Groups Geom. Dyn. 9 (2015), 1047–1129
DOI 10.4171/GGD/335

Groups, Geometry, and Dynamics

Random walks on nilpotent groups driven by measures supported on powers of generators

Laurent Saloff-Coste and Tianyi Zheng[1]

Abstract. We study the decay of convolution powers of a large family $\mu_{S,a}$ of measures on finitely generated nilpotent groups. Here, $S = (s_1, \dots, s_k)$ is a generating k-tuple of group elements and $a = (\alpha_1, \dots, \alpha_k)$ is a k-tuple of reals in the interval $(0, 2)$. The symmetric measure $\mu_{S,a}$ is supported by $S^* = \{s_i^m, 1 \leq i \leq k, m \in \mathbb{Z}\}$ and gives probability proportional to $(1+m)^{-\alpha_i - 1}$ to $s_i^{\pm m}$, $i = 1, \dots, k$, $m \in \mathbb{N}$. We determine the behavior of the probability of return $\mu_{S,a}^{(n)}(e)$ as n tends to infinity. This behavior depends in somewhat subtle ways on interactions between the k-tuple a and the positions of the generators s_i within the lower central series $G_j = [G_{j-1}, G]$, $G_1 = G$.

Mathematics Subject Classification (2010). 20F18, 20F65, 60J10, 60J51.

Keywords. Random walk, stable laws, nilpotent groups.

Contents

[1] Both authors partially supported by NSF grant DMS 1004771.

1. Introduction

Generating sets play an essential role in the theory of countable groups. This is obvious when a group is defined by generators and relations or when a group is defined as the subgroup generated by a given finite subset of elements in a much larger group. In this context, the larger ambient group serves as a sort of "black box" that encodes the law of the group.

Given a group G with finite symmetric generating set A, the simple random walk on G can be interpreted as a way to randomly explore the group G. Starting at the identity element e, the position of the walk at time n is the product $\xi_1 \ldots \xi_n$ where the G-valued random variables ξ_i are independent equidistributed with law given by the uniform probability on the set A. More generally, given a probability measure μ on G, the random walk driven by μ corresponds to taking the sequence (ξ_i) to be i.i.d. with law μ and the position at time n has distribution $\mu^{(n)}$, the n-fold convolution product of μ with itself. In particular, $\mathbf{P}_e(\xi_1 \ldots \xi_n = e) = \mu^{(n)}(e)$. In the case of the simple random walk based on the generating set A, $\mu = |A|^{-1}\mathbf{1}_A$.

Not surprisingly, many aspects of the behavior of these random processes are closely related to the algebraic and geometric property of the underlying group G. Harry Kesten introduced this question in his Ph.D. thesis published in 1958. One of Kesten's fundamental results states that, for a random walk driven by a symmetric measure with generating support, the probability of return, $\mathbf{P}_e(\xi_1 \ldots \xi_n = e)$, decays exponentially fast if and only the group G is non-amenable. See [16, 15].

1.1. The measures $\mu_{S,a}$. This is the first of a series of papers where we study a natural family of random walks driven by measures $\mu_{S,a}$ which are defined as follows. The letter S represents a finite generating tuple, i.e., a list $S = (s_1, s_2, \ldots, s_k)$ of generators (repetitions are permitted). In addition, we are given a k-tuple a of (extended) positive reals $a = (\alpha_1, \alpha_2, \ldots, \alpha_k)$, $\alpha_i \in (0, \infty]$. The measure $\mu_{S,a}$ allows long steps along any of the one-parameter group $\langle s_i \rangle = \{s_i^n : n \in \mathbb{Z}\}$, $1 \le i \le k$. The probability of such a long step along $\langle s_i \rangle$ is given by a power law whose exponent α_i is the i-th entry of the tuple a. Namely, we set,

$$\mu_{S,a}(g) = \frac{1}{k} \sum_{i=1}^{k} c(\alpha_i) \sum_{m \in \mathbb{Z}} (1 + |m|)^{-\alpha_i - 1} \mathbf{1}_{s_i^m}(g) \tag{1.1}$$

where

$$c(\alpha)^{-1} = \sum_{\mathbb{Z}} (1 + |m|)^{-\alpha - 1}.$$

We make the somewhat arbitrary convention that if $\alpha = \infty$ then $(1+|m|)^{-\alpha-1} = 0$ unless $m = 0, \pm 1$ in which case $(1 + |m|)^{-\alpha-1} = 1$. Note that $\mu_{S,a}$ is symmetric, that is, satisfies $\mu_{S,a}(g^{-1}) = \mu_{S,a}(g)$. We can also describe $\mu_{S,a}$ as the push-forward of the probability measure μ_a on the free group $\mathbf{F}_k$ on k generators $\mathbf{s}_i$, $1 \le i \le k$, which gives probability

$$\mu_a(\mathbf{s}_i^{\pm m}) = k^{-1} c(\alpha_i)(1+|m|)^{-\alpha_i - 1} \quad \text{to } \mathbf{s}_i^{\pm m}.$$

Indeed, if π is the projection from $\mathbf{F}_k$ onto G which sends $\mathbf{s}_i$ to s_i,

$$\mu_{S,a}(g) = \mu_a(\pi^{-1}(g)).$$

On $\mathbb{Z}$, the power laws $\mu_\alpha(\pm k) = c(\alpha)(1+|k|)^{-\alpha-1}$ are very natural probability measures. For $\alpha \in (0,2)$, μ_α can be viewed as a discrete version of the symmetric stable laws which is the probability distribution on $\mathbb{R}$ whose Laplace transform is $e^{-|y|^\alpha}$.

The main result of this paper, Theorem 1.2 below, describes the behavior of

$$n \longmapsto \mu_{S,a}^{(n)}(e)$$

when G is any given finitely generated nilpotent group, S is any given finite generating tuple of elements of G and the entries of the tuple a are in $(0,2)$. What makes this problem interesting is the interaction between the nature of the long jumps allowed in the directions of each generators and the non-commutative structure of the group. As we shall see, the behaviors of the random walks driven by the measures $\mu_{S,a}$ capture a wealth of information on the algebraic structure of G.

Because of the results of [19] – in particular, Theorem 1.9 stated below – the very precise form of the measure $\mu_{S,a}$ defined at (1.1) is not really essential in determining the behavior of $n \mapsto \mu_{S,a}^{(n)}(e)$. Indeed, any symmetric measure ν on G such that $c\nu \le \mu_{S,a} \le C\nu$ will satisfy

$$\nu^{(kn)}(e) \le K\mu_{S,a}^{(n)}(e) \quad \text{and} \quad \mu_{S,a}^{(kn)}(e) \le K\nu^{(n)}(e)$$

for some k, K independent of n.

