Electronic Journal of Probability

Electron. J. Probab. **25** (2020), article no. 86, 1–46.
ISSN: 1083-6489 https://doi.org/10.1214/20-EJP486

Central limit theorems for non-symmetric random walks on nilpotent covering graphs: Part I

Satoshi Ishiwata* Hiroshi Kawabi† Ryuya Namba‡

Dedicated to the memory of Kazumasa Kuwada

Abstract

In the present paper, we study central limit theorems (CLTs) for non-symmetric random walks on nilpotent covering graphs from a point of view of discrete geometric analysis developed by Kotani and Sunada. We establish a semigroup CLT for a non-symmetric random walk on a nilpotent covering graph. Realizing the nilpotent covering graph into a nilpotent Lie group through a discrete harmonic map, we give a geometric characterization of the limit semigroup on the nilpotent Lie group. More precisely, we show that the limit semigroup is generated by the sub-Laplacian with a non-trivial drift on the nilpotent Lie group equipped with the Albanese metric. The drift term arises from the non-symmetry of the random walk and it vanishes when the random walk is symmetric. Furthermore, by imposing the "centered condition", we establish a functional CLT (i.e., Donsker-type invariance principle) in a Hölder space over the nilpotent Lie group. The functional CLT is extended to the case where the realization is not necessarily harmonic. We also obtain an explicit representation of the limiting diffusion process on the nilpotent Lie group and discuss a relation with rough path theory. Finally, we give an example of random walks on nilpotent covering graphs with explicit computations.

Keywords: central limit theorem; non-symmetric random walk; nilpotent covering graph; discrete geometric analysis; modified harmonic realization; Albanese metric; rough path theory.
AMS MSC 2010: 60F17; 60G50; 60J10; 22E25.
Submitted to EJP on October 25, 2019, final version accepted on June 23, 2020.
Supersedes arXiv:1806.03804.

*Department of Mathematical Sciences, Faculty of Science, Yamagata University, 1-4-12, Kojirakawa, Yamagata 990-8560, Japan. E-mail: ishiwata@sci.kj.yamagata-u.ac.jp

†Department of Mathematics, Hiyoshi Campus, Keio University, 4-1-1, Hiyoshi, Kohoku-ku, Yokohama 223-8521, Japan. E-mail: kawabi@keio.jp

‡Department of Mathematical Sciences, College of Science and Engineering, Ritsumeikan University, 1-1-1, Noji-higashi, Kusatsu, Shiga 525-8577, Japan. E-mail: rnamba@fc.ritsumei.ac.jp

1 Introduction

There are many interests in the study of random walks on infinite graphs in many branches of mathematics such as probability theory, harmonic analysis, geometry, graph theory and group theory. Among these branches, the long time behavior of random walks on infinite graphs is one of the major themes. For instance, a central limit theorem (CLT), that is, a generalization of the Laplace–de Moivre theorem, has been studied intensively and extensively in various settings. These mathematical backgrounds basically motivate our study. For basic results on random walks, we refer to Spitzer [58], Woess [70], Lawler–Limic [40] and references therein.

In these studies of random walks on infinite graphs, many authors have also discussed what kinds of structures of underlying graphs affect the long time behavior of random walks. It is known that geometric structures such as the *periodicity* of underlying graphs play important roles in them (cf. Spitzer [58]). A typical example of periodic infinite graphs is a *crystal lattice*, that is, a covering graph X of a finite graph X_0 whose covering transformation group Γ is abelian. It is a generalization of the square lattice, the triangular lattice, the hexagonal lattice, the dice lattice and so on (see Figure 1). We remark that the crystal lattice has inhomogeneous local structures though it has a

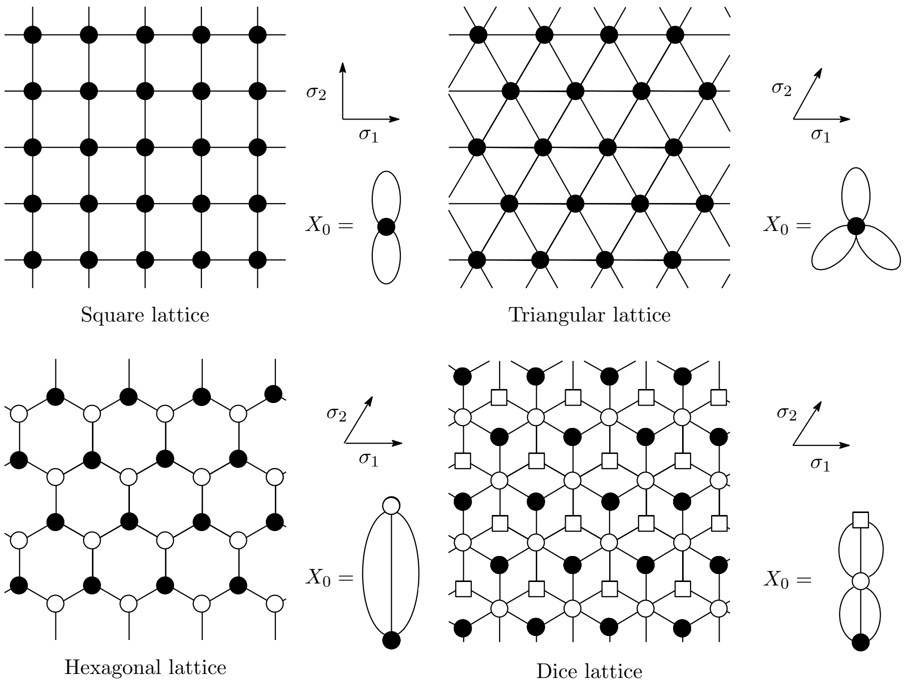

Figure 1: Crystal lattices with the covering transformation group $\Gamma = \langle \sigma_1, \sigma_2 \rangle \cong \mathbb{Z}^2$

periodic global structure. Kotani and Sunada [32] introduced the notion of *standard realization* of a crystal lattice X, which is a discrete harmonic map from X into the Euclidean space $\Gamma \otimes \mathbb{R}$ equipped with the *Albanese metric*, to characterize an equilibrium configuration of X. In a series of papers Kotani–Shirai–Sunada [34], Kotani [29] and Kotani–Sunada [31, 32, 33], they developed a hybrid field of several traditional disciplines including graph theory, geometry, discrete group theory and probability theory. Since this new field, called *discrete geometric analysis*, was introduced by Sunada, it has been making new interactions with many other fields. For example, Le Jan employs discrete geometric analysis effectively in a series of recent studies of Markov loops (see e.g., [41, 42]). We refer to Sunada [63, 64] for recent developments of discrete geometric analysis. Especially, in [31], a geometric characterization of the diffusion semigroup appeared in the CLT-scaling limit of the symmetric random walk on X was given in terms of discrete geometric analysis. Ishiwata, Kawabi and Kotani [25] generalized these results to the non-symmetric case and established two kinds of functional CLTs (i.e., Donsker-type invariance principles) for non-symmetric random walks on crystal lattices. We also refer to Guivar'ch [21] and Kramli–Szasz [36] for related early works, Kotani [30] and Kotani–Sunada [33] for a large deviation principle (LDP) and Namba [49] for yet another functional CLT for non-symmetric random walks on crystal lattices.

On the other hand, long time behaviors of symmetric or centered random walks on groups have been studied intensively and extensively. In particular, the notion of *volume growth* of groups plays a key role in the interface between probability theory and group theory. Generally speaking, it is difficult to characterize a finitely generated group itself in terms of its volume growth. We refer to Saloff-Coste [57] for basic problems and results for random walks on such groups including ones of superpolynomial volume growth. On the contrary, there is a remarkable theorem on a group of polynomial volume growth due to Gromov, which asserts that it is essentially characterized as a nilpotent group (cf. Gromov [20] and Ozawa [51]). Hence, we find a large number of papers on long time behaviors of random walks on nilpotent groups. See e.g., Wehn [69], Tutubalin [67], Stroock–Varadhan [60], Raugi [55], Watkins [68], Pap [52] and Alexopoulos [3] for related results on CLTs on nilpotent Lie groups, and Breuillard [6] for an overview of random walks on Lie groups. We also refer to Alexopoulos [1, 2], Breuillard [7], Diaconis–Hough [13] and Hough [22] for local CLTs on nilpotent Lie groups.

In view of these developments, we study the long time behavior of random walks on a covering graph X whose covering transformation group Γ is a finitely generated group of polynomial volume growth. It is regarded as a generalization of a crystal lattice or the Cayley graph of a finitely generated group of polynomial volume growth. A typical example of such a group is the 3-dimensional discrete Heisenberg group $\Gamma = \mathbb{H}^3(\mathbb{Z})$ (see Figure 2). Thanks to Gromov's theorem mentioned above, Γ has a finite extension of a torsion free nilpotent subgroup $\widetilde{\Gamma} \lhd \Gamma$. Therefore, X is regarded as a covering graph of the finite quotient graph $\widetilde{\Gamma}\backslash X$ whose covering transformation group is $\widetilde{\Gamma}$. Throughout the present paper, we may assume that X is a covering graph of a finite graph X_0 whose covering transformation group Γ is a finitely generated, torsion free nilpotent group of step r ($r \in \mathbb{N}$) without loss of generality. We now mention a few related works. Ishiwata [23] discussed symmetric random walks on nilpotent covering graphs and extended the notion of standard realization of crystal lattices to the nilpotent case. Besides, in [23, 24], a semigroup CLT and a local CLT for symmetric random walks were obtained by realizing the nilpotent covering graph X into a nilpotent Lie group G such that Γ is isomorphic to a cocompact lattice in G (cf. Malcév [48]). We notice that, in spite of such developments, long time behaviors of non-symmetric random walks on nilpotent covering graphs have not been studied sufficiently though an LDP on nilpotent covering graphs was obtained in Tanaka [65].

Under these circumstances, we establish CLTs for non-symmetric random walks on a Γ-nilpotent covering graph X. As an extension of the notion of standard realization introduced in [23] to the non-symmetric case, we define the *modified standard realization* Φ_0 from X into a nilpotent Lie group $G = G_\Gamma$ whose Lie algebra is equipped with the Albanese metric. Through the map Φ_0, we obtain a semigroup CLT (Theorem 2.1), which means that the n-th iteration of the "transition shift operator" converges to a diffusion semigroup on G as $n \to \infty$ with a suitable scale change on G. The infinitesimal generator $-\mathcal{A}$ of the diffusion semigroup is the sub-Laplacian with a non-trivial drift $\beta(\Phi_0)$ affected by the non-symmetry of the given random walk. Furthermore, by imposing an additional natural condition **(A3)**, we prove a functional CLT in a Hölder space over

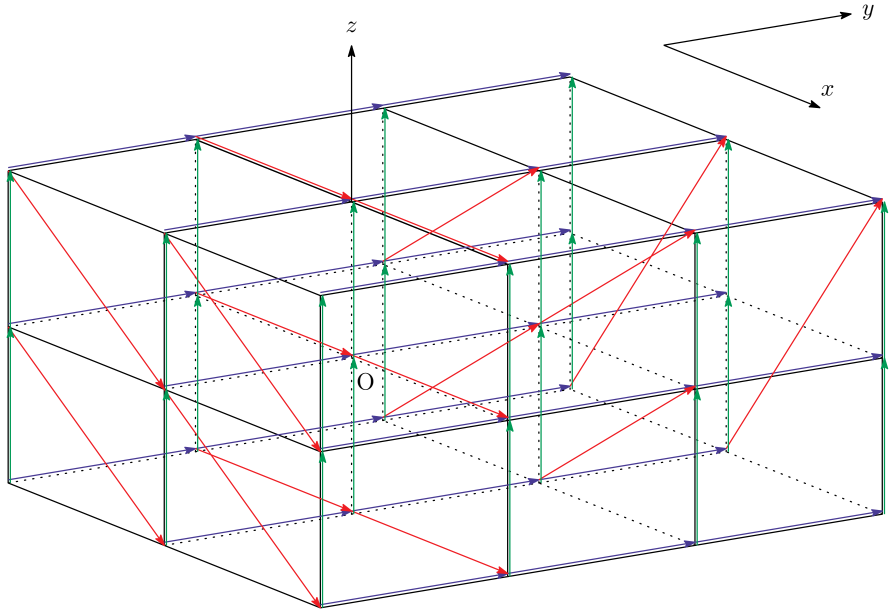

Figure 2: A part of the Cayley graph of $\Gamma = \mathbb{H}^3(\mathbb{Z})$

G (Theorem 2.2). Roughly speaking, we capture a G-valued diffusion process associated with $-\mathcal{A}$ through the CLT-scaling limit of the non-symmetric random walk on X. We call the condition **(A3)** the *centered condition*. The functional CLT is also extended to the case where the realization $\Phi : X \to G$ is not necessarily harmonic (Theorem 2.3) under **(A3)**. In this case, several technical difficulties appear in the proof of the functional CLT. To overcome them,we take a modified harmonic realization $\Phi_0 : X \to G$ and show that the ($\mathfrak{g}^{(1)}$-)*corrector*, the difference between Φ and Φ_0 in the $\mathfrak{g}^{(1)}$-direction, is not so big. This approach is the so-called *corrector method* in the context of stochastic homogenization theory, and it is effectively used in the study of random walks in random environments (see e.g., Papanicolaou–Varadhan [53], Kozlov [35] and Kumagai [37]).We then obtain that a sequence $\{\tau_{n^{-1/2}}(\Phi(w_{[nt]})); 0 \le t \le 1\}_{n=1}^{\infty}$ also converges in law to the diffusion process $(Y_t)_{0\le t\le 1}$ as $n \to \infty$. In a subsequent paper [26], we will consider the *weakly asymmetric case* and establish another kind of CLTs for a family of random walks on the nilpotent covering graph X which interpolates between the original non-symmetric random walk and the symmetrized one. We also capture a G-valued diffusion process different from the one obtained in the present paper. The comparison between these two diffusions will be given in Remark 5.3.

Let us give another motivation of the present paper from rough path theory. It is known that rough path theory was first initiated by Lyons in [46] to discuss line integrals and ordinary differential equations (ODEs) driven by an irregular path such as a sample path of Brownian motion $B = (B_t)_{0\le t\le 1}$ on $\mathbb{R}^d$. Rough path theory makes us possible to handle a Stratonovich type stochastic differential equation (SDE) driven by Brownian motion B as a deterministic ODE driven by standard *Brownian rough path* (i.e., Stratonovich enhanced Brownian motion) $\mathbf{B} = (B, \mathbb{B})$, where $\mathbf{B}$ is a couple of Brownian motion B itself and its Stratonovich iterated integral $\mathbb{B}$. Thus rough path theory provides a new insight to the usual SDE-theory and it has developed rapidly in stochastic analysis. For more details on an overview of rough path theory and its applications to stochastic analysis, see Lyons–Qian [47], Friz–Victoir [18] and Friz–Hairer [15]. In the rough path framework, several authors have studied Donsker-type invariance principles. Among them, Breuillard–Friz–Huesmann [8] first studied this problem for Brownian rough path. Namely, they captured Stratonovich enhanced Brownian motion $\mathbf{B} = (B, \mathbb{B})$ on $\mathbb{R}^d$ as the usual CLT-scaling limit of the natural rough path lift of an $\mathbb{R}^d$-valued random walk with the centered condition. We also refer to Bayer–Friz [4] for applications to cubature and Chevyrev [10] for a recent study on an extension to the case of Lévy processes. Here we should note that there are good approximations to Brownian motion which do not converge to $\mathbf{B}$ but instead to a *distorted Brownian rough path* $\overline{\mathbf{B}} = (B, \mathbb{B} + \boldsymbol{\beta})$, where $\boldsymbol{\beta}$ is an anti-symmetric perturbation of $\mathbb{B}$. For example, Friz–Gassiat–Lyons [14] constructed such a rough path called *magnetic Brownian rough path* as the small mass limit of the natural rough path lift of a physical Brownian motion on $\mathbb{R}^d$ in a magnetic field. Through this approximation, they showed an effect of the magnetic field appears explicitly in the anti-symmetric perturbation term $\boldsymbol{\beta}$. See also e.g., Lejay–Lyons [43] and Friz–Oberhauser [16] for related results on this topic.

In view of the background described above, we discuss a random walk approximation of the distorted Brownian rough path $\overline{\mathbf{B}}$ from a perspective of discrete geometric analysis. Since the unique *Lyons extension* of $\overline{\mathbf{B}}$ of order r ($r \geq 2$) can be regarded as a diffusion process on a free step-r nilpotent Lie group $\mathbb{G}^{(r)}(\mathbb{R}^d)$ (see Section 5 below for definition), we obtain such a diffusion process in Corollary 5.5 through the CLT-scaling limit of a non-symmetric random walk on a nilpotent covering graph X as a direct application of Theorem 2.2. Besides, we observe that the non-symmetry of the random walk on X affects the anti-symmetric perturbation term of $\overline{\mathbb{B}}$ explicitly. Recently, Lopusanschi–Simon [45] proved a similar invariance principle for $\overline{\mathbf{B}}$ to ours. However, they did not discuss an explicit relation between the perturbation term, called the *area anomaly*, and the non-symmetry of the given random walk. See also Lopusanschi–Orenshtein [44] and Deuschel–Orenshtein–Perkowski [12] for related results. In view of that, Corollary 5.5 gives a new approach to such an invariance principle in that we pay much attention to the non-symmetry of random walks on X.

The rest of the present paper is organized as follows: We introduce our framework and state the main results in Section 2. We make a preparation from nilpotent Lie groups, the Carnot–Carathéodory metric, homogeneous norms and discrete geometric analysis in Section 3. A relation between the G-valued random walk and the notion of modified harmonicity is also discussed. In Section 4.1, we give a brief outline of the proof of main results through a simple example. In Section 4.2, we prove the first main result (Theorem 2.1) and give several properties of the non-trivial drift $\beta(\Phi_0)$ (Proposition 4.4). Trotter's approximation theorem plays a crucial role in the proof of Theorem 2.1. We then prove a functional CLT (Theorem 2.2) for the non-symmetric random walk under the centered condition **(A3)** in Section 4.3. We show the tightness of the family of probability measures induced by the G-valued stochastic processes given by the geodesic interpolation of the given random walk (Lemma 4.5). In the case $r = 2$, we prove it by combining the modified harmonicity of Φ_0 with standard martingale techniques. On the other hand, the same argument is insufficient in the case $r \geq 3$. To handle the higher-step terms, we employ a novel pathwise argument inspired by the proof of the Lyons extension theorem (cf. Lyons [46]) in rough path theory. However, we need a careful examination of the proof of Lyons' extension theorem since rough path theory is build on free nilpotent Lie groups and our nilpotent Lie group G is not necessarily free. As a consequence of Theorem 2.1, the convergence of the finite dimensional distribution of the stochastic process (Lemma 4.8) is proved. Moreover, in Section 4.4, a functional CLT in the case where the realization is non-harmonic (Theorem 2.3) is also proved by applying the corrector method described above. An explicit representation of the limiting diffusion process is given in Section 5. We also discuss a relation between this diffusion process and rough path theory by using this representation formula in the case where Γ is the free discrete nilpotent group over $\mathbb{Z}^d$. We give two examples of non-symmetric random walks on nilpotent covering graphs with explicit calculations in Section 6. Finally, we give a comment on another approach to CLTs in the non-centered case in Appendix A (see Theorems A.2 and A.3).

Throughout the present paper, C denotes a positive constant that may change from line to line and $O(\cdot)$ stands for the Landau symbol. If the dependence of C and $O(\cdot)$ are significant, we denote them like $C(N)$ and $O_N(\cdot)$, respectively.

2 Framework and results

We introduce our framework and state the main results in this section. Let Γ be a torsion free, finitely generated nilpotent group and $X = (V, E)$ a Γ-nilpotent covering graph, where V is the set of all vertices and E is the set of all oriented edges. The graph X possibly have multiple edges or loops and is equipped with the discrete topology induced by the graph distance. For an edge $e \in E$, we denote by $o(e)$ and $t(e)$ the origin and the terminus of e, respectively. The inverse edge of $e \in E$ is defined by an edge, say $\overline{e}$, satisfying $o(\overline{e}) = t(e)$ and $t(\overline{e}) = o(e)$. We set $E_x = \{e \in E \,|\, o(e) = x\}$ for $x \in V$. A path c in X of length n is a sequence $c = (e_1, e_2, \dots, e_n)$ of n edges $e_1, e_2, \dots, e_n \in E$ with $o(e_{i+1}) = t(e_i)$ for $i = 1, 2, \dots, n-1$. We denote by $\Omega_{x,n}(X)$ the set of all paths in X of length $n \in \mathbb{N} \cup \{\infty\}$ starting from $x \in V$. Put $\Omega_x(X) = \Omega_{x,\infty}(X)$ for simplicity.

We introduce a *transition probability*, that is, a function $p : E \to [0, 1]$ satisfying

$$\sum_{e \in E_x} p(e) = 1 \qquad (x \in V) \quad \text{and} \quad p(e) + p(\overline{e}) > 0 \qquad (e \in E).$$

Moreover, we impose that p is invariant under the Γ-action, that is, $p(\gamma e) = p(e)$ for $\gamma \in \Gamma$ and $e \in E$. The random walk associated with p is the X-valued time-homogeneous Markov chain $(\Omega_x(X), \mathbb{P}_x, \{w_n\}_{n=0}^{\infty})$, where $\mathbb{P}_x$ is the probability measure on $\Omega_x(X)$ satisfying

$$\mathbb{P}_x\big(\{c = (e_1, e_2, \dots, e_n, *, *, \dots)\}\big) = p(e_1)p(e_2)\cdots p(e_n) \qquad \big(c \in \Omega_x(X)\big)$$

and $w_n(c) := o(e_{n+1})$ for $n \in \mathbb{N} \cup \{0\}$ and $c = (e_1, e_2, \dots, e_n, \dots) \in \Omega_x(X)$.

We define the *transition operator* L associated with the transition probability p by

$$Lf(x) := \sum_{e \in E_x} p(e) f\big(t(e)\big) \qquad (x \in V,\ f : V \to \mathbb{R})$$

and the n-step transition probability $p(n, x, y)$ by $p(n, x, y) := L^n \delta_y(x)$ for $n \in \mathbb{N}$ and $x, y \in V$, where δ_y stands for the Dirac delta function at y and $p(c) = p(e_1)p(e_2)\cdots p(e_n)$ for $c = (e_1, e_2, \dots, e_n) \in \Omega_{x,n}(X)$. Let $X_0 = (V_0, E_0) = \Gamma\backslash X$ be the finite quotient graph.Then the random walk on X_0 is induced through the covering map $\pi : X \to X_0$. We write $p : E_0 \to [0, 1]$ the transition probability on X_0, by abuse of notation. For $n \in \mathbb{N}$ and $x, y \in V_0$, we also denote by $p(n, x, y)$ the n-step transition probability of the random walk on X_0.

Throughout the present paper, we impose the following two conditions.

(A1): The random walk on X_0 is *irreducible*. Namely, for $x, y \in V_0$, there exists $n = n(x, y) \in \mathbb{N}$ such that $p(n, x, y) > 0$.

(A2): There exists some $e_* \in E_0$ such that $p(e_*) > 0$ and $p(\overline{e}_*) > 0$.

We mention that the condition **(A2)** is called the *mixed traffic condition* in Sunada [61]. Thanks to **(A1)** and the Perron–Frobenius theorem, we find a unique positive function

$m : V_0 \to (0,1]$ which is called the *invariant measure* on X_0 satisfying

$$\sum_{x \in V_0} m(x) = 1 \quad \text{and} \quad m(x) = \sum_{e \in (E_0)_x} p(\overline{e}) m\bigl(t(e)\bigr) \qquad (x \in V_0).$$

We set $\widetilde{m}(e) := p(e)m\bigl(o(e)\bigr)$ for $e \in E_0$. The random walk on X_0 is called (m-)*symmetric* if $\widetilde{m}(e) = \widetilde{m}(\overline{e})$ for $e \in E_0$. Otherwise, it is called (m-)*non-symmetric*. We also write $m : V \to (0,1]$ for the Γ-invariant lift of $m : V_0 \to (0,1]$. We denote by $\mathrm{H}_1(X_0, \mathbb{R})$ and $\mathrm{H}^1(X_0, \mathbb{R})$ the first homology group and the first cohomology group of X_0, respectively. We define the *homological direction* of the given random walk on X_0 by

$$\gamma_p := \sum_{e \in E_0} \widetilde{m}(e) e \in \mathrm{H}_1(X_0, \mathbb{R}).$$

It is clear that the random walk on X_0 is (m-)symmetric if and only if $\gamma_p = 0$. In this sense, γ_p gives the homological drift of given random walk on X_0.

On the other hand, we provide a continuous state space in which the Γ-nilpotent covering graph X is properly realized. There exists a connected and simply connected nilpotent Lie group $(G, \cdot) = G_\Gamma$ such that Γ is isomorphic to a cocompact lattice in G by applying Malcév's theorem (cf. Malcév [48]). A piecewise smooth Γ-equivariant map $\Phi : X \to G$ is called a *periodic realization* of X. Let $(\mathfrak{g}, [\cdot,\cdot])$ be the Lie algebra of G. Since the exponential map $\exp : \mathfrak{g} \to G$ is a diffeomorphism, global coordinate systems on G are induced through the exponential map. We write $\log : G \to \mathfrak{g}$ for the inverse map of $\exp : \mathfrak{g} \to G$.

We construct a new product $*$ on G in the following manner. Set $\mathfrak{n}_1 := \mathfrak{g}$ and $\mathfrak{n}_{k+1} := [\mathfrak{g}, \mathfrak{n}_k]$ for $k \in \mathbb{N}$. Since $\mathfrak{g}$ is nilpotent, we find an integer $r \in \mathbb{N}$ such that $\mathfrak{g} = \mathfrak{n}_1 \supset \cdots \supset \mathfrak{n}_r \supsetneq \mathfrak{n}_{r+1} = \{\mathbf{0}_\mathfrak{g}\}$. The integer r is called the *step number* of $\mathfrak{g}$ or G. We define the subspace $\mathfrak{g}^{(k)}$ of $\mathfrak{g}$ by $\mathfrak{n}_k = \mathfrak{g}^{(k)} \oplus \mathfrak{n}_{k+1}$ for $k = 1, 2, \dots, r$. Then the Lie algebra $\mathfrak{g}$ is decomposed as $\mathfrak{g} = \mathfrak{g}^{(1)} \oplus \mathfrak{g}^{(2)} \oplus \cdots \oplus \mathfrak{g}^{(r)}$ and each $Z \in \mathfrak{g}$ is uniquely written as $Z = Z^{(1)} + Z^{(2)} + \cdots + Z^{(r)}$, where $Z^{(k)} \in \mathfrak{g}^{(k)}$ for $k = 1, 2, \dots, r$. Define a map $\tau_\varepsilon^{(\mathfrak{g})} : \mathfrak{g} \to \mathfrak{g}$ by

$$\tau_\varepsilon^{(\mathfrak{g})}(Z) := \varepsilon Z^{(1)} + \varepsilon^2 Z^{(2)} + \cdots + \varepsilon^r Z^{(r)} \qquad (\varepsilon \geq 0,\ Z \in \mathfrak{g})$$

and also define a Lie bracket product $[\![\cdot,\cdot]\!]$ on $\mathfrak{g}$ by

$$[\![Z_1, Z_2]\!] := \lim_{\varepsilon \searrow 0} \tau_\varepsilon^{(\mathfrak{g})} \bigl[\tau_{1/\varepsilon}^{(\mathfrak{g})}(Z_1), \tau_{1/\varepsilon}^{(\mathfrak{g})}(Z_2)\bigr] \qquad (Z_1, Z_2 \in \mathfrak{g}).$$

We introduce a map $\tau_\varepsilon : G \to G$, called the *dilation operator* on G, by

$$\tau_\varepsilon(g) := \exp\bigl(\tau_\varepsilon^{(\mathfrak{g})}\bigl(\log(g)\bigr)\bigr) \qquad (\varepsilon \geq 0,\ g \in G),$$

which gives scalar multiplications on G. We note that τ_ε may not be a group homomorphism, though it is a diffeomorphism on G. By making use of the dilation map τ_ε, a Lie group product $*$ on G is defined as follows:

$$g * h := \lim_{\varepsilon \searrow 0} \tau_\varepsilon\bigl(\tau_{1/\varepsilon}(g) \cdot \tau_{1/\varepsilon}(h)\bigr) \qquad (g, h \in G).$$

The Lie group $G_\infty = (G, *)$ is called the *limit group* of $(G, \cdot)$. It is a *stratified Lie group* of step r in the sense that $(\mathfrak{g}, [\![\cdot,\cdot]\!])$ is decomposed as $\mathfrak{g} = \bigoplus_{k=1}^r \mathfrak{g}^{(k)}$ satisfying $[\![\mathfrak{g}^{(k)}, \mathfrak{g}^{(\ell)}]\!] \subset \mathfrak{g}^{(k+\ell)}$ unless $k + j > r$ and the subspace $\mathfrak{g}^{(1)}$ generates $\mathfrak{g}$. The relation between these two Lie group products is given in the next section. We endow G with the so-called *Carnot–Carathéodory metric* d_{CC}, which is an intrinsic metric defined by

$$d_{\mathrm{CC}}(g, h) := \inf\left\{ \int_0^1 \|\dot{w}_t\|_{g_0}\, dt \,\middle|\, \begin{array}{c} w \in \mathrm{Lip}([0,1]; G),\ w_0 = g,\ w_1 = h, \\ w \text{ is tangent to } \mathfrak{g}^{(1)} \end{array} \right\} \tag{2.1}$$

for $g, h \in G$, where we write $\mathrm{Lip}([0,1]; G)$ for the set of all Lipschitz continuous paths and $\|\cdot\|_{g_0}$ for the norm on $\mathfrak{g}^{(1)}$ induced by the Albanese metric (see Section 3 for details).

Let $\pi_1(X_0)$ be the fundamental group of X_0. Then we have a canonical surjective homomorphism $\rho : \pi_1(X_0) \to \Gamma$ by the general theory of covering spaces. This map gives rise to a surjective homomorphism $\rho : \mathrm{H}_1(X_0, \mathbb{Z}) \to \Gamma/[\Gamma, \Gamma]$ and we have a surjective linear map $\rho_{\mathbb{R}} : \mathrm{H}_1(X_0, \mathbb{R}) \to \mathfrak{g}^{(1)}$ by extending it linearly. We call $\rho_{\mathbb{R}}(\gamma_p) \in \mathfrak{g}^{(1)}$ the *asymptotic direction*. Note that $\gamma_p = 0$ implies $\rho_{\mathbb{R}}(\gamma_p) = \mathbf{0}_{\mathfrak{g}}$. However, the converse does not always hold. We induce a special flat metric g_0 on $\mathfrak{g}^{(1)}$, which is called the *Albanese metric* associated with the transition probability p by using the discrete Hodge-Kodaira theorem (cf. Kotani–Sunada [33, Lemma 5.2]). The construction of the Albanese metric is given in the next section. A periodic realization $\Phi_0 : X \to G$ is called (p-)*modified harmonic* if

$$\sum_{e \in E_x} p(e) \log\Big(\Phi_0\big(o(e)\big)^{-1} \cdot \Phi_0\big(t(e)\big)\Big)\Big|_{\mathfrak{g}^{(1)}} = \rho_{\mathbb{R}}(\gamma_p) \qquad (x \in V). \tag{2.2}$$

Such Φ_0 is uniquely determined up to $\mathfrak{g}^{(1)}$-translation. The modified harmonicity describes the most natural realization of the nilpotent covering graph X in the geometric point of view. If we equip $\mathfrak{g}^{(1)}$ with the Albanese metric g_0, the modified harmonic realization $\Phi_0 : X \to G$ is called the *modified standard realization*.

