2}\varepsilon)$, we obtain

$$P_tf^2 - (P_tf)^2 \leq \varepsilon P_{t-\varepsilon}\Gamma_1(f) + \gamma(\varepsilon)P_t\Gamma(f)$$

After passing with time to infinity we obtain

$$\nu(f-\nu f)^2 \leq \varepsilon\nu\Gamma_1(f) + \gamma(\varepsilon)\nu\Gamma(f)$$

which after optimisation with respect to the free parameter ε implies a generalised Nash type inequality.

References

[1] C. Ané, S. Blachère, D. Chafaï, P. Fougères, I. Gentil, F. Malrieu, C. Roberto, G. Scheffer, Sur les inégalités de Sobolev logarithmiques, No. 10 in Panoramas et Synthèses, Soc. Math. France, Paris, 2000.

[2] D. Bakry and M. Emery, Diffusions hypercontractives, pp. 177-206 in Sém. de Probab. XIX, Lecture Notes in Math., Vol. 1123, Springer-Verlag, Berlin, 1985.

[3] D. Bakry, F. Baudoin, M. Bonnefont and D. Chafai, On gradient bounds for the heat kernel on the Heisenberg group, J. Func. Analysis 255 (2008), 1905–1938.

[4] F. Baudoin, M. Hairer, J. Teichmann, Ornstein-Uhlenbeck processes on Lie groups, J. Func. Analysis 255 (2008), 877–890.

[5] F. Baudoin, J. Teichmann, Hypoellipticity in infinite dimensions and an application in interest rate theory, Ann. Appl. Probab. 15 (2005), 1765-1777.

[6] A. Bendikov, L. Saloff-Coste, On the hypoellipticity of sub-Laplacians on infnite dimensional compact groups, Forum Mathematicum 15:1 (2003), 135-163.

[7] A. Bonfiglioli, E. Lanconelli, F. Uguzzoni, Stratified Lie Groups and Potential Theory for their Sub-Laplacians, Springer Monographs in Mathematics, Springer, 2007.

[8] G. Da Prato and J. Zabczyk, Stochastic Equations in Infinite Dimensions. Cambridge Univ. Press (1992).

[9] B.K. Driver, T. Melcher, Hypoelliptic heat kernel inequalities on Lie groups, Stoch. Process. Appl. 118 (2008), 368–388.

[10] B.K. Driver, T. Melcher, Hypoelliptic heat kernel inequalities on the Heisenberg group, J. Func. Analysis 221 (2005), 340 –365.

[11] W.H. Fleming, H.M. Soner, Controlled Markov Processes and Viscosity Solutions, Springer, New York, 1993.

[12] M.I. Freidlin, A.D. Wentzell, Random Perturbations of Dynamical Systems, 2nd Ed., Springer, New York, 1998.

[13] M. Grayson, R. Grossman, Models for Free Nilpotent Lie Algebras , Journal Algebra, 35 (1990), 177–191.

[14] A. Guionnet, B. Zegarliński, Lectures on logarithmic Sobolev inequalities, in: Séminaire de Probabilités, XXXVI, No. 1801 in Lecture Notes in Math., Springer-Verlag, 2003, pp. 1–134.

[15] M. Hairer, Hypoellipticity in Infinite Dimensions, in Proceedings of 7th ISAAC Congress. Imperial College London 2009, World Scientific, 2010.

[16] H.-Q. Li, Estimation optimale du gradient du semi-groupe de la chaleur sur le groupe de Heisenberg, J. Funct. Analysis 236 (2006), 369–394.

[17] W. Hebisch and A. Sikora, A smooth subadditive homogeneous norm on a homogeneous group, Studia Math. 96 (1990), no. 3, 231–236.

[18] W. Hebisch and B. Zegarliński, Coercive inequalities on metric measure spaces, J. Funct. Anal. 258 (2010), 814–851.

[19] J. Inglis, PhD Thesis, Imperial College 2010.

[20] J. Inglis and I. Papageorgiou, Logarithmic Sobolev inequalities for infinite dimensional Hörmander type generators on the Heisenberg group, J. Pot. Anal., 31 (2009), 79–102.

[21] J. Inglis, V. Kontis, B. Zegarliński, From U-bounds to isoperimetry with applications to H-type groups, J. Funct. Anal. 260 (2011), 76–116.

[22] P. Ługiewicz and B. Zegarliński, Coercive inequalities for Hörmander type generators in infinite dimensions, J. Funct. Anal. 247 (2007), 438–476.

[23] T. Melcher, Hypoelliptic heat kernel inequalities on Lie groups, PhD thesis, University of California at San Diego.

[24] C. Villani, Hypocoercivity, vol. 202, Nr.950 of Memoirs of the American Mathematical Society, 2009.

[25] B. Zegarliński, The strong decay to equilibrium for the stochastic dynamics of unbounded spin systems on a lattice, Comm. Math. Phys. 175 (1996), 401-432.