1.2. The case of $\mathbb{Z}^d$. In the $G = \mathbb{Z}$, one can apply classical Fourier analysis techniques or the results from [10] to find that if $a = (\alpha_i)_1^k \in (0,\infty]^k$ and we set

$$\alpha = \min\{\alpha_i\} \quad \text{and} \quad \tilde{\alpha} = \min\{\alpha, 2\}$$

then

$$\mu_{S,a}^{(n)}(0) \simeq \begin{cases} n^{-1/\tilde{\alpha}} & \text{if } \alpha \ne 2, \\ (n \log n)^{-1/2} & \text{if } \alpha = 2. \end{cases}$$

Here and in the rest of this paper $\sim$ and $\simeq$ are used with the following meaning. For two functions f, g defined either over the positive reals or the natural numbers, we write

$$f \sim g$$

(usually, at 0 or infinity), if

$$\lim f/g = 1.$$

We write

$$f \simeq g$$

if there are constants c_1 such that

$$c_1 f(c_2 t) \leq g(t) \leq c_3 f(c_4 t)$$

(in a neighborhood of the relevant value, usually 0 or infinity). We recommend to restrict the use of $\simeq$ to cases where one of the two functions f or g is monotone.

In the next simplest case where

- $G = \mathbb{Z}^2 = \{(x, y) : x, y \in \mathbb{Z}\}$,
- $S = \{(1,0), (0,1)\}$, and
- $a = (\alpha_1, \alpha_2) \in (0, \infty]^2$,

it is not hard to see that $\mu_{S,a}^{(n)}(e)$, $e = (0,0)$, behaves as follows. Set

$$\tilde{\alpha} = \min\{\alpha, 2\},$$

$$\frac{1}{\beta} = \frac{1}{\tilde{\alpha}_1} + \frac{1}{\tilde{\alpha}_2}$$

and

$$\gamma = \#\{i : \alpha_i = 2\}.$$

(1) If $2 \notin \{\alpha_1, \alpha_2\}$, then

$$\mu_{S,a}^{(n)}(e) \sim c(\alpha_1, \alpha_2) n^{-1/\beta} p$$

(2) If $2 \in \{\alpha_1, \alpha_2\}$, then

$$\mu_{S,a}^{(n)}(e) \simeq n^{-1/\beta} (\log n)^{-\gamma/2}.$$

We briefly review what happens when $G = \mathbb{Z}^d$ and $S = (s_1, \dots, s_k)$, $k \geq d$. By hypothesis, S is generating. Given $a = (\alpha_1, \dots, \alpha_k)$, we extract from S a d-tuple $\Sigma = (\sigma_1, \dots, \sigma_d)$ using the following algorithm. Set $\Sigma_1 = \{\sigma_1 = s_{i_1}\}$ where $\alpha_{i_1} = \min\{\alpha_i : 1 \leq i \leq k\}$. For $t \geq 1$, if

$$\Sigma_t = (\sigma_1, \dots, \sigma_t), \quad \sigma_1 = s_{i_1}, \dots, \sigma_t = s_{i_t}$$

have been chosen, pick $\sigma_{t+1} = s_{i_{t+1}}$ in $\{s_i : 1 \leq i \leq k\}$ with the properties that $\alpha_{i_{t+1}} = \min\{\alpha_j : j \notin \{i_1, \dots, i_t\}\}$ and the rank of the lattice generated by $\Sigma_{t+1} = \Sigma_t \cup \{\sigma_{t+1}\}$ is (strictly) greater than the rank of the lattice generated by Σ_t. Note that the final d-tuple Σ might not generates $\mathbb{Z}^d$ but does generate a lattice of finite index in $\mathbb{Z}^d$. Set $a(\Sigma) = (\alpha_{i_1}, \dots, \alpha_{i_d})$.

Theorem 1.1. *Let $G = \mathbb{Z}^d$. Let $S = (s_i)_1^k$ be a generating k-tuple. Let $a = (\alpha_i)_1^k \in (0, \infty]^k$. Let $\Sigma = (\sigma_i)_1^d$ and $a(\Sigma)$ be obtained from (S, a) by the algorithm described above. Set*

$$\gamma = \#\{j \in \{1, \dots, d\} : \alpha_{i_j} = 2\}$$

and

$$\frac{1}{\beta} = \sum_{s=1}^{d} \frac{1}{\tilde{\alpha}_{i_s}},$$

where $\tilde{\alpha} = \min\{\alpha, 2\}$. Then we have

$$\mu_{S,a}^{(n)}(e) \simeq \mu_{\Sigma,a(\Sigma)}^{(n)}(e) \simeq n^{-1/\beta}[\log n]^{-\gamma/2}$$

With some work, this result can be extracted from [9] which treats a much larger class of examples. More precisely, Griffin explains in [9] how to construct what he call a *minimal orthonormal basis* (MONB) of $\mathbb{R}^d$ adapted to a given probability distribution (that drives a random walk on $\mathbb{R}^d$ but can be supported on $\mathbb{Z}^d$ as in our case). In the general case treated by Griffin, the choice of the (MONB) may change with the number of steps n taken by the random walk but in our case, the (MONB) will stay essentially the same for all n. Indeed, in the case of interest to us, Griffin algorithm [9, page 239] to pick the (MONB) is, in a sense, equivalent to the algorithm described above to pick Σ and the vector subspace spanned by the first i vectors on the (MNOB) essentially coincides with the vector subspace spanned by the first i generators in Σ. Further, each vector in the (MONB) comes equipped with a scaling coefficient that captures the natural scaling of the random walk in that particular direction. These scaling coefficients are directly related to the entries in $a(\Sigma)$. For each entry $\alpha(\sigma_i)$ of $a(\Sigma)$, the associated scaling coefficient is $n^{-1/\alpha(\sigma_i)}$ if $\alpha(\sigma_i) \in (0,2)$, $n^{-1/2}$ is $\alpha(\sigma_i) > 2$ and $(n \log n)^{-1/2}$ if $\alpha(\sigma_i) = 2$. The computation of these scaling coefficients (in [9] as well as in what follows) is directly related to the (directional) truncated second moments associated to the probability measure driving the random walk in question.

1.3. The main result in its simplest form.

The goal of this paper is to prove the following theorem together with more sophisticated assorted results.

Theorem 1.2. *Let G be a nilpotent group equipped with a generating k-tuple $S = (s_i)_1^k$ and $a = (\alpha_i)_1^k \in (0,\infty]^k$. Assume that the subgroup generated by $\{s_i : \alpha_i < 2\}$ is of finite index in G. Then there exists a real $D \geq 0$ depending on (G, S, a) such that*

$$\mu_{S,a}^{(n)}(e) \simeq n^{-D}.$$

This statement suggests further questions including the following three.

- Can we compute D? how does it depends on S, a and G?
- What happen if the subgroup generated by $\{s_i : \alpha_i < 2\}$ is not of finite index in G?
- What happens on other groups? In particular, how does Theorem 1.2 generalize to finitely generated groups of polynomial volume growth?

The first question will be answered completely in this paper. Indeed, we would not be able to prove the above theorem without a detailed understanding of how to compute the real D. In particular, D belongs to the subring of the reals generated by the inverses of the α_i with $\alpha_i < 2$. The exact value of D depends in an intricate and interesting way on (a) the commutator structure of G, (b) the position of the generators s_i in the commutator structure of G and (c) the values of the parameters α_i. See Theorem 1.8 in the next subsection.

The second question is rather subtle and will not be completely elucidated in this paper although some partial results will be obtain in this direction including what we believe is a sharp upper-bound. See Theorem 4.8. A technically very challenging goal would be to extend the abelian results of Griffin [9] in the context of nilpotent groups.