For a metric space $\mathcal{T}$, we denote by $C_\infty(\mathcal{T})$ the Banach space of continuous functions $f : \mathcal{T} \to \mathbb{R}$ vanishing at infinity with the uniform topology $\|\cdot\|_\infty^{\mathcal{T}}$. For $q > 1$, we define

$$C_{\infty,q}(X \times \mathbb{Z}) := \big\{f = f(x,z) : X \times \mathbb{Z} \to \mathbb{R} \,\big|\, f(\cdot, z) \in C_\infty(X),\, \|f\|_{\infty,q} < \infty\big\},$$

where $\|f\|_{\infty,q}$ is a norm on $C_{\infty,q}(X \times \mathbb{Z})$ given by

$$\|f\|_{\infty,q} := \frac{1}{C_q} \sum_{z \in \mathbb{Z}} \frac{\|f(\cdot, z)\|_\infty^X}{1 + |z|^q}, \qquad C_q := \sum_{z \in \mathbb{Z}} \frac{1}{1 + |z|^q} < \infty.$$

Then we see that $(C_{\infty,q}(X \times \mathbb{Z}), \|\cdot\|_{\infty,q})$ is a Banach space. We introduce the *transition-shift operator* $\mathcal{L}_p : C_{\infty,q}(X \times \mathbb{Z}) \to C_{\infty,q}(X \times \mathbb{Z})$ by

$$\mathcal{L}_p f(x,z) := \sum_{e \in E_x} p(e) f\big(t(e), z+1\big) \qquad (x \in V,\, z \in \mathbb{Z}) \tag{2.3}$$

and the *approximation operator* $\mathcal{P}_\varepsilon : C_\infty(G) \to C_{\infty,q}(X \times \mathbb{Z})$ by

$$\mathcal{P}_\varepsilon f(x,z) := f\Big(\tau_\varepsilon\big(\Phi_0(x) * \exp(-z\rho_{\mathbb{R}}(\gamma_p))\big)\Big) \qquad (0 \le \varepsilon \le 1,\, x \in V,\, z \in \mathbb{Z}). \tag{2.4}$$

We extend each $Z \in \mathfrak{g}$ as a left invariant vector field Z_* on G as follows:

$$Z_* f(g) = \frac{d}{d\varepsilon}\Big|_{\varepsilon=0} f\big(g * \exp(\varepsilon Z)\big) \qquad \big(f \in C^\infty(G),\, g \in G\big).$$

We put

$$\beta(\Phi_0) := \sum_{e \in E_0} \widetilde{m}(e) \log\Big(\Phi_0\big(o(\widetilde{e})\big)^{-1} \cdot \Phi_0\big(t(\widetilde{e})\big) \cdot \exp(-\rho_{\mathbb{R}}(\gamma_p))\Big)\Big|_{\mathfrak{g}^{(2)}}, \tag{2.5}$$

where $\widetilde{e}$ stands for a lift of $e \in E_0$ to X. We note that $\gamma_p = 0$ implies $\beta(\Phi_0) = \mathbf{0}_{\mathfrak{g}}$. However, even if $\rho_{\mathbb{R}}(\gamma_p) = \mathbf{0}_{\mathfrak{g}}$, the quantity $\beta(\Phi_0)$ does not vanish in general. Furthermore, $\beta(\Phi_0)$ does not depend on $\mathfrak{g}^{(2)}$-components of the modified harmonic realization $\Phi_0 : X \to G$, though it has the ambiguity in the components corresponding to $\mathfrak{g}^{(2)} \oplus \mathfrak{g}^{(3)} \oplus \cdots \oplus \mathfrak{g}^{(r)}$. See Proposition 4.4 for details and Section 6 for a concrete example.

Then the first main result is as follows:

Theorem 2.1. For $q > 4r+1$, the following hold:
(1) For $0 \le s \le t$ and $f \in C_\infty(G)$, we have

$$\lim_{n\to\infty} \Big\| \mathcal{L}_p^{[nt]-[ns]} \mathcal{P}_{n^{-1/2}} f - \mathcal{P}_{n^{-1/2}} \mathrm{e}^{-(t-s)\mathcal{A}} f \Big\|_{\infty,q} = 0, \tag{2.6}$$

where $(\mathrm{e}^{-t\mathcal{A}})_{t\ge 0}$ is the C_0-semigroup with the infinitesimal generator $\mathcal{A}$ on $C_0^\infty(G)$ defined by

$$\mathcal{A} := -\frac{1}{2}\sum_{i=1}^{d_1} V_{i*}^2 - \beta(\Phi_0)_*, \tag{2.7}$$

where $\{V_1, V_2, \dots, V_{d_1}\}$ denotes an orthonormal basis of $(\mathfrak{g}^{(1)}, g_0)$.

(2) Let μ be a Haar measure on G. Fix $z \in \mathbb{Z}$. Then, for any sequence $\{x_n\}_{n=1}^\infty \subset V$ with

$$\lim_{n\to\infty} \tau_{n^{-1/2}} \Big(\Phi_0(x_n) * \exp\big(-z\rho_{\mathbb{R}}(\gamma_p)\big)\Big) = g \in G$$

and for any $f \in C_\infty(G)$, we have

$$\lim_{n\to\infty} \mathcal{L}_p^{[nt]} \mathcal{P}_{n^{-1/2}} f(x_n, z) = \mathrm{e}^{-t\mathcal{A}} f(g) := \int_G \mathcal{H}_t(h^{-1} * g) f(h)\, \mu(dh) \qquad (t>0), \tag{2.8}$$

where $\mathcal{H}_t(g)$ is a fundamental solution to $\partial u/\partial t + \mathcal{A}u = 0$.

Fix a reference point $x_* \in V$ with $\Phi_0(x_*) = \mathbf{1}_G$ and put $\xi_n(c) := \Phi_0\big(w_n(c)\big)$ for $n \in \mathbb{N}\cup\{0\}$ and $c \in \Omega_{x_*}(X)$. We then have a G-valued random walk $(\Omega_{x_*}(X), \mathbb{P}_{x_*}, \{\xi_n\}_{n=0}^\infty)$ starting from $\mathbf{1}_G$. For $t \ge 0$, we define a map $\mathcal{X}_t^{(n)} : \Omega_{x_*}(X) \to G$ by

$$\mathcal{X}_t^{(n)}(c) := \tau_{n^{-1/2}} \Big(\xi_{[nt]}(c) * \exp\big(-[nt]\rho_{\mathbb{R}}(\gamma_p)\big)\Big) \qquad \big(n \in \mathbb{N},\, c \in \Omega_{x_*}(X)\big).$$

Denote by $\mathcal{D}_n$ the partition $\{t_k = k/n \,|\, k = 0, 1, \dots, n\}$ of $[0,1]$ for $n \in \mathbb{N}$. We define a G-valued continuous stochastic process $(\mathcal{Y}_t^{(n)})_{0\le t\le 1}$ by the geodesic interpolation of $\{\mathcal{X}_{t_k}^{(n)}\}_{k=0}^n$ with respect to d_{CC}. It is worth noting that (2.8) implies

$$\lim_{n\to\infty} \sum_{c \in \Omega_{x_*}(X)} p(c) f\big(\mathcal{X}_t^{(n)}(c)\big) = \int_G \mathcal{H}_t(h^{-1}) f(h)\, \mu(dh) \qquad \big(f \in C_\infty(G)\big). \tag{2.9}$$

We now consider an SDE

$$dY_t = \sum_{i=1}^{d_1} V_{i*}(Y_t) \circ dB_t^i + \beta(\Phi_0)_*(Y_t)\, dt, \qquad Y_0 = \mathbf{1}_G, \tag{2.10}$$

where $(B_t)_{0\le t\le 1} = (B_t^1, B_t^2, \dots, B_t^{d_1})_{0\le t\le 1}$ is an $\mathbb{R}^{d_1}$-valued standard Brownian motion with $B_0 = \mathbf{0}$. Let $(Y_t)_{0\le t\le 1}$ be the G-valued diffusion process which solves (2.10). In Proposition 5.4 below, we prove that the infinitesimal generator of $(Y_t)_{0\le t\le 1}$ coincides with $-\mathcal{A}$ defined by (2.7). Let $C_{\mathbf{1}_G}([0,1]; G)$ be the set of all continuous paths $w : [0,1] \to G$ such that $w_0 = \mathbf{1}_G$ and $\mathrm{Lip}([0,1]; G) \subset C_{\mathbf{1}_G}([0,1]; G)$ the set of all Lipschitz continuous paths. For $\alpha < 1/2$, we define the α-Hölder distance ρ_α on $C_{\mathbf{1}_G}([0,1]; G)$ by

$$\rho_\alpha(w^1, w^2) := \sup_{0\le s<t\le 1} \frac{d_{\mathrm{CC}}(u_s, u_t)}{|t-s|^\alpha}, \qquad u_t := (w_t^1)^{-1} * w_t^2 \qquad (0 \le t \le 1).$$

We set $C_{\mathbf{1}_G}^{0,\alpha\text{-Höl}}([0,1]; G) := \overline{\mathrm{Lip}([0,1]; G)}^{\rho_\alpha}$, which is a Polish space (cf. Friz–Victoir [18, Section 8]). Let $\mathbf{P}^{(n)}$ be the image measure on $C_{\mathbf{1}_G}^{0,\alpha\text{-Höl}}([0,1]; G)$ induced by $\mathcal{Y}_\cdot^{(n)}$ for $n \in \mathbb{N}$.

We now in a position to present a functional CLT, the second main theorem, for the non-symmetric random walk $\{w_n\}_{n=0}^\infty$ on X.

Theorem 2.2. We assume the centered condition **(A3):** $\rho_{\mathbb{R}}(\gamma_p) = \mathbf{0}_{\mathfrak{g}}$. Then the sequence $(\mathcal{Y}_t^{(n)})_{0\le t\le 1}$ converges in law to the G-valued diffusion process $(Y_t)_{0\le t\le 1}$ in $C^{0,\alpha\text{-Höl}}_{\mathbf{1}_G}([0,1];G)$ as $n\to\infty$ for all $\alpha<1/2$.

Finally, we generalize Theorem 2.2 to the case where the realization is not necessarily modified harmonic. We take a periodic realizations $\Phi : X \to G$ such that $\Phi(x_*) = \mathbf{1}_G$ for some base point $x_* \in V$. On the other hand, we may take the modified harmonic realization $\Phi_0 : X \to G$ such that $\Phi_0(x)^{(i)} = \Phi(x)^{(i)}$ for $x \in V$ and $i = 2, 3, \dots, r$ without loss of generality. We now define the ($\mathfrak{g}^{(1)}$-)*corrector* $\mathrm{Cor}_{\mathfrak{g}^{(1)}} : X \to \mathfrak{g}^{(1)}$ by

$$\mathrm{Cor}_{\mathfrak{g}^{(1)}}(x) := \log\big(\Phi(x)\big)\big|_{\mathfrak{g}^{(1)}} - \log\big(\Phi_0(x)\big)\big|_{\mathfrak{g}^{(1)}} \qquad (x\in V). \tag{2.11}$$

By periodicities of Φ and Φ_0, we easily see that the set $\{\mathrm{Cor}_{\mathfrak{g}^{(1)}}(x) \,|\, x\in V\}$ is finite. In particular, we find a positive constant $M>0$ such that $\max_{x\in\mathcal{F}} \|\mathrm{Cor}_{\mathfrak{g}^{(1)}}(x)\|_{\mathfrak{g}^{(1)}} \le M$.

Let $(\overline{\mathcal{Y}}_t^{(n)})_{0\le t\le 1}$ ($n\in\mathbb{N}$) be the G-valued stochastic processes defined by just replacing Φ_0 by Φ in the definition of $(\mathcal{Y}_t^{(n)})_{0\le t\le 1}$. Thanks to several properties of $\mathrm{Cor}_{\mathfrak{g}^{(1)}}$, we establish the following functional CLT. Note that the information of the modified harmonic realization Φ_0 still remains in the drift term of the limiting diffusion even if we replace $\mathcal{Y}^{(n)}$ by $\overline{\mathcal{Y}}^{(n)}$.

Theorem 2.3. Assume the centered condition **(A3)**. The sequence $\{\overline{\mathcal{Y}}_t^{(n)}\}_{n=1}^{\infty}$ converges in law to the G-valued diffusion process $(Y_t)_{0\le t\le 1}$ in $C^{0,\alpha\text{-Höl}}_{\mathbf{1}_G}([0,1];G)$ as $n\to\infty$ for $\alpha<1/2$.

Let us make comments on our main theorems. As is emphasized in Breuillard [6, Section 6], the situation of the non-centered case $\rho_{\mathbb{R}}(\gamma_p) \ne \mathbf{0}_{\mathfrak{g}}$ is quite different from the centered case $\rho_{\mathbb{R}}(\gamma_p) = \mathbf{0}_{\mathfrak{g}}$ and thus some technical difficulties arise to obtain CLTs. That is why there are few papers which discuss CLTs for non-centered random walks on nilpotent Lie groups. We obtain, in Theorem 2.1, a semigroup CLT for the non-centered random walk $\{\xi_n\}_{n=0}^{\infty}$ on G with a canonical dilation $\tau_{n^{-1/2}}$, while Crépel–Raugi [11] and Raugi [55] proved similar CLTs for the random walk to (2.9) with spatial scalings whose orders are higher than $\tau_{n^{-1/2}}$. On the other hand, in the present paper, we need to assume the centered condition **(A3)** to obtain a functional CLT (Theorem 2.2) for $\{\xi_n\}_{n=0}^{\infty}$ in the Hölder topology, stronger than the uniform topology in $C_{\mathbf{1}_G}([0,1];G)$. In Appendix A, we mention a method to reduce the non-centered case $\rho_{\mathbb{R}}(\gamma_p) \ne \mathbf{0}_{\mathfrak{g}}$ to the centered case by employing a measure-change technique based on Alexopoulos [2].

3 Preparations

3.1 Limit groups

Let us review some properties of the limit group. For more details, see e.g., Alexopoulos [1] and Ishiwata [23]. We also refer to Crépel–Raugi [11] and Goodman [19] for related topics. Let $(G,\cdot)$ be a connected and simply connected nilpotent Lie group of step r and $(\mathfrak{g},[\cdot,\cdot])$ the corresponding Lie algebra. Then the limit group $G_\infty = (G,*)$ of $(G,\cdot)$ is a stratified Lie group of step r and its Lie algebra coincides with $(\mathfrak{g},[\![\cdot,\cdot]\!])$. Namely, the Lie algebra $\mathfrak{g} = \mathfrak{g}^{(1)}\oplus\mathfrak{g}^{(2)}\oplus\cdots\oplus\mathfrak{g}^{(r)}$ satisfies that $[\![\mathfrak{g}^{(i)},\mathfrak{g}^{(j)}]\!] \subset \mathfrak{g}^{(i+j)}$ whenever $i+j\le r$ and the subspace $\mathfrak{g}^{(1)}$ generates $\mathfrak{g}$. It should be noted that the dilation map $\tau_\varepsilon : G\to G$ is a group automorphism on $(G,*)$ (see [23, Lemma 2.1]). We also note that the exponential map $\exp : \mathfrak{g}_\infty \to G_\infty$ coincides with the original exponential map $\exp : \mathfrak{g}\to G$. Furthermore, for any $g\in G$, the inverse element of g in $(G,\cdot)$ coincides with the inverse element in $(G,*)$.

We set $d_k = \dim_{\mathbb{R}}\mathfrak{g}^{(k)}$ for $k=1,2,\dots,r$ and $N = d_1+d_2+\cdots+d_r$. For $k = 1,2,\dots,r$, we denote by $\{X_1^{(k)}, X_2^{(k)},\dots,X_{d_k}^{(k)}\}$ a basis of the subspace $\mathfrak{g}^{(k)}$. We introduce several kinds of global coordinate systems in G through $\exp : \mathfrak{g} \to G$. We write $g^{(k)} = (g_1^{(k)}, g_2^{(k)}, \dots, g_{d_k}^{(k)}) \in \mathbb{R}^{d_k}$ for $k = 1, 2, \dots, r$. We identify the nilpotent Lie group G with $\mathbb{R}^N$ as a differentiable manifold by

• *canonical* $(\cdot)$*-coordinates of the first kind*:

$$\mathbb{R}^N \ni (g^{(1)}, g^{(2)}, \dots, g^{(r)}) \longmapsto g = \exp\Big(\sum_{k=1}^{r}\sum_{i=1}^{d_k} g_i^{(k)} X_i^{(k)}\Big) \in G,$$

• *canonical* $(\cdot)$*-coordinates of the second kind*:

$$\begin{aligned}\mathbb{R}^N \ni &(g^{(1)}, g^{(2)}, \dots, g^{(r)})\\ \longmapsto g &= \exp\big(g_{d_r}^{(r)} X_{d_r}^{(r)}\big) \cdot \exp\big(g_{d_r-1}^{(r)} X_{d_r-1}^{(r)}\big) \cdot \dots \cdot \exp\big(g_1^{(r)} X_1^{(r)}\big)\\ &\cdot \exp\big(g_{d_{r-1}}^{(r-1)} X_{d_{r-1}}^{(r-1)}\big) \cdot \exp\big(g_{d_{r-1}-1}^{(r-1)} X_{d_{r-1}-1}^{(r-1)}\big) \cdot \dots \cdot \exp\big(g_1^{(r-1)} X_1^{(r-1)}\big)\\ &\cdots \exp\big(g_{d_1}^{(1)} X_{d_1}^{(1)}\big) \cdot \exp\big(g_{d_1-1}^{(1)} X_{d_1-1}^{(1)}\big) \cdot \dots \cdot \exp\big(g_1^{(1)} X_1^{(1)}\big) \in G,\end{aligned}$$

• *canonical* $(*)$*-coordinates of the second kind*:

$$\begin{aligned}\mathbb{R}^N \ni &(g_*^{(1)}, g_*^{(2)}, \dots, g_*^{(r)})\\ \longmapsto g &= \exp\big(g_{d_r*}^{(r)} X_{d_r}^{(r)}\big) * \exp\big(g_{d_r-1*}^{(r)} X_{d_r-1}^{(r)}\big) * \dots * \exp\big(g_{1*}^{(r)} X_1^{(r)}\big)\\ &* \exp\big(g_{d_{r-1}*}^{(r-1)} X_{d_{r-1}}^{(r-1)}\big) * \exp\big(g_{d_{r-1}-1*}^{(r-1)} X_{d_{r-1}-1}^{(r-1)}\big) * \dots * \exp\big(g_{1*}^{(r-1)} X_1^{(r-1)}\big)\\ &* \dots * \exp\big(g_{d_1*}^{(1)} X_{d_1}^{(1)}\big) * \exp\big(g_{d_1-1*}^{(1)} X_{d_1-1}^{(1)}\big) * \dots * \exp\big(g_{1*}^{(1)} X_1^{(1)}\big) \in G_\infty.\end{aligned}$$

We give the relations between the deformed product and the given product on G as an easy application of the Campbell–Baker–Hausdorff (CBH) formula

$$\log\big(\exp(Z_1)\cdot\exp(Z_2)\big) = Z_1 + Z_2 + \frac{1}{2}[Z_1, Z_2] + \cdots \qquad (Z_1, Z_2 \in \mathfrak{g}). \tag{3.1}$$

The following is straightforward from the definition of the deformed product.

$$\log(g * h)\big|_{\mathfrak{g}^{(k)}} = \log(g \cdot h)\big|_{\mathfrak{g}^{(k)}} \qquad (g, h \in G,\ k = 1, 2). \tag{3.2}$$

We notice that the relation above does not hold in general for $k = 3, 4, \dots, r$. The following identities give us a comparison between $(\cdot)$-coordinates and $(*)$-coordinates. For $g \in G$, we have the following.

$$g_{i*}^{(k)} = g_i^{(k)} \quad (i = 1, 2, \dots, d_k,\ k = 1, 2), \tag{3.3}$$

$$g_{i*}^{(k)} = g_i^{(k)} + \sum_{0<|K|\le k-1} C_K \mathcal{P}^K(g) \qquad (i = 1, 2, \dots, d_k,\ k = 3, 4, \dots, r) \tag{3.4}$$

for some constant C_K, where K stands for a multi-index $\big((i_1, k_1), (i_2, k_2), \dots, (i_\ell, k_\ell)\big)$ with length $|K| := k_1 + k_2 + \cdots + k_\ell$ and $\mathcal{P}^K(g) := g_{i_1}^{(k_1)} \cdot g_{i_2}^{(k_2)} \cdots g_{i_\ell}^{(k_\ell)}$. The invariances (3.2) and (3.3) play an important role to obtain main results. For $g, h \in G$, we also have

$$(g * h)_{i*}^{(k)} = (g \cdot h)_i^{(k)} \qquad (i = 1, 2, \dots, d_k,\ k = 1, 2), \tag{3.5}$$

$$(g * h)_{i*}^{(k)} = (g \cdot h)_i^{(k)} + \sum_{\substack{|K_1|+|K_2|\le k-1\\ |K_2|>0}} C_{K_1,K_2} \mathcal{P}_*^{K_1}(g) \mathcal{P}^{K_2}(g \cdot h)$$
$$(i = 1, 2, \dots, d_k,\ k = 3, 4, \dots, r) \tag{3.6}$$

by using (3.3) and (3.4). See [23, Section 2] for more details.

3.2 Carnot–Carathéodory metric and homogeneous norms

As is well-known, a nilpotent Lie group G is a candidate of the typical sub-Riemannian manifolds, which is a certain generalization of a Riemannian manifold. The notion of the Carnot–Carathéodory metric is known as an important intrinsic metric in the context sub-Riemannian geometry. We discuss several properties of the Carnot–Carathéodory metric on a nilpotent Lie group G in this subsection.

Recall that the *Carnot–Carathéodory metric* d_{CC} on G is defined by (2.1). We know that the subspace $\mathfrak{g}^{(1)}$ satisfies the so-called *Hörmander condition* in $\mathfrak{g}$, that is, $L_{\mathfrak{g}^{(1)}}(g) = T_gG$ for any $g \in G$, where $L_{\mathfrak{g}^{(1)}}(g)$ denotes the evaluation of $\mathfrak{g}^{(1)}$ at $g \in G$. The Carnot-Carathéodory metric is then well-defined in the sense that $d_{\mathrm{CC}}(g,h) < \infty$ for every $g, h \in G$, thanks to the Hörmander condition on $\mathfrak{g}^{(1)}$. Furthermore, the topology induced by d_{CC} coincides with the original one of G. We emphasize that d_{CC} behaves well under dilations. More precisely, we have

$$d_{\mathrm{CC}}\big(\tau_\varepsilon(g), \tau_\varepsilon(h)\big) = \varepsilon d_{\mathrm{CC}}(g,h) \qquad (\varepsilon \geq 0,\ g, h \in G). \tag{3.7}$$

We now present the notion of homogeneous norm on G. The one-parameter group of dilations $(\tau_\varepsilon)_{\varepsilon \geq 0}$ allows us to consider scalar multiplications on nilpotent Lie groups. We replace the usual Euclidean norms by the following functions. A continuous function $\|\cdot\| : G \to [0,\infty)$ is called a *homogeneous norm* on G if (i) $\|g\| = 0$ if and only if $g = \mathbf{1}_G$ and (ii) $\|\tau_\varepsilon g\| = \varepsilon\|g\|$ for $\varepsilon \geq 0$ and $g \in G$. One of the typical examples of homogeneous norms is given in terms of d_{CC}. We define a continuous function $\|\cdot\|_{\mathrm{CC}} : G \to [0,\infty)$ by $\|g\|_{\mathrm{CC}} := d_{\mathrm{CC}}(\mathbf{1}_G, g)$ for $g \in G$. Then $\|\cdot\|_{\mathrm{CC}}$ is a homogeneous norm on G in view of (3.7). Another basic homogeneous norm is given in the following way. We denote by $\{X_1^{(k)}, X_2^{(k)}, \dots, X_{d_k}^{(k)}\}$ a basis in $\mathfrak{g}^{(k)}$ for $k = 1, 2, \dots, r$. We introduce a norm $\|\cdot\|_{\mathfrak{g}^{(k)}}$ on $\mathfrak{g}^{(k)}$ by the usual Euclidean norm. If $Z \in \mathfrak{g}$ is decomposed as $Z = Z^{(1)} + Z^{(2)} + \cdots + Z^{(r)}$ ($Z^{(k)} \in \mathfrak{g}^{(k)}$), we define a function $\|\cdot\|_{\mathfrak{g}} : \mathfrak{g} \to [0,\infty)$ by $\|Z\|_{\mathfrak{g}} := \sum_{k=1}^r \|Z^{(k)}\|_{\mathfrak{g}^{(k)}}^{1/k}$. We set $\|g\|_{\mathrm{Hom}} := \|\log(g)\|_{\mathfrak{g}}$ for $g \in G$. We then observe that $\|\cdot\|_{\mathrm{Hom}}$ is a homogeneous norm on G. The homogenuity (ii) leads to the most important fact that all homogeneous norms on G are equivalent, which is similar to the case of norms on the Euclidean space. More precisely, we have the following, which plays a crucial role to obtain Theorem 2.2.

Proposition 3.1. (cf. Goodman [19]) If $\|\cdot\|_1$ and $\|\cdot\|_2$ are two homogeneous norms on G, then there exists a constant $C > 0$ such that $C^{-1}\|g\|_1 \leq \|g\|_2 \leq C\|g\|_1$ for $g \in G$.

3.3 Discrete geometric analysis

We present some basics of discrete geometric analysis on graphs due to Kotani–Sunada [33] or Sunada [62, 63, 64]. We consider a finite graph $X_0 = (V_0, E_0)$ and an irreducible random walk on X_0 associated with a non-negative transition probability $p : E_0 \to [0,1]$. We define the 0-chain group, 1-chain group, 0-cochain group and 1-cochain group by

$$C_0(X_0,\mathbb{R}) := \Big\{ \sum_{x \in V_0} a_x x \,\Big|\, a_x \in \mathbb{R} \Big\}, \quad C_1(X_0,\mathbb{R}) := \Big\{ \sum_{e \in E_0} a_e e \,\Big|\, a_e \in \mathbb{R},\ \overline{e} = -e \Big\},$$
$$C^0(X_0,\mathbb{R}) := \{ f : V_0 \to \mathbb{R} \}, \qquad C^1(X_0,\mathbb{R}) := \{ \omega : E_0 \to \mathbb{R} \,|\, \omega(\overline{e}) = -\omega(e) \},$$

respectively. An element of $C^1(X_0,\mathbb{R})$ is called a 1-form on X_0. The boundary operator $\partial : C_1(X_0,\mathbb{R}) \to C_0(X_0,\mathbb{R})$ and the difference operator $d : C^0(X_0,\mathbb{R}) \to C^1(X_0,\mathbb{R})$ are defined by $\partial(e) = t(e) - o(e)$ for $e \in E_0$ and $df(e) = f\big(t(e)\big) - f\big(o(e)\big)$ for $e \in E_0$, respectively. Then, the first homology group $\mathrm{H}_1(X_0,\mathbb{R})$ and the first cohomology group $\mathrm{H}^1(X_0,\mathbb{R})$ are defined by $\mathrm{Ker}\,(\partial) \subset C_1(X_0,\mathbb{R})$ and $C^1(X_0,\mathbb{R})/\mathrm{Im}\,(d)$, respectively. We write $L : C^0(X_0,\mathbb{R}) \to C^0(X_0,\mathbb{R})$ for the transition operator associated with p. We define a special 1-chain by

$$\gamma_p := \sum_{e\in E_0} \widetilde{m}(e)e \in C_1(X_0,\mathbb{R}).$$

It is easily seen that $\partial(\gamma_p) = 0$ so that $\gamma_p \in \mathrm{H}_1(X_0,\mathbb{R})$. Furthermore, it is clear that the random walk on X_0 is (m-)symmetric if and only if $\gamma_p = 0$. The 1-cycle γ_p is called the *homological direction* of the given random walk on X_0. A simple application of the ergodic theorem leads to the law of large numbers on $C_1(X_0,\mathbb{R})$.

$$\lim_{n\to\infty}\frac{1}{n}(e_1+e_2+\cdots+e_n) = \gamma_p, \quad \mathbb{P}_x\text{-a.e. } c = (e_1,e_2,\dots,e_n,\dots)\in\Omega_x(X_0).$$

A 1-form $\omega \in C^1(X_0,\mathbb{R})$ is said to be *modified harmonic* if

$$\sum_{e\in(E_0)_x} p(e)\omega(x) - \langle\gamma_p,\omega\rangle = 0 \qquad (x\in V_0), \tag{3.8}$$

where $\langle\gamma_p,\omega\rangle := {}_{C_1(X_0,\mathbb{R})}\langle\gamma_p,\omega\rangle_{C^1(X_0,\mathbb{R})}$ is constant as a function on V_0. We denote by $\mathcal{H}^1(X_0)$ the space of modified harmonic 1-forms and equip it with the inner product

$$\langle\!\langle\omega_1,\omega_2\rangle\!\rangle_p := \sum_{e\in E_0}\widetilde{m}(e)\omega_1(e)\omega_2(e) - \langle\gamma_p,\omega_1\rangle\langle\gamma_p,\omega_2\rangle \qquad \big(\omega_1,\omega_2\in\mathcal{H}^1(X_0)\big)$$

associated with the transition probability p. We should emphasize that the condition **(A2)** plays a crucial role in proving that $\langle\!\langle\omega,\omega\rangle\!\rangle_p = 0$ implies $\omega = 0$. Indeed, $\langle\!\langle\omega,\omega\rangle\!\rangle_p = 0$ implies

$$\sum_{e\in E_0}\widetilde{m}(e)\Big(\omega(e) - \langle\gamma_p,\omega\rangle\Big)^2 = 0.$$

By using **(A2)**, we find an edge $e_* \in E_0$ satisfying $p(e_*) > 0$ and $p(\overline{e}_*) > 0$, so that we have $\omega(e_*) - \langle\gamma_p,\omega\rangle = \omega(\overline{e}_*) - \langle\gamma_p,\omega\rangle = 0$. Hence, we know $\langle\gamma_p,\omega\rangle = 0$ and this completes the proof of $\omega = 0$ since

$$\sum_{e\in E_0}\widetilde{m}(e)\omega(e)^2 = 0 \quad \text{and} \quad \widetilde{m}(e)+\widetilde{m}(\overline{e}) > 0 \quad (e\in E_0).$$

We may identify $\mathrm{H}^1(X_0,\mathbb{R})$ with $\mathcal{H}^1(X_0)$ by the discrete Hodge-Kodaira theorem (cf. [33, Lemma 5.2]). We induce an inner product from $\mathrm{H}^1(X_0,\mathbb{R})$ by using this identification.

Let Γ be a torsion free, finitely generated nilpotent group of step r. Then a Γ-nilpotent covering graph $X = (V,E)$ is defined by the Γ-covering of X_0. Let $p : E \to [0,1]$ and $m : V \to (0,1]$ be the Γ-invariant lifts of $p : E_0 \to [0,1]$ and $m : V_0 \to (0,1]$, respectively. Denote by $\widehat{\pi} : G \to G/[G,G]$ the canonical projection. Since Γ is a cocompact lattice in G, the subset $\widehat{\pi}(\Gamma) \subset G/[G,G]$ is also a lattice in $G/[G,G] \cong \mathfrak{g}^{(1)}$ (cf. Malcév [48] and Raghunathan [54]). We take the canonical surjective homomorphism $\rho : \mathrm{H}_1(X_0,\mathbb{Z}) \to \widehat{\pi}(\Gamma) \cong \Gamma/[\Gamma,\Gamma]$ and its realification is denoted by $\rho_{\mathbb{R}} : \mathrm{H}_1(X_0,\mathbb{R}) \to \widehat{\pi}(\Gamma)\otimes\mathbb{R}$. We identify $\mathrm{Hom}(\widehat{\pi}(\Gamma),\mathbb{R})$ with a subspace of $\mathrm{H}^1(X_0,\mathbb{R})$ through ${}^t\rho_{\mathbb{R}}$. We restrict $\langle\!\langle\cdot,\cdot\rangle\!\rangle_p$ on $\mathrm{H}^1(X_0,\mathbb{R})$ to the subspace $\mathrm{Hom}(\widehat{\pi}(\Gamma),\mathbb{R})$ and take it up the dual inner product $\langle\cdot,\cdot\rangle_{alb}$ on $\widehat{\pi}(\Gamma)\otimes\mathbb{R}$. Then, a flat metric g_0 on $\mathfrak{g}^{(1)}$ is induced and we call it the *Albanese metric* on $\mathfrak{g}^{(1)}$. This procedure can be summarized as follows:

$$\begin{array}{ccccc}
(\mathfrak{g}^{(1)},g_0) & \cong & \widehat{\pi}(\Gamma)\otimes\mathbb{R} & \overset{\rho_{\mathbb{R}}}{\twoheadleftarrow} & \mathrm{H}_1(X_0,\mathbb{R}) \\
\updownarrow{\scriptstyle \text{dual}} & & \updownarrow{\scriptstyle \text{dual}} & & \updownarrow{\scriptstyle \text{dual}} \\
\mathrm{Hom}(\mathfrak{g}^{(1)},\mathbb{R}) & \cong & \mathrm{Hom}(\widehat{\pi}(\Gamma),\mathbb{R}) & \underset{{}^t\rho_{\mathbb{R}}}{\hookrightarrow} & \mathrm{H}^1(X_0,\mathbb{R}) \cong \big(\mathcal{H}^1(X_0),\langle\!\langle\cdot,\cdot\rangle\!\rangle_p\big).
\end{array}$$

A map $\Phi : X \to G$ is said to be a *periodic realization* of X if it satisfies $\Phi(\gamma x) = \gamma \cdot \Phi(x)$ for $\gamma \in \Gamma$ and $x \in V$. Fix a reference point $x_* \in V$ and define a special realization $\Phi_0 : X \to G$ by

$$ {}_{\mathrm{Hom}(\mathfrak{g}^{(1)},\mathbb{R})}\big\langle \omega, \log\big(\Phi_0(x)\big)\big|_{\mathfrak{g}^{(1)}}\big\rangle_{\mathfrak{g}^{(1)}} = \int_{x_*}^{x} \widetilde{\omega} \qquad \big(\omega \in \mathrm{Hom}(\mathfrak{g}^{(1)}, \mathbb{R}),\, x \in V\big), \tag{3.9} $$

where $\widetilde{\omega}$ is the lift of $\omega = {}^t\rho_{\mathbb{R}}(\omega) \in \mathrm{H}^1(X_0, \mathbb{R})$ to X. Here $\int_{x_*}^{x} \widetilde{\omega} = \int_c \widetilde{\omega} = \sum_{i=1}^{n} \widetilde{\omega}(e_i)$ for a path $c = (e_1, e_2, \ldots, e_n)$ with $o(e_1) = x_*$ and $t(e_n) = x$. We note that this line integral does not depend on the choice of a path c. By following the arguments in [33, Section 4], [23, Section 5] and [25, Section 3], we obtain that such Φ_0 enjoys the *modified harmonicity* in the sense of (2.2). See Namba [50] for the proof.

Lemma 3.2 (cf. Namba [50, Lemma 2.4.8])**.** The periodic realization $\Phi_0 : X \to G$ defined by (3.9) is the modified harmonic realization, that is,

$$ \sum_{e \in E_x} p(e) \log\Big(\Phi_0\big(o(e)\big)^{-1} \cdot \Phi_0\big(t(e)\big)\Big)\Big|_{\mathfrak{g}^{(1)}} = \rho_{\mathbb{R}}(\gamma_p) \qquad (x \in V). $$

Let us consider a random walk $(\Omega_x(X), \mathbb{P}_x, \{w_n\}_{n=0}^{\infty})$ with values in a Γ-nilpotent covering graph X. We denote by $\Phi : X \to G$ a Γ-equivariant realization of X. We then have the G-valued Markov chain $(\Omega_x(X), \mathbb{P}_x, \{\xi_n\}_{n=0}^{\infty})$ defined by $\xi_n(c) := \Phi\big(w_n(c)\big)$ for $n \in \mathbb{N} \cup \{0\}$ and $c \in \Omega_x(X)$, through the map Φ. This gives rise to the $\mathfrak{g}$-valued random walk $\Xi_n(c) := \log\big(\xi_n(c)\big) = \log\big(\Phi\big(w_n(c)\big)\big)$ for $n \in \mathbb{N} \cup \{0\}$ and $c \in \Omega_x(X)$. We obtain the following law of large numbers on $\mathfrak{g}^{(1)}$ by the ergodic theorem.

$$ \lim_{n\to\infty} \frac{1}{n} \Xi_n(\cdot)\big|_{\mathfrak{g}^{(1)}} = \rho_{\mathbb{R}}(\gamma_p), \quad \mathbb{P}_x\text{-a.s.} \tag{3.10} $$

It is known that the notion of martingales plays a crucial role in the theory of stochastic processes. We give a certain characterization of modified harmonic realizations in view of martingale theory. Let $\pi_n : \Omega_x(X) \to \Omega_{x,n}(X)\, (n \in \mathbb{N} \cup \{0\})$ be a projection defined by $\pi_n(c) := (e_1, e_2, \ldots, e_n)$ for $c = (e_1, e_2, \ldots, e_n, \ldots) \in \Omega_x(X)$. Denote by $\{\mathcal{F}_n\}_{n=0}^{\infty}$ the filtration such that $\mathcal{F}_0 = \{\emptyset, \Omega_x(X)\}$ and $\mathcal{F}_n := \sigma\big(\pi_n^{-1}(A) \,\big|\, A \subset \Omega_{x,n}(X)\big)$ for $n \in \mathbb{N}$. We mention that $\mathcal{F}_n$ is a sub-σ-algebra of $\mathcal{F}_\infty := \bigvee_{n=0}^{\infty} \mathcal{F}_n$ for $n \in \mathbb{N}$. We use the following in the proof of Lemma 4.6.

Lemma 3.3 (cf. Namba [50, Lemma 2.5.3])**.** Let $\{X_1^{(1)}, X_2^{(1)}, \ldots, X_{d_1}^{(1)}\}$ be a basis of $\mathfrak{g}^{(1)}$. Then a periodic realization $\Phi_0 : X \to G$ is the modified harmonic realization if and only if the stochastic process

$$ \big\{\Xi_n\big|_{X_i^{(1)}} - n\rho_{\mathbb{R}}(\gamma_p)\big|_{X_i^{(1)}}\big\}_{n=0}^{\infty} \qquad (i = 1, 2, \ldots, d_1), $$

with values in $\mathbb{R}$, is an $\{\mathcal{F}_n\}$-martingale.

4 Proof of main results

The aim of this section is to prove Theorems 2.1, 2.2 and 2.3. Before going into details, we give a brief outline of the proof for the readers' convenience.

4.1 A brief outline of the proof through a simple example

While the theory of discrete geometric analysis plays a crucial role throughout the present paper, some readers may regard it as too complicated. Hence, we give a brief outline of the proofs in the case where Γ is the 3-dimensional discrete Heisenberg group in order to help readers get a bird's eye view of them.

It goes without saying that the most typical but non-trivial example of nilpotent groups of step 2 is the 3-*dimensional discrete Heisenberg group* defined by $\Gamma = \mathbb{H}^3(\mathbb{Z}) := (\mathbb{Z}^3, \star)$, where the product $\star$ on $\mathbb{Z}^3$ is given by

$$(x, y, z) \star (x', y', z') = (x + x', y + y', z + z' + xy').$$

Then, we see that $G = \mathbb{H}^3(\mathbb{R})$ is the corresponding connected and simply connected nilpotent Lie group of step 2 in which Γ is isomorphic to a cocompact lattice. Furthermore, the corresponding Lie algebra $\mathfrak{g}$ is given by $\mathfrak{g} = (\mathbb{R}^3, [\cdot,\cdot])$ with $[X_1, X_2] = X_3$ and $[X_1, X_3] = [X_2, X_3] = \mathbf{0}_{\mathfrak{g}}$, where $\{X_1, X_2, X_3\}$ be the standard basis of $\mathfrak{g}$. We note that the Lie algebra $\mathfrak{g}$ is decomposed as $\mathfrak{g} = \mathfrak{g}^{(1)} \oplus \mathfrak{g}^{(2)}$, where $\mathfrak{g}^{(1)} := \mathrm{span}_{\mathbb{R}}\{X_1, X_2\}$ and $\mathfrak{g}^{(2)} := \mathrm{span}_{\mathbb{R}}\{X_3\}$. The nilpotent Lie group $G = \mathbb{H}^3(\mathbb{R})$ is free of step 2, which implies that the limit group G_∞ coincides with G itself. In general, the difference between G and G_∞ may appear in more general step cases.

Let $X = (V, E)$ be a Cayley graph of Γ with a generating set

$$S = \{(\pm 1, 0, 0),\ (0, \pm 1, 0),\ (0, 0, \pm 1)\}.$$

Namely, X is an oriented graph with $V = \Gamma$ and for $\gamma, \eta \in \Gamma$, γ is adjacent to η if $\gamma^{-1} \star \eta \in S$. The quotient $X_0 = (V_0, E_0) = \Gamma\backslash X$ is a 3-*bouquet graph* with $V_0 = \{x_0\}$ and E_0 consisting of three loops e_1, e_2, e_3 and their inverse loops as well. Consider a non-symmetric random walk on X_0 defined by

$$p(e_1) = \alpha, \quad p(e_2) = \beta, \quad p(e_3) = \gamma, \quad p(\overline{e}_1) = \alpha', \quad p(\overline{e}_2) = \beta', \quad p(\overline{e}_3) = \gamma',$$

where $\alpha, \beta, \gamma, \alpha', \beta', \gamma' > 0$ and $\alpha + \beta + \gamma + \alpha' + \beta' + \gamma' = 1$. Then we easily have $\rho_{\mathbb{R}}(\gamma_p) = (\alpha - \alpha')X_1 + (\beta - \beta')X_2$. We fix an equivariant realization $\Phi_0 : X \to G$. Since X_0 has only one vertex, the realization Φ always enjoy the *modified harmonicity* in the sense of (2.2). Then the random walk $\big\{\log\big(\Phi(w_n)\big)\big|_{\mathfrak{g}^{(1)}} - n\rho_{\mathbb{R}}(\gamma_p)\big\}_{n=0}^{\infty}$ is a $\mathfrak{g}^{(1)}$-valued martingale, which will play a key role in the proof. Here, $\{w_n\}_{n=0}^{\infty}$ is the random walk on X associated with the transition probability p.

Step 1 (To show Theorem 2.1): In showing the first main theorem (Theorem 2.1), there are several difficulties related to the asymptotic direction $\rho_{\mathbb{R}}(\gamma_p)$. In order to overcome them, we need to use the *transition-shift operator* $\mathcal{L}_p$ on a Banach space $C_\infty(X \times \mathbb{Z})$ (see Section 2. We have to note that such an operator has not been introduced in any early works discussing CLTs under non-centered settings (cf. Raugi [55]).

At first, we prove the convergence of the infinitesimal generator under the CLT-scaling (Lemma 4.2). We use the Taylor formula to $(I - \mathcal{L}_p^N)\mathcal{P}_\varepsilon f$ in ε. Then, the first order terms vanish due to the modified harmonicity of Φ so that we formally have, for $x \in V$ and $f \in C_0^\infty(G)$,

$$\frac{1}{N\varepsilon^2}(I - \mathcal{L}_p^N)\mathcal{P}_\varepsilon f(x) = \mathcal{P}_\varepsilon\Big(\Delta_G - \beta(\Phi)\Big)f(x) + O\Big(\frac{1}{N}\Big) + O(N^2\varepsilon)$$

as $N \to \infty$, $\varepsilon \searrow 0$ with $N^2\varepsilon \searrow 0$, where Δ_G is a sub-elliptic operator on G and $\beta(\Phi) \in \mathfrak{g}^{(2)}$ is a quantity given by (2.5), that is,

$$\begin{aligned}\beta(\Phi) &= \sum_{e \in E_0} p(e) \log\Big(\Phi\big(o(\widetilde{e})\big)^{-1} \star \Phi\big(t(\widetilde{e})\big) \star \exp\big(-\rho_{\mathbb{R}}(\gamma_p)\big)\Big)\Big|_{\mathfrak{g}^{(2)}} \\ &= \Big\{(\gamma - \gamma') + \frac{1}{2}(\alpha - \alpha')(\beta - \beta')\Big\}X_3.\end{aligned}$$

This leads to Lemma 4.2. Finally, we combine the *Trotter approximation theorem* (cf. [66]) with Lemma 4.2. Then we obtain Theorem 2.1 by letting $n \to \infty$. Moreover, if we endow $\mathfrak{g}^{(1)}$ with the *Albanese metric* g_0 associated with p, the sub-elliptic operator Δ_G coincides with $-\frac{1}{2}(V_1^2 + V_2^2)$ for some orthonormal basis $\{V_1, V_2\}$ of $(\mathfrak{g}^{(1)}, g_0)$.

Step 2 (To show the functional CLT): We put an additional assumption **(A3)**. Let $\{\mathcal{Y}^{(n)}\}_{n=1}^{\infty}$ be a sequence of stochastic processes defined by the geodesic interpolation of the underlying random walk. In order to show the functional CLT, it is sufficient to prove the following two items:

- the convergence of the finite-dimensional distribution of $\{\mathcal{Y}^{(n)}\}_{n=1}^{\infty}$, and
- the tightness of the image measures $\{\mathbf{P}^{(n)} = \mathbb{P}_{x_*} \circ (\mathcal{Y}^{(n)})^{-1}\}_{n=1}^{\infty}$.

Since the former item is obtained by an easy application of Theorem 2.1, we mention the proof of the latter item. In the proof of the latter one, we may show that there is a positive constant $C > 0$ independent of $n \in \mathbb{N}$ such that

$$\mathbb{E}^{\mathbb{P}_{x_*}}\Big[d_{\mathrm{CC}}\big(\mathcal{Y}_s^{(n)}, \mathcal{Y}_t^{(n)}\big)^{4m}\Big] \leq C(t-s)^{2m} \qquad (m \in \mathbb{N},\ 0 \leq s \leq t \leq 1). \tag{4.1}$$

Indeed, (4.1) is established by applying several martingale inequalities such as the Birkholder–Davis–Gundy inequality for the martingale $\{\log(\Phi(w_n))|_{\mathfrak{g}^{(1)}}\}_{n=0}^{\infty}$. (We mention that this kind of tightness argument in the rough path framework was performed by e.g., Breuillard–Friz–Huesmann [8] and Bayer–Friz [4].) Consequently, we obtain that $\mathcal{Y}^{(n)}$ converges in law to a G-valued diffusion process Y which solves the SDE

$$dY_t = V_1^2(Y_t) \circ dB_t^1 + V_2^2(Y_t) \circ dB_t^2 + \beta(\Phi)(Y_t)\,dt, \qquad Y_0 = \mathbf{1}_G$$

in the Hölder space $C^{0,\alpha\text{-Höl}}([0,1]; G)$ for $\alpha < 1/2$, where $(B_t^1, B_t^2)_{0\leq t\leq 1}$ is a 2-dimensional standard Brownian motion starting from the origin.

Nevertheless, we emphasize that several essential difficulties arise in the case of more general covering graphs. The one difficulty appears when the number of vertices of X_0 is large. If it is larger than one, then each Γ-equivariant realization Φ is *not* always modified harmonic. Since **Step 2** basically uses the modified harmonicity, an additional step should be inserted in the proof of the functional CLT *for modified harmonic realizations*, which is given as follows: We first show the functional CLT for the modified harmonic realization Φ_0 in the same way as **Step 2**. Next, let Φ be a Γ-equivariant realization. Let $\{\overline{\mathcal{Y}}^{(n)}\}_{n=1}^{\infty}$ be a sequence of stochastic processes defined by the geodesic interpolation of the underlying random walk $\{\Phi(w_n)\}_{n=0}^{\infty}$. We here introduce the $(\mathfrak{g}^{(1)})$-corrector $\mathrm{Cor}_{\mathfrak{g}^{(1)}}$ defined by (2.11). Thanks to a nice estimation of the $\mathfrak{g}^{(1)}$-corrector, we have the similar moment estimate to (4.1) for $\{\overline{\mathcal{Y}}^{(n)}\}_{n=1}^{\infty}$ (see Lemma 4.9). Finally, we obtain that the sequence $\{\overline{\mathcal{Y}}^{(n)}\}_{n=1}^{\infty}$ also converges in law to the G-valued diffusion process $(Y_t)_{0\leq t\leq 1}$ in $C^{0,\alpha\text{-Höl}}([0,1]; G)$ for $\alpha < 1/2$. This completes the proof of Theorem 2.3.

Another difficulty arises when the nilpotency of Γ is greater than 2. If the nilpotent Lie group G is of step $r \geq 3$, the dilation operators lack their nice propeties, which makes the analysis on G much difficult. In order to recover such nice properties, we consider everything on the correponding limit group G_∞ intead of G. Nevertheless, there is a further difficulty in the proof of the tightness of $\{\mathbf{P}^{(n)}\}_{n=1}^{\infty}$, since it is difficult to give (4.1) in higher-step cases directly. Therefore, we extend a novel pathwise argument inspired by the proof of Lyons' extension theorem to the cases of general nilpotent Lie groups (Lemma 4.9). Such an extension enables us to prove the tightness of $\{\mathbf{P}^{(n)}\}_{n=1}^{\infty}$ as well as the case mentioned in **Step 2**, which is one of remarkable contributions of the present paper.

4.2 Proof of Theorem 2.1

In what follows, we set

$$d\Phi_0(e) = \Phi_0\big(o(e)\big)^{-1} \cdot \Phi_0\big(t(e)\big) \qquad (e \in E),$$
$$\|d\Phi_0\|_\infty = \max_{e \in E_0} \Big\{ \big\| \log\big(d\Phi_0(\widetilde{e})\big)\big|_{\mathfrak{g}^{(1)}} \big\|_{\mathfrak{g}^{(1)}} + \big\| \log\big(d\Phi_0(\widetilde{e})\big)\big|_{\mathfrak{g}^{(2)}} \big\|_{\mathfrak{g}^{(2)}}^{1/2} \Big\},$$

where $\widetilde{e}$ stands for a lift of $e \in E_0$ to X. We should mention that

$$\Big(\Phi_0(x)^{-1} \cdot \Phi_0\big(t(c)\big)\Big)_i^{(k)} = O(N^k) \tag{4.2}$$

for $x \in V, c \in \Omega_{x,N}(X)$, $i = 1, 2, \dots, d_k$ and $k = 1, 2, \dots, r$. We also write $\rho = \rho_{\mathbb{R}}(\gamma_p)$ and $e^{z\rho} = \exp\big(z\rho_{\mathbb{R}}(\gamma_p)\big)$ $(z \in \mathbb{R})$ for brevity. We give an important property of the family of approximation operators $(\mathcal{P}_\varepsilon)_{0 \le \varepsilon \le 1}$ defined by (2.4).

Lemma 4.1. Let $q > 1$. Then $\big((C_{\infty,q}(X \times \mathbb{Z}), \|\cdot\|_{\infty,q}; \mathcal{P}_\varepsilon)\big)_{0 \le \varepsilon \le 1}$ is a family of Banach spaces approximating to the Banach space $\big(C_\infty(G), \|\cdot\|_\infty^G\big)$ in the sense of Trotter [66]:

$$\|\mathcal{P}_\varepsilon f\|_{\infty,q} \le \|f\|_\infty^G \quad \text{and} \quad \lim_{\varepsilon \searrow 0} \|\mathcal{P}_\varepsilon f\|_{\infty,q} = \|f\|_\infty^G \qquad \big(f \in C_\infty(G)\big).$$

Proof. The former assertion follows from

$$\|\mathcal{P}_\varepsilon f\|_{\infty,q} = \frac{1}{C_q} \sum_{z \in \mathbb{Z}} \frac{\|f(\cdot, z)\|_\infty}{1 + |z|^q} \le \frac{1}{C_q} \sum_{z \in \mathbb{Z}} \frac{\|f\|_\infty}{1 + |z|^q} = \|f\|_\infty.$$

We prove the latter one. Let $g_0 \in G$ be an element which attains $\|f\|_\infty = \sup_{g \in G} |f(g)|$. We fix $z \in \mathbb{Z}$. Then we have

$$\|\mathcal{P}_\varepsilon f(\cdot, z)\|_\infty \ge |f(g_0)| - \inf_{x \in X} \Big| f(g_0) - f\Big(\tau_\varepsilon\big(\Phi_0(x) * \exp(-z\rho_{\mathbb{R}}(\gamma_p))\big)\Big)\Big|.$$

On the other hand, we have

$$\begin{aligned} &\inf_{x \in X} d_{\mathrm{CC}}\Big(g_0, \tau_\varepsilon\big(\Phi_0(x) * \exp(-z\rho_{\mathbb{R}}(\gamma_p))\big)\Big) \\ &= \varepsilon \inf_{x \in X} d_{\mathrm{CC}}\Big(\tau_{1/\varepsilon}(g_0), \Phi_0(x) * \exp(-z\rho_{\mathbb{R}}(\gamma_p))\Big) < \varepsilon M \end{aligned}$$

for some $M = M(z) > 0$. From the continuity of f, for any $\delta > 0$, there exists $\delta' > 0$ such that $d_{\mathrm{CC}}(g_0, h) < \delta'$ implies $|f(g_0) - f(h)| < \delta$. By choosing a sufficiently small $\varepsilon > 0$, we have

$$d_{\mathrm{CC}}\Big(g_0, \tau_\varepsilon\big(\Phi_0(x_*) * \exp(-z\rho_{\mathbb{R}}(\gamma_p))\big)\Big) < \delta'$$

for some $x_* \in X$. Then we have

$$\begin{aligned} &\inf_{x \in X} \Big| f(g_0) - f\Big(\tau_\varepsilon\big(\Phi_0(x) * \exp(-z\rho_{\mathbb{R}}(\gamma_p))\big)\Big)\Big| \\ &\le \Big| f(g_0) - f\Big(\tau_\varepsilon\big(\Phi_0(x_*) * \exp(-z\rho_{\mathbb{R}}(\gamma_p))\big)\Big)\Big| < \delta \end{aligned}$$

and this implies $\lim_{\varepsilon \searrow 0} \|\mathcal{P}_\varepsilon f(\cdot, z)\|_\infty = \|f\|_\infty$ for $z \in \mathbb{Z}$. By using the dominated convergence theorem, we obtain $\lim_{\varepsilon \searrow 0} \|\mathcal{P}_\varepsilon f\|_{\infty,r} = \|f\|_\infty$. This completes the proof. ☐

The following lemma is significant to prove Theorem 2.1.

Lemma 4.2. Let $f \in C_0^\infty(G)$ and $q > 4r + 1$. Then we have

$$\Big\| \frac{1}{N\varepsilon^2}\big(I - \mathcal{L}_p^N\big)\mathcal{P}_\varepsilon f - \mathcal{P}_\varepsilon \mathcal{A} f \Big\|_{\infty,q} \to 0$$

as $N \to \infty$ and $\varepsilon \searrow 0$ with $N^2\varepsilon \searrow 0$, where $\mathcal{L}_p$ is the transition-shift operator defined by (2.3) and $\mathcal{A}$ is the sub-elliptic operator defined by (2.7).

Proof. We divide the proof into several steps.

Step 1. We first apply Taylor's formula (cf. Alexopoulos [2, Lemma 5.3]) for the $(*)$-coordinates of the second kind to $f \in C_0^\infty(G)$ at $\tau_\varepsilon\big(\Phi_0(x) * e^{-z\rho}\big) \in G$. By recalling that $(G,*)$ is a stratified Lie group, we have

$$\begin{aligned}
&\frac{1}{N\varepsilon^2}(I-\mathcal{L}_p^N)\mathcal{P}_\varepsilon f(x,z)\\
&= -\sum_{(i,k)} \frac{\varepsilon^{k-2}}{N} X_{i*}^{(k)} f\big(\tau_\varepsilon\big(\Phi_0(x) * e^{-z\rho}\big)\big) \sum_{c\in\Omega_{x,N}(X)} p(c)\big(\mathcal{B}_N(x,z,c)\big)_{i*}^{(k)}\\
&\quad -\Big(\sum_{(i_1,k_1)\geq(i_2,k_2)} \frac{\varepsilon^{k_1+k_2-2}}{2N} X_{i_1*}^{(k_1)} X_{i_2*}^{(k_2)} + \sum_{(i_2,k_2)>(i_1,k_1)} \frac{\varepsilon^{k_1+k_2-2}}{2N} X_{i_2*}^{(k_2)} X_{i_1*}^{(k_1)}\Big)\\
&\qquad \times f\Big(\tau_\varepsilon\big(\Phi_0(x) * e^{-z\rho}\big)\Big) \sum_{c\in\Omega_{x,N}(X)} p(c)\big(\mathcal{B}_N(x,z,c)\big)_{i_1*}^{(k_1)}\big(\mathcal{B}_N(x,z,c)\big)_{i_2*}^{(k_2)}\\
&\quad - \sum_{(i_1,k_1),(i_2,k_2),(i_3,k_3)} \frac{\varepsilon^{k_1+k_2+k_3-2}}{6N} \frac{\partial^3 f}{\partial g_{i_1*}^{(k_1)}\partial g_{i_2*}^{(k_2)}\partial g_{i_3*}^{(k_3)}}(\theta) \sum_{c\in\Omega_{x,N}(X)} p(c)\big(\mathcal{B}_N(x,z,c)\big)_{i_1*}^{(k_1)}\\
&\qquad \times \big(\mathcal{B}_N(x,z,c)\big)_{i_2*}^{(k_2)}\big(\mathcal{B}_N(x,z,c)\big)_{i_3*}^{(k_3)} \qquad (x\in V,\ z\in\mathbb{Z}),
\end{aligned}\tag{4.3}$$

for some $\theta \in G$ with $|\theta_{i_*}^{(k)}| \leq \varepsilon^k\big|\big(\mathcal{B}_N(x,z,c)\big)_{i*}^{(k)}\big|$ for $i = 1,2,\dots,d_k$ and $k = 1,2,\dots,r$, where the summation $\sum_{(i_1,k_1)\geq(i_2,k_2)}$ runs over all (i_1,k_1) and (i_2,k_2) with $k_1 > k_2$ or $k_1 = k_2$, $i_1 \geq i_2$. We put

$$\mathcal{B}_N(x,z,c) := e^{z\rho} * \Phi_0(x)^{-1} * \Phi_0\big(t(c)\big) * e^{-(z+N)\rho} \quad \big(N\in\mathbb{N},\ x\in V,\ z\in\mathbb{Z},\ c\in\Omega_{x,N}(X)\big).$$

We denote by $\mathrm{Ord}_\varepsilon(k)$ the terms of the right-hand side of (4.3) whose order of ε equals just k. Then (4.3) is rewritten as

$$\frac{1}{N\varepsilon^2}(I-\mathcal{L}_p^N)\mathcal{P}_\varepsilon f(x,z) = \mathrm{Ord}_\varepsilon(-1) + \mathrm{Ord}_\varepsilon(0) + \sum_{k\geq 1}\mathrm{Ord}_\varepsilon(k) \qquad (x\in V,\ z\in\mathbb{Z}),$$

where

$$\begin{aligned}
\mathrm{Ord}_\varepsilon(-1) &= -\frac{1}{N\varepsilon}\sum_{i=1}^{d_1} X_{i*}^{(1)} f\big(\tau_\varepsilon\big(\Phi_0(x) * e^{-z\rho}\big)\big) \sum_{c\in\Omega_{x,N}(X)} p(c)\big(\mathcal{B}_N(x,z,c)\big)_{i*}^{(1)},\\
\mathrm{Ord}_\varepsilon(0) &= -\frac{1}{N}\sum_{i=1}^{d_2} X_{i*}^{(2)} f\big(\tau_\varepsilon\big(\Phi_0(x) * e^{-z\rho}\big)\big) \sum_{c\in\Omega_{x,N}(X)} p(c)\Big\{\big(\mathcal{B}_N(x,z,c)\big)_{i*}^{(2)}\\
&\qquad - \frac{1}{2}\sum_{1\leq\lambda<\nu\leq d_1} \big(\mathcal{B}_N(x,z,c)\big)_{\lambda*}^{(1)}\big(\mathcal{B}_N(x,z,c)\big)_{\nu*}^{(1)} [\![X_\lambda^{(1)}, X_\nu^{(1)}]\!]\big|_{X_i^{(2)}}\Big\}\\
&\quad - \frac{1}{2N}\sum_{1\leq i,j\leq d_1} X_{i*}^{(1)} X_{j*}^{(1)} f\big(\tau_\varepsilon\big(\Phi_0(x) * e^{-z\rho}\big)\big)\\
&\qquad \times \sum_{c\in\Omega_{x,N}(X)} p(c)\big(\mathcal{B}_N(x,z,c)\big)_{i*}^{(1)}\big(\mathcal{B}_N(x,z,c)\big)_{j*}^{(1)}
\end{aligned}$$

and $\sum_{k\geq 1}\mathrm{Ord}_\varepsilon(k)$ is given by the sum of the following three parts:

$$\mathcal{I}_1(\varepsilon,N)=-\sum_{k\geq 3}\sum_{i=1}^{d_k}\frac{\varepsilon^{k-2}}{N}X_{i*}^{(k)}f\big(\tau_\varepsilon\big(\Phi_0(x)*e^{-z\rho}\big)\big)\sum_{c\in\Omega_{x,N}(X)}p(c)\big(\mathcal{B}_N(x,z,c)\big)_{i*}^{(k)},$$

$$\begin{aligned}\mathcal{I}_2(\varepsilon,N)=-\Big(&\sum_{\substack{(i_1,k_1)\geq(i_2,k_2)\\ k_1+k_2\geq 3}}\frac{\varepsilon^{k_1+k_2-2}}{2N}X_{i_1*}^{(k_1)}X_{i_2*}^{(k_2)}+\sum_{\substack{(i_2,k_2)>(i_1,k_1)\\ k_1+k_2\geq 3}}\frac{\varepsilon^{k_1+k_2-2}}{2N}X_{i_2*}^{(k_2)}X_{i_1*}^{(k_1)}\Big)\\ &\times f\big(\tau_\varepsilon\big(\Phi_0(x)*e^{-z\rho}\big)\big)\sum_{c\in\Omega_{x,N}(X)}p(c)\big(\mathcal{B}_N(x,z,c)\big)_{i_1*}^{(k_1)}\big(\mathcal{B}_N(x,z,c)\big)_{i_2*}^{(k_2)},\end{aligned}$$

$$\begin{aligned}\mathcal{I}_3(\varepsilon,N)=-&\sum_{(i_1,k_1),(i_2,k_2),(i_3,k_3)}\frac{\varepsilon^{k_1+k_2+k_3-2}}{6N}\frac{\partial^3 f}{\partial g_{i_1*}^{(k_1)}\partial g_{i_2*}^{(k_2)}\partial g_{i_3*}^{(k_3)}}(\theta)\\ &\times\sum_{c\in\Omega_{x,N}(X)}p(c)\big(\mathcal{B}_N(x,z,c)\big)_{i_1*}^{(k_1)}\big(\mathcal{B}_N(x,z,c)\big)_{i_2*}^{(k_2)}\big(\mathcal{B}_N(x,z,c)\big)_{i_3*}^{(k_3)}.\end{aligned}$$

To complete the proof of Lemma 4.2, it is sufficient to show the following items:
(1) $\mathrm{Ord}_\varepsilon(-1)=0$.
(2) We have

$$\mathrm{Ord}_\varepsilon(0)=-\mathcal{A}f\big(\tau_\varepsilon\big(\Phi_0(x)*e^{-z\rho}\big)\big)+O\Big(\frac{1}{N}\Big). \tag{4.4}$$

(3) As $N\to\infty$ and $\varepsilon\searrow 0$ with $N^2\varepsilon\searrow 0$, we have

$$\|\mathcal{I}_i(\varepsilon,N)\|_{\infty,q}\to 0\qquad(i=1,2,3). \tag{4.5}$$

Step 2. We here show (1). We fix $i=1,2,\dots,d_1$. By recalling (2.2) and (3.2), we have inductively

$$\begin{aligned}&\sum_{c\in\Omega_{x,N}(X)}p(c)\big(\mathcal{B}_N(x,z,c)\big)_{i*}^{(1)}\\ &=\sum_{c'\in\Omega_{x,N-1}(X)}p(c')\sum_{e\in E_{t(c')}}p(e)\Big\{\log\Big(\Phi_0(x)^{-1}\cdot\Phi_0\big(t(c')\big)\cdot e^{-(N-1)\rho}\Big)\Big|_{X_i^{(1)}}\\ &\qquad+\log\Big(\Phi_0\big(o(e)\big)^{-1}\cdot\Phi_0\big(t(e)\big)\cdot e^{-\rho}\Big)\Big|_{X_i^{(1)}}\Big\}\\ &=\sum_{c'\in\Omega_{x,N-1}(X)}p(c')\log\Big(\Phi_0(x)^{-1}\cdot\Phi_0\big(t(c')\big)\cdot e^{-(N-1)\rho}\Big)\Big|_{X_i^{(1)}}=0\qquad(x\in V,\,z\in\mathbb{Z}).\end{aligned}$$