In its full generality, the third question is too wide ranging to be discussed here in details. Partial results for various classes of groups (e.g., some classes of solvable groups and free groups) will be discussed elsewhere. The question regarding groups of polynomial growth is tantalizing but appears surprisingly difficult to attack.

1.4. Why studying the measures $\mu_{S,a}$? Since this is a long paper devoted to the random walks driven by the measures $\mu_{S,a}$, the question of why these measures are interesting should be addressed. The theory of random walk is about understanding how the algebraic structure of the underlying group G and the properties of the measure driving the random walk interact and are reflected in the behavior of the random walk itself. If the measure is symmetric finitely supported then the central stage is taken by the algebraic structure of G (some may prefer to say, its geometry). Next, suppose that G is equipped with a "norm" $\|\cdot\|$ ($\|g^{-1}\| = \|g\| \geq 0$, $\|g_1 g_2\| \leq \|g_1\| + \|g_2\|$ and $\|g\| = 0$ only if $g = e$). Then one may want to understand the behavior of the random walks driven by "radial" measures of the type $\nu(g) = 1/\psi(\|g\|)$. See Section 1.7 and Theorem 5.1 for results in this direction. In this case, the role played by the underlying group structure is very weak and may, in some cases, disappear entirely. See [1, 2] and the references therein. Further these radial measures are rather abstract objects. Take for instance the case when $\|\cdot\|$ is the word-length $|\cdot|$ associated with a finite set of generators. If a group element g is given to us in one way or another, it is rather unlikely we can actually compute precisely $|g|$ (this, of course, is related to question about (geodesic) normal forms). Unless we can determine $\|g\|$, we just do not know what exactly $\nu(g)$ is. How could we, for example, simulate the random walk driven by ν? Even taken at a rhetorical level, this is a compelling question.

From this point of view, the measures $\mu_{S,a}$ (and there natural generalizations based on other classes of one dimensional measures) are among the rare natural examples of explicitly defined probability measures on a group. The present work demonstrates how, in the case of nilpotent groups, the behavior of the associated random walks is determined by a rich interaction between the algebraic structure of the group and the parameters defining $\mu_{S,a}$. Further evidence of the interest of these walks are provided in [24], especially, the proof of Theorem 1.5 in [24].

1.5. Weight systems and the value of D. The goal of this section is to give the reader a clear idea of the key ingredients that enter the exact computation of the real D governing the behavior of $\mu_{S,a}^{(n)}(e)$ in Theorem 1.2.

Consider $S = (s_1, \dots, s_k)$ as a formal alphabet equipped with a weight system $\mathfrak{w}$ which assigns weight $w_i \in (0, \infty)$ to the letter s_i, $1 \leq i \leq k$. We extend our alphabet by adjoining to each s_i its formal inverse s_i^{-1}. Using this alphabet, we build the set $\mathfrak{C}(S, m)$ of all formal commutators of length m by induction on m. Commutators of length 1 are the letters in $S^{\pm 1}$. Commutators of length m are the formal expression c of the form $c = [c_1, c_2]$ where c_1, c_2 are commutators of length $m_1, m_2 \geq 1$ with $m_1 + m_2 = m$.

The commutators of length 2 are (the ± 1 must be understood here as independent of each other)

$$[s_i^{\pm 1}, s_j^{\pm 1}], \quad 1 \le i, j \le k.$$

The commutators of length 3 are

$$[[s_i^{\pm 1}, s_j^{\pm 1}], s_\ell^{\pm 1}], \; [s_i^{\pm 1}, [s_j^{\pm 1}, s_\ell^{\pm 1}]], \quad 1 \le i, j, \ell \le k.$$

For $1 \le i_1, i_2, i_3, i_4 \le k$, the commutators of length 4 are

$$[[[s_{i_1}^{\pm 1}, s_{i_2}^{\pm 1}], s_{i_3}^{\pm 1}], s_{i_4}^{\pm 1}], \quad [[s_{i_1}^{\pm 1}, [s_{i_2}^{\pm 1}, s_{i_3}^{\pm 1}]], s_{i_4}^{\pm 1}], \quad [[s_{i_1}^{\pm 1}, s_{i_2}^{\pm 1}], [s_{i_3}^{\pm 1}, s_{i_4}^{\pm 1}]]$$

$$[s_{i_1}^{\pm 1}, [[s_{i_2}^{\pm 1}, s_{i_3}^{\pm 1}], s_{i_4}^{\pm 1}]], \quad [s_{i_1}^{\pm 1}, [s_{i_2}^{\pm 1}, [s_{i_3}^{\pm 1}, s_{i_4}^{\pm 1}]]].$$

To any formal commutators we can associate its build-word and its group-word. The build-word of a commutator c is the word over S that list the entries of c in order after one removes brackets and ± 1. So, the build-word of

$$c = [[s_{i_1}^{\pm 1}, s_{i_2}^{\pm 1}], [s_{i_3}^{\pm 1}, s_{i_4}^{\pm 1}]]$$

is

$$s_{i_1} s_{i_2} s_{i_3} s_{i_4}.$$

The group word is the word on $S^{\pm 1}$ obtained by applying repeatedly the group rules

$$[c_1, c_2]^{-1} = [c_2, c_1] \quad \text{and} \quad [c_1, c_2] = c_1^{-1} c_2^{-1} c_1 c_2.$$

So the group-word of

$$c = [[s_i, s_j^{-1}], s_\ell]$$

is

$$s_j s_i^{-1} s_j^{-1} s_i s_\ell^{-1} s_i^{-1} s_j s_i s_j^{-1} s_\ell.$$

Definition 1.3 (power weight systems). Given a k-tuple $(s_1, \ldots, s_k)$ of formal letters and a k-tuple $(w_1, \ldots, w_k)$ of positive reals, define the weight system $\mathfrak{w}$ on $\mathfrak{C}(S)$ by setting (inductively)

$$w(c) = w(c_1) + w(c_2) \text{ if } c = [c_1, c_2].$$

Let

$$\bar{w}_1 < \bar{w}_2 < \cdots < \bar{w}_j < \cdots$$

be the increasing sequence of the weight values of the weight system $\mathfrak{w}$. For $j = 1, 2, \ldots$, let $\mathfrak{C}_j^{\mathfrak{w}}$ be the set of all commutators c with $w(c) \ge \bar{w}_j$.

Clearly, the weight of a formal commutator is the sum of the weights of the letters appearing in its build-word. If $S = (s_1, s_2)$ and

$$w_1 = 3, \quad w_2 = 13/2,$$

then the weight-value sequence is

$$\bar{w}_1 = 3, \quad \bar{w}_2 = 6, \quad \bar{w}_3 = 13/2, \quad \bar{w}_4 = 9,$$
$$\bar{w}_5 = 12, \quad \bar{w}_6 = 25/2, \quad \bar{w}_7 = 13, \quad \ldots.$$

Given a group G generated by a k-tuple $S = (s_1, \ldots, s_k)$, any finite word ω on the alphabet $S^{\pm 1}$ has a well defined image $\pi_G(\omega)$ in G. Similarly, any formal commutator c on the alphabet $S^{\pm 1}$ has an image in G given by its group-word representation.