Step 3. We prove the item (2). First consider the coefficient of $X_{i*}^{(2)}f\big(\tau_\varepsilon\big(\Phi_0(x)*e^{-z\rho}\big)\big)$ which is given by

$$\begin{aligned}&-\frac{1}{N}\sum_{c\in\Omega_{x,N}(X)}p(c)\Big\{\big(\mathcal{B}_N(x,z,c)\big)_{i*}^{(2)}\\ &\qquad-\frac{1}{2}\sum_{1\leq\lambda<\nu\leq d_1}\big(\mathcal{B}_N(x,z,c)\big)_{\lambda*}^{(1)}\big(\mathcal{B}_N(x,z,c)\big)_{\nu*}^{(1)}[\![X_\lambda^{(1)},X_\nu^{(1)}]\!]\big|_{X_i^{(2)}}\Big\}\\ &=-\frac{1}{N}\sum_{c\in\Omega_{x,N}(X)}p(c)\log\big(\mathcal{B}_N(x,z,c)\big)\big|_{X_i^{(2)}}\qquad(x\in V,\,i=1,2,\dots,d_2).\end{aligned}$$

Let us fix $i = 1, 2, \dots, d_2$. We then deduce from (2.2) and (3.2) that, for $x \in V$ and $z \in \mathbb{Z}$,

$$\begin{aligned}
&-\frac{1}{N}\sum_{c\in\Omega_{x,N}(X)} p(c)\log\big(\mathcal{B}_N(x,z,c)\big)\Big|_{X_i^{(2)}}\\
&= -\frac{1}{N}\sum_{c'\in\Omega_{x,N-1}(X)} p(c')\sum_{e\in E_{t(c')}} p(e)\log\Big(\big(e^{z\rho}*\Phi_0(x)^{-1}*\Phi_0\big(t(c')\big)*e^{-(z+N-1)\rho}\big)\\
&\qquad *\big(e^{(z+N-1)\rho}*\Phi_0\big(o(e)\big)^{-1}*\Phi_0\big(t(e)\big)*e^{-(z+N)\rho}\big)\Big)\Big|_{X_i^{(2)}}\\
&= -\frac{1}{N}\sum_{c'\in\Omega_{x,N-1}(X)} p(c')\log\Big(e^{z\rho}\cdot\Phi_0(x)^{-1}\cdot\Phi_0(t(c'))\cdot e^{-(z+N-1)\rho}\Big)\Big|_{X_i^{(2)}}\\
&\qquad +\sum_{c'\in\Omega_{x,N-1}(X)} p(c')\sum_{e\in E_{t(c')}} p(e)\log\Big(e^{(z+N-1)\rho}\cdot d\Phi_0(e)\cdot e^{-(z+N)\rho}\Big)\Big|_{X_i^{(2)}}\\
&= -\frac{1}{N}\sum_{k=0}^{N-1}\sum_{c\in\Omega_{x,k}(X)} p(c)\sum_{e\in E_{t(c)}} p(e)\log\Big(e^{(z+k)\rho}\cdot d\Phi_0(e)\cdot e^{-(z+k+1)\rho}\Big)\Big|_{X_i^{(2)}}.
\end{aligned}$$

For $g, h \in G$, we denote by $[g,h] := g\cdot h\cdot g^{-1}\cdot h^{-1}$ the commutator of g and h. Then we have

$$\begin{aligned}
&\sum_{e\in E_{t(c)}} p(e)\log\Big(e^{(z+k)\rho}\cdot d\Phi_0(e)\cdot e^{-(z+k+1)\rho}\Big)\Big|_{X_i^{(2)}}\\
&= \sum_{e\in E_{t(c)}} p(e)\log\Big(\big[e^{(z+k)\rho}, d\Phi_0(e)\big]\cdot d\Phi_0(e)\cdot e^{-\rho}\Big)\Big|_{X_i^{(2)}}\\
&= \sum_{e\in E_{t(c)}} p(e)\log\Big(\big[e^{(z+k)\rho}, d\Phi_0(e)\big]\Big)\Big|_{X_i^{(2)}} + \sum_{e\in E_{t(c)}} p(e)\log\big(d\Phi_0(e)\cdot e^{-\rho}\big)\Big|_{X_i^{(2)}}\\
&= \sum_{e\in E_{t(c)}} p(e)\log\big(d\Phi_0(e)\cdot e^{-\rho}\big)\Big|_{X_i^{(2)}} \qquad (z\in\mathbb{Z},\ k=0,1,\dots,N-1)
\end{aligned}$$

by again using (2.2). It should be noted that this is the most important equality in the proof. Since the function

$$M_i(x) := \sum_{e\in E_x} p(e)\log\big(d\Phi_0(e)\cdot e^{-\rho}\big)\Big|_{X_i^{(2)}} \qquad (i=1,2,\dots,d_2,\ x\in V)$$

satisfies $M_i(\gamma x) = M_i(x)$ for $\gamma \in \Gamma$ and $x \in V$ due to the Γ-invariance of p and the Γ-equivariance of Φ_0, there exists a function $\mathcal{M}_i : V_0 \to \mathbb{R}$ such that $\mathcal{M}_i\big(\pi(x)\big) = M_i(x)$ for $i = 1, 2, \dots, d_2$ and $x \in V$. Moreover, we have $L^k\mathcal{M}_i\big(\pi(x)\big) = L^k M_i(x)$ for $k \in \mathbb{N}$, $i = 1, 2, \dots, d_2$ and $x \in V$ by using the Γ-invariance of p. Then the ergodic theorem (cf. [25, Theorem 3.2]) for the transition operator L gives

$$\begin{aligned}
&-\frac{1}{N}\sum_{c\in\Omega_{x,N}(X)} p(c)\log\big(\mathcal{B}_N(x,z,c)\big)\Big|_{X_i^{(2)}}\\
&= -\frac{1}{N}\sum_{k=0}^{N-1} L^k M_i(x)\\
&= -\sum_{x\in V_0} m(x)\mathcal{M}_i(x) + O\Big(\frac{1}{N}\Big) = -\beta(\Phi_0)\Big|_{X_i^{(2)}} + O\Big(\frac{1}{N}\Big) \quad (x\in V,\ z\in\mathbb{Z}).
\end{aligned} \tag{4.6}$$

We next consider the coefficient of $X_{i*}^{(1)}X_{j*}^{(1)}f\big(\tau_\varepsilon\big(\Phi_0(x)*e^{-z\rho}\big)\big)$ which is given by

$$-\frac{1}{2N}\sum_{c\in\Omega_{x,N}(X)} p(c)\big(\mathcal{B}_N(x,z,c)\big)_{i*}^{(1)}\big(\mathcal{B}_N(x,z,c)\big)_{j*}^{(1)} \qquad (x\in V,\ z\in\mathbb{Z},\ i,j=1,2,\dots,d_1).$$

Fix $i,j=1,2,\dots,d_1$. Then (2.2) and (3.2) imply

$$
\begin{aligned}
&-\frac{1}{2N}\sum_{c\in\Omega_{x,N}(X)} p(c)\big(\mathcal{B}_N(x,z,c)\big)^{(1)}_{i*}\big(\mathcal{B}_N(x,z,c)\big)^{(1)}_{j*}\\
&=-\frac{1}{2N}\sum_{c'\in\Omega_{x,N-1(X)}} p(c')\sum_{e\in E_{t(c')}} p(e)\Big\{\log\big(\mathcal{B}_{N-1}(x,z,c')\big)\big|_{X^{(1)}_i}+\log\big(d\Phi_0(e)\cdot e^{-\rho}\big)\big|_{X^{(1)}_i}\Big\}\\
&\qquad\times\Big\{\log\big(\mathcal{B}_{N-1}(x,z,c')\big)\big|_{X^{(1)}_j}+\log\big(d\Phi_0(e)\cdot e^{-\rho}\big)\big|_{X^{(1)}_j}\Big\}\\
&=-\frac{1}{2N}\Big\{\sum_{c'\in\Omega_{x,N-1(X)}} p(c')\log\big(\mathcal{B}_{N-1}(x,z,c')\big)\big|_{X^{(1)}_i}\log\big(\mathcal{B}_{N-1}(x,z,c')\big)\big|_{X^{(1)}_j}\\
&\qquad+\sum_{e\in E_{t(c')}} p(e)\log\big(d\Phi_0(e)\cdot e^{-\rho}\big)\big|_{X^{(1)}_i}\log\big(d\Phi_0(e)\cdot e^{-\rho}\big)\big|_{X^{(1)}_j}\Big\}\\
&=-\frac{1}{2N}\sum_{k=0}^{N-1}\sum_{c\in\Omega_{x,N}(X)} p(c)\sum_{e\in E_{t(c)}} p(e)\log\big(d\Phi_0(e)\cdot e^{-\rho}\big)\big|_{X^{(1)}_i}\log\big(d\Phi_0(e)\cdot e^{-\rho}\big)\big|_{X^{(1)}_j}
\end{aligned}
$$

for $x\in V$ and $z\in\mathbb{Z}$. Since the function $N_{ij}:V\to\mathbb{R}$ defined by

$$
N_{ij}(x):=\sum_{e\in E_x} p(e)\log\big(d\Phi_0(e)\cdot e^{-\rho}\big)\big|_{X^{(1)}_i}\log\big(d\Phi_0(e)\cdot e^{-\rho}\big)\big|_{X^{(1)}_j}\quad (i,j=1,2,\dots,d_1,\ x\in V)
$$

is Γ-invariant, by the same argument as above, we obtain

$$
\begin{aligned}
&-\frac{1}{2N}\sum_{c\in\Omega_{x,N}(X)} p(c)\big(\mathcal{B}_N(x,z,c)\big)^{(1)}_{i*}\big(\mathcal{B}_N(x,z,c)\big)^{(1)}_{j*}\\
&=-\frac{1}{2}\sum_{e\in E_0}\widetilde{m}(e)\log\big(d\Phi_0(\widetilde{e})\cdot e^{-\rho}\big)\big|_{X^{(1)}_i}\log\big(d\Phi_0(\widetilde{e})\cdot e^{-\rho}\big)\big|_{X^{(1)}_j}+O\Big(\frac{1}{N}\Big).
\end{aligned}\tag{4.7}
$$

Recall that $\{V_1,V_2,\dots,V_{d_1}\}$ denotes an orthonormal basis of $(\mathfrak{g}^{(1)},g_0)$. We especially put $X^{(1)}_i=V_i$ for $i=1,2,\dots,d_1$. Let $\{\omega_1,\omega_2,\dots,\omega_{d_1}\}\subset\mathrm{Hom}(\mathfrak{g}^{(1)},\mathbb{R})\hookrightarrow\mathrm{H}^1(X_0,\mathbb{R})$ be the dual basis of $\{V_1,V_2,\dots,V_{d_1}\}$. Namely, $\omega_i(V_j)=\delta_{ij}$ for $i,j=1,2,\dots,d_1$. It follows from the modified harmonicity and (3.9) that

$$
\begin{aligned}
&\sum_{e\in E_0}\widetilde{m}(e)\log\big(d\Phi_0(\widetilde{e})\cdot e^{-\rho}\big)\big|_{V_i}\log\big(d\Phi_0(\widetilde{e})\cdot e^{-\rho}\big)\big|_{V_j}\\
&=\sum_{e\in E_0}\widetilde{m}(e)\log\big(d\Phi_0(\widetilde{e})\big)\big|_{V_i}\log\big(d\Phi_0(\widetilde{e})\big)\big|_{V_j}-\rho_{\mathbb{R}}(\gamma_p)\big|_{V_i}\rho_{\mathbb{R}}(\gamma_p)\big|_{V_j}\\
&=\sum_{e\in E_0}\widetilde{m}(e)^t\rho_{\mathbb{R}}(\omega_i)(e)^t\rho_{\mathbb{R}}(\omega_j)(e)-\omega_i(\rho_{\mathbb{R}}(\gamma_p))\omega_j(\rho_{\mathbb{R}}(\gamma_p))\\
&=\sum_{e\in E_0}\widetilde{m}(e)\omega_i(e)\omega_j(e)-\langle\gamma_p,\omega_i\rangle\langle\gamma_p,\omega_j\rangle=\langle\!\langle\omega_i,\omega_j\rangle\!\rangle_p=\delta_{ij}.
\end{aligned}\tag{4.8}
$$

Hence, we obtain (4.4) by combining (4.6) with (4.7) and (4.8).

Step 4. We show (3) at the last step. We first discuss the estimate of $\mathcal{I}_1(\varepsilon,N)$. By using (3.6) and (4.2), we have

$$
\begin{aligned}
\Big|\Big(\Phi_0(x)^{-1}*\Phi_0\big(t(c)\big)\Big)^{(k)}_{i*}\Big|&\le C\sum_{\substack{|K_1|+|K_2|\le k\\|K_2|>0}}\Big|\mathcal{P}^{K_1}_*\Big(\Phi_0(x)^{-1}\Big)\Big|\Big|\mathcal{P}^{K_2}\Big(\Phi_0(x)^{-1}\cdot\Phi_0\big(t(c)\big)\Big)\Big|\\
&\le C\sum_{\substack{|K_1|+|K_2|\le k\\|K_2|>0}} N^{|K_2|}\Big|\mathcal{P}^{K_1}_*\Big(e^{-z\rho}*\big(\Phi_0(x)*e^{-z\rho}\big)^{-1}\Big)\Big|
\end{aligned}
$$

for $i = 1, 2, \dots, d_k$ and $k = 1, 2, \dots, r$. Then (3.1) implies that there is a continuous function $Q_1 : G \to \mathbb{R}$ such that

$$\Big|\Big(\Phi_0(x)^{-1} * \Phi_0\big(t(c)\big)\Big)^{(k)}_{i*}\Big| \le |z|^{k-1} Q_1\big(\tau_\varepsilon\big(\Phi_0(x) * e^{-z\rho}\big)\big) \sum_{\substack{|K_1|+|K_2| \le k \\ |K_2|>0}} \varepsilon^{-|K_1|} N^{|K_2|} \tag{4.9}$$

for $i = 1, 2, \dots, d_k$ and $k = 1, 2, \dots, r$. Thus, (3.1) and (4.9) yields

$$\begin{aligned} \big|\big(\mathcal{B}_N(x, 0, c)\big)^{(k)}_{i*}\big| &\le C \sum_{\substack{|L_1|+|L_2|=k \\ |L_1|,|L_2| \ge 0}} \Big|\mathcal{P}^{L_1}_*\Big(\Phi_0(x)^{-1} * \Phi_0\big(t(c)\big)\Big)\Big| \Big|\mathcal{P}^{L_2}_*\big(e^{-N\rho}\big)\Big| \\ &\le C|z|^k Q_2\big(\tau_\varepsilon\big(\Phi_0(x) * e^{-z\rho}\big)\big) \sum_{\substack{|L_1|+|L_2|=k \\ |L_1|,|L_2| \ge 0}} N^{|L_2|} \sum_{\substack{|K_1|+|K_2| \le |L_1| \\ |K_2|>0}} \varepsilon^{-|K_1|} N^{|K_2|} \\ &= C|z|^k Q_2\big(\tau_\varepsilon\big(\Phi_0(x) * e^{-z\rho}\big)\big) F(\varepsilon, N) \end{aligned} \tag{4.10}$$

for some continuous function $Q_2 : G \to \mathbb{R}$, where $F(\varepsilon, N)$ denotes the polynomial of ε and N which satisfies $\varepsilon^{k-2}N^{-1}F(\varepsilon, N) \to 0$ as $N \to \infty$ and $\varepsilon \searrow 0$ with $N^2\varepsilon \searrow 0$.

On the other hand, combining (4.10) with $\rho_{\mathbb{R}}(\gamma_p) \in \mathfrak{g}^{(1)}$, there is some continuous function $Q_3 : G \to \mathbb{R}$ such that

$$\begin{aligned} \frac{\varepsilon^{k-2}}{N}\big|\big(\mathcal{B}_N(x, z, c)\big)^{(k)}_{i*}\big| &= \frac{\varepsilon^{k-2}}{N}\Big|\Big(\big[e^{z\rho}, \mathcal{B}_N(x, 0, c)\big]_* * \mathcal{B}_N(x, 0, c)\Big)^{(k)}_{i*}\Big| \\ &\le C\frac{\varepsilon^{k-2}}{N} \sum_{\substack{|K_1|+|K_2|=k \\ |K_1|,|K_2| \ge 0}} \Big|\mathcal{P}^{K_1}_*\Big(\big[e^{z\rho}, \mathcal{B}_N(x, 0, c)\big]_*\Big)\Big| \Big|\mathcal{P}^{K_2}_*\big(\mathcal{B}_N(x, 0, c)\big)\Big| \\ &\le C|z|^{2k}\frac{\varepsilon^{k-2}}{N} Q_3\big(\tau_\varepsilon\big(\Phi_0(x) * e^{-z\rho}\big)\big) F(\varepsilon, N) \end{aligned} \tag{4.11}$$

for $i = 1, 2, \dots, d_k$, $k = 3, 4, \dots, r$, $x \in V$, $z \in \mathbb{Z}$ and $c \in \Omega_{x,N}(X)$. Hence, we obtain that $\|\mathcal{I}_1(\varepsilon, N)\|_{\infty,q} \to 0$ as $N \to \infty$ and $\varepsilon \searrow 0$ with $N^2\varepsilon \searrow 0$ in $C_{\infty,q}(X \times \mathbb{Z})$ by using (4.11). This follows from $2k < 2r < q$. In the same argument as above, we also obtain $\|\mathcal{I}_2(\varepsilon, N)\|_{\infty,q} \to 0$ as $N \to \infty$ and $\varepsilon \searrow 0$ with $N^2\varepsilon \searrow 0$ in $C_{\infty,q}(X \times \mathbb{Z})$-topology since the order of $|z|$ in $\mathcal{I}_2(\varepsilon, N)$ satisfies $2 \times 2k < 4r < q$.

Finally, we study the estimate of $\mathcal{I}_3(\varepsilon, N)$. We recall that $f \in C_0^\infty(G)$ and the support of the function $\partial^3 f/(\partial g^{(k_1)}_{i_1*}\partial g^{(k_2)}_{i_2*}\partial g^{(k_3)}_{i_3*})$ is included in $\mathrm{supp}\, f$. Therefore, it suffices to show by induction on $k = 1, 2, \dots, r$ that, if $\varepsilon N < 1$,

$$\varepsilon^k\big|\big(\mathcal{B}_N(x, z, c)\big)^{(k)}_{i*}\big| \le |z|^k Q^{(k)}\big(\tau_\varepsilon(\Phi_0(x) * e^{-z\rho}) * \theta\big) \times \varepsilon N \tag{4.12}$$

for some continuous function $Q^{(k)} : G \to \mathbb{R}$, where $\theta \in G$ appears in the remainder term of (4.3). The cases $k = 1$ and $k = 2$ are obvious. Suppose that (4.12) holds for less than k. Then we have

$$\varepsilon^k\big|\big(\mathcal{B}_N(x, z, c)\big)^{(k)}_{i*}\big| \le C\varepsilon^k \sum_{\substack{|K_1|+|K_2| \le k \\ |K_2|>0}} \Big|\mathcal{P}^{K_1}_*\Big(\Phi_0(x)^{-1}\Big)\Big| \Big|\mathcal{P}^{K_2}\Big(\Phi_0(x)^{-1} \cdot \Phi_0\big(t(c)\big)\Big)\Big|$$

by using (3.6). Since

$$\big(\Phi_0(x)^{-1}\big)^{(k_1)}_{i_1*} = \Big(e^{-z\rho} * (\tau_{\varepsilon^{-1}}\theta) * \big(\tau_{\varepsilon^{-1}}(\tau_\varepsilon(\Phi_0(x) * e^{-z\rho}) * \theta)^{-1}\big)\Big)^{(k_1)}_{i_1*} \qquad (k_1 \le k-1),$$

we have inductively

$$\big|\big(\Phi_0(x)^{-1}\big)^{(k_1)}_{i_1*}\big| \le |z|^{k_1} Q\big(\tau_\varepsilon(\Phi_0(x) * e^{-z\rho}) * \theta\big)$$

for a continuous function $Q : G \to \mathbb{R}$ and $k_1 \le k-1$. We thus obtain

$$
\begin{aligned}
&\varepsilon^k \big| \big(\mathcal{B}_N(x,z,c)\big)^{(k)}_{i*} \big| \\
&\le C\varepsilon^k \sum_{\substack{|K_1|+|K_2|\le k \\ |K_2|>0}} N^{|K_2|} \Big| \mathcal{P}^{K_1}_{*} \Big(e^{-z\rho} * (\tau_{\varepsilon^{-1}}\theta) * \big(\tau_{\varepsilon^{-1}}(\tau_\varepsilon(\Phi_0(x) * e^{-z\rho}) * \theta)^{-1}\big)\Big)\Big| \\
&\le C|z|^k Q\big(\tau_\varepsilon(\Phi_0(x) * e^{-z\rho}) * \theta\big) \sum_{\substack{|K_1|+|K_2|\le k \\ |K_2|>0}} \varepsilon^{k-|K_1|+1} N^{|K_2|+1} \\
&\le |z|^k Q^{(k)}\big(\tau_\varepsilon(\Phi_0(x) * e^{-z\rho}) * \theta\big) \times \varepsilon N
\end{aligned}
$$

for some continuous function $Q^{(k)} : G \to \mathbb{R}$. Therefore, (4.12) holds for $k = 1, 2, \dots, r$ and this implies that $\|\mathcal{I}_3(\varepsilon, N)\|_{\infty,q} \to 0$ as $N \to \infty$ and $\varepsilon \searrow 0$ with $N^2\varepsilon \searrow 0$ in $C_{\infty,q}(X \times \mathbb{Z})$ since the order of $|z|$ in $\mathcal{I}_3(\varepsilon, N)$ satisfies $3k < 3r < q$. This completes the proof. □

We now give the proof of Theorem 2.1 by using this lemma. We note that the infinitesimal operator $\mathcal{A}$ in Lemma 4.2 enjoys the following property.

Lemma 4.3 (cf. Robinson [56, page 304])**.** The range of $\lambda - \mathcal{A}$ is dense in $C_\infty(G)$ for some $\lambda > 0$. Namely, $(\lambda - \mathcal{A})\big(C_0^\infty(G)\big)$ is dense in $C_\infty(G)$.

Proof of Theorem 2.1. (1) We follow the argument in Kotani [29, Theorem 4]. Let $N = N(n)$ be the integer satisfying $n^{1/5} \le N < n^{1/5} + 1$ and k_N and r_N be the quotient and the remainder of $([nt] - [ns])/N(n)$, respectively. Note that $r_N < N$. We put $\varepsilon_N := n^{-1/2}$ and $h_N := N\varepsilon_N^2$. Then we have $N = N(n) \to \infty$, $r_N^2\varepsilon_N < N^2\varepsilon_N \le (1 + n^{1/5})^2 \cdot n^{-1/2} \to 0$ and $h_N \le (1 + n^{1/5}) \cdot n^{-1} \to 0$ as $n \to \infty$. We also see that $r_N\varepsilon_N^2 < N\varepsilon_N^2 \le (1 + n^{1/5}) \cdot n^{-1} \to 0$ as $n \to \infty$. Hence, we have

$$
k_N h_N = \frac{[nt] - [ns] - r_N}{N} \cdot N\varepsilon_N^2 = \big([nt] - [ns] - r_N\big)\varepsilon_N^2 \to t - s \quad (n \to \infty).
$$

Since $C_0^\infty(G) \subset \mathrm{Dom}(\mathcal{A}) \subset C_\infty(G)$ and $C_0^\infty(G)$ is dense in $C_\infty(G)$, the operator $\mathcal{A}$ is densely defined in $C_\infty(G)$. We use this fact and Lemma 4.3 to apply Trotter's approximation theorem (cf. Trotter [66] and Kurtz [39]). We obtain, for $f \in C_0^\infty(G)$,

$$
\lim_{n\to\infty} \Big\| \mathcal{L}_p^{Nk_N} \mathcal{P}_{n^{-1/2}} f - \mathcal{P}_{n^{-1/2}} \mathrm{e}^{-(t-s)\mathcal{A}} f \Big\|_{\infty,q} = 0. \tag{4.13}
$$

Then Lemma 4.2 implies

$$
\lim_{n\to\infty} \Big\| \frac{1}{r_N\varepsilon_N^2}\big(I - \mathcal{L}_p^{r_N}\big)\mathcal{P}_{n^{-1/2}} f - \mathcal{P}_{n^{-1/2}}\mathcal{A} f \Big\|_{\infty,q} = 0 \tag{4.14}
$$

for all $f \in C_0^\infty(G)$. We thus have

$$
\begin{aligned}
&\Big\| \mathcal{L}_p^{[nt]-[ns]} P_{n^{-1/2}} f - P_{n^{-1/2}} \mathrm{e}^{-(t-s)\mathcal{A}} f \Big\|_{\infty,q} \\
&\le \Big\| \big(I - \mathcal{L}_p^{r_N}\big)\mathcal{P}_{n^{-1/2}} f \Big\|_{\infty,q} + \Big\| \mathcal{L}_p^{Nk_N} \mathcal{P}_{n^{-1/2}} f - \mathcal{P}_{n^{-1/2}} \mathrm{e}^{-(t-s)\mathcal{A}} f \Big\|_{\infty,q}.
\end{aligned} \tag{4.15}
$$

On the other hand, we have

$$
\begin{aligned}
&\Big\| \big(I - \mathcal{L}_p^{r_N}\big)\mathcal{P}_{n^{-1/2}} f \Big\|_{\infty,q} \\
&\le r_N\varepsilon_N^2 \Big\| \frac{1}{r_N\varepsilon_N^2}\big(I - \mathcal{L}_p^{r_N}\big)\mathcal{P}_{n^{-1/2}} f - \mathcal{P}_{n^{-1/2}}\mathcal{A} f \Big\|_{\infty,q} + r_N\varepsilon_N^2 \big\| \mathcal{P}_{n^{-1/2}}\mathcal{A} f \big\|_{\infty,q} \\
&\le r_N\varepsilon_N^2 \Big\| \frac{1}{r_N\varepsilon_N^2}\big(I - \mathcal{L}_p^{r_N}\big)\mathcal{P}_{n^{-1/2}} f - \mathcal{P}_{n^{-1/2}}\mathcal{A} f \Big\|_{\infty,q} + r_N\varepsilon_N^2 \big\| \mathcal{A} f \big\|_\infty^G.
\end{aligned} \tag{4.16}
$$

We obtain (2.6) for $f \in C_0^\infty(G)$ by combining (4.14), (4.15) and (4.16) with $r_N \varepsilon_N^2 \to 0\,(n \to \infty)$. For $f \in C_\infty(G)$, we also obtain the convergence (2.6) by following the same argument as [25, Theorem 2.1].

(2) For $t > 0$ and $z \in \mathbb{Z}$, we have

$$\begin{aligned}
&\big|\mathcal{L}_p^{[nt]}\mathcal{P}_{n^{-1/2}}f(x_n,z) - \mathrm{e}^{-t\mathcal{A}}f(g)\big| \\
&\leq \big|\mathcal{L}_p^{[nt]}\mathcal{P}_{n^{-1/2}}f(x_n,z) - \mathcal{P}_{n^{-1/2}}\mathrm{e}^{-t\mathcal{A}}f(x_n,z)\big| + \big|\mathcal{P}_{n^{-1/2}}\mathrm{e}^{-t\mathcal{A}}f(x_n,z) - \mathrm{e}^{-t\mathcal{A}}f(g)\big| \\
&\leq (1+|z|^q)\Big\|\mathcal{L}_p^{[nt]}\mathcal{P}_{n^{-1/2}}f - \mathcal{P}_{n^{-1/2}}\mathrm{e}^{-t\mathcal{A}}f\Big\|_{\infty,q} \\
&\qquad + \Big|\mathrm{e}^{-t\mathcal{A}}f\Big(\tau_{n^{-1/2}}\big(\Phi_0(x_n) * \exp(-z\rho_{\mathbb{R}}(\gamma_p))\big)\Big) - \mathrm{e}^{-t\mathcal{A}}f(g)\Big|.
\end{aligned}$$

We thus obtain (2.8) by (2.6) and the continuity of the function $\mathrm{e}^{-t\mathcal{A}}f : G \to \mathbb{R}$. This completes the proof of Theorem 2.1. □

We now give several properties of $\beta(\Phi_0)$.

Proposition 4.4. (1) If the random walk on X is m-symmetric, then $\beta(\Phi_0) = \mathbf{0}_{\mathfrak{g}}$.
(2) Let Φ_0, $\widehat{\Phi}_0 : X \to G$ be two modified harmonic realizations. Then

$$\beta(\Phi_0) = \beta(\widehat{\Phi}_0) - \big[\rho_{\mathbb{R}}(\gamma_p), \log\big(\Phi_0(x)^{-1}\cdot\widehat{\Phi}_0(x)\big)\big]\big|_{\mathfrak{g}^{(2)}} \qquad (x \in V).$$

In particular, if either

- $\log \Phi_0(x_*)\big|_{\mathfrak{g}^{(1)}} = \log\widehat{\Phi}_0(x_*)\big|_{\mathfrak{g}^{(1)}}$ for some reference point $x_* \in V$, or
- $\rho_{\mathbb{R}}(\gamma_p) = \mathbf{0}_{\mathfrak{g}}$

holds, then we have $\beta(\Phi_0) = \beta(\widehat{\Phi}_0)$.

Proof. Assertion (1) is easily obtained as follows:

$$\beta(\Phi_0) = \frac{1}{2}\sum_{e \in E_0}\big(\widetilde{m}(e) - \widetilde{m}(\overline{e})\big)\log\big(d\Phi_0(\widetilde{e})\big)\big|_{\mathfrak{g}^{(2)}} = \mathbf{0}_{\mathfrak{g}}.$$

Next we show Assertion (2). We set $\Psi(x) := \Phi_0(x)^{-1}\cdot\widehat{\Phi}_0(x)$ for $x \in V$. We note that the map $\Psi : X \to G$ is Γ-invariant. Since the $\mathfrak{g}^{(1)}$-components of Φ_0 and $\widehat{\Phi}_0$ are uniquely determined up to $\mathfrak{g}^{(1)}$-translation, there exists a constant vector $C \in \mathfrak{g}^{(1)}$ such that $\log\big(\Psi(x)\big)\big|_{\mathfrak{g}^{(1)}} = C$ for $x \in V$. Define a function $F_i : V \to \mathbb{R}$ by $F_i(x) := \log\big(\Psi(x)\big)\big|_{X_i^{(2)}}$ for $i = 1, 2, \dots, d_2$ and $x \in V$. Then we see that the function F_i is Γ-invariant. Hence, there is a function $\widehat{F}_i : V_0 \to \mathbb{R}$ satisfying $\widehat{F}_i\big(\pi(x)\big) = F_i(x)$ for $x \in V$. Then we obtain

$$\begin{aligned}
\beta(\Phi_0) &= \sum_{e \in E_0}\widetilde{m}(e)\log\Big(\Psi\big(o(\widetilde{e})\big)\cdot\big(d\widehat{\Phi}_0(\widetilde{e})\cdot e^{-\rho}\big)\cdot e^{\rho}\cdot\Psi\big(t(\widetilde{e})\big)^{-1}\cdot e^{-\rho}\Big)\Big|_{\mathfrak{g}^{(2)}} \\
&= \beta(\widehat{\Phi}_0) - \sum_{e \in E_0}\widetilde{m}(e)\Big\{\log\Big(\Psi\big(t(\widetilde{e})\big)\Big)\Big|_{\mathfrak{g}^{(2)}} - \log\Big(\Psi\big(o(\widetilde{e})\big)\Big)\Big|_{\mathfrak{g}^{(2)}}\Big\} - [\rho_{\mathbb{R}}(\gamma_p), C]\big|_{\mathfrak{g}^{(2)}} \\
&= \beta(\widehat{\Phi}_0) - \sum_{i=1}^{d_2}\big({}_{C_1(X_0,\mathbb{R})}\langle\gamma_p, d\widehat{F}_i\rangle_{C^1(X_0,\mathbb{R})}\big)X_i^{(2)} - [\rho_{\mathbb{R}}(\gamma_p), C]\big|_{\mathfrak{g}^{(2)}} \\
&= \beta(\widehat{\Phi}_0) - \sum_{i=1}^{d_2}\big({}_{C_0(X_0,\mathbb{R})}\langle\partial(\gamma_p), \widehat{F}_i\rangle_{C^0(X_0,\mathbb{R})}\big)X_i^{(2)} - [\rho_{\mathbb{R}}(\gamma_p), C]\big|_{\mathfrak{g}^{(2)}} \\
&= \beta(\widehat{\Phi}_0) - [\rho_{\mathbb{R}}(\gamma_p), C]\big|_{\mathfrak{g}^{(2)}},
\end{aligned}$$

where we used (3.1) for the second line and $\gamma_p \in \mathrm{H}_1(X_0,\mathbb{R})$ for the fourth line. This completes the proof. □

4.3 Proof of Theorem 2.2

We now assume the centered condition **(A3)**: $\rho_{\mathbb{R}}(\gamma_p) = \mathbf{0}_{\mathfrak{g}}$, throughout this subsection. For $k = 1, 2, \dots, r$, we denote by $(G^{(k)}, \cdot)$ and $(G^{(k)}, *)$ the connected and simply connected nilpotent Lie group of step k and the corresponding limit group whose Lie algebras are $\big(\mathfrak{g}^{(1)} \oplus \mathfrak{g}^{(2)} \oplus \cdots \oplus \mathfrak{g}^{(k)}, [\cdot,\cdot]\big)$ and $\big(\mathfrak{g}^{(1)} \oplus \mathfrak{g}^{(2)} \oplus \cdots \oplus \mathfrak{g}^{(k)}, [\![\cdot,\cdot]\!]\big)$, respectively. For the piecewise smooth stochastic process $(\mathcal{Y}_t^{(n)})_{0 \le t \le 1} = (\mathcal{Y}_t^{(n),1}, \mathcal{Y}_t^{(n),2}, \dots, \mathcal{Y}_t^{(n),r})_{0 \le t \le 1}$ defined in Section 2, we define its truncated process by

$$\mathcal{Y}_t^{(n;\,k)} = \big(\mathcal{Y}_t^{(n),1}, \mathcal{Y}_t^{(n),2}, \dots, \mathcal{Y}_t^{(n),k}\big) \in G^{(k)}$$

for $0 \le t \le 1$ and $k = 1, 2, \dots, r$ in the $(\cdot)$-coordinate system. To complete the proof of Theorem 2.2, it is sufficient to show the tightness of $\{\mathbf{P}^{(n)}\}_{n=1}^\infty$ (Lemma 4.5) and the convergence of the finite dimensional distribution of $\{\mathcal{Y}_\cdot^{(n)}\}_{n=1}^\infty$ (Lemma 4.8).