Definition 1.4 (group filtration associated to $\mathfrak{w}$). Let G be a nilpotent group equipped with a generating k-tuple $S = (s_1, \ldots, s_k)$ and a weight system $\mathfrak{w}$ generated by $(w_1, \ldots, w_k) \in (0, \infty)^k$. Set

$$G_j^{\mathfrak{w}} = \langle \mathfrak{C}_j^{\mathfrak{w}} \rangle.$$

That is, $G_j^{\mathfrak{w}}$ is the subgroup of G generated by the images of all formal commutators of weight greater or equal to $\bar{w}_j$. Let $j_* = j_*(G, S, \mathfrak{w})$ be the smallest integer such that $G_{j_*+1}^{\mathfrak{w}} = \{e\}$.

Example 1.1. Let G be the discrete Heisenberg group

$$G = \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} : x, y, z \in \mathbb{Z} \right\}.$$

Let

$$s_1 = X = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$s_2 = Y = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix},$$

$$s_3 = Z^5 = \begin{pmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

and

$$w_1 = 1,\ w_2 = 3/2,\ w_3 = 3.$$

In this case, the increasing sequence $\bar{w}_j$ is given by

$$\bar{w}_1 = 1, \quad \bar{w}_2 = 3/2, \quad \bar{w}_3 = 2, \quad \bar{w}_4 = 5/2,$$
$$\bar{w}_5 = 3, \quad \bar{w}_6 = 7/2, \quad \ldots,$$

and we have

$$G_6^{\mathfrak{w}} = \{e\},$$
$$G_5^{\mathfrak{w}} = \{s_3^k : k \in \mathbb{Z}\},$$
$$G_4^{\mathfrak{w}} = G_3^{\mathfrak{w}} = \left\{ \begin{pmatrix} 1 & 0 & z \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} : z \in \mathbb{Z} \right\},$$
$$G_2^{\mathfrak{w}} = \left\{ \begin{pmatrix} 1 & 0 & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} : y, z \in \mathbb{Z} \right\},$$
$$G_1^{\mathfrak{w}} = G.$$

Proposition 1.5. *Referring to the setting and notation of Definition 1.4, for all $j = 1, 2 \ldots,$ we have $G_j^{\mathfrak{w}} \supseteq G_{j+1}^{\mathfrak{w}}$ and $[G, G_j^{\mathfrak{w}}] \subseteq G_{j+1}^{\mathfrak{w}}$. In particular,*

$$G = G_1^{\mathfrak{w}} \supseteq G_2^{\mathfrak{w}} \supseteq \cdots \supseteq G_j^{\mathfrak{w}} \supseteq \cdots \supseteq G_{j_*}^{\mathfrak{w}} \supset G_{j_*+1}^{\mathfrak{w}} = \{e\}$$

is a descending normal series with $[G_j^{\mathfrak{w}}, G_j^{\mathfrak{w}}] \subseteq G_{j+1}^{\mathfrak{w}}$.

Proof. Recall that if X, Y are subsets of G, $[X, Y]$ denotes the subgroup generated by $\{[x, y] : x \in X, y \in Y\}$. Recall further that

$$[\langle X\rangle, \langle Y\rangle] = [X, Y]^{\langle X\rangle\langle Y\rangle}$$

where the right-hand side is the group generated by all conjugates of $[X, Y]$ by elements of the form $g = xy$, $x \in \langle X\rangle$, $y \in \langle Y\rangle$. Since $[f_1, f_j] \in \mathfrak{C}_{j+1}^{\mathfrak{w}}$ for all $f_1 \in \mathfrak{C}_1^{\mathfrak{w}}$, $f_j \in \mathfrak{C}_j^{\mathfrak{w}}$ and

$$[G, G_j^{\mathfrak{w}}] = [\mathfrak{C}_1^{\mathfrak{w}}, \mathfrak{C}_j^{\mathfrak{w}}]^G$$

it follows that

$$[G, G_j^{\mathfrak{w}}] \subseteq (G_{j+1}^{\mathfrak{w}})^G.$$

Thus a descending induction on j shows that the groups $G_j^{\mathfrak{w}}$ are all normal subgroups of G and that

$$[G, G_j^{\mathfrak{w}}] \subseteq G_{j+1}^{\mathfrak{w}}.$$

Note that it may happen that $G_j^{\mathfrak{w}} = G_{j+1}^{\mathfrak{w}}$ for some values of j, $1 < j < j_*$. For instance, it may happen that all formal commutators of a certain weight are trivial in G. In Example 1.1, $G_3^{\mathfrak{w}} = G_4^{\mathfrak{w}}$ because all commutators of weight $\bar{w}_3 = 2$ are obviously trivial. □

Definition 1.6. Referring to the setting and notation of Definition 1.4, let

$$R_j^{\mathfrak{w}} = \operatorname{rank}(G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}})$$

be the torsion free rank of the abelian group $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$.

By construction, the images of the formal commutators of weight $\bar{w}_j$ form a generating subset of $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$, $\quad j = 1, 2, \dots, j_*$. By definition, the torsion free rank of this abelian group is the minimal number of elements needed to generates $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$ modulo torsion.

Definition 1.7. Referring to the setup and notation of Definition 1.4, set

$$D(S, \mathfrak{w}) = \sum_1^{j_*} \bar{w}_j \operatorname{rank}(G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}).$$

Note that $D(S, \mathfrak{w})$ depends on the weights values $\bar{w}_j$ as well as on the algebraic relations between elements of S in G (via the rank of the group $G_j^{\mathfrak{w}}$).

Example 1.1 (continued). In Example 1.1, we have $j_* = 5$,

$$G_5^{\mathfrak{w}}/G_6^{\mathfrak{w}} = \mathbb{Z}, \quad G_4^{\mathfrak{w}}/G_5^{\mathfrak{w}} = \mathbb{Z}/5\mathbb{Z}, \quad G_3^{\mathfrak{w}}/G_4^{\mathfrak{w}} = \{0\},$$
$$G_2^{\mathfrak{w}}/G_3^{\mathfrak{w}} = \mathbb{Z} \quad G_1^{\mathfrak{w}}/G_2^{\mathfrak{w}} = \mathbb{Z}.$$

Hence

$$\operatorname{rank}(G_5^{\mathfrak{w}}/G_6^{\mathfrak{w}}) = 1,$$
$$\operatorname{rank}(G_4^{\mathfrak{w}}/G_5^{\mathfrak{w}}) = \operatorname{rank}(G_3^{\mathfrak{w}}/G_4^{\mathfrak{w}}) = 0,$$
$$\operatorname{rank}(G_2^{\mathfrak{w}}/G_3^{\mathfrak{w}}) = \operatorname{rank}(G_1^{\mathfrak{w}}/G_2^{\mathfrak{w}}) = 1$$

and

$$D(S, \mathfrak{w}) = 1 + 3/2 + 3 = 11/2$$

since

$$\bar{w}_1 = 1, \quad \bar{w}_2 = 3/2, \quad \bar{w}_3 = 2, \quad \bar{w}_4 = 5/2,$$
$$\bar{w}_5 = 3, \quad \bar{w}_6 = 7/2, \quad \dots.$$

Example 1.2. Assume that the weight w_i are all equal, namely,

$$w_i = v, \quad i = 1, \dots, k.$$

Then the weight-value sequence is given by

$$\bar{w}_j = jv$$

and j_* is equal to the nilpotency class of G. In this case, the descending normal series $G_j^{\mathfrak{w}}$ is the lower central series defined inductively by

- $G_1 = G$,
- $G_j = [G, G_{j-1}]$, $j \geq 2$,

and

$$D(S, \mathfrak{w}) = vD(G),$$

where

$$D(G) = \sum_1^{j_*} j \operatorname{rank}(G_j/G_{j+1}). \tag{1.2}$$

Theorem 1.8. *Let G be a nilpotent group equipped with a generating k-tuple $S = (s_i)_1^k$ and $a = (\alpha_i)_1^k \in (0, \infty]^k$. Assume that the subgroup generated by $\{s_i : \alpha_i < 2\}$ is of finite index in G. Consider the weight system $\mathfrak{w}(a) = \mathfrak{w}$ induced by setting $w_i = 1/\tilde{\alpha}_i$ where $\tilde{\alpha} = \min\{2, \alpha\}$. Then*

$$\mu_{S,a}^{(n)}(e) \simeq n^{-D(S,\mathfrak{w})}$$

with $D(S, \mathfrak{w})$ as in Definition 1.7.

Example 1.3. Let G be the discrete Heisenberg group equipped with the generating triple $S = (s_i)_1^3$ has in Example 1.1. Let $a = (\alpha_i)_1^3$. In this case, the condition that $\{s_i : \alpha_i < 2\}$ generates a subgroup of finite index is equivalent to $\alpha_1, \alpha_2 \in (0, 2)$. Let $\mathfrak{w}$ be as defined in Theorem 1.8. Then

$$D(S, \mathfrak{w}) = \frac{1}{\alpha_1} + \frac{1}{\alpha_2} + \max\left\{\frac{1}{\alpha_1} + \frac{1}{\alpha_2}, \frac{1}{\alpha_3}\right\}.$$

1.6. Some background on random walks. Given a finite symmetric generating set A, we set

$$|x|_A = \inf\{k : x \in A^k\}$$

(since $A^0 = \{e\}$, by convention, $|e| = 0$). This is called the *word-length* of x (with respect to the generating set A). With some abuse of notation, if $S = (s_1, \dots, s_k)$ is a generating k-tuple, we write $|\cdot|_S$ for the word-length associated with the symmetric generating set $\{s_i^{\pm 1}, 1 \le i \le k\}$. The *volume growth* of G (with respect to A) is the function

$$V_A(r) = \#\{g : |g|_A \le r\}.$$

The $\simeq$-equivalence class of the function V_A is independent of the choice of A. It is a group invariant called the growth function of G.

We say that a probability measure ϕ is *symmetric* if

$$\check{\phi} = \phi$$

where

$$\check{\phi}(x) = \phi(x^{-1}), \quad x \in G.$$

The *Dirichlet form* associated with ϕ is the quadratic form

$$\mathcal{E}_\phi(f, f) = \frac{1}{2} \sum_{x,y \in G} |f(xy) - f(x)|^2 \phi(y).$$

This form is fundamental in the study of random walks because of the following basic result.

Theorem 1.9 ([19]). *Assume that ϕ, ψ are two symmetric probability measures on a countable group G. If $\mathcal{E}_\phi \le C\,\mathcal{E}_\psi$ then*

$$\psi^{(2kn)}(e) \le 2\phi^{(2n)}(e) + 2e^{-2kn}, \quad k = [C] + 2.$$

This theorem will be use extensively in the present paper. In [19], it is used to prove that the long time asymptotic behavior of the probability of return is roughly the same for all random walks driven by symmetric measures with generating support and finite second moment.

Theorem 1.10 ([19]). *Assume that ϕ is a symmetric probability measure on a finitely generated group G with finite symmetric generating set A. Let u_A be the uniform probability measure on A. If ϕ satisfies*

$$\sum_{g \in G} |g|_A^2 \phi(g) < \infty \tag{1.3}$$

then there are constants k, C *such that*

$$u_A^{(2kn)}(e) \le C\phi^{(2n)}(e).$$

Further, if ϕ *satisfies* (1.3) *and* $\phi > 0$ *on a finite generating set then*

$$\phi^{(2n)}(e) \simeq u_A^{(2n)}(e).$$

This theorem implies that, if A and B are two symmetric finite generating sets of the group G, we have $u_A^{(2n)}(e) \simeq u_B^{(2n)}(e)$. Further, for any symmetric ϕ with finite second moment and generating support, $\phi^{(2n)}(e) \simeq u_A^{(n)}(e)$. In this sense, the equivalence class of the function $n \mapsto u_A^{(2n)}(e)$ under the equivalence relation $\simeq$ is a group invariant. This group invariant, which we denote by Φ_G, i.e.,

$$\Phi_G(n) \simeq u_A^{(2n)}(e), \tag{1.4}$$

has been studied extensively ([19] shows that Φ_G is invariant under quasi-isometries). In particular,

$$\Phi_G(n) \simeq \begin{cases} n^{-D/2} & \text{if } G \text{ has volume growth } V(r) \simeq r^D, \\ \exp(-n^{1/3}) & \text{if } G \text{ is polycyclic with exponential volume growth,} \\ \exp(-n) & \text{if } G \text{ is non-amenable.} \end{cases}$$

Nilpotent groups belong to the first category and have $D = D(G)$ given explicitly by (1.2). Many other behaviors beyond the three mentioned above are known to occurs and their are many groups for which Φ_G is unknown. See, e.g., [22, 23] and the references therein.

To explain how Theorem 1.10 applies to the measures $\mu_{S,a}$ defined at (1.1), we need the following definition.

Definition 1.11. Let G be a nilpotent group with descending lower central series G_j. The commutator length $\ell(g)$ of an element g of G is the supremum of the integers ℓ such that $g^m \in G_\ell$ for some integer m. In particular, by definition, torsion elements have infinite commutator length.

Corollary 1.12. *On any finitely generated group* G *equipped with a generating* k*-tuple* S*, we have*

$$\mu_{S,a}^{(n)}(e) \simeq \Phi_G(n) \simeq n^{-D(G)/2}$$

for all k*-tuple* $a = (\alpha_1, \dots, \alpha_k)$ *such that* $\alpha_i \ell(s_i) > 2$ *for all* $i = 1, \dots, k$.

Proof. It is well known that for any fixed $g \in G$, we have $|g^n|_S \simeq n^{1/\ell(g)}$ (see also Proposition 2.17 where a more general version of this fact is proved). It follows that, as long as the k-tuple a satisfies the condition stated in the corollary, $\mu_{S,a}$ has finite second moment. Hence, Theorem 1.10 implies $\mu_{S,a}^{(n)}(e) \simeq \Phi_G(n)$ as desired. $\square$

As a consequence of the more detailed results proved in this paper, we can state the following complementary result.