In the former part of this subsection, we aim to show the following.

Lemma 4.5. Under **(A3)**, the family $\{\mathbf{P}^{(n)}\}_{n=1}^\infty$ is tight in $C^{0,\alpha\text{-Höl}}_{\mathbf{1}_G}([0,1]; G)$, where α is an arbitrary real number less than $1/2$.

As the first step of the proof of Lemma 4.5, we prepare the following lemma.

Lemma 4.6. Let m, n be positive integers. Then there exists a constant $C > 0$ which is independent of n (however, it may depend on m) such that

$$\mathbb{E}^{\mathbb{P}_{x_*}}\Big[d_{\mathrm{CC}}(\mathcal{Y}_s^{(n;\,2)}, \mathcal{Y}_t^{(n;\,2)})^{4m}\Big] \le C(t-s)^{2m} \qquad (0 \le s \le t \le 1). \tag{4.17}$$

Proof. The proof is partially based on Bayer–Friz [4, Proposition 4.3]. We emphasize that the relation between the modified harmonicity of Φ_0 and martingale theory such as Lemma 3.3 is significant in this proof. We split the proof into several steps.

Step 1. At the beginning, we show

$$\mathbb{E}^{\mathbb{P}_{x_*}}\Big[d_{\mathrm{CC}}(\mathcal{Y}_{t_k}^{(n;\,2)}, \mathcal{Y}_{t_\ell}^{(n;\,2)})^{4m}\Big] \le C\Big(\frac{\ell-k}{n}\Big)^{2m} \qquad \big(n, m \in \mathbb{N},\ t_k, t_\ell \in \mathcal{D}_n\ (k \le \ell)\big) \tag{4.18}$$

for some $C > 0$ independent of n (depending on m). By recalling the equivalence of two homogeneous norms $\|\cdot\|_{\mathrm{CC}}$ and $\|\cdot\|_{\mathrm{hom}}$ (cf. Proposition 3.1), we readily see that (4.18) is equivalent to the existence of positive constants $C^{(1)}$ and $C^{(2)}$ independent of n such that

$$\mathbb{E}^{\mathbb{P}_{x_*}}\Big[\big\|\log\big((\mathcal{Y}_{t_k}^{(n)})^{-1}\cdot\mathcal{Y}_{t_\ell}^{(n)}\big)\big|_{\mathfrak{g}^{(1)}}\big\|_{\mathfrak{g}^{(1)}}^{4m}\Big] \le C^{(1)}\Big(\frac{\ell-k}{n}\Big)^{2m}, \tag{4.19}$$

$$\mathbb{E}^{\mathbb{P}_{x_*}}\Big[\big\|\log\big((\mathcal{Y}_{t_k}^{(n)})^{-1}\cdot\mathcal{Y}_{t_\ell}^{(n)}\big)\big|_{\mathfrak{g}^{(2)}}\big\|_{\mathfrak{g}^{(2)}}^{2m}\Big] \le C^{(2)}\Big(\frac{\ell-k}{n}\Big)^{2m}. \tag{4.20}$$

Step 2. We now show (4.19). We see

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}_{x_*}}\Big[\big\|\log\big((\mathcal{Y}_{t_k}^{(n)})^{-1}\cdot\mathcal{Y}_{t_\ell}^{(n)}\big)\big|_{\mathfrak{g}^{(1)}}\big\|_{\mathfrak{g}^{(1)}}^{4m}\Big] \\
&= \Big(\frac{1}{\sqrt{n}}\Big)^{4m}\mathbb{E}^{\mathbb{P}_{x_*}}\Big[\Big(\sum_{i=1}^{d_1}\log(\xi_k^{-1}\cdot\xi_\ell)\big|_{X_i^{(1)}}^2\Big)^{2m}\Big] \\
&\le \Big(\frac{1}{\sqrt{n}}\Big)^{4m}\cdot d_1^{2m}\max_{i=1,2,\dots,d_1}\max_{x\in\mathcal{F}}\Big\{\sum_{c\in\Omega_{x,\ell-k}(X)} p(c)\log\Big(\Phi_0(x)^{-1}\cdot\Phi_0\big(t(c)\big)\Big)\Big|_{X_i^{(1)}}^{4m}\Big\},
\end{aligned} \tag{4.21}$$

where $\mathcal{F}$ stands for the fundamental domain in X containing the reference point $x_* \in V$. For $i = 1, 2, \dots, d_1$, $x \in \mathcal{F}$, $N \in \mathbb{N}$ and $c = (e_1, e_2, \dots, e_N) \in \Omega_{x,N}(X)$, we put

$$\mathcal{M}_N^{(i,x)}(c) = \mathcal{M}_N^{(i,x)}(\Phi_0; c) := \log\Big(\Phi_0(x)^{-1}\cdot\Phi_0\big(t(c)\big)\Big)\Big|_{X_i^{(1)}} = \sum_{j=1}^N \log\big(d\Phi_0(e_j)\big)\big|_{X_i^{(1)}}.$$

By Lemma 3.3, $\{\mathcal{M}_N^{(i,x)}\}_{N=1}^{\infty}$ is an $\mathbb{R}$-valued martingale for every $i = 1, 2, \dots, d_1$ and $x \in \mathcal{F}$. Therefore, we apply the Burkholder–Davis–Gundy inequality with the exponent $4m$ to obtain

$$\begin{aligned}\sum_{c\in\Omega_{x,N}(X)} p(c)\big(\mathcal{M}_N^{(i,x)}(c)\big)^{4m} &= \sum_{c\in\Omega_{x,N}(X)} p(c)\Big(\sum_{j=1}^{N}\log\big(d\Phi_0(e_j)\big)\big|_{X_i^{(1)}}\Big)^{4m} \\ &\le \mathcal{C}_{(4m)}^{4m}\sum_{c\in\Omega_{x,N}(X)} p(c)\Big(\sum_{j=1}^{N}\log\big(d\Phi_0(e_j)\big)\big|_{X_i^{(1)}}^2\Big)^{2m} \\ &\le \mathcal{C}_{(4m)}^{4m}\|d\Phi_0\|_\infty^{4m}N^{2m} \qquad (4.22)\end{aligned}$$

for $i = 1, 2, \dots, d_1$, $x \in \mathcal{F}$ and $N \in \mathbb{N}$, where $\mathcal{C}_{(4m)}$ stands for the positive constant which appears in the Burkholder–Davis–Gundy inequality with the exponent $4m$. In particular, by putting $N = \ell - k$, (4.22) leads to

$$\sum_{c\in\Omega_{x,\ell-k}(X)} p(c)\log\Big(\Phi_0(x)^{-1}\cdot\Phi_0\big(t(c)\big)\Big)\Big|_{X_i^{(1)}}^{4m} \le \mathcal{C}_{(4m)}^{4m}\|d\Phi_0\|_\infty^{4m}(\ell-k)^{2m}. \qquad (4.23)$$

Thus, we obtain

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}_{x*}}\Big[\big\|\log\big((\mathcal{Y}_{t_k}^{(n)})^{-1}\cdot\mathcal{Y}_{t_\ell}^{(n)}\big)\big|_{\mathfrak{g}^{(1)}}\big\|_{\mathfrak{g}^{(1)}}^{4m}\Big] \\ &\le d_1^{2m}\mathcal{C}_{(4m)}^{4m}\|d\Phi_0\|_\infty^{4m}\cdot\Big(\frac{\ell-k}{n}\Big)^{2m} = C^{(1)}\Big(\frac{\ell-k}{n}\Big)^{2m}\end{aligned}$$

by combining (4.21) with (4.23), which is the desired estimate (4.19).

Step 3. Next we prove (4.20). In the similar way to (4.21), we also have

$$\begin{aligned}&\mathbb{E}^{\mathbb{P}_{x*}}\Big[\big\|\log\big((\mathcal{Y}_{t_k}^{(n)})^{-1}\cdot\mathcal{Y}_{t_\ell}^{(n)}\big)\big|_{\mathfrak{g}^{(2)}}\big\|_{\mathfrak{g}^{(2)}}^{2m}\Big] \\ &\le\Big(\frac{1}{n}\Big)^{2m}\cdot d_2^{2m}\max_{i=1,2,\dots,d_2}\max_{x\in\mathcal{F}}\Big\{\sum_{c\in\Omega_{x,\ell-k}(X)} p(c)\log\Big(\Phi_0(x)^{-1}\cdot\Phi_0\big(t(c)\big)\Big)\Big|_{X_i^{(2)}}^{2m}\Big\}. \qquad (4.24)\end{aligned}$$

An elementary inequality $(a_1 + a_2 + \cdots + a_K)^{2m} \le K^{2m-1}(a_1^{2m} + a_2^{2m}\cdots + a_K^{2m})$ yields

$$\begin{aligned}&\log\Big(\Phi_0(x)^{-1}\cdot\Phi_0\big(t(c)\big)\Big)\Big|_{X_i^{(2)}}^{2m} \\ &= \log\Big(\Phi_0\big(o(e_1)\big)^{-1}\cdot\Phi_0\big(t(e_1)\big)\cdots\cdot\Phi_0\big(o(e_{\ell-k})\big)^{-1}\cdot\Phi_0\big(t(e_{\ell-k})\big)\Big)\Big|_{X_i^{(2)}}^{2m} \\ &= \Big(\sum_{j=1}^{\ell-k}\log\big(d\Phi_0(e_j)\big)\big|_{X_i^{(2)}} - \frac{1}{2}\sum_{1\le j_1<j_2\le\ell-k}\ \sum_{1\le\lambda<\nu\le d_1}[\![X_\lambda^{(1)}, X_\nu^{(1)}]\!]\big|_{X_i^{(2)}} \\ &\qquad\times\Big\{\log\big(d\Phi_0(e_{j_1})\big)\big|_{X_\lambda^{(1)}}\log\big(d\Phi_0(e_{j_2})\big)\big|_{X_\nu^{(1)}} \\ &\qquad - \log\big(d\Phi_0(e_{j_1})\big)\big|_{X_\nu^{(1)}}\log\big(d\Phi_0(e_{j_2})\big)\big|_{X_\lambda^{(1)}}\Big\}\Big)^{2m} \\ &\le 3^{2m-1}\Big\{\Big(\sum_{j=1}^{\ell-k}\log\big(d\Phi_0(e_j)\big)\big|_{X_i^{(2)}}\Big)^{2m} \\ &\qquad + L\max_{1\le\lambda<\nu\le d_1}\Big(\sum_{1\le j_1<j_2\le\ell-k}\log\big(d\Phi_0(e_{j_1})\big)\big|_{X_\lambda^{(1)}}\log\big(d\Phi_0(e_{j_2})\big)\big|_{X_\nu^{(1)}}\Big)^{2m} \\ &\qquad + L\max_{1\le\lambda<\nu\le d_1}\Big(\sum_{1\le j_1<j_2\le\ell-k}\log\big(d\Phi_0(e_{j_1})\big)\big|_{X_\nu^{(1)}}\log\big(d\Phi_0(e_{j_2})\big)\big|_{X_\lambda^{(1)}}\Big)^{2m}\Big\}, \qquad (4.25)\end{aligned}$$

where we put

$$L := \frac{1}{2} \max_{i=1,2,\dots,d_2} \max_{1 \le \lambda < \nu \le d_1} \big| [\![X_\lambda^{(1)}, X_\nu^{(1)}]\!] \big|_{X_i^{(2)}} \big|.$$

We fix $i = 1, 2, \dots, d_2$. Then the Jensen inequality gives

$$\begin{aligned} \Big(\sum_{j=1}^{\ell-k} \log \big(d\Phi_0(e_j) \big) \big|_{X_i^{(2)}} \Big)^{2m} &= (\ell-k)^{2m} \Big(\sum_{j=1}^{\ell-k} \frac{1}{\ell-k} \log \big(d\Phi_0(e_j) \big) \big|_{X_i^{(2)}} \Big)^{2m} \\ &\le (\ell-k)^{2m} \sum_{j=1}^{\ell-k} \frac{1}{\ell-k} \log \big(d\Phi_0(e_j) \big) \big|_{X_i^{(2)}}^{2m} \\ &\le (\ell-k)^{2m} \| d\Phi_0 \|_\infty^{4m}. \end{aligned} \tag{4.26}$$

For $1 \le \lambda < \nu \le d_1$, $x \in \mathcal{F}$, $N \in \mathbb{N}$ and $c = (e_1, e_2, \dots, e_N) \in \Omega_{x,N}(X)$, we put

$$\begin{aligned} \widetilde{\mathcal{M}}_N^{(\lambda,\nu,x)}(c) = \widetilde{\mathcal{M}}_N^{(\lambda,\nu,x)}(\Phi_0; c) &:= \sum_{1 \le j_1 < j_2 \le N} \log \big(d\Phi_0(e_{j_1}) \big) \big|_{X_\lambda^{(1)}} \log \big(d\Phi_0(e_{j_2}) \big) \big|_{X_\nu^{(1)}} \\ &= \sum_{j_2=2}^{N} \log \big(d\Phi_0(e_{j_2}) \big) \big|_{X_\nu^{(1)}} \Big(\sum_{j_1=1}^{j_2-1} \log \big(d\Phi_0(e_{j_1}) \big) \big|_{X_\lambda^{(1)}} \Big). \end{aligned}$$

We clearly observe that $\{\widetilde{\mathcal{M}}_N^{(\lambda,\nu,x)}\}_{N=1}^\infty$ is an $\mathbb{R}$-valued martingale for every $1 \le \lambda < \nu \le d$ and $x \in \mathcal{F}$ due to Lemma 3.3. Hence, we apply the Burkholder–Davis–Gundy inequality with the exponent $2m$ to obtain

$$\begin{aligned} &\sum_{c \in \Omega_{x,N}(X)} p(c) \big(\widetilde{\mathcal{M}}_N^{(\lambda,\nu,x)}(c) \big)^{2m} \\ &\le \mathcal{C}_{(2m)}^{2m} \sum_{c \in \Omega_{x,N}(X)} p(c) \Big\{ \sum_{j_2=2}^{N} \log \big(d\Phi_0(e_{j_2}) \big) \big|_{X_\nu^{(1)}}^2 \Big(\sum_{j_1=1}^{j_2-1} \log \big(d\Phi_0(e_{j_1}) \big) \big|_{X_\lambda^{(1)}} \Big)^2 \Big\}^m \\ &\le \mathcal{C}_{(2m)}^{2m} \sum_{c \in \Omega_{x,N}(X)} p(c) N^m \sum_{j_2=2}^{N} \frac{1}{N-1} \log \big(d\Phi_0(e_{j_2}) \big) \big|_{X_\nu^{(1)}}^{2m} \Big(\sum_{j_1=1}^{j_2-1} \log \big(d\Phi_0(e_{j_1}) \big) \big|_{X_\lambda^{(1)}} \Big)^{2m} \\ &\le \mathcal{C}_{(2m)}^{2m} N^m \sum_{j_2=2}^{N} \frac{1}{N-1} \Big\{ \sum_{c \in \Omega_{x,N}(X)} p(c) \log \big(d\Phi_0(e_{j_2}) \big) \big|_{X_\nu^{(1)}}^{4m} \Big\}^{1/2} \\ &\qquad \times \Big\{ \sum_{c \in \Omega_{x,N}(X)} p(c) \Big(\sum_{j_1=1}^{j_2-1} \log \big(d\Phi_0(e_{j_1}) \big) \big|_{X_\lambda^{(1)}} \Big)^{4m} \Big\}^{1/2} \\ &\le \mathcal{C}_{(2m)}^{2m} \| d\Phi_0 \|_\infty^{2m} N^m \sum_{j_2=2}^{N} \frac{1}{N-1} \Big\{ \sum_{c \in \Omega_{x,N}(X)} p(c) \Big(\sum_{j_1=1}^{j_2-1} \log \big(d\Phi_0(e_{j_1}) \big) \big|_{X_\lambda^{(1)}} \Big)^{4m} \Big\}^{1/2}, \end{aligned} \tag{4.27}$$

where we used Jensen's inequality for the third line and Schwarz' inequality for the fourth line. Then we have

$$\begin{aligned}
&\sum_{c\in\Omega_{x,N}(X)} p(c)\Big(\sum_{j_1=1}^{j_2-1} \log\big(d\Phi_0(e_{j_1})\big)\big|_{X_\lambda^{(1)}}\Big)^{4m}\\
&\le \mathcal{C}_{(4m)}^{4m} \sum_{c\in\Omega_{x,N}(X)} p(c)\Big(\sum_{j_1=1}^{j_2-1} \log\big(d\Phi_0(e_{j_1})\big)\big|_{X_\lambda^{(1)}}^2\Big)^{2m}\\
&= \mathcal{C}_{(4m)}^{4m}(j_2-1)^{2m} \sum_{c\in\Omega_{x,N}(X)} p(c)\Big(\sum_{j_1=1}^{j_2-1} \frac{1}{j_2-1}\log\big(d\Phi_0(e_{j_1})\big)\big|_{X_\lambda^{(1)}}^2\Big)^{2m}\\
&\le \mathcal{C}_{(4m)}^{4m} j_2^{2m} \sum_{c\in\Omega_{x,N}(X)} p(c)\sum_{j_1=1}^{j_2-1} \frac{1}{j_2-1}\log\big(d\Phi_0(e_{j_1})\big)\big|_{X_\lambda^{(1)}}^{4m} \le \mathcal{C}_{(4m)}^{4m}\|d\Phi_0\|_\infty^{4m} j_2^{2m} \qquad (4.28)
\end{aligned}$$

by applying the Burkholder–Davis–Gundy inequality with the exponent $4m$. It follows from (4.27) and (4.28) that

$$\begin{aligned}
\sum_{c\in\Omega_{x,N}(X)} p(c)\big(\widetilde{\mathcal{M}}_N^{(\lambda,\nu,x)}(c)\big)^{2m} &\le \mathcal{C}_{(2m)}^{2m}\|d\Phi_0\|_\infty^{2m}N^m\sum_{j_2=2}^{N}\frac{1}{N-1}\Big(\mathcal{C}_{(4m)}^{4m}\|d\Phi_0\|_\infty^{4m} j_2^{2m}\Big)^{1/2}\\
&\le \mathcal{C}_{(2m)}^{2m}\mathcal{C}_{(4m)}^{2m}\|d\Phi_0\|_\infty^{4m}N^m\sum_{j_2=2}^{N}\frac{1}{N-1}\cdot N^m\\
&= \mathcal{C}_{(2m)}^{2m}\mathcal{C}_{(4m)}^{2m}\|d\Phi_0\|_\infty^{4m}N^{2m}. \qquad (4.29)
\end{aligned}$$

We now put $N=\ell-k$. Then (4.29) implies

$$\begin{aligned}
&\sum_{c\in\Omega_{x,\ell-k}(X)} p(c)\Big\{\Big(\sum_{1\le j_1<j_2\le \ell-k}\log\big(d\Phi_0(e_{j_1})\big)\big|_{X_\lambda^{(1)}}\log\big(d\Phi_0(e_{j_2})\big)\big|_{X_\nu^{(1)}}\Big)^{2m}\\
&\qquad+\Big(\sum_{1\le j_1<j_2\le \ell-k}\log\big(d\Phi_0(e_{j_1})\big)\big|_{X_\nu^{(1)}}\log\big(d\Phi_0(e_{j_2})\big)\big|_{X_\lambda^{(1)}}\Big)^{2m}\Big\}\\
&\le 2\mathcal{C}_{(2m)}^{2m}\mathcal{C}_{(4m)}^{2m}\|d\Phi_0\|_\infty^{4m}(\ell-k)^{2m} \qquad (1\le\lambda<\nu\le d_1). \qquad (4.30)
\end{aligned}$$

By combining (4.24) with (4.25), (4.26) and (4.30), we obtain

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}_{x*}}\Big[\big\|\log\big((\mathcal{Y}_{t_k}^{(n)})^{-1}\cdot\mathcal{Y}_{t_\ell}^{(n)}\big)\big|_{\mathfrak{g}^{(2)}}\big\|_{\mathfrak{g}^{(2)}}^{2m}\Big]\\
&\le\Big(\frac{1}{n}\Big)^{2m}d_2^{2m}3^{2m-1}\|d\Phi_0\|_\infty^{4m}\Big\{1+2L\mathcal{C}_{(2m)}^{2m}\mathcal{C}_{(4m)}^{2m}\Big\}(\ell-k)^{2m}=C^{(2)}\Big(\frac{\ell-k}{n}\Big)^{2m}.
\end{aligned}$$

This is the desired estimate (4.20), and thus we have shown (4.18).

Step 4. We finally prove (4.17). Suppose that $t_k\le s\le t_{k+1}$ and $t_\ell\le t\le t_{\ell+1}$ for some $1\le k\le\ell\le n$. Since the process $\mathcal{Y}^{(n)}$ is given by the d_{CC}-geodesic interpolation, we have

$$\begin{aligned}
d_{\mathrm{CC}}(\mathcal{Y}_s^{(n;\,2)},\mathcal{Y}_{t_{k+1}}^{(n;\,2)})&=(k-ns)d_{\mathrm{CC}}(\mathcal{Y}_{t_k}^{(n;\,2)},\mathcal{Y}_{t_{k+1}}^{(n;\,2)}),\\
d_{\mathrm{CC}}(\mathcal{Y}_{t_\ell}^{(n;\,2)},\mathcal{Y}_t^{(n;\,2)})&=(nt-\ell)d_{\mathrm{CC}}(\mathcal{Y}_{t_\ell}^{(n;\,2)},\mathcal{Y}_{t_{\ell+1}}^{(n;\,2)}).
\end{aligned}$$

By using (4.18) and the triangle inequality, we have

$$\begin{aligned}
&\mathbb{E}^{\mathbb{P}_{x*}}\Big[d_{\mathrm{CC}}(\mathcal{Y}_s^{(n;\,2)},\mathcal{Y}_t^{(n;\,2)})^{4m}\Big]\\
&\le 3^{4m-1}\Big\{(k+1-ns)^{4m}\cdot C\Big(\frac{1}{n}\Big)^{2m}+C\Big(\frac{\ell-k-1}{n}\Big)^{2m}+(nt-\ell)^{4m}\cdot C\Big(\frac{1}{n}\Big)^{2m}\Big\}\\
&\le C\Big\{(t_{k+1}-s)^{2m}+(t_\ell-t_{k+1})^{2m}+(t-t_\ell)^{2m}\Big\}\le C(t-s)^{2m},
\end{aligned}$$

which is the desired estimate (4.17) and we have proved Lemma 4.6. □

In what follows, we write $d\mathcal{Y}_{s,t}^{(n)*} := (\mathcal{Y}_s^{(n)})^{-1} * \mathcal{Y}_t^{(n)}$ for $n \in \mathbb{N}$ and $0 \le s \le t \le 1$. By using Lemma 4.6, we obtain the following.

Lemma 4.7. For $m, n \in \mathbb{N}$, $k = 1, 2, \dots, r$ and $\alpha < \frac{2m-1}{4m}$, there exist an $\mathcal{F}_\infty$-measurable set $\Omega_k^{(n)} \subset \Omega_{x_*}(X)$ and a non-negative random variable $\mathcal{K}_k^{(n)} \in L^{4m}(\Omega_{x_*}(X) \to \mathbb{R}; \mathbb{P}_{x_*})$ such that $\mathbb{P}_{x_*}(\Omega_k^{(n)}) = 1$ and

$$d_{\mathrm{CC}}\big(\mathcal{Y}_s^{(n;k)}(c), \mathcal{Y}_t^{(n;k)}(c)\big) \le \mathcal{K}_k^{(n)}(c)(t-s)^\alpha \qquad (c \in \Omega_k^{(n)},\, 0 \le s \le t \le 1). \tag{4.31}$$

Proof. We partially follow the original proof of Lyons extension theorem (cf. [46, Theorem 2.2.1]) in rough path theory. As mentioned in Section 2, his proof is not available in the case where the nilpotent Lie group G is not necessarily free due to technical reasons. Hence, we need to make a careful modification of it. We show (4.31) by induction on the step number $k = 1, 2, \dots, r$.

Step 1. In the cases $k = 1, 2$, we have already obtained (4.31) in Lemma 4.6. Indeed, (4.31) for $k = 1, 2$ are readily obtained by a simple application of the Kolmogorov–Chentov criterion with the bound

$$\|\mathcal{K}_k^{(n)}\|_{L^{4m}(\mathbb{P}_{x_*})} \le \frac{5C}{(1-2^{-\theta})(1-2^{\alpha-\theta})} \qquad (n, m \in \mathbb{N},\, k = 1, 2), \tag{4.32}$$

where $\theta = (2m-1)/4m$ and C is a constant independent of n, which appears in the right-hand side of (4.17). See e.g., Stroock [59, Theorem 4.3.2] for details.

Step 2. Suppose that (4.31) holds up to step k. Then, for $n \in \mathbb{N}$, there are $\mathcal{F}_\infty$-measurable sets $\{\widehat{\Omega}_j^{(n)}\}_{j=1}^k$ and non-negative random variables $\{\widehat{\mathcal{K}}_j^{(n)}\}_{j=1}^k$ satisfying that $\mathbb{P}_{x_*}(\widehat{\Omega}_j^{(n)}) = 1$ for $j = 1, 2, \dots, k$ and

$$\big\|\big(d\mathcal{Y}_{s,t}^{(n)*}(c)\big)^{(j)}\big\|_{\mathbb{R}^{d_j}} \le \widehat{\mathcal{K}}_j^{(n)}(c)(t-s)^{j\alpha} \quad (j = 1, 2, \dots, k,\, c \in \widehat{\Omega}_j^{(n)},\, 0 \le s \le t \le 1) \tag{4.33}$$

with $\widehat{\mathcal{K}}_j^{(n)} \in L^{4m/j}(\Omega_{x_*}(X) \to \mathbb{R}; \mathbb{P}_{x_*})$ for $m \in \mathbb{N}$ and $j = 1, 2, \dots, k$.

We fix $0 \le s \le t \le 1$ and $n \in \mathbb{N}$. Set $\widehat{\Omega}_{k+1}^{(n)} = \bigcap_{j=1}^k \widehat{\Omega}_j^{(n)}$. We denote by Δ the partition $\{s = t_0 < t_1 < \cdots < t_N = t\}$ of the time interval $[s, t]$ independent of $n \in \mathbb{N}$. We define two $G^{(k+1)}$-valued random variables $\mathcal{Z}_{s,t}^{(n)}$ and $\mathcal{Z}(\Delta)_{s,t}^{(n)}$ by

$$\big(\mathcal{Z}_{s,t}^{(n)}\big)^{(j)} := \begin{cases} \big(d\mathcal{Y}_{s,t}^{(n)*}\big)^{(j)}, & (j = 1, 2, \dots, k) \\ \mathbf{0} & (j = k+1) \end{cases}, \quad \mathcal{Z}(\Delta)_{s,t}^{(n)} := \mathcal{Z}_{t_0,t_1}^{(n)} * \mathcal{Z}_{t_1,t_2}^{(n)} * \cdots * \mathcal{Z}_{t_{N-1},t_N}^{(n)},$$

respectively. For $i = 1, 2, \dots, d_{k+1}$, (3.1) and (4.33) imply

$$\begin{aligned}
&\Big|\big(\mathcal{Z}(\Delta)_{s,t}^{(n)}(c)\big)_{i*}^{(k+1)} - \big(\mathcal{Z}(\Delta \setminus \{t_{N-1}\})_{s,t}^{(n)}(c)\big)_{i*}^{(k+1)}\Big| \\
&= \Big|\big(\mathcal{Z}_{t_{N-2},t_{N-1}}^{(n)}(c) * \mathcal{Z}_{t_{N-1},t_N}^{(n)}(c)\big)_{i*}^{(k+1)} - \big(\mathcal{Z}_{t_{N-2},t_N}^{(n)}(c)\big)_{i*}^{(k+1)}\Big| \\
&= \Bigg|\sum_{\substack{|K_1|+|K_2|=k+1 \\ |K_1|,|K_2|\ge 0}} C_{K_1,K_2} \mathcal{P}_*^{K_1}\big(\mathcal{Z}_{t_{N-2},t_{N-1}}^{(n)}(c)\big)\mathcal{P}_*^{K_2}\big(\mathcal{Z}_{t_{N-1},t_N}^{(n)}(c)\big)\Bigg| \\
&\le C \sum_{\substack{|K_1|+|K_2|=k+1 \\ |K_1|,|K_2|\ge 0}} \Big|\mathcal{P}_*^{K_1}\big(d\mathcal{Y}_{t_{N-2},t_{N-1}}^{(n)*}(c)\big)\Big|\Big|\mathcal{P}_*^{K_2}\big(d\mathcal{Y}_{t_{N-1},t_N}^{(n)*}(c)\big)\Big| \\
&\le \widehat{\mathcal{K}}_{k+1}^{(n)}(c)(t_N - t_{N-2})^{(k+1)\alpha} \le \widehat{\mathcal{K}}_{k+1}^{(n)}(c)\Big(\frac{2}{N-1}(t-s)\Big)^{(k+1)\alpha} \qquad (c \in \widehat{\Omega}_{k+1}^{(n)}),
\end{aligned}$$

where the random variable $\widehat{\mathcal{K}}^{(n)}_{k+1} : \Omega_{x_*}(X) \to \mathbb{R}$ is given by

$$\widehat{\mathcal{K}}^{(n)}_{k+1}(c) := C \sum_{\substack{|K_1|+|K_2|=k+1 \\ |K_1|,|K_2|\geq 0}} \mathcal{Q}^{(n,K_1)}(c)\mathcal{Q}^{(n,K_2)}(c),$$

$$\mathcal{Q}^{(n,K)}(c) := \widehat{\mathcal{K}}^{(n)}_{k_1}(c)\widehat{\mathcal{K}}^{(n)}_{k_2}(c)\cdots\widehat{\mathcal{K}}^{(n)}_{k_\ell}(c) \qquad \big(K = \big((i_1,k_1),(i_2,k_2),\dots,(i_\ell,k_\ell)\big)\big).$$

Note that $\widehat{\mathcal{K}}^{(n)}_{k+1}$ is non-negative and it has the following integrability:

$$\begin{aligned}\mathbb{E}^{\mathbb{P}_{x_*}}\big[(\widehat{\mathcal{K}}^{(n)}_{k+1})^{4m/(k+1)}\big] &\leq C \sum_{\substack{k_1,\dots,k_\ell>0 \\ k_1+k_2+\cdots+k_\ell=k+1}} \mathbb{E}^{\mathbb{P}_{x_*}}\Big[\big(\widehat{\mathcal{K}}^{(n)}_{k_1}\widehat{\mathcal{K}}^{(n)}_{k_2}\cdots\widehat{\mathcal{K}}^{(n)}_{k_\ell}\big)^{4m/(k+1)}\Big] \\ &\leq C \sum_{\substack{k_1,\dots,k_\ell>0 \\ k_1+k_2+\cdots+k_\ell=k+1}} \prod_{\lambda=1}^{\ell} \mathbb{E}^{\mathbb{P}_{x_*}}\Big[\big(\widehat{\mathcal{K}}^{(n)}_{k_\lambda}\big)^{4m/k_\lambda}\Big]^{k_\lambda/(k+1)} < \infty,\end{aligned}$$

where we used the generalized Hölder inequality for the second line. By removing points in Δ successively until the partition Δ coincides with $\{s,t\}$, we have

$$\begin{aligned}\Big|\big(\mathcal{Z}(\Delta)^{(n)}_{s,t}(c)\big)^{(k+1)}_{i_*}\Big| &\leq \Big|\big(\mathcal{Z}(\Delta\setminus\{t_{N-1}\})^{(n)}_{s,t}(c)\big)^{(k+1)}_{i_*}\Big| + \widehat{\mathcal{K}}^{(n)}_{k+1}(c)\Big(\frac{2}{N-1}(t-s)\Big)^{(k+1)\alpha} \\ &\leq \Big|\big(\mathcal{Z}(\{s,t\})^{(n)}_{s,t}(c)\big)^{(k+1)}_{i_*}\Big| + \sum_{\ell=1}^{N-2}\widehat{\mathcal{K}}^{(n)}_{k+1}(c)\Big(\frac{2}{N-\ell}\Big)^{(k+1)\alpha}(t-s)^{(k+1)\alpha} \\ &\leq \Big|\big(\mathcal{Z}^{(n)}_{s,t}(c)\big)^{(k+1)}_{i_*}\Big| + \widehat{\mathcal{K}}^{(n)}_{k+1}(c)2^{(k+1)\alpha}\zeta\big((k+1)\alpha\big)(t-s)^{(k+1)\alpha} \\ &\leq \widehat{\mathcal{K}}^{(n)}_{k+1}(c)(t-s)^{(k+1)\alpha} \qquad (i=1,2,\dots,d_{k+1},\, c\in\widehat{\Omega}^{(n)}_{k+1}),\end{aligned} \tag{4.34}$$

where $\zeta(z)$ denotes the Riemann zeta function $\zeta(z) := \sum_{n=1}^{\infty}(1/n^z)$ for $z \in \mathbb{R}$.