Theorem 1.13. *Let G be a nilpotent group equipped with a generating k-tuple S. Let $a \in (0,\infty]^k$. If there exists $i \in \{1,\dots,k\}$ such that $(\alpha_i, \ell(s_i)) = (2,1)$ or $\alpha_i \ell(s_i) < 2$ then we have*

$$\lim_{n\to\infty} [n^{D(G)/2} \mu_{S,a}^{(n)}(e)] = 0. \tag{1.5}$$

Regarding (1.5), we conjecture but are not able to prove that the sufficient condition provided by Theorem 1.13 is also necessary. See Theorems 5.11–5.12.

1.7. Radial stable laws. Let G be a finitely generated group with symmetric finite generating set A. Set $B_m = \{g \colon |g|_A \le m\}$. Define the radially symmetric "stable law" on G with index $\alpha \in (0,2)$ to be probability measure

$$\mu_\alpha(g) = c_\alpha \sum_{m=0}^{\infty} (1+m)^{-\alpha-1} \frac{\mathbf{1}_{B_m}(g)}{V_A(m)}, \quad c_\alpha^{-1} = \sum_0^\infty (1+m)^{-\alpha-1}$$

Note that μ_α is well defined for all $\alpha > 0$ and that

$$\sum_g |g|_A^\beta \mu_\alpha(g) < \infty \quad \text{for all } 0 < \beta < \alpha < \infty.$$

It is observed in [20, 21, 27] that

$$V_A(n) \ge cn^D, \text{ for all } n \implies \mu_\alpha^{(n)}(e) \le Cn^{-D/\alpha}, \text{ for all } n.$$

In addition, by [13, 4], for a given group G and for some/any $\alpha \ne 2$,

$$V_A(n) \simeq cn^D \iff \mu_\alpha^{(n)}(e) \simeq Cn^{-D/\tilde\alpha}, \quad \tilde\alpha = \min\{2,\alpha\}. \tag{1.6}$$

In fact, if we assume that the group G has polynomial volume growth $V(n) \simeq n^D$ then

$$\mu_\alpha(g) \simeq (1+|g|_A)^{-D-\alpha}.$$

Further, it follows from [13] that, for any $\alpha \in (0,2)$, there are constants $c_1(\alpha), c_2(\alpha)$ such that

$$c_1(\alpha)\mu_\alpha \le \nu_\alpha \le c_2(\alpha)\mu_\alpha$$

where ν_α denotes the measure that is α-subordinated to u_A in the sense of ([5]), that is,

$$\nu_\alpha = \sum_1^\infty \frac{\Gamma(n-\alpha)}{\Gamma(1-\alpha)\Gamma(n+1)} u_A^{(n)}.$$

Moreover, for any $\alpha \in (0,2)$,

$$\mu_\alpha^{(n)}(e) \simeq \nu_\alpha^{(n)}(e) \simeq n^{-D/\alpha} \quad \text{for all } n \in \mathbb{N}.$$

In [24], motivated by applications given below, the authors prove the following complementary statement regarding the behavior of μ_2.

Proposition 1.14 (Special case of [24, Theorem 1.5]). *Assume that G has polynomial volume growth $V_S(n) \simeq n^D$. Then we have*

$$\mu_2^{(n)}(e) \simeq (n \log n)^{-D/2}.$$

The lower bounds on $\mu_{S,a}^{(n)}(e)$ obtained in this paper are proved by establishing Dirichlet form comparisons involving appropriate generalization of the above radially symmetric stable measures and using Theorem 1.9.

1.8. Background on nilpotent groups. The classical setting for the study of random walks is the lattice $\mathbb{Z}^d$. See [25]. Since this work is concerned with random walks on nilpotent groups, we briefly discuss some of the similarities and differences between the lattice $\mathbb{Z}^d$ and finitely generated nilpotent groups. We also describe three basic examples.

The most fundamental similarity between a finitely generated nilpotent group G and the lattice $\mathbb{Z}^d$ is that, assuming that G is torsion free, there exists a real nilpotent Lie group $\mathbb{G}$ such that G can be identified with a discrete subgroup of $\mathbb{G}$ with compact quotient $\mathbb{G}/G$. In other words, G is a (co-compact) lattice in $\mathbb{G}$ in exactly the same way that $\mathbb{Z}^d$ is a lattice in $\mathbb{R}^d$ (except that the quotient is not a group, in general). This is a fundamental result of Malcev. See, e.g., Philip Hall famous notes [12]. However, simply connected real nilpotent Lie groups and their lattices are classified only in very small dimensions. See [8]. For instance, there are essentially 5 distinct "irreducible" simply connected real nilpotent Lie groups of dimension 5. In dimension 6, there are 34. In high dimension, the classification is unknown and there are continuously many different isomorphic classes.

From a technical viewpoint, the study of random walks on abelian groups is mostly based on the use of the Fourier transform (see [25]). Although the representation theory of (real) nilpotent Lie groups is well developed, it has proved very hard to use this theory to study random walks (except in some very particular cases). For these reasons, the study of random walks on nilpotent groups is often based on techniques that are rather different from the classical techniques used in the abelian case. This is certainly the case for the present work.

Example 1.4. Let $U(d)$ be the group of all upper triangular $d \times d$ matrices over $\mathbb{Z}$ with diagonal entries equal to 1. This group is a lattice in the nilpotent real Lie group $\mathbb{U}(d)$ of all upper triangular $d \times d$ matrices over the reals with diagonal entries equal to 1. Let $E_{i,j}$, $1 \le i < j \le d$, be the matrix in $\mathbb{U}(d)$ with all non-diagonal entries equal to 0 except for the entry in the i-th row and j-th column which equals 1. These elements are related by

$$E_{i,j} E_{\ell,m} = \delta_{j,\ell} E_{i,m}.$$

Further,

$$E_{i,j} = [E_{i,i+1}, [E_{i+1,i+2}, \dots, [E_{j-2,j-1}, E_{j-1,j}] \cdots]].$$

In particular, the $(d-1)$-tuple $S = (E_{i,i+1})_1^{d-1}$ is generating. For any $m = 1, \dots, d-1$, the elements $\{E_{i,i+m} \colon 1 \le i \le d-m\}$ can be expressed as commutators of length m on $S^{\pm 1}$ and form a minimal generating set for the subgroup

$$U(d)_m = [U(d), U(d)_{m-1}]$$

in the lower central series of $U(d)$. The nilpotency class of $U(d)$ is $d-1$, that is, any commutator of length greater than $d-1$ equals the identity in $U(d)$.

Any matrix $M = (m_{i,j})$ in $U(d)$ can (obviously) be written uniquely (order matters!)

$$M = \prod_{k=1}^{d-1} \Big(\prod_{i=0}^{k-1} E_{k-i,d-i}^{m_{k-i,d-i}} \Big)$$

where the $m_{i,j}$ are simply the entry of the matrix M. Much less trivially, there is also a unique expression of the form

$$M = \prod_{k=1}^{d-1} \Big(\prod_{i=k}^{d-1} E_{i-k+1,i+1}^{m'_{i-k+1,i+1}} \Big)$$

where $(m'_{i,j})_{1 \le i < j \le n}$ is obtained from $(m_{i,j})_{1 \le i < j \le n}$ by a polynomial bijective transformation with polynomial inverse.

Since $A = \{E_{i,i+1}^{\pm 1}, 1 \leq i \leq d-1\}$ generates $U(d)$, it is of great interest to describe the word length $|M|_A$ of a matrix $M \in U(d)$ in terms of the coordinate systems $(m_{i,j})_{1\leq i<j\leq d}$ and $(m'_{i,j})_{1\leq i<j\leq d}$. The answer is essentially the same in both cases, namely,

$$|M|_A \simeq \sum_{1\leq i<j\leq d} |m_{i,j}|^{1/|j-i|} \simeq \sum_{1\leq i<j\leq d} |m'_{i,j}|^{1/|j-i|}.$$

This well known (but non-trivial) result is the key to the volume growth estimate

$$V_{U(d),A}(r) \simeq r^{D(U(d))}, \quad D(U(d)) = \sum_{i=1}^{d-1} i(d-i)$$

and to the assorted random walk result (see, e.g., [28]) $\Phi_{U(d)}(n) \simeq n^{-D(U(d))/2}$. If we set

$$S = (s_i = E_{i,i+1})_1^{d-1},$$

then for any $a = (\alpha_i)_1^{d-1} \in (0,2)^{d-1}$ our main result yields

$$\mu_{S,a}^{(n)}(e) \simeq n^{-D}, \quad D = \sum_{1\leq i<j\leq d} \sum_{m=i}^{j-1} \frac{1}{\alpha_m}.$$

Example 1.5. The free nilpotent group of nilpotency class ℓ on k generators , $N(k,\ell)$, can be defined as the quotient of the free group on k generators by the normal subgroup generated by the images of all formal commutators of length greater than ℓ. This group has the (universal) property that it covers any k generated nilpotent group G of nilpotency class ℓ with a covering homomorphism sending the canonical generating k-tuple of $N(k,\ell)$ to the given generating k-tuple of G.

Marshal Hall gave a description of $N(k,\ell)$ in terms of the so-called "basic commutators". See [11, Chapter 11]. Let $(s_1,\dots,s_k)$ be the canonical generators of $N(k,\ell)$. Define the ordered set of all basic commutators $c_1 < \cdots < c_t$ using the following inductive procedure.

(1) $s_1,\dots,s_k$ are the basic commutators of length 1 and $s_1 < s_2 < \cdots < s_k$ by definition;

(2) for each m the basic commutators of length m are all commutators of the form $c = [c', c'']$ with c', c'' basic commutators of length m', m'' with $m' + m'' = m$ such that $c' > c''$ and, if $c' = [d', d'']$ $(d, d'$ basic commutators) then $c'' \geq d''$;

(3) commutators of length m come after commutators of length $m-1$ and are ordered arbitrary with respect to each other.

By a theorem of Witt (e.g., [11, Theorem 11.2.2]), the number of basic commutators of length m on k generators is

$$M_k(m) = m^{-1} \sum_{d|m} \mu(d) k^{m/d},$$

where μ denotes the classical Möbius function. M. Hall proved that the basic commutators of length m form a basis of the abelian group $N(k,\ell)_m/N(k,\ell)_{m+1}$ for $1 \le m \le \ell$ and that any element g of $N(k,\ell)$ can be written uniquely

$$g = \prod_1^t c_i^{x_i}, \quad x_i \in \mathbb{Z}.$$

Moreover, the length of g with respect to the generating set $A = \{s_i^{\pm 1}\}$ satisfies

$$|g|_A \simeq \sum_1^t |x_i|^{1/m_i},$$

where m_i is the commutator length of c_i. This gives the volume growth estimate

$$V_A(r) \simeq r^{D(N(k,\ell))}, \quad D(N(k,\ell)) = \sum_{m=1}^{\ell} m M_k(m) = \sum_{m=1}^{\ell} \sum_{d|m} \mu(d) k^{m/d}$$

and the assorted random walk estimate $\Phi_{N(k,\ell)}(n) \simeq n^{-D(N(k,\ell))/2}$.

In this case, the main result of the present work, together with Witt's theorem (e.g., [11, Theorem 11.2.2]), gives that for any k-tuple $a = (\alpha_i)_1^k \in (0,2)^k$, we have

$$\mu_{S,a}^{(n)}(e) \simeq n^{-D}$$

where

$$D = \sum_{m=1}^{\ell} \sum_{(m_1,\dots,m_k)\vdash m} \frac{1}{m} \Big(\sum_1^k \frac{m_i}{\alpha_i} \Big) \sum_{d|m_1,\dots,m_k} \mu(d) \binom{m/d}{m_1/d, \cdots, m_k/d}.$$

Example 1.6. Let G be the group

$$G = \langle u_1, \dots, u_\ell, t | [u_i, u_j] = 1; [u_i, t] = u_{i+1}, i < \ell; [u_\ell, t] = 1 \rangle$$

defined by generators and relations. This group is nilpotent of nilpotency class ℓ and it is generated by $S = (s_1 = u_1, s_2 = t)$ with G_m generated by $\{u_i : i \ge m\}$.

In this case, we have

$$\Phi_G(n) \simeq n^{-D(G)/2},$$

with

$$D(G) = 1 + \ell(\ell+1)/2.$$

If we let $a = (\alpha_1, \alpha_2) \in (0,2)^2$, our main result yields $\mu^{(n)}_{S,a}(e) \simeq n^{-D}$ with

$$D = \frac{\ell}{\alpha_1} + \frac{1 + (\ell-1)\ell/2}{\alpha_2}.$$

In any of the above examples, we can also consider other choices of generating tuples. For instance, in the current example, we can fix $j \in \{1, \dots, \ell - 1\}$ and consider the generating 3-tuple $S_j = (s_1 = u_1, s_2 = t, s_3 = u_{j+1})$ with $a' = (\alpha'_1, \alpha'_2, \alpha'_3) \in (0,2)^3$. In this case, our main result yields

$$\mu^{(n)}_{S_j,a'}(e) \simeq n^{-D}$$

with

$$D = \begin{cases} \dfrac{\ell}{\alpha'_1} + \dfrac{1 + (\ell-1)\ell/2}{\alpha'_2} & \text{if } \dfrac{1}{\alpha'_3} \le \dfrac{1}{\alpha'_1} + \dfrac{j}{\alpha'_2}, \\[2ex] \dfrac{j}{\alpha'_1} + \dfrac{1 + (j-1)j/2}{\alpha'_2} + \dfrac{\ell - j}{\alpha'_3} + \dfrac{(\ell - j - 1)(\ell - j)/2}{\alpha'_2} & \text{if } \dfrac{1}{\alpha'_3} > \dfrac{1}{\alpha'_1} + \dfrac{j}{\alpha'_2}. \end{cases}$$

2. Quasi-norms and approximate coordinates

This section describes results of an algebraic and geometric nature that play a key role in our study to the random walks driven by the measures $\mu_{S,a}$ defined at (1.1). One of the basic idea in the study of simple random walks on groups (i.e., the collection of random walks driven by the uniform probability measures u_A where A is a finite symmetric generating set) is that the notion of "volume growth" of the group leads to basic upper bounds on $u^{(2n)}_A(e)$: the faster the volume growth, the faster the decay of the probability of return. In the case of nilpotent group, this heuristic leads to sharp bounds. Indeed, for any given $D \ge 0$, $V_A(n) \simeq n^D$ if and only if $u^{(2n)}_A(e) \simeq n^{-D/2}$. See [28].