We will show that the family $\{\mathcal{Z}(\Delta)^{(n)}_{s,t}\}$ satisfies the Cauchy convergence principle. Let $\delta > 0$ and take two partitions $\Delta = \{s = t_0 < t_1 \cdots < t_N = t\}$ and Δ' of $[s,t]$ independent of $n \in \mathbb{N}$ satisfying $|\Delta|, |\Delta'| < \delta$. We set $\widehat{\Delta} := \Delta \cup \Delta'$ and write

$$\widehat{\Delta}_\ell = \widehat{\Delta} \cap [t_\ell, t_{\ell+1}] = \{t_\ell = s_{\ell 0} < s_{\ell 1} < \cdots < s_{\ell L_\ell} = t_{\ell+1}\} \qquad (\ell = 0,1,\dots,N-1).$$

By using (4.34), we have

$$\begin{aligned}&\Big|\big(\mathcal{Z}(\Delta)^{(n)}_{s,t}(c)\big)^{(k+1)}_{i_*} - \big(\mathcal{Z}(\widehat{\Delta})^{(n)}_{s,t}(c)\big)^{(k+1)}_{i_*}\Big| \\ &= \Big|\big(\mathcal{Z}^{(n)}_{t_0,t_1}(c) * \cdots * \mathcal{Z}^{(n)}_{t_{N-1},t_N}(c)\big)^{(k+1)}_{i_*} - \big(\mathcal{Z}(\widehat{\Delta}_0)^{(n)}_{t_0,t_1}(c) * \cdots * \mathcal{Z}(\widehat{\Delta}_{N-1})^{(n)}_{t_{N-1},t_N}(c)\big)^{(k+1)}_{i_*}\Big| \\ &= \Big|\big(\mathcal{Z}^{(n)}_{t_0,t_1}(c)\big)^{(k+1)}_{i_*} + \big(\mathcal{Z}^{(n)}_{t_1,t_2}(c) * \cdots * \mathcal{Z}^{(n)}_{t_{N-1},t_N}(c)\big)^{(k+1)}_{i_*} \\ &\qquad - \big(\mathcal{Z}(\widehat{\Delta}_0)^{(n)}_{t_0,t_1}(c)\big)^{(k+1)}_{i_*} - \big(\mathcal{Z}(\widehat{\Delta}_1)^{(n)}_{t_1,t_2}(c) * \cdots * \mathcal{Z}(\widehat{\Delta}_{N-1})^{(n)}_{t_{N-1},t_N}(c)\big)^{(k+1)}_{i_*}\Big| \\ &\leq \widehat{\mathcal{K}}^{(n)}_{k+1}(c)(t_1-t_0)^{(k+1)\alpha} + \Big|\big(\mathcal{Z}^{(n)}_{t_1,t_2}(c) * \cdots * \mathcal{Z}^{(n)}_{t_{N-1},t_N}(c)\big)^{(k+1)}_{i_*} \\ &\qquad - \big(\mathcal{Z}(\widehat{\Delta}_0)^{(n)}_{t_1,t_2}(c) * \cdots * \mathcal{Z}(\widehat{\Delta}_{N-1})^{(n)}_{t_{N-1},t_N}(c)\big)^{(k+1)}_{i_*}\Big| \qquad (i=1,2,\dots,d_{k+1},\, c\in\widehat{\Omega}^{(n)}_{k+1}).\end{aligned}$$

Repeating this kind of estimate and recalling $(k+1)\alpha > 1$ yield

$$\begin{aligned}&\Big|\big(\mathcal{Z}(\Delta)^{(n)}_{s,t}(c)\big)^{(k+1)}_{i_*} - \big(\mathcal{Z}(\widehat{\Delta})^{(n)}_{s,t}(c)\big)^{(k+1)}_{i_*}\Big| \\ &\leq \sum_{\ell=0}^{N-1}\widehat{\mathcal{K}}^{(n)}_{k+1}(c)(t_{\ell+1}-t_\ell)^{(k+1)\alpha} \\ &\leq \widehat{\mathcal{K}}^{(n)}_{k+1}(c)\Big(\max_{\Delta}(t_{\ell+1}-t_\ell)^{(k+1)\alpha-1}\Big)\sum_{\ell=0}^{N-1}(t_{\ell+1}-t_\ell) \leq \widehat{\mathcal{K}}^{(n)}_{k+1}(c)(t-s)\times\delta^{(k+1)\alpha-1}\end{aligned} \tag{4.35}$$

for $i = 1, 2, \ldots, d_{k+1}$ and $c \in \widehat{\Omega}^{(n)}_{k+1}$. We thus obtain

$$\begin{aligned}
&\Big|\big(\mathcal{Z}(\Delta)^{(n)}_{s,t}(c)\big)^{(k+1)}_{i_*} - \big(\mathcal{Z}(\Delta')^{(n)}_{s,t}(c)\big)^{(k+1)}_{i_*}\Big| \\
&\le \Big|\big(\mathcal{Z}(\Delta)^{(n)}_{s,t}(c)\big)^{(k+1)}_{i_*} - \big(\mathcal{Z}(\widehat{\Delta})^{(n)}_{s,t}(c)\big)^{(k+1)}_{i_*}\Big| + \Big|\big(\mathcal{Z}(\widehat{\Delta})^{(n)}_{s,t}(c)\big)^{(k+1)}_{i_*} - \big(\mathcal{Z}(\widetilde{\Delta})^{(n)}_{s,t}(c)\big)^{(k+1)}_{i_*}\Big| \\
&\le 2\widehat{\mathcal{K}}^{(n)}_{k+1}(c)(t-s) \times \delta^{(k+1)\alpha - 1} \to 0 \qquad (i = 1, 2, \ldots, d_{k+1},\, c \in \widehat{\Omega}^{(n)}_{k+1})
\end{aligned}$$

as $\delta \searrow 0$ uniformly in $0 \le s \le t \le 1$ by (4.35). Therefore, noting the estimate (4.34), there exists a random variable

$$\overline{\mathcal{Z}}^{(n)}_{s,t}(c) := \begin{cases} \lim\limits_{|\Delta| \searrow 0} \mathcal{Z}(\Delta)^{(n)}_{s,t}(c) & (c \in \widehat{\Omega}^{(n)}_{k+1}), \\ \mathbf{1}_G & (c \in \Omega_{x_*}(X) \setminus \widehat{\Omega}^{(n)}_{k+1}), \end{cases} \qquad (0 \le s \le t \le 1)$$

satisfying

$$\big\|\big(\overline{\mathcal{Z}}^{(n)}_{s,t}(c)\big)^{(k+1)}\big\|_{\mathbb{R}^{d_{k+1}}} \le \widehat{\mathcal{K}}^{(n)}_{k+1}(c)(t-s)^{(k+1)\alpha} \qquad (c \in \widehat{\Omega}^{(n)}_{k+1}).$$

Our final goal is to show

$$\overline{\mathcal{Z}}^{(n)}_{s,t}(c) = \mathcal{Y}^{(n;\,k+1)}_s(c) * \mathcal{Y}^{(n;\,k+1)}_t(c) \qquad (0 \le s \le t \le 1,\, c \in \widehat{\Omega}^{(n)}_{k+1}).$$

However, it suffices to check that

$$\big(\overline{\mathcal{Z}}^{(n)}_{s,t}(c)\big)^{(k+1)} = \big(d\mathcal{Y}^{(n)*}_{s,t}(c)\big)^{(k+1)} \qquad (0 \le s \le t \le 1,\, c \in \widehat{\Omega}^{(n)}_{k+1}) \tag{4.36}$$

by the definition of $\overline{\mathcal{Z}}^{(n)}_{s,t}$. We fix $i = 1, 2, \ldots, d_{k+1}$ and $c \in \widehat{\Omega}^{(n)}_{k+1}$. Put

$$\Psi^i_{s,t}(c) := \big(d\mathcal{Y}^{(n)*}_{s,t}(c)\big)^{(k+1)}_{i*} - \big(\overline{\mathcal{Z}}^{(n)}_{s,t}(c)\big)^{(k+1)}_{i_*} \quad (0 \le s \le t \le 1).$$

Then we easily see that $\Psi^i_{s,t}(c)$ is additive in the sense that

$$\Psi^i_{s,t}(c) = \Psi^i_{s,u}(c) + \Psi^i_{u,t}(c) \qquad (0 \le s \le u \le t \le 1). \tag{4.37}$$

Since the piecewise smooth stochastic process $(\mathcal{Y}^{(n)}_t)_{0 \le t \le 1}$ is defined by the d_{CC}- geodesic interpolation of $\{\mathcal{X}^{(n)}_{t_k}\}_{k=0}^n$, we know

$$\big\|\big(d\mathcal{Y}^{(n)*}_{s,t}(c)\big)^{(k+1)}\big\|_{\mathbb{R}^{d_{k+1}}} \le \widetilde{\mathcal{K}}^{(n)}_{k+1}(c)(t-s)^{(k+1)\alpha} \qquad (c \in \widetilde{\Omega}^{(n)}_{k+1})$$

for some set $\widetilde{\Omega}^{(n)}_{k+1}$ with $\mathbb{P}_{x_*}(\widetilde{\Omega}^{(n)}_{k+1}) = 1$ and random variable $\widetilde{\mathcal{K}}^{(n)}_{k+1} : \Omega_{x_*}(X) \to \mathbb{R}$. Then we have

$$\Big|\Psi^i_{s,t}(c)\Big| \le \big(\widetilde{\mathcal{K}}^{(n)}_{k+1}(c) + \widehat{\mathcal{K}}^{(n)}_{k+1}(c)\big)(t-s)^{(k+1)\alpha} \qquad (0 \le s \le t \le 1,\, c \in \widetilde{\Omega}^{(n)}_{k+1} \cap \widehat{\Omega}^{(n)}_{k+1}).$$

We may write $\widehat{\Omega}^{(n)}_{k+1}$ instead of $\widetilde{\Omega}^{(n)}_{k+1} \cap \widehat{\Omega}^{(n)}_{k+1}$ by abuse of notation, because its probability is equal to one. For any $\varepsilon > 0$, there is a sufficiently large $N \in \mathbb{N}$ such that $1/N < \varepsilon$. Then, as $\varepsilon \searrow 0$,

$$\begin{aligned}
\Big|\Psi^i_{0,t}(c)\Big| &= \Big|\Psi^i_{0,1/N}(c) + \Psi^i_{1/N,2/N}(c) + \cdots + \Psi^i_{[Nt]/N,t}(c)\Big| \\
&\le \big(\widetilde{\mathcal{K}}^{(n)}_{k+1}(c) + \widehat{\mathcal{K}}^{(n)}_{k+1}(c)\big)\varepsilon^{(k+1)\alpha-1}\Big\{\underbrace{\frac{1}{N} + \frac{1}{N} + \cdots + \frac{1}{N}}_{[Nt]\text{-times}} + \Big(t - \frac{[Nt]}{N}\Big)\Big\} \\
&= \big(\widetilde{\mathcal{K}}^{(n)}_{k+1}(c) + \widehat{\mathcal{K}}^{(n)}_{k+1}(c)\big)\varepsilon^{(k+1)\alpha-1} t \to 0 \qquad (0 \le t \le 1,\, c \in \widehat{\Omega}^{(n)}_{k+1})
\end{aligned}$$

by (4.37) and $(k+1)\alpha - 1 > 0$. This implies that $\Psi^i_{0,t}(c) = 0$ for $0 \le t \le 1$ and $c \in \widehat{\Omega}^{(n)}_{k+1}$. Therefore, it follows from (4.36) that

$$\Psi^i_{s,t}(c) = \Psi^i_{0,t}(c) - \Psi^i_{0,s}(c) = 0 \qquad (0 \le s \le t \le 1,\, c \in \widehat{\Omega}^{(n)}_{k+1}),$$

which means (4.35). Consequently, there exist a $\mathcal{F}_\infty$-measurable set $\Omega^{(n)}_{k+1} \subset \Omega_{x_*}(X)$ with probability one and a non-negative random variable $\mathcal{K}^{(n)}_{k+1} \in L^{4m}(\Omega_{x_*}(X) \to \mathbb{R};\, \mathbb{P}_{x_*})$ satisfying

$$d_{\mathrm{CC}}\big(\mathcal{Y}^{(n;\,k+1)}_s(c), \mathcal{Y}^{(n;\,k+1)}_t(c)\big) \le \mathcal{K}^{(n)}_{k+1}(c)(t-s)^\alpha \quad (0 \le s \le t \le 1,\, c \in \Omega^{(n)}_{k+1}).$$

This completes the proof of Lemma 4.7. □

Proof of Lemma 4.5. For $m, n \in \mathbb{N}$ and $\widehat{\alpha} < \frac{2m-1}{4m}$, Lemma 4.7 implies that

$$\mathbb{E}^{\mathbb{P}_{x_*}}\Big[d_{\mathrm{CC}}\big(\mathcal{Y}^{(n;\,r)}_s, \mathcal{Y}^{(n;\,r)}_t\big)^{4m}\Big] \le \mathbb{E}^{\mathbb{P}_{x_*}}\big[\big(\mathcal{K}^{(n)}_r\big)^{4m}\big](t-s)^{4m\widehat{\alpha}}$$

for $0 \le s \le t \le 1$. Thus, it follows from (4.32) that

$$\mathbb{E}^{\mathbb{P}_{x_*}}\Big[d_{\mathrm{CC}}\big(\mathcal{Y}^{(n;\,r)}_s, \mathcal{Y}^{(n;\,r)}_t\big)^{4m}\Big] \le C(t-s)^{4m\widehat{\alpha}} \qquad (0 \le s \le t \le 1).$$

for a positive constant $C > 0$ independent of $n \in \mathbb{N}$. By applying the Kolmogorov criterion (cf. Friz–Hairer [15, Section 3.1]), we know that the family $\{\mathbf{P}^{(n)}\}_{n=1}^\infty$ is tight in $C^{0,\alpha\text{-Höl}}_{\mathbf{1}_G}([0,1]; G)$ for $\alpha < \frac{4m\widehat{\alpha}-1}{4m} < \frac{1}{2} - \frac{1}{2m}$. Since $m \in \mathbb{N}$ is taken arbitrarily, we complete the proof. □

We conclude Theorem 2.2 by showing the following convergence of the finite dimensional distribution.

Lemma 4.8. Let $\ell \in \mathbb{N}$. For fixed $0 \le s_1 < s_2 < \cdots < s_\ell \le 1$, we have

$$(\mathcal{Y}^{(n)}_{s_1}, \mathcal{Y}^{(n)}_{s_2}, \dots, \mathcal{Y}^{(n)}_{s_\ell}) \stackrel{(d)}{\to} (Y_{s_1}, Y_{s_2}, \dots, Y_{s_\ell}) \quad (n \to \infty).$$

Proof. We only prove the convergence for $\ell = 2$. General cases ($\ell \ge 3$) can be also proved by repeating the same argument. Put $s = s_1$ and $t = s_2$. Then, by applying Theorem 2.1, we obtain $(\mathcal{X}^{(n)}_s, \mathcal{X}^{(n)}_t) \stackrel{(d)}{\to} (Y_s, Y_t)$ as $n \to \infty$ in the same way as [25, Lemma 4.2]. On the other hand, Lemma 4.7 tells us that there exists a non-negative random variable $\mathcal{K}^{(n)}_r \in L^{4m}(\Omega_{x_*}(X) \to \mathbb{R};\, \mathbb{P}_{x_*})$ such that

$$d_{\mathrm{CC}}\big(\mathcal{Y}^{(n)}_s(c), \mathcal{Y}^{(n)}_t(c)\big) \le \mathcal{K}^{(n)}_r(c)(t-s)^\alpha \quad \mathbb{P}_{x_*}\text{-a.s.} \qquad (0 \le s \le t \le 1).$$

Now suppose that $t_k \le t \le t_{k+1}$ for some $k = 0, 1, \dots, n-1$. For all $\varepsilon > 0$ and sufficiently large $m \in \mathbb{N}$, by using Chebyshev's inequality, we have

$$\begin{aligned}
&\mathbb{P}_{x_*}\Big(d_{\mathrm{CC}}\big(\mathcal{X}^{(n)}_t, \mathcal{Y}^{(n)}_t\big) > \varepsilon\Big)\\
&\le \frac{1}{\varepsilon^{4m}}\mathbb{E}^{\mathbb{P}_{x_*}}\Big[d_{\mathrm{CC}}\big(\mathcal{X}^{(n)}_t, \mathcal{Y}^{(n)}_t\big)^{4m}\Big]\\
&\le \frac{1}{\varepsilon^{4m}}\mathbb{E}^{\mathbb{P}_{x_*}}\Big[d_{\mathrm{CC}}\big(\mathcal{Y}^{(n)}_{t_k}, \mathcal{Y}^{(n)}_{t_{k+1}}\big)^{4m}\Big]\\
&\le \frac{1}{\varepsilon^{4m}}\mathbb{E}^{\mathbb{P}_{x_*}}\Big[(\mathcal{K}^{(n)}_r)^{4m}(t_{k+1}-t_k)^{4m\alpha}\Big] = \frac{1}{n^{2m-1}\varepsilon^{4m}}\mathbb{E}^{\mathbb{P}_{x_*}}\big[(\mathcal{K}^{(n)}_r)^{4m}\big] \to 0 \quad (n \to \infty).
\end{aligned}$$

Thus, Slutzky's theorem (cf. Klenke [28, Theorem 13.8]) allows us to obtain the desired convergence $(\mathcal{Y}^{(n)}_s, \mathcal{Y}^{(n)}_t) \stackrel{(d)}{\to} (Y_s, Y_t)$ as $n \to \infty$. This completes the proof. □

4.4 Proof of Theorem 2.3

In this section, we show Theorem 2.3, which is a generalization of Theorem 2.2 to non-harmonic cases. Our first aim is to show that the same pathwise Hölder estimate as Lemma 4.7 holds for the stochastic process $\{\overline{\mathcal{Y}}^{(n)}\}_{n=1}^{\infty}$.

Lemma 4.9. For $m, n \in \mathbb{N}$ and $\alpha < \frac{2m-1}{4m}$, there exist an $\mathcal{F}_\infty$-measurable set $\overline{\Omega}_r^{(n)} \subset \Omega_{x_*}(X)$ and a non-negative random variable $\overline{\mathcal{K}}_r^{(n)} \in L^{4m}\big(\Omega_{x_*}(X) \to \mathbb{R}; \mathbb{P}_{x_*}\big)$ such that $\mathbb{P}_{x_*}(\overline{\Omega}_r^{(n)}) = 1$ and

$$d_{\mathrm{CC}}\big(\overline{\mathcal{Y}}_s^{(n)}(c), \overline{\mathcal{Y}}_t^{(n)}(c)\big) \le \overline{\mathcal{K}}_r^{(n)}(c)(t-s)^\alpha \qquad (c \in \overline{\Omega}_r^{(n)},\ 0 \le s < t \le 1). \tag{4.38}$$

Proof. Fix $n \in \mathbb{N}$ and $1 \le k \le \ell \le n$. We then have

$$d_{\mathrm{CC}}(\overline{\mathcal{Y}}_{t_k}^{(n)}, \overline{\mathcal{Y}}_{t_\ell}^{(n)}) \le d_{\mathrm{CC}}(\overline{\mathcal{Y}}_{t_k}^{(n)}, \mathcal{Y}_{t_k}^{(n)}) + d_{\mathrm{CC}}(\mathcal{Y}_{t_k}^{(n)}, \mathcal{Y}_{t_\ell}^{(n)}) + d_{\mathrm{CC}}(\mathcal{Y}_{t_\ell}^{(n)}, \overline{\mathcal{Y}}_{t_\ell}^{(n)}).$$

Set set $\mathcal{Z}_t^{(n)} := (\mathcal{Y}_t^{(n)})^{-1} * \overline{\mathcal{Y}}_t^{(n)}$ for $0 \le t \le 1$ and $n \in \mathbb{N}$. We then have

$$\log\big(\mathcal{Z}_{t_k}^{(n)}\big)|_{\mathfrak{g}^{(1)}} = \frac{1}{\sqrt{n}}\mathrm{Cor}_{\mathfrak{g}^{(1)}}(w_k) \qquad (n \in \mathbb{N},\ k = 0, 1, \dots, n)$$

so that there is a constant $C > 0$ such that $\big\| \log\big(\mathcal{Z}_{t_k}^{(n)}\big)|_{\mathfrak{g}^{(1)}}\big\|_{\mathfrak{g}^{(1)}} \le Cn^{-1/2}$ for $n \in \mathbb{N}$ and $k = 0, 1, 2, \dots, n$. By the choice of the components of $\Phi_0(x)$ for $x \in V$, we have $\big\| \log\big(\mathcal{Z}_{t_k}^{(n)}\big)|_{\mathfrak{g}^{(i)}}\big\|_{\mathfrak{g}^{(i)}} \le Cn^{-i/2}$ for $n \in \mathbb{N}$ and $k = 0, 1, 2, \dots, n$. Then Proposition 3.1 leads to

$$d_{\mathrm{CC}}(\overline{\mathcal{Y}}_{t_k}^{(n)}, \mathcal{Y}_{t_k}^{(n)}) \le C\big\|\mathcal{Z}_{t_k}^{(n)}\big\|_{\mathrm{Hom}} = C\sum_{i=1}^{r} \big\| \log\big(\mathcal{Z}_{t_k}^{(n)}\big)|_{\mathfrak{g}^{(i)}}\big\|_{\mathfrak{g}^{(i)}}^{1/i} \le \frac{C}{\sqrt{n}} \tag{4.39}$$

for $n \in \mathbb{N}$ and $k = 0, 1, 2, \dots, n$. Then Lemma 4.7 and (4.39) imply that there exist an $\mathcal{F}_\infty$-measurable set $\overline{\Omega}_r^{(n)} \subset \Omega_{x_*}(X)$ and a non-negative random variable $\overline{\mathcal{K}}_r^{(n)} \in L^{4m}\big(\Omega_{x_*}(X) \to \mathbb{R}; \mathbb{P}_{x_*}\big)$ such that $\mathbb{P}_{x_*}^{(n^{-1/2})}(\overline{\Omega}_r^{(n)}) = 1$ and

$$d_{\mathrm{CC}}\big(\overline{\mathcal{Y}}_{t_k}^{(n)}(c), \overline{\mathcal{Y}}_{t_\ell}^{(n)}(c)\big) \le \frac{C}{\sqrt{n}} + \mathcal{K}_r^{(n)}(c)\Big(\frac{\ell-k}{n}\Big)^\alpha + \frac{C}{\sqrt{n}} \le \overline{\mathcal{K}}_r^{(n)}(c)\Big(\frac{\ell-k}{n}\Big)^\alpha \tag{4.40}$$

for $c \in \overline{\Omega}_r^{(n)}$ and $0 \le k \le \ell \le n$. We now take $0 \le k \le \ell \le n$ such that $k/n \le s < (k+1)/n$ and $\ell/n \le t < (\ell+1)/n$. By the definition of $(\overline{\mathcal{Y}}_t^{(n)})_{0 \le t \le 1}$, we have

$$\begin{aligned} d_{\mathrm{CC}}\big(\overline{\mathcal{Y}}_s^{(n)}, \overline{\mathcal{Y}}_{t_{k+1}}^{(n)}\big) &= (k - ns)d_{\mathrm{CC}}\big(\overline{\mathcal{Y}}_{t_k}^{(n)}, \overline{\mathcal{Y}}_{t_{k+1}}^{(n)}\big), \\ d_{\mathrm{CC}}\big(\overline{\mathcal{Y}}_{t_\ell}^{(n)}, \overline{\mathcal{Y}}_t^{(n)}\big) &= (nt - \ell)d_{\mathrm{CC}}\big(\overline{\mathcal{Y}}_{t_\ell}^{(n)}, \overline{\mathcal{Y}}_{t_{\ell+1}}^{(n)}\big). \end{aligned}$$

We then use the triangular inequality and (4.40) to obtain

$$\begin{aligned} &d_{\mathrm{CC}}\big(\overline{\mathcal{Y}}_s^{(n)}(c), \overline{\mathcal{Y}}_t^{(n)}(c)\big) \\ &\le (k - ns)\overline{\mathcal{K}}_r^{(n)}(c)\Big(\frac{1}{n}\Big)^\alpha + \overline{\mathcal{K}}_r^{(n)}(c)\Big(\frac{\ell-k-1}{n}\Big)^\alpha + (nt-\ell)\overline{\mathcal{K}}_r^{(n)}(c)\Big(\frac{1}{n}\Big)^\alpha \\ &\le \overline{\mathcal{K}}_r^{(n)}(c)\Big\{\Big(\frac{k+1}{n} - s\Big)^\alpha + \Big(\frac{\ell-k-1}{n}\Big)^\alpha + \Big(t - \frac{\ell}{n}\Big)^\alpha\Big\} \le \overline{\mathcal{K}}_r^{(n)}(c)(t-s)^\alpha \qquad (c \in \overline{\Omega}_r^{(n)}). \end{aligned}$$

This completes the proof. □

Proof of Theorem 2.3. The proof is split into two steps.

Step 1. We show that $\{\overline{\mathcal{Y}}^{(n)}\}_{n=1}^{\infty}$ converges in law to $(Y_t)_{0\le t\le 1}$ in $C_{\mathbf{1}_G}([0,1];G)$ as $n\to\infty$. For $0\le t\le 1$, take an integer $0\le k\le n$ such that $k/n\le t<(k+1)/n$. Then (4.31), (4.38) and (4.39) imply, $\mathbb{P}_{x_*}$-almost surely,

$$\begin{aligned} d_{\mathrm{CC}}(\mathcal{Y}_t^{(n)},\overline{\mathcal{Y}}_t^{(n)}) &\le d_{\mathrm{CC}}(\mathcal{Y}_{t_k}^{(n)},\mathcal{Y}_t^{(n)})+d_{\mathrm{CC}}(\mathcal{Y}_{t_k}^{(n)},\overline{\mathcal{Y}}_{t_k}^{(n)})+d_{\mathrm{CC}}(\overline{\mathcal{Y}}_{t_k}^{(n)},\overline{\mathcal{Y}}_t^{(n)})\\ &\le \mathcal{K}_r^{(n)}\Big(t-\frac{k}{n}\Big)^{\alpha}+\frac{C}{\sqrt{n}}+\overline{\mathcal{K}}_r^{(n)}\Big(t-\frac{k}{n}\Big)^{\alpha}\\ &\le \big\{\mathcal{K}_r^{(n)}+\overline{\mathcal{K}}_r^{(n)}+C\big\}\Big(\frac{1}{\sqrt{n}}\Big)^{\alpha}\qquad \Big(m\in\mathbb{N},\,\alpha<\frac{2m-1}{4m}\Big). \end{aligned}\tag{4.41}$$

Let ρ be a metric on $C_{\mathbf{1}_G}([0,1];G)$ defined by $\rho(w^{(1)},w^{(2)}):=\max_{0\le t\le 1}d_{\mathrm{CC}}(w_t^{(1)},w_t^{(2)})$. By applying the Chebyshev inequality and (4.41), we have, for $\varepsilon>0$ and $m\in\mathbb{N}$,

$$\begin{aligned} &\mathbb{P}_{x_*}\Big(\rho(\mathcal{Y}^{(n)},\overline{\mathcal{Y}}^{(n)})>\varepsilon\Big)\\ &\le\Big(\frac{1}{\varepsilon}\Big)^{4m}\mathbb{E}^{\mathbb{P}_{x_*}}\Big[\rho(\mathcal{Y}^{(n)},\overline{\mathcal{Y}}^{(n)})^{4m}\Big]\\ &\le\Big(\frac{1}{\varepsilon}\Big)^{4m}\mathbb{E}^{\mathbb{P}_{x_*}}\Big[\max_{0\le t\le 1}d_{\mathrm{CC}}(\mathcal{Y}_t^{(n)},\overline{\mathcal{Y}}_t^{(n)})^{4m}\Big]\\ &\le 3^{4m-1}\Big(\frac{1}{\varepsilon}\Big)^{4m}\Big(\frac{1}{\sqrt{n}}\Big)^{4m\alpha}\Big\{\mathbb{E}^{\mathbb{P}_{x_*}}\big[(\mathcal{K}_r^{(n)})^{4m}\big]+\mathbb{E}^{\mathbb{P}_{x_*}}\big[(\overline{\mathcal{K}}_r^{(n)})^{4m}\big]+\mathbb{E}^{\mathbb{P}_{x_*}}\big[C^{4m}\big]\Big\}\to 0 \end{aligned}$$

as $n\to\infty$. Therefore, by Slutzky's theorem, the convergence in law of $\{\overline{\mathcal{Y}}^{(n)}\}_{n=1}^{\infty}$ to the diffusion process $(Y_t)_{0\le t\le 1}$ in $C_{\mathbf{1}_G}([0,1];G)$ as $n\to\infty$ is obtained.

Step 2. By the previous step, we see that the convergence of the finite-dimensional distribution of $\{\overline{\mathcal{Y}}^{(n)}\}_{n=1}^{\infty}$ holds. On the other hand, we can prove that the sequence $\{\overline{\mathbf{P}}^{(n)}:=\mathbb{P}_{x_*}\circ(\overline{\mathcal{Y}}^{(n)})^{-1}\}_{n=1}^{\infty}$ is tight in $C_{\mathbf{1}_G}^{0,\alpha\text{-Höl}}([0,1];G)$, by applying Lemma 4.9 and by following the same argument as the proof of Lemma 4.5. We complete the proof by combining these two. □

5 An explicit representation of the limiting diffusions and a relation with rough path theory

5.1 An explicit representation of the limiting diffusion

Let us consider an SDE on $\mathbb{R}^N$

$$d\xi_t=\sum_{i=1}^{d}U_i(\xi_t)\circ dB_t^i+U_0(\xi_t)\,dt,\qquad \xi_0=x_0\in\mathbb{R}^N,\tag{5.1}$$

where $U_0,U_1,\dots,U_d$ are C^∞-vector fields on $\mathbb{R}^d$ and $(B_t)_{0\le t\le 1}=(B_t^1,B_t^2,\dots,B_t^d)_{0\le t\le 1}$ is a d-dimensional standard Brownian motion. As is well-known, a number of authors have studied explicit representations of the unique solution to (5.1) as a functional of Itô/Stratonovich iterated integrals under some assumptions on vector fields $U_0,U_1,\dots,U_d$. In particular, Kunita [38] has obtained the explicit formula by using the CBH formula in the case where the Lie algebra generated by $U_0,U_1,\dots,U_d$ is nilpotent or solvable. Castell [9] gave a universal representation formula, which contains the above results in the nilpotent case and extends the study of Ben Arous [5] to more general diffusions.

We now recall the result in [9] when the Lie algebra generated by $U_0,U_1,\dots,U_d$ is nilpotent of step r. We first introduce several notations of multi-indices. Set $\mathcal{I}^{(k)}=\{0,1,\dots,d\}^k$ and let $I=(i_1,i_2,\dots,i_k)\in\mathcal{I}^{(k)}$ be a multi-index of length $|I|=k$. For vector fields $U_0,U_1,\dots,U_d$ on $\mathbb{R}^d$ and $I=(i_1,i_2,\dots,i_k)\in\mathcal{I}^{(k)}$, we denote by U^I the vector field of the form $U^I = [U_{i_1}, [U_{i_2}, \cdots, [U_{i_{k-1}}, U_{i_k}] \cdots]]$. For a multi-index $I = (i_1, i_2, \dots, i_k) \in \mathcal{I}^{(k)}$, we define B_t^I by

$$B_t^I := \int_{\Delta^{(k)}[0,t]} \circ dB_{t_1}^{i_1} \circ dB_{t_2}^{i_2} \cdots \circ dB_{t_k}^{i_k},$$

where $\Delta^{(k)}[0,t] := \{(t_1, t_2, \dots, t_k) \in [0,t]^k \mid 0 \le t_1 \le t_2 \le \cdots \le t_k \le t\}$ for $0 \le t \le 1$ and $B_t^0 = t$ for convention. Next we introduce notations of the permutations. Denote by $\mathfrak{S}_k$ be the symmetric group of degree k. For a permutation $\sigma \in \mathfrak{S}_k$, we write $e(\sigma)$ for the cardinality of the set $\{i \in \{1, 2, \dots, k-1\} \mid \sigma(i) > \sigma(i+1)\}$, which we call the number of inversions of σ. For $I = (i_1, i_2, \dots, i_k) \in \mathcal{I}^{(k)}$ and $\sigma \in \mathfrak{S}_k$, we put $I_\sigma := (i_{\sigma(1)}, i_{\sigma(2)}, \dots, i_{\sigma(k)}) \in \mathcal{I}^{(k)}$.

Proposition 5.1. (cf. [9]) Let $U_0, U_1, \dots, U_d$ be bounded C^∞-vector fields on $\mathbb{R}^N$ such that the Lie algebra generated by $U_0, U_1, \dots, U_d$ is nilpotent of step r. We consider the solution $(\xi_t)_{0 \le t \le 1}$ of (5.1). Then we have

$$\xi_t = \exp\Big(\sum_{k=1}^{r} \sum_{I \in \mathcal{I}^{(k)}} c_t^I U^I\Big)(x_0) \qquad (0 \le t \le 1) \quad a.s.,$$

where

$$c_t^I := \sum_{\sigma \in \mathfrak{S}_{|I|}} \frac{(-1)^{e(\sigma)}}{|I|^2 \binom{|I|-1}{e(\sigma)}} B_t^{I_{\sigma^{-1}}}.$$

We now provide an explicit representation of $(Y_t)_{0 \le t \le 1}$, the solution to the SDE (2.10). As mentioned in Section 3.1, since G is identified with $\mathbb{R}^N$ $(N = d_1 + d_2 + \cdots + d_r)$, we may apply Proposition 5.1 by replacing $U_0, U_1, \dots, U_d$ by $V_0, V_1, \dots, V_{d_1}$, where $V_0 = \beta(\Phi_0)_*$. By recalling several concrete computations of $c_t^I U^I$ mentioned in e.g., Kunita [38], we obtain the following.

Theorem 5.2. The limiting diffusion process $(Y_t)_{0 \le t \le 1}$ is explicitly represented as

$$\begin{aligned} Y_t = \exp\Big(t\beta(\Phi_0)_* &+ \sum_{i=1}^{d_1} B_t^i V_{i*} \\ &+ \sum_{0 \le i < j \le d_1} \frac{1}{2} \int_0^t (B_s^i dB_s^j - B_s^j dB_s^i) [\![V_{i*}, V_{j*}]\!] + \sum_{k=3}^{r} \sum_{I \in \mathcal{I}^{(k)}} c_t^I V_*^I \Big)(\mathbf{1}_G), \end{aligned} \tag{5.2}$$

where $V_*^I = [\![V_{i_1 *}, [\![V_{i_2 *}, \cdots, [\![V_{i_k - 1 *}, V_{i_k *}]\!] \cdots]\!]]\!]$ for $I = (i_1, i_2, \dots, i_k) \in \mathcal{I}^{(k)}$.

We should note that some of $[\![V_{i*}, V_{j*}]\!]$ $(0 \le i < j \le d_1)$ in (5.2) may vanish because $\{[\![V_{i*}, V_{j*}]\!]\}_{1 \le i < j \le d}$ is not always linearly independent.

Remark 5.3. We will discuss yet another functional CLT for non-symmetric random walks on a nilpotent covering graph in a subsequent paper [26]. We also obtain a G-valued diffusion process whose infinitesimal generator differs from $-\mathcal{A}$ of $(Y_t)_{0 \le t \le 1}$ through another CLT. Precisely, the generator is the sub-Laplacian plus drift of the asymptotic direction $\rho_{\mathbb{R}}(\gamma_p) \in \mathfrak{g}^{(1)}$, and the corresponding diffusion process $(\widehat{Y}_t)_{0 \le t \le 1}$ is given by

$$\begin{aligned} \widehat{Y}_t = \exp\Big(t\rho_{\mathbb{R}}(\gamma_p)_* &+ \sum_{i=1}^{d_1} B_t^i V_{i*} \\ &+ \sum_{0 \le i < j \le d_1} \frac{1}{2} \int_0^t (B_s^i dB_s^j - B_s^j dB_s^i) [\![V_{i*}, V_{j*}]\!] + \sum_{k=3}^{r} \sum_{I \in \mathcal{I}^{(k)}} c_t^I V_*^I \Big)(\mathbf{1}_G), \end{aligned}$$

where $V_0 = \rho_{\mathbb{R}}(\gamma_p) \in \mathfrak{g}^{(1)}$. We see that these two diffusions are completely same when the random walk on X is m-symmetric. However, the difference between them appears in the case where $\gamma_p \neq 0$ and $\rho_{\mathbb{R}}(\gamma_p) = \mathbf{0}_{\mathfrak{g}}$. Namely, $(Y_t)_{0\le t\le 1}$ is still given by (5.2), while $(\widehat{Y}_t)_{0\le t\le 1}$ is nothing but the "Brownian motion on G" given by

$$\widehat{Y}_t = \exp\Big(\sum_{i=1}^{d_1} B_t^i V_{i*} + \sum_{1\le i<j\le d_1} \frac{1}{2}\int_0^t (B_s^i dB_s^j - B_s^j dB_s^i)[\![V_{i*}, V_{j*}]\!] + \cdots\Big)(\mathbf{1}_G).$$

Before closing this subsection, we prove the following, which was mentioned in Section 2.

Proposition 5.4. The C_0-semigroup $(e^{-t\mathcal{A}})_{0\le t\le 1}$ coincides with the C_0-semigroup $(T_t)_{0\le t\le 1}$ on $C_\infty(G)$ defined by $T_t f(g) = \mathbb{E}[f(Y_t^g)]$ for $g \in G$, where $(Y_t^g)_{0\le t\le 1}$ is a solution to the stochastic differential equation

$$dY_t^g = \sum_{i=1}^{d_1} V_{i*}(Y_t^g) \circ dB_t^i + \beta(\Phi_0)_*(Y_t^g)\, dt, \qquad Y_0^g = g \in G. \tag{5.3}$$

Proof. By recalling Lemma 4.3, the linear operator $\mathcal{A}$ satisfies the maximal dissipativity, that is, $\lambda - \mathcal{A}$ is surjective for some $\lambda > 0$. Therefore, the Lumer–Fillips theorem implies that $(\mathrm{e}^{-t\mathcal{A}})_{0\le t\le 1}$ is the unique Feller semigroup on $C_\infty(G)$ whose infinitesimal generator extends $\big(-\mathcal{A}, C_0^\infty(G)\big)$. By applying Itô's formula to (5.3), we easily see that the generator of $(Y_t)_{0\le t\le 1}$ coincides with $-\mathcal{A}$ on $C_0^\infty(G)$. Therefore, it suffices to show that the semigroup $(T_t)_{0\le t\le 1}$ enjoys the Feller property, that is, $T_t\big(C_\infty(G)\big) \subset C_\infty(G)$ for $0 \le t \le 1$.

Suppose $f \in C_\infty(G)$. For any $\varepsilon > 0$, we choose a sufficiently large $R > 0$ such that $|f(g)| < \varepsilon$ for $g \in B_R(\mathbf{1}_G)^c$, where $B_R(\mathbf{1}_G) := \{g \in G \,|\, d_{\mathrm{CC}}(\mathbf{1}_G, g) < R\}$. Then, for $g \in B_{2R}(\mathbf{1}_G)^c$, we have

$$\begin{aligned} |T_t f(g)| &\le \mathbb{E}\big[|f(Y_t^g)| \,:\, d_{\mathrm{CC}}(g, Y_t^g) < R\big] + \mathbb{E}\big[|f(Y_t^g)| \,:\, d_{\mathrm{CC}}(g, Y_t^g) \ge R\big] \\ &\le \varepsilon + \|f\|_\infty^G \mathbb{P}\big(d_{\mathrm{CC}}(g, Y_t^g) \ge R\big). \end{aligned}$$

By combining Proposition 3.1 and the Chebyshev inequality with Theorem 5.2,

$$\begin{aligned} \mathbb{P}\big(d_{\mathrm{CC}}(g, Y_t^g) \ge R\big) &= \mathbb{P}\big(d_{\mathrm{CC}}(\mathbf{1}_G, Y_t) \ge R\big) \\ &\le \mathbb{P}\big(C\|Y_t\|_{\mathrm{Hom}} \ge R\big) \le \frac{C}{R^2}\mathbb{E}\Big[\Big(\sum_{k=1}^r \Big\| \sum_{I\in\mathcal{I}^{(k)}} c_t^I V_*^I \Big\|_{\mathfrak{g}^{(k)}}^{1/k}\Big)^2\Big]. \end{aligned}$$

Now we recall the following fact (cf. Friz–Riedel [17, Lemma 2]): For a multi-index $I = (i_1, i_2, \dots, i_k) \in \mathcal{I}^{(k)}$, there exists a positive constant C depending only on k such that

$$\mathbb{E}\Big[\Big(\int_{\Delta^{(k)}[0,t]} \circ dB_{t_1}^{i_1} \circ dB_{t_2}^{i_2} \cdots \circ dB_{t_k}^{i_k}\Big)^2\Big] \le Ct^k \qquad (0 \le t \le 1).$$

In view of this bound, we obtain $\mathbb{P}\big(d_{\mathrm{CC}}(g, Y_t^g) \ge R\big) \le CR^{-2}t$. Taking a sufficiently large $R > 0$ such that $C\|f\|_\infty^G tR^{-2} < \varepsilon$, we conclude $|T_t f(g)| < 2\varepsilon$ for $g \in B_{2R}(\mathbf{1}_G)^c$. This implies that $T_t\big(C_\infty(G)\big) \subset C_\infty(G)$ for $0 \le t \le 1$. $\square$

5.2 The free case: a relation with rough path theory

Consider the step-r non-commutative tensor algebra $T^{(r)}(\mathbb{R}^d) = \mathbb{R} \oplus \big(\bigoplus_{k=1}^r (\mathbb{R}^d)^{\otimes k}\big)$. The tensor product on $T^{(r)}(\mathbb{R}^d)$ is defined by

$$(g_0, g_1, \dots, g_r) \otimes_r (h_0, h_1, \dots, h_r) = \Big(g_0 h_0, g_0 h_1 + g_1 h_0, \dots, \sum_{k=0}^r g_k \otimes h_{r-k}\Big).$$

An element $g = (g_0, g_1, \dots, g_r) \in T^{(r)}(\mathbb{R}^d)$ is occasionally written as $g = g_0 + g_1 + \cdots + g_r$. We define two subsets of $T^{(r)}(\mathbb{R}^d)$ by $T_1^{(r)}(\mathbb{R}^d) := \{g \in T^{(r)}(\mathbb{R}^d) \,|\, g_0 = 1\}$ and $T_0^{(r)}(\mathbb{R}^d) := \{A \in T^{(r)}(\mathbb{R}^d) \,|\, A_0 = 0\}$. It is easy to see that $T_1^{(r)}(\mathbb{R}^d)$ is a Lie group under the tensor product $\otimes_r$. In fact, $\mathbf{1} = (1, 0, 0, \dots, 0)$ is the unit element of $T_1^{(r)}(\mathbb{R}^d)$ and the inverse element of $g \in T_1^{(r)}(\mathbb{R}^d)$ is given by $g^{-1} = \sum_{k=1}^{r}(-1)^k (g - \mathbf{1})^{\otimes_r k}$. The Lie bracket on $T_0^{(r)}(\mathbb{R}^d)$ is defined by $[A, B] = A \otimes_r B - B \otimes_r A$ for $A, B \in T_0^{(r)}(\mathbb{R}^d)$. Note that $T_0^{(r)}(\mathbb{R}^d)$ is the Lie algebra of the Lie group $T_1^{(r)}(\mathbb{R}^d)$, that is, $T_0^{(r)}(\mathbb{R}^d)$ is the tangent space of $T_1^{(r)}(\mathbb{R}^d)$ at $\mathbf{1}$. The map $\exp : T_0^{(r)}(\mathbb{R}^d) \to T_1^{(r)}(\mathbb{R}^d)$ is defined by

$$\exp(A) := 1 + \sum_{k=1}^{r} \frac{1}{k!} A^{\otimes_r k} \qquad \big(A \in T_0^{(r)}(\mathbb{R}^d)\big).$$

Let $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_d\}$ be the standard basis of $\mathbb{R}^d$. We introduce a discrete subgroup $\mathfrak{g}^{(r)}(\mathbb{Z}^d) \subset T_0^{(r)}(\mathbb{R}^d)$ by the set of $\mathbb{Z}$-linear combinations of $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_d$ together with all possible commutators $[\mathbf{e}_{i_1}, [\mathbf{e}_{i_2}, \cdots, [\mathbf{e}_{i_{k-1}}, \mathbf{e}_{i_k}] \cdots]]$ for $i_1, i_2, \dots, i_k = 1, 2, \dots, d$ and $k = 2, 3, \dots, r$.

We now set $\Gamma = \mathbb{G}^{(r)}(\mathbb{Z}^d) := \exp\big(\mathfrak{g}^{(r)}(\mathbb{Z}^d)\big)$. We also define $\mathfrak{g}^{(r)}(\mathbb{R}^d)$ and $\mathbb{G}^{(r)}(\mathbb{R}^d)$ analogously. Then we see that $\big(\mathbb{G}^{(r)}(\mathbb{R}^d), \otimes_r\big)$ is the nilpotent Lie group in which Γ is included as its cocompact lattice and the corresponding limit group coincides with $\big(\mathbb{G}^{(r)}(\mathbb{R}^d), \otimes_r\big)$ itself. We call $\big(\mathbb{G}^{(r)}(\mathbb{R}^d), \otimes_r\big)$ the *free nilpotent Lie group of step* r and $\big(\mathfrak{g}^{(r)}(\mathbb{R}^d), [\cdot, \cdot]\big)$ the *free nilpotent Lie algebra of step* r. Let $\mathfrak{g}^{(1)} = \mathbb{R}^d$ and $\mathfrak{g}^{(k)} = [\mathbb{R}^d, [\mathbb{R}^d, \cdots, [\mathbb{R}^d, \mathbb{R}^d] \cdots]]$ (k-times) for $k = 2, 3, \dots, r$. Then we see that the Lie algebra $\mathfrak{g}^{(r)}(\mathbb{R}^d)$ is decomposed into $\mathfrak{g}^{(1)} \oplus \mathfrak{g}^{(2)} \oplus \cdots \oplus \mathfrak{g}^{(r)}$. It is known that the free nilpotent Lie group $\mathbb{G}^{(r)}(\mathbb{R}^d)$ is highly related to rough path theory, as is seen in e.g., Friz–Victoir [18].

Let $\Gamma = \mathbb{G}^{(r)}(\mathbb{Z}^d)$ and X be a Γ-nilpotent covering graph. Then we see that X is realized into the free nilpotent Lie group $G = \mathbb{G}^{(r)}(\mathbb{R}^d)$ through the modified harmonic realization $\Phi_0 : X \to G$, because Γ is a cocompact lattice in G. Then Theorem 5.2 reads in terms of rough path theory. Precisely speaking, the $\mathbb{G}^{(r)}(\mathbb{R}^d)$-valued diffusion process $(Y_t)_{0 \le t \le 1}$ which solves (2.10) is represented as the Lyons extension of the so-called *distorted Brownian rough path* of order r.

Corollary 5.5. Let $\{V_1, V_2, \dots, V_d\}$ be an orthonormal basis of $\mathfrak{g}^{(1)}$ with respect to the Albanese metric g_0. We write

$$\beta(\Phi_0) = \sum_{1 \le i < j \le d} \beta(\Phi_0)^{ij} [V_i, V_j] \in \mathfrak{g}^{(2)},$$

where we note that $\{[V_i, V_j] \,:\, 1 \le i < j \le d\} \subset \mathfrak{g}^{(2)}$ forms a basis of $\mathfrak{g}^{(2)}$. Let $\overline{\boldsymbol{\beta}}(\Phi_0) = \big(\overline{\boldsymbol{\beta}}(\Phi_0)^{ij}\big)_{i,j=1}^{d}$ be an anti-symmetric matrix defined by

$$\overline{\boldsymbol{\beta}}(\Phi_0)^{ij} := \begin{cases} \beta(\Phi_0)^{ij} & (1 \le i < j \le d), \\ -\beta(\Phi_0)^{ij} & (1 \le j < i \le d), \\ 0 & (i = j). \end{cases}$$

Then the $\mathbb{G}^{(r)}(\mathbb{R}^d)$-valued diffusion process $(Y_t)_{0 \le t \le 1}$ coincides with the Lyons extension of the distorted Brownian rough path $\overline{\mathbf{B}}_t = 1 + \overline{\mathbf{B}}_t^1 + \overline{\mathbf{B}}_t^2 \in \mathbb{G}^{(2)}(\mathbb{R}^d)$ of order r, where

$$\overline{\mathbf{B}}_t^1 := \sum_{i=1}^{d} B_t^i V_i \in \mathbb{R}^d, \quad \overline{\mathbf{B}}_t^2 := \int_0^t \int_0^s \circ dB_u \otimes \circ dB_s + t\overline{\boldsymbol{\beta}}(\Phi_0) \in \mathbb{R}^d \otimes \mathbb{R}^d.$$

6 Example

In this section, we discuss an example of the modified harmonic realizations associated with non-symmetric random walks on a nilpotent covering graph whose covering transformation group Γ is the 3-dimensional discrete Heisenberg group $\mathbb{H}^3(\mathbb{Z})$.

We now consider the 3-*dimensional Heisenberg dice lattice*. This graph is defined by a covering graph of a finite graph consisting of three vertices with a covering transformation group $\Gamma = \mathbb{H}^3(\mathbb{Z})$ (see Figure 3). We emphasize that it is regarded as an extension of the *dice graph* discussed in Namba [49] to the nilpotent case.

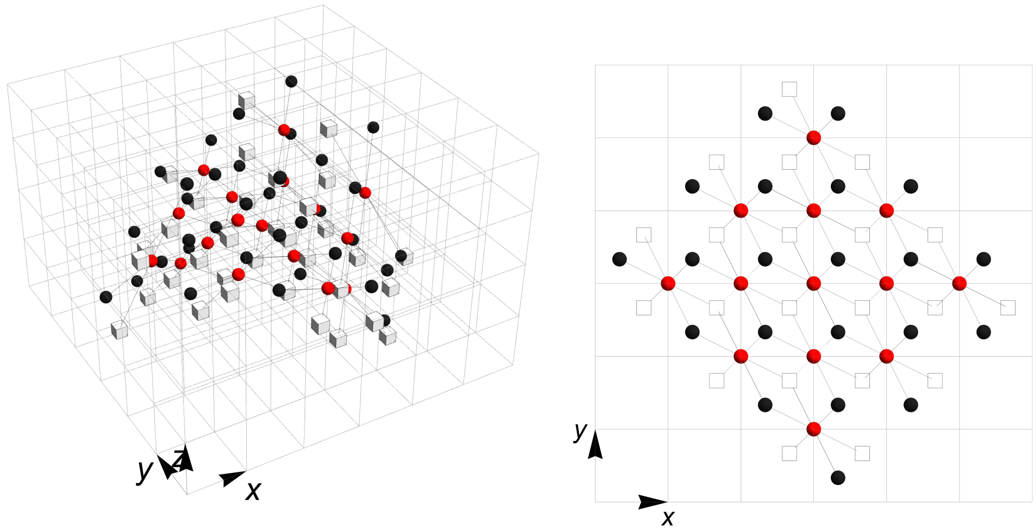

Figure 3: A part of 3-dimensional Heisenberg dice lattice and the projection of it on the xy-plane

Suppose that $\Gamma = \mathbb{H}^3(\mathbb{Z})$ is generated by two elements $\gamma_1 = (1,0,0)$ and $\gamma_2 = (0,1,0)$. We also set two elements $\mathbf{g}_1 := (1/3, 1/3, 1/3)$, $\mathbf{g}_2 := (-1/3, -1/3, -1/3)$ in $G = \mathbb{H}^3(\mathbb{R})$. We put

$$V_1 := \big\{g = \gamma_{i_1}^{\varepsilon_1} \star \cdots \star \gamma_{i_\ell}^{\varepsilon_\ell} \star \mathbf{1}_G \,\big|\, i_k \in \{1,2\},\, \varepsilon_k = \pm 1\, (1 \le k \le \ell),\, \ell \in \mathbb{N} \cup \{0\}\big\},$$
$$V_2 := \big\{g = \gamma_{i_1}^{\varepsilon_1} \star \cdots \star \gamma_{i_\ell}^{\varepsilon_\ell} \star \mathbf{g}_1 \,\big|\, i_k \in \{1,2\},\, \varepsilon_k = \pm 1\, (1 \le k \le \ell),\, \ell \in \mathbb{N} \cup \{0\}\big\},$$
$$V_3 := \big\{g = \gamma_{i_1}^{\varepsilon_1} \star \cdots \star \gamma_{i_\ell}^{\varepsilon_\ell} \star \mathbf{g}_2 \,\big|\, i_k \in \{1,2\},\, \varepsilon_k = \pm 1\, (1 \le k \le \ell),\, \ell \in \mathbb{N} \cup \{0\}\big\}.$$

We consider an $\mathbb{H}^3(\mathbb{Z})$-nilpotent covering graph $X = (V, E)$ defined by $V = V_1 \sqcup V_2 \sqcup V_3$ and $E = E_1 \sqcup E_2$, where

$$E_1 := \big\{(g,h) \in V_1 \times V_2 \,|\, g^{-1} \star h = \mathbf{g}_1,\, \gamma_1^{-1} \star \mathbf{g}_1,\, \gamma_2^{-1} \star \mathbf{g}_1\big\},$$
$$E_2 := \big\{(g,h) \in V_1 \times V_3 \,|\, g^{-1} \star h = \mathbf{g}_2,\, \gamma_1 \star \mathbf{g}_2,\, \gamma_2 \star \mathbf{g}_2\big\}.$$

We note that X is invariant under the actions γ_1 and γ_2. Its quotient graph $X_0 = (V_0, E_0) = \Gamma\backslash X$ is given by $V_0 = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ and $E_0 = \{e_i, \overline{e}_i \,|\, 1 \le i \le 6\}$ (cf. Figure 4).

From now on we define a non-symmetric random walk on X. We define the transition probability $p : E \to (0,1]$ by

$$p\big((g, g \star \mathbf{g}_1)\big) = \xi, \quad p\big((g, g \star \gamma_1^{-1} \star \mathbf{g}_1)\big) = \eta, \quad p\big((g, g \star \gamma_2^{-1} \star \mathbf{g}_1)\big) = \zeta,$$
$$p\big((g, g \star \mathbf{g}_2)\big) = \zeta, \qquad p\big((g, g \star \gamma_1 \star \mathbf{g}_2)\big) = \eta, \qquad p\big((g, g \star \gamma_2 \star \mathbf{g}_2)\big) = \xi,$$
$$p\big(\overline{(g, g \star \mathbf{g}_1)}\big) = \gamma, \quad p\big(\overline{(g, g \star \gamma_1^{-1} \star \mathbf{g}_1)}\big) = \beta, \quad p\big(\overline{(g, g \star \gamma_2^{-1} \star \mathbf{g}_1)}\big) = \alpha,$$
$$p\big(\overline{(g, g \star \mathbf{g}_2)}\big) = \alpha, \qquad p\big(\overline{(g, g \star \gamma_1 \star \mathbf{g}_2)}\big) = \beta, \qquad p\big(\overline{(g, g \star \gamma_2 \star \mathbf{g}_2)}\big) = \gamma,$$

for every $g \in V_1$, where $\xi, \eta, \zeta, \alpha, \beta, \gamma > 0$, $2(\xi+\eta+\zeta) = 1$ and $\alpha+\beta+\gamma = 1$. The invariant measure $m : V_0 = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\} \to (0, 1]$ is given by $m(\mathbf{x}) = 1/2$ and $m(\mathbf{y}) = m(\mathbf{z}) = 1/4$. Note that this random walk is (m-)symmetric if and only if $\alpha = 2\zeta$, $\beta = 2\eta$ and $\gamma = 2\xi$.

The first homology group $\mathrm{H}_1(X_0, \mathbb{R})$ is spanned by the four 1-cycles

$$[c_1] := [e_1 * \overline{e}_2], \quad [c_2] := [e_1 * \overline{e}_3], \quad [c_3] := [e_4 * \overline{e}_5], \quad [c_4] := [e_4 * \overline{e}_6].$$

Then the homological direction is calculated as

$$\gamma_p = \frac{\beta - 2\eta}{4}[c_1] + \frac{\alpha - 2\zeta}{4}[c_2] + \frac{\beta - 2\eta}{4}[c_3] + \frac{\gamma - 2\xi}{4}[c_4].$$

The canonical surjective linear map $\rho_{\mathbb{R}} : \mathrm{H}_1(X_0, \mathbb{R}) \to \mathfrak{g}^{(1)}$ is given by

$$\rho_{\mathbb{R}}([c_1]) = X_1, \quad \rho_{\mathbb{R}}([c_2]) = X_2, \quad \rho_{\mathbb{R}}([c_3]) = -X_1, \quad \rho_{\mathbb{R}}([c_4]) = -X_2.$$

Then we obtain

$$\rho_{\mathbb{R}}(\gamma_p) = \frac{(\alpha - \gamma) - 2(\zeta - \xi)}{4} X_2. \tag{6.1}$$

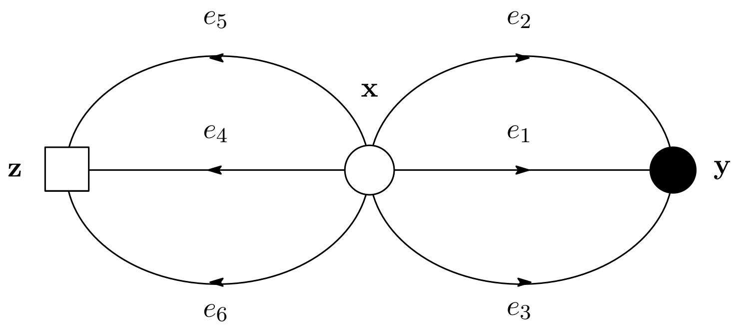

Figure 4: The quotient $X_0 = (V_0, E_0)$ of the 3D-Heisenberg dice graph $X = (V, E)$

We write $\{u_1, u_2\} \subset \mathrm{Hom}(\mathfrak{g}^{(1)}, \mathbb{R})$ for the dual basis of $\{X_1, X_2\} \subset \mathfrak{g}^{(1)}$. We also denote by $\{\omega_1, \omega_2, \omega_3, \omega_4\} \subset \big(\mathrm{H}^1(X_0, \mathbb{R}), \langle\!\langle \cdot, \cdot \rangle\!\rangle_p\big)$ the dual basis of $\{[c_1], [c_2], [c_3], [c_4]\} \subset \mathrm{H}_1(X_0, \mathbb{R})$. Namely, $\omega_i([c_j]) = \delta_{ij}$ for $1 \le i, j \le 4$. Then the modified harmonicity (3.8) yields

$$\begin{aligned}
&\omega_1(e_1) = \beta - \frac{\beta - 2\eta}{4}, && \omega_1(e_2) = -(1-\beta) - \frac{\beta - 2\eta}{4}, && \omega_1(e_3) = \beta - \frac{\beta - 2\eta}{4}, \\
&\omega_1(e_4) = -\frac{\beta - 2\eta}{4}, && \omega_1(e_5) = -\frac{\beta - 2\eta}{4}, && \omega_1(e_6) = -\frac{\beta - 2\eta}{4}, \\
&\omega_2(e_1) = \alpha - \frac{\alpha - 2\zeta}{4}, && \omega_2(e_2) = \alpha - \frac{\alpha - 2\zeta}{4}, && \omega_2(e_3) = -(1-\alpha) - \frac{\alpha - 2\zeta}{4}, \\
&\omega_2(e_4) = -\frac{\alpha - 2\zeta}{4}, && \omega_2(e_5) = -\frac{\alpha - 2\zeta}{4}, && \omega_2(e_6) = -\frac{\alpha - 2\zeta}{4}, \\
&\omega_3(e_1) = -\frac{\beta - 2\eta}{4}, && \omega_3(e_2) = -\frac{\beta - 2\eta}{4}, && \omega_3(e_3) = -\frac{\beta - 2\eta}{4}, \\
&\omega_3(e_4) = \beta - \frac{\beta - 2\eta}{4}, && \omega_3(e_5) = -(1-\beta) - \frac{\beta - 2\eta}{4}, && \omega_3(e_6) = \beta - \frac{\beta - 2\eta}{4}, \\
&\omega_4(e_1) = -\frac{\gamma - 2\xi}{4}, && \omega_4(e_2) = -\frac{\gamma - 2\xi}{4}, && \omega_4(e_3) = -\frac{\gamma - 2\xi}{4}, \\
&\omega_4(e_4) = \gamma - \frac{\gamma - 2\xi}{4}, && \omega_4(e_5) = \gamma - \frac{\gamma - 2\xi}{4}, && \omega_4(e_6) = -(1-\gamma) - \frac{\gamma - 2\xi}{4}.
\end{aligned}$$

By direct computation, we have

$$
\begin{aligned}
&\langle\!\langle \omega_1,\omega_1\rangle\!\rangle_p=\frac{\beta+2\eta}{4}-\frac{(\beta+2\eta)^2}{8}, && \langle\!\langle \omega_1,\omega_2\rangle\!\rangle_p=-\frac{(\alpha+2\zeta)(\beta+2\eta)}{8},\\
&\langle\!\langle \omega_1,\omega_3\rangle\!\rangle_p=-\frac{(\beta-2\eta)^2}{8}, && \langle\!\langle \omega_1,\omega_4\rangle\!\rangle_p=-\frac{(\beta-2\eta)(\gamma-2\xi)}{8},\\
&\langle\!\langle \omega_2,\omega_2\rangle\!\rangle_p=\frac{\alpha+2\zeta}{4}-\frac{(\alpha+2\zeta)^2}{8}, && \langle\!\langle \omega_2,\omega_3\rangle\!\rangle_p=-\frac{(\alpha-2\zeta)(\beta-2\eta)}{8},\\
&\langle\!\langle \omega_2,\omega_4\rangle\!\rangle_p=-\frac{(\alpha-2\zeta)(\gamma-2\xi)}{8}, && \langle\!\langle \omega_3,\omega_3\rangle\!\rangle_p=\frac{\beta+2\eta}{4}-\frac{(\beta+2\eta)^2}{8},\\
&\langle\!\langle \omega_3,\omega_4\rangle\!\rangle_p=-\frac{(\beta+2\eta)(\gamma+2\xi)}{8}, && \langle\!\langle \omega_4,\omega_4\rangle\!\rangle_p=\frac{\gamma+2\xi}{4}-\frac{(\gamma+2\xi)^2}{8}.
\end{aligned}
\tag{6.2}
$$

Since the linear space $\mathrm{Hom}(\mathfrak{g}^{(1)},\mathbb{R})$ can be seen as a 2-dimensional subspace of $\mathrm{H}^1(X_0,\mathbb{R})$ through the injection ${}^t\rho_{\mathbb{R}}$, we see that $u_1={}^t\rho_{\mathbb{R}}(u_1)=\omega_1-\omega_3$ and $u_2={}^t\rho_{\mathbb{R}}(u_2)=\omega_2-\omega_4$ form a $\mathbb{Z}$-basis in $\mathrm{Hom}(\mathfrak{g}^{(1)},\mathbb{R})$. We then obtain

$$
\langle\!\langle u_1,u_1\rangle\!\rangle_p=\frac{\beta+2\eta-4\beta\eta}{2},\quad \langle\!\langle u_1,u_2\rangle\!\rangle_p=-\frac{\beta+2\eta-4\beta\eta}{4},
$$
$$
\langle\!\langle u_2,u_2\rangle\!\rangle_p=\frac{(\beta+2\eta)(2-\beta-2\eta)+4\alpha\gamma+16\xi\zeta}{8}.
$$

by (6.2). Thus the volume of the Albanese torus is computed as

$$
\mathrm{vol}(\mathrm{Alb}^{\Gamma})^{-1}=\frac{1}{4}\sqrt{(\beta+2\eta-4\beta\eta)\big\{(\beta+2\eta)-(\beta^2+4\eta^2)+4\alpha\gamma+16\xi\zeta\big\}}.
$$

Furthermore, the Albanese metric g_0 on $\mathfrak{g}^{(1)}$ is given by

$$
\langle X_1,X_1\rangle_{g_0}=\frac{(\beta+2\eta)(2-\beta-2\eta)+4\alpha\gamma+16\xi\zeta}{8}\mathrm{vol}(\mathrm{Alb}^{\Gamma}),
$$
$$
\langle X_1,X_2\rangle_{g_0}=\frac{\beta+2\eta-4\beta\eta}{4}\mathrm{vol}(\mathrm{Alb}^{\Gamma}),\quad \langle X_2,X_2\rangle_{g_0}=\frac{\beta+2\eta-4\beta\eta}{2}\mathrm{vol}(\mathrm{Alb}^{\Gamma}).
$$

We now determine the modified standard realization $\Phi_0: X\to G=\mathbb{H}^3(\mathbb{R})$. Let $\widetilde{e}_i$ $(i=1,2,3,4,5,6)$ be a lift of $e_i\in E_0$ to X and put $\Phi_0\big(o(\widetilde{e}_i)\big)=\mathbf{1}_G$. Then it follows from (2.2) and (6.1) that the Γ-equivariant realization $\Phi_0: X\to G$ satisfying

$$
\begin{aligned}
&\Phi_0\big(t(\widetilde{e}_1)\big)=\Big(\beta,\frac{(3\alpha+\gamma)+2(\zeta-\xi)}{4},\kappa_1\Big),\\
&\Phi_0\big(t(\widetilde{e}_2)\big)=\Big(\beta-1,\frac{(3\alpha+\gamma)+2(\zeta-\xi)}{4},\kappa_1-\frac{(3\alpha+\gamma)+2(\zeta-\xi)}{4}\Big),\\
&\Phi_0\big(t(\widetilde{e}_3)\big)=\Big(\beta,\frac{(3\alpha+\gamma)+2(\zeta-\xi)}{4}-1,\kappa_1\Big),\\
&\Phi_0\big(t(\widetilde{e}_4)\big)=\Big(-\beta,\frac{-(\alpha+3\gamma)+2(\zeta-\xi)}{4},-\kappa_2\Big),\\
&\Phi_0\big(t(\widetilde{e}_5)\big)=\Big(1-\beta,\frac{-(\alpha+3\gamma)+2(\zeta-\xi)}{4},-\kappa_2+\frac{-(\alpha+3\gamma)+2(\zeta-\xi)}{4}\Big),\\
&\Phi_0\big(t(\widetilde{e}_6)\big)=\Big(-\beta,\frac{-(\alpha+3\gamma)+2(\zeta-\xi)}{4}+1,-\kappa_2\Big)
\end{aligned}
$$

is the modified harmonic realization, where κ_1,κ_2 is two real parameters which indicates the ambiguity of the realization corresponding to $\mathfrak{g}^{(2)}$. Let $\{v_1,v_2\}$ be the Gram–Schmidt orthonormalization of the basis $\{u_1,u_2\}$, that is,

$$
v_1=\langle\!\langle u_1,u_1\rangle\!\rangle_p^{-1/2}u_1,\quad v_2=\langle\!\langle u_1,u_1\rangle\!\rangle_p^{1/2}\mathrm{vol}(\mathrm{Alb}^{\Gamma})\Big(u_2-\frac{\langle\!\langle u_1,u_2\rangle\!\rangle_p}{\langle\!\langle u_1,u_1\rangle\!\rangle_p}u_1\Big),
$$

and $\{V_1, V_2\} \subset \mathfrak{g}^{(1)}$ its dual basis. We write $V_3 := [V_1, V_2] = V_1V_2 - V_2V_1$. Then we obtain

$$v_1 = \Big(\frac{\beta + 2\eta - 4\beta\eta}{2}\Big)^{-1/2} u_1, \quad v_2 = \Big(\frac{\beta + 2\eta - 4\beta\eta}{2}\Big)^{1/2} \mathrm{vol}(\mathrm{Alb}^\Gamma)\Big(u_2 + \frac{1}{2}u_1\Big)$$

by (6.2). Moreover, we have

$$\begin{aligned} V_1 &= \Big(\frac{\beta + 2\eta - 4\beta\eta}{2}\Big)^{1/2} X_1 - \frac{1}{2}\Big(\frac{\beta + 2\eta - 4\beta\eta}{2}\Big)^{1/2} X_2, \\ V_2 &= \Big(\frac{\beta + 2\eta - 4\beta\eta}{2}\Big)^{-1/2} \mathrm{vol}(\mathrm{Alb}^\Gamma)^{-1} X_2, \\ V_3 &= \mathrm{vol}(\mathrm{Alb}^\Gamma)^{-1} X_3. \end{aligned}$$

Finally, we see that $\beta(\Phi_0) \in \mathfrak{g}^{(2)}$ and the infinitesimal generator $\mathcal{A}$ are calculated as

$$\begin{aligned} \beta(\Phi_0) &= \sum_{i=1}^{6} \big(\widetilde{m}(e_i) - \widetilde{m}(\overline{e}_i)\big) \log\Big(d\Phi_0(\widetilde{e}_i) \cdot \exp\big(-\rho_{\mathbb{R}}(\gamma_p)\big)\Big)\Big|_{\mathfrak{g}^{(2)}} = \frac{\beta - 2\eta}{8} \mathrm{vol}(\mathrm{Alb}^\Gamma) V_3, \\ \mathcal{A} &= -\frac{1}{2}(V_1^2 + V_2^2) - \frac{\beta - 2\eta}{8} \mathrm{vol}(\mathrm{Alb}^\Gamma) V_3. \end{aligned}$$

We should observe that the coefficient of $\beta(\Phi_0)$ does not include the parameters κ_1 and κ_2, though the realization Φ_0 has the ambiguity of $\mathfrak{g}^{(2)}$-components.

A A comment on CLTs in the non-centered case

As was already mentioned, the centered condition **(A3)** is crucial to establish the functional CLT (Theorem 2.2). We present a method to reduce the non-centered case $\rho_{\mathbb{R}}(\gamma_p) \neq \mathbf{0}_{\mathfrak{g}}$ to the centered case by employing a measure-change technique based on Alexopoulos [2]. See also Namba [49] for this kind of technique in the case where X is a crystal lattice.

We consider a *positive* transition probability $p : E \to (0, 1]$ to avoid several technical difficulties. Then the random walk on X associated with p is automatically irreducible. Let $\Phi_0 : X \to G$ be the (p-)modified harmonic realization. We define a function $F = F_x(\lambda) : V_0 \times \mathrm{Hom}(\mathfrak{g}^{(1)}, \mathbb{R}) \to \mathbb{R}$ by

$$F_x(\lambda) := \sum_{e \in (E_0)_x} p(e) \exp\Big({}_{\mathrm{Hom}(\mathfrak{g}^{(1)},\mathbb{R})}\big\langle \lambda, \log\big(d\Phi_0(\widetilde{e})\big)\big|_{\mathfrak{g}^{(1)}}\big\rangle_{\mathfrak{g}^{(1)}}\Big) \tag{A.1}$$

for $x \in V_0$ and $\lambda \in \mathrm{Hom}(\mathfrak{g}^{(1)}, \mathbb{R})$. Since the lemma below is obtained by following the argument in [49, Lemma 3.1], we omit the proof.

Lemma A.1. For every $x \in V_0$, the function $F_x(\cdot) : \mathrm{Hom}(\mathfrak{g}^{(1)}, \mathbb{R}) \to (0, \infty)$ has a unique minimizer $\lambda_* = \lambda_*(x) \in \mathrm{Hom}(\mathfrak{g}^{(1)}, \mathbb{R})$.

We now define a positive function $\mathfrak{p} : E_0 \to (0, 1]$ by

$$\mathfrak{p}(e) := \frac{\exp\Big({}_{\mathrm{Hom}(\mathfrak{g}^{(1)},\mathbb{R})}\big\langle \lambda_*\big(o(e)\big), \log\big(d\Phi_0(\widetilde{e})\big)\big|_{\mathfrak{g}^{(1)}}\big\rangle_{\mathfrak{g}^{(1)}}\Big)}{F_{o(e)}\big(\lambda_*\big(o(e)\big)\big)} p(e) \qquad (e \in E_0). \tag{A.2}$$

It is straightforward to check that the function $\mathfrak{p}$ also gives a positive transition probability on X_0 and it yields an irreducible Markov chain $(\Omega_x(X), \widehat{\mathbb{P}}_x, \{w_n^{(\mathfrak{p})}\}_{n=0}^\infty)$ with values in X. We then find a unique positive normalized invariant measure $\mathfrak{m} : V_0 \to (0, 1]$ by applying the Perron-Frobenius theorem again. We set $\widetilde{\mathfrak{m}}(e) := \mathfrak{p}(e)\mathfrak{m}\big(o(e)\big)$ for $e \in E_0$. We also denote by $\mathfrak{p} : E \to (0, 1]$ and $\mathfrak{m} : V \to (0, 1]$ the Γ-invariant lifts of $\mathfrak{p} : E_0 \to (0, 1]$ and $\mathfrak{m} : V_0 \to (0, 1]$ to X, respectively. The Albanese metric on $\mathfrak{g}^{(1)}$ associated with the transition probability $\mathfrak{p}$ is denoted by $g_0^{(\mathfrak{p})}$. We write $\{V_1^{(\mathfrak{p})}, V_2^{(\mathfrak{p})}, \dots, V_{d_1}^{(\mathfrak{p})}\}$ for an orthonormal basis of $(\mathfrak{g}^{(1)}, g_0^{(\mathfrak{p})})$.

Let $L_{(\mathfrak{p})} : C_\infty(X) \to C_\infty(X)$ be the transition operator associated with the transition probability $\mathfrak{p}$. By virtue of Lemma A.1, we have

$$\sum_{e \in (E_0)_x} p(e) \exp\Big({}_{\mathrm{Hom}(\mathfrak{g}^{(1)},\mathbb{R})}\big\langle \lambda_*, \log\big(d\Phi_0(\widetilde{e})\big)\big|_{\mathfrak{g}^{(1)}} \big\rangle_{\mathfrak{g}^{(1)}} \Big) \log\big(d\Phi_0(\widetilde{e})\big)\big|_{\mathfrak{g}^{(1)}} = \mathbf{0}_{\mathfrak{g}} \qquad (x \in V_0).$$

Hence, we conclude

$$(L_{(\mathfrak{p})} - I)(\log \Phi_0\big|_{\mathfrak{g}^{(1)}})(x) = \sum_{e \in E_x} \mathfrak{p}(e) \log\big(d\Phi_0(e)\big)\big|_{\mathfrak{g}^{(1)}} = \mathbf{0}_{\mathfrak{g}} \qquad (x \in V). \tag{A.3}$$

This means that the (p-)modified harmonic realization $\Phi_0 : X \to G$ in the sense of (2.2) is regarded as the ($\mathfrak{p}$-)harmonic realization and $\rho_{\mathbb{R}}(\gamma_{\mathfrak{p}}) = \mathbf{0}_{\mathfrak{g}}$.

We fix a reference point $x_* \in V$ such that $\Phi_0(x_*) = \mathbf{1}_G$ and put

$$\xi_n^{(\mathfrak{p})}(c) := \Phi_0\big(w_n^{(\mathfrak{p})}(c)\big) \qquad \big(n \in \mathbb{N} \cup \{0\},\, c \in \Omega_{x_*}(X)\big).$$

This yields a G-valued random walk $(\Omega_{x_*}(X), \widehat{\mathbb{P}}_{x_*}, \{\xi_n^{(\mathfrak{p})}\}_{n=0}^\infty)$. We define

$$\mathcal{Y}_{t_k}^{(n;\mathfrak{p})}(c) := \tau_{n^{-1/2}}\big(\xi_{nt_k}^{(\mathfrak{p})}(c)\big) = \tau_{n^{-1/2}}\big(\Phi_0(w_k^{(\mathfrak{p})}(c))\big)$$

for $k = 0, 1, \dots, n$, $t_k \in \mathcal{D}_n$ and $c \in \Omega_{x_*}(X)$. We consider a G-valued stochastic process $(\mathcal{Y}_t^{(n;\,\mathfrak{p})})_{0 \le t \le 1}$ defined by the d_{CC}-geodesic interpolation of $\{\mathcal{Y}_{t_k}^{(n;\,\mathfrak{p})}\}_{k=0}^n$. Let $(\widetilde{Y}_t)_{0 \le t \le 1}$ be the G-valued diffusion process which solves the SDE

$$d\widetilde{Y}_t = \sum_{i=1}^{d_1} V_{i*}^{(\mathfrak{p})}(\widetilde{Y}_t) \circ dB_t^i + \beta^{(\mathfrak{p})}(\Phi_0)_*(\widetilde{Y}_t)\, dt, \qquad \widetilde{Y}_0 = \mathbf{1}_G,$$

where

$$\beta^{(\mathfrak{p})}(\Phi_0) := \sum_{e \in E_0} \widetilde{\mathfrak{m}}(e) \log\Big(\Phi_0\big(o(\widetilde{e})\big)^{-1} \cdot \Phi_0\big(t(\widetilde{e})\big)\Big)\Big|_{\mathfrak{g}^{(2)}}.$$

The following two theorems are CLTs for non-symmetric random walks associated with the changed transition probability $\mathfrak{p}$. We remark that the proofs of these theorems below are done by combining the ones of Theorems 2.1 and 2.2 with the argument in [49, Theorem 1.3].

Theorem A.2. Let $P_\varepsilon : C_\infty(G) \to C_\infty(X)$ be the approximation operator defined by $P_\varepsilon f(x) := f\big(\tau_\varepsilon\big(\Phi_0(x)\big)\big)$ for $0 \le \varepsilon \le 1$ and $x \in V$. Then we have, for $0 \le s \le t$ and $f \in C_\infty(G)$,

$$\lim_{n \to \infty} \Big\| L_{(\mathfrak{p})}^{[nt]-[ns]} P_{n^{-1/2}} f - P_{n^{-1/2}} \mathrm{e}^{-(t-s)\mathcal{A}_{(\mathfrak{p})}} f \Big\|_\infty^X = 0, \tag{A.4}$$

where $(\mathrm{e}^{-t\mathcal{A}_{(\mathfrak{p})}})_{t \ge 0}$ is the C_0-semigroup with the infinitesimal generator $\mathcal{A}_{(\mathfrak{p})}$ on $C_0^\infty(G)$ defined by

$$\mathcal{A}_{(\mathfrak{p})} := -\frac{1}{2} \sum_{i=1}^{d_1} (V_{i*}^{(\mathfrak{p})})^2 - \beta^{(\mathfrak{p})}(\Phi_0)_*. \tag{A.5}$$

Theorem A.3. The sequence $(\mathcal{Y}_t^{(n;\,\mathfrak{p})})_{0 \le t \le 1}$ converges in law to the G-valued diffusion process $(\widetilde{Y}_t)_{0 \le t \le 1}$ in $C_{\mathbf{1}_G}^{0,\alpha\text{-Höl}}([0,1]; G)$ as $n \to \infty$ for all $\alpha < 1/2$.

We emphasize that the transition probability $\mathfrak{p}$ coincides with the given one p under the centered condition **(A3)**. Therefore, Theorems A.2 and A.3 are regarded as extensions of Theorems 2.1 (under the centered condition **(A3)**) and 2.2 to the non-centered case. We might prove Theorem 2.2 without the centered condition **(A3)** via Theorem A.3. We will discuss this problem in the future.

References

[1] G. Alexopoulos: *Convolution powers on discrete groups of polynomial volume growth*, Canad. Math. Soc. Conf. Proc. **21** (1997), pp. 31–57. MR-1472777

[2] G. Alexopoulos: *Random walks on discrete groups of polynomial growth*, Ann. Probab. **30** (2002), pp. 723–801. MR-1905856

[3] G. Alexopoulos: *Sub-Laplacians with drift on Lie groups of polynomial volume growth*, Mem. Amer. Math. Soc. **155** (2002), no. 739. MR-1878341

[4] C. Bayer and P.K. Friz: *Cubature on Wiener space: pathwise convergence*, Appl. Math. Optim. **67** (2013), pp. 261–278. MR-3027591

[5] G. Ben Arous: *Flots et Taylor stochastiques*, Probab. Theory Relat. Fields **81** (1989), pp. 29–77. MR-0981567

[6] E. Breuillard: *Random walks on Lie groups*, preprint (2006), 45 pages. Available at the author's webpage (https://www.math.u-psud.fr/~breuilla/publ.html). MR-2705764

[7] E. Breuillard: *Local limit theorems and equidistribution of random walks on the Heisenberg group*, Geom. Funct. Anal. **15** (2005), pp. 35–82. MR-2140628

[8] E. Breuillard, P. Friz and M. Huesmann: *From random walks to rough paths*, Proc. Amer. Math. Soc. **137** (2009), pp. 3487–3496. MR-2515418

[9] F. Castell: *Asymptotic expansion of stochastic flows*, Probab. Theory Relat. Fields **96** (1993), pp. 225–239. MR-1227033

[10] I. Chevyrev: *Random walks and Lévy processes as rough paths*, Probab. Theory Relat. Fields **170** (2018), pp. 891–932. MR-3773803

[11] P. Crépel and A. Raugi: *Théorème central limite sur les groupes nilpotents*, Ann. Inst. Henri. Poincaré, Sect. B (N.S.) **14** (1978), pp.145–164. MR-0507730

[12] J.-D. Deuschel, T. Orenshtein and N. Perkowski: *Additive functionals as rough paths*, preprint (2019), arXiv:1912.09819.

[13] P. Diaconis and B. Hough: *Random walk on unipotent matrix groups*, preprint (2017), 42 pages, arXiv:1512.06304.

[14] P.K. Friz, P. Gassiat and T. Lyons: *Physical Brownian motion in a magnetic field as a rough path*, Trans. Amer. Math. Soc. **367** (2015), pp. 7939–7955. MR-3391905

[15] P.K. Friz and M. Hairer: A Course on Rough Paths, With an introduction to regularity structures, Universitext, Springer, 2014. MR-3289027

[16] P.K. Friz and H. Oberhauser: *Rough path limits of the Wong-Zakai type with a modified drift term*, J. Funct. Anal. **256** (2009), pp. 3236–3256. MR-2504524

[17] P.K. Friz and M. Riedel: *Convergence rates for the full Brownian rough paths with applications to limit theorems for stochastic flows*, Bull. Sci. Math. **135** (2011), pp. 613–628. MR-2838093

[18] P.K. Friz and N.B. Victoir: Multidimensional Stochastic Processes as Rough Paths, Theory and Applications, Cambridge Studies in Advanced Mathematics **120**, Cambridge Univ. Press, Cambridge, 2010. MR-2604669

[19] R.W. Goodman: Nilpotent Lie Groups, Structure and Applications to Analysis, LNM **562**, Springer-Verlag, Berlin Heidelberg, New York, 1976. MR-0442149

[20] M. Gromov: *Groups of polynomial growth and expanding maps*, IHES. Publ. Math. **53** (1981), pp. 53–73. MR-0623534

[21] Y. Guivar'ch: *Applications d'un theoreme limite local a la transience et a la recurrence de marches de Markov*, Colloquede Théorie du Potientiel, Lecture Notes in Math. **1096**, Springer, Berlin, 1984, pp. 301–332. MR-0890364

[22] R. Hough: *The local limit theorem on nilpotent Lie groups*, Probab. Theory Relat. Fields. **174** (2019), pp. 761–786. MR-3980304

[23] S. Ishiwata: *A central limit theorem on a covering graph with a transformation group of polynomial growth*, J. Math. Soc. Japan **55** (2003), pp. 837–853. MR-1978225

[24] S. Ishiwata: *A Berry-Esseen type theorem on nilpotent covering graphs*, Canad. J. Math. **56** (2004), pp. 963–982. MR-2085630

[25] S. Ishiwata, H. Kawabi and M. Kotani: *Long time asymptotics of non-symmetric random walks on crystal lattices*, J. Funct. Anal. **272** (2017), pp. 1553–1624. MR-3590246

[26] S. Ishiwata, H. Kawabi and R. Namba: *Central limit theorems for non-symmetric random walks on nilpotent covering graphs: Part II*, To appear in Potential Analysis.

[27] I. Karatzas and S. Shreve: Brownian Motion and Stochastic Calculus, Second Edition. GTM **113**, Springer, 1991. MR-1121940

[28] A. Klenke: Probability Theory, A Comprehensive Course. Universitext. Springer-Verlag London, Ltd., London, 2008. MR-2372119

[29] M. Kotani: *A central limit theorem for magnetic transition operators on a crystal lattice*, J. London Math. Soc. **65** (2002), pp. 464–482. MR-1883195

[30] M. Kotani: *An asymptotic of the large deviation for random walks on a crystal lattice*, Contemp. Math. **347** (2004), pp. 141–152. MR-2077035

[31] M. Kotani and T. Sunada: *Albanese maps and off diagonal long time asymptotics for the heat kernel*, Comm. Math. Phys. **209** (2000), pp. 633–670. MR-1743611

[32] M. Kotani and T. Sunada: *Standard realizations of crystal lattices via harmonic maps*, Trans. Amer. Math. Soc. **353** (2000), pp. 1–20. MR-1783793

[33] M. Kotani and T. Sunada: *Large deviation and the tangent cone at infinity of a crystal lattice*, Math. Z. **254** (2006), pp. 837–870. MR-2253471

[34] M. Kotani, T. Shirai and T. Sunada: *Asymptotic behavior of the transition probability of a random walk on an infinite graph*, J. Funct. Anal. **159** (1998), pp. 664–689. MR-1658100

[35] S.M. Kozlov: *The averaging method and walks in inhomogeneous environments*, Russian Math. Surveys **40** (1985), pp. 73–145. MR-0786087

[36] A. Kramli and D. Szasz: *Random walks with integral degrees of freedom*, I, *Local limit theorem*, Z. Wahrsch. Verw. Gebiete **63** (1983), pp. 85–95. MR-0699788

[37] T. Kumagai: Random Walks on Disordered Media and their Scaling Limits, École d'Été de Probabilités de Saint-Flour XL-2010, LNM **2101**, Springer, Cham, 2014. MR-3156983

[38] H. Kunita: *On the representation of solutions of stochastic differential equations*, in Séminaire de Probabilités XIV, LNM **784**, pp.282–304. MR-0580134

[39] T.G. Kurtz: *Extensions of Trotter's semigroup approximation theorems*, J. Funct. Anal. **3** (1969), pp. 354–375. MR-0242016

[40] G. Lawler, V. Limic: Ramdom Walk: A Modern Introduction, Cambridge Studies in Advanced Mathematics **123**, 2010. MR-2677157

[41] Y. Le Jan: *Markov loops, free field and Eulerian networks*, J. Math. Soc. Japan **67** (2015), pp. 1671–1680. MR-3417508

[42] Y. Le Jan: Markov Paths, Loops and Fields, LNM **2026**, Springer Heidelberg, 2011. MR-2815763

[43] A. Lejay and T. Lyons: *On the importance of the Lévy area for studying the limits of functions of converging stochastic processes. Application to homogenization*, In: Current Trends in Potential Theory, Theta Ser. Adv. Math. **4** (2005), pp. 63–84. MR-2243956

[44] O. Lopusanschi, T. Orenshtein: *Ballistic random walks in random environment as rough paths: convergence and area anomaly*, preprint (2018), arXiv:1812.01403.

[45] O. Lopusanschi, D. Simon: *Lévy area with a drift as a renormalization limit of Markov chains on periodic graphs*, Stoch. Proc. Appl. **128** (2018), pp. 2404–2426. MR-3804798

[46] T. Lyons: *Differential equations driven by rough signals*, Rev. Math. Iberoamericana **14** (1998), pp. 215–310. MR-1654527

[47] T. Lyons and Z. Qian: System Control and Rough Paths, Oxford Mathematical Monographs. Oxford Univ. Press, Oxford, 2002. MR-2036784

[48] A.I. Malćev: *On a class of homogeneous spaces*, Amer. Math. Soc. Transl. **39** (1951), pp. 276–307. MR-0039734

[49] R. Namba: *A remark on a central limit theorem for non-symmetric random walks on crystal lattices*, Math. J. Okayama. Univ. **60** (2018), pp. 109–135. MR-3728433

[50] R. Namba: *Central limit theorems for non-symmetric random walks on covering graphs*, Ph.D. thesis at Okayama University (2019), available at http://ousar.lib.okayama-u.ac.jp/files/public/5/56797/20190625084049750154/K0005956fulltext.pdf. MR-3728433

[51] N. Ozawa: *A functional analysis proof of Gromov's polynomial growth theorem*, Ann. Sci. Ec. Norm. Super. **51** (2018), pp. 551–558. MR-3831031

[52] G. Pap: *Central limit theorems on stratified Lie groups*, In: Probability Theory and Mathematical Statistics (Vilnius, 1993), pp. 613–627, TEV, Vilnius, 1994. MR-1649603

[53] G.C. Papanicolaou and S.R.S. Varadhan: *Boundary value problem with rapidly oscillating random coefficients*, In: Random Fields, Vol. I, II (Esztergom, 1979), pp. 835–873, Colloq. Math. Soc. János Bolyai, **27**, North-Holland, Amsterdam-New York, 1981. MR-0712714

[54] M.S. Raghunathan: Discrete Subgroups of Lie Groups, Springer-Verlag Berlin, 1972. MR-0507234

[55] A. Raugi: Thèoréme de la limite centrale sur les groupes nilpotents, Z. Wahrsch. Verw. Gebiete. **43** (1978), pp. 149–172. MR-0499429

[56] D.W. Robinson: Elliptic Operators and Lie Groups, Oxford Mathematical Monographs, Oxford Univ. Press, New York, 1991. MR-1144020

[57] L. Saloff-Coste: *Probability on groups: random walks and invariant diffusions*, Notices Amer. Math. Soc. **48** (2001), pp. 968–977. MR-1854532

[58] F. Spitzer: Principles of Random Walks, D. Van Nostrand, Princeton, NJ, 1964. MR-0171290

[59] D.W. Stroock: Probability Theory. An Analytic View, Cambridge Univ. Press, 1993. MR-1267569

[60] D.W. Stroock and S.R.S. Varadhan: *Limit theorems for random walks on Lie groups*, Sankhya Ser. A **35** (1973), pp. 277–294. MR-0517406

[61] T. Sunada: *Large deviation for non-symmetric random walks on crystal lattices*, Lecture slide at Humboldt University, https://www.researchgate.net/publication/268218756_Large_deviation_for_non-symmetric_random_walks_on_crystal_lattices, 2006.

[62] T. Sunada: *Discrete geometric analysis*, In: Analysis on Graphs and its Applications, pp. 51–83, Proc. Sympos. Pure Math. **77**, Amer. Math. Soc., Providence, RI, 2008. MR-2459864

[63] T. Sunada: Topological Crystallography with a View Towards Discrete Geometric Analysis, Surveys and Tutorials in the Applied Mathematical Sciences **6**, Springer Japan, 2013. MR-3014418

[64] T. Sunada: *Topics in mathematical crystallography*, In: Proceedings of the Symposium "Groups, Graphs and Random Walks", London Math. Soc., Lecture Note Series **436**, Cambridge Univ. Press, 2017, pp. 473–513. MR-3644018

[65] R. Tanaka: *Large deviation on a covering graph with group of polynomial growth*, Math. Z. **267** (2011), pp. 803–833. MR-2776059

[66] H.F. Trotter: *Approximation of semi-groups of operators*, Pacific J. Math. **8** (1958), pp. 887–919. MR-0103420

[67] V.N. Tutubalin: *Composition of measures on the simplest nilpotent group*, Theory Probab. Appl. **9** (1964), pp. 479–487. MR-0169287

[68] J.C. Watkins: *Donsker's invariance principle for Lie groups*, Ann. Probab. **17** (1989), pp. 1220–1242. MR-1009453

[69] D. Wehn: *Probabilities on Lie groups*, Proc. Nat. Acad. Sci. USA **48** (1962), pp. 791–795. MR-0153042

[70] W. Woess: Random Walks on Infinite Graphs and Groups, Cambridge Tracts in Mathematics **138**, Cambridge University Press, Cambridge, 2000. MR-1743100

Acknowledgments. The authors are grateful to Professor Shoichi Fujimori for making pictures of the 3-dimensional Heisenberg dice lattice and kindly allowing them to use these pictures in the present paper. They would also like to thank Professors Takafumi Amaba, Takahiro Aoyama, Peter Friz, Naotaka Kajino, Atsushi Katsuda, Takashi Kumagai, Seiichiro Kusuoka, Kazumasa Kuwada, Laurent Saloff-Coste and Ryokichi Tanaka for helpful discussions and encouragement. A part of this work was done during the stay of R. N. at Hausdorff Center for Mathematics, Universität Bonn in March 2017 with the support of research fund of Research Institute for Interdisciplinary Science, Okayama University. He would like to thank Professor Massimiliano Gubinelli for warm hospitality and helpful discussions. S. I. was supported by JSPS Grant-in-Aid for Young Scientists (B) No. 25800034 and JSPS Grant-in-Aid for Scientific Research (C) No. 17K05215. H. K. was supported by JSPS Grant-in-Aid for Scientific Research (C) No. 26400134 and (C) No. 17K05300. R. N. was supported by JSPS Research Fellowships for Young Scientists No. 18J10225.