The estimates of $\mu_{S,a}^{(n)}(e)$ obtained in this work are based on a similar heuristic which requires us to define appropriate geometries associated with the different choices of S and a. This section defines these geometries and develop the needed key results.

2.1. Weight systems and weight-functions systems. We refer the reader to subsection 1.5 for notation regarding words and formal commutators over a finite alphabet $S^{\pm 1}$, $S = (s_1, \dots, s_k)$.

Definition 2.1 (multidimensional weight system). Given a k-tuple $(w_1, \dots, w_k)$ with

$$w_i \in (0, \infty) \times \mathbb{R}^{d-1}, \quad 1 \leq i \leq k,$$

let $\mathfrak{w}$ be the weight system

$$\mathfrak{w} \colon \mathfrak{C}(S) \ni c \longmapsto w(c) \in (0, \infty) \times \mathbb{R}^{d-1}$$

on the set $\mathfrak{C}(S)$ of all formal commutators on $S^{\pm 1}$ defined by

$$w(s_i^{\pm 1}) = w_i$$

and

$$w(c) = w(c_1) + w(c_2),$$

if $c = [c_1, c_2]$. Let

$$\bar{w}_1 < \bar{w}_2 < \cdots < \bar{w}_j < \cdots$$

be the ordered sequence of the values $w(c)$ when c runs over all formal commutators and $(0, \infty) \times \mathbb{R}^{d-1}$ is given the usual lexicographic order.

Note that we always have $w([c_1, c_2]) > \max\{w(c_1), w(c_2)\}$.

Definition 2.2. For each $j = 1, \dots,$ let $\mathfrak{C}_j(S)$ be the set of all formal commutators of weight at least $\bar{w}_j$. If G is a group generated by a k-tuple $S = (s_1, \dots, s_k)$, let

$$G_j^{\mathfrak{w}} = \langle \mathfrak{C}_j(S) \rangle$$

be the subgroup of G generated by the image in G of $\mathfrak{C}_j(S)$. Assuming that G is nilpotent, let $j_* = j_*(\mathfrak{w})$ be the smallest integer such that

$$G_{j_*+1}^{\mathfrak{w}} = \{e\}.$$

The proof of the following proposition is the same as that of Proposition 1.5.

Proposition 2.3. *Referring to the setting and notation of Definition 2.2, assume that G is nilpotent. Then, for all $j = 1, 2 \dots$, we have $G_j^{\mathfrak{w}} \supseteq G_{j+1}^{\mathfrak{w}}$ and $[G, G_j^{\mathfrak{w}}] \subseteq G_{j+1}^{\mathfrak{w}}$. In particular,*

$$G = G_1^{\mathfrak{w}} \supseteq G_2^{\mathfrak{w}} \supseteq \cdots \supseteq G_j^{\mathfrak{w}} \supseteq \cdots \supseteq G_{j*}^{\mathfrak{w}} \supset G_{j*+1}^{\mathfrak{w}} = \{e\}$$

is a descending normal series with $[G_j^{\mathfrak{w}}, G_j^{\mathfrak{w}}] \subseteq G_{j+1}^{\mathfrak{w}}$. We let $R_j^{\mathfrak{w}}$ be the torsion free rank of the abelian group $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$.

Definition 2.4 (Weight-function system). Given increasing functions

$$F_i : [1, \infty) \longrightarrow [1, \infty),$$

we define the weight-function system $\mathfrak{F}$ to be the collection of functions

$$F_c : [1, \infty) \longrightarrow [1, \infty),\ c \in \mathfrak{C}(S),$$

by setting inductively

$$F_{s_i^{\pm 1}} = F_i, \qquad 1 \leq i \leq k,$$

and

$$F_c = F_{c_1} F_{c_2} \quad \text{if } c = [c_1, c_2].$$

Remark 2.5. According to Definitions 2.1-2.4, if the build-sequence of the commutators c of length ℓ is $(u_1, \dots, u_\ell) \in S^\ell$ then

$$w(c) = \sum_1^\ell w_i, \quad F_c(r) = \prod_1^\ell F_i(r).$$

Remark 2.6. A key collection of examples of weight systems are the (one-dimensional) power-weight systems introduced in 1.3 where $w_i \in (0, \infty)$. Such a weight system is naturally associated with the weight-function system of power functions where $F_i(r) = r^{w_i}$. In the context of the study of the random walks driven by the measures $\mu_{S,a}$, these power weight systems and associated power function systems are relevant to the case when $a = (\alpha_i)_1^k \in (0, 2)^k$.

Example 2.1. In order to study the measures $\mu_{S,a}$ with tuples a with $\alpha_j = 2$ for some j, it is necessary to introduce weight functions of the type $r^2 \log r$. To allow for such functions, one can consider the two-dimensional weight systems built on

$$w_i = (u_i, v_i) \quad \text{with } u_i > 0 \text{ and } v_i \in \mathbb{R}, 1 \leq i \leq k.$$

In this case a natural compatible weight-function system would be

$$F_i(r) = r^{u_i}[\log(e+r)]^{v_i}, \quad 1 \le i \le k.$$

Example 2.2. When dealing with more general measures than $\mu_{S,a}$, it makes sense to consider multiparameter weight functions such that

$$f_{v_1,v_2,v_3}(r) = r^{v_1}[\log(e+r)]^{v_2}[\log(e+\log(e+r))]^{v_3}, \quad v_1 \in (0,\infty), v_2, v_3 \in \mathbb{R},$$

together with the natural associated lexicographical order on the parameter space (v_1, v_2, v_3).

Example 2.3. The functions considered in the previous example are special case of regularly varying functions. Recall that a slowly varying function (at infinity) is a positive measurable function $\ell\colon (0,\infty) \to (0,\infty)$ such that, for any $\lambda > 0$, $\ell(\lambda x)/\ell(x) \to 1$ as x tends to infinity. A regularly varying function of index $\alpha \in \mathbb{R}$ is a function of the type $t \mapsto t^\alpha \ell(t)$ where ℓ is slowly varying. We refer the reader to the classical treaty [6] for details on the basic definitions and the many classical properties and techniques related to the notion of regular variation. Functions of regular variation appear in our context in several different ways including via natural generalizations of the definition of $\mu_{S,a}$ at (1.1).

In what follows we will mostly use weight-function systems $\mathfrak{F}$ such that there exists $C \ge 1$ such that

$$2F_i(r) \le F_i(Cr), \ F_i(2r) \le CF_i(r) \quad \text{for all } i \in \{1,\dots,k\}, r \ge 1. \tag{2.1}$$

Further, we will often make the assumption that we are given a weight system $\mathfrak{w}$ and a weight-function system $\mathfrak{F}$ that are compatible in the sense that there exists $C \ge 1$ such that

$$w(c) \preceq w(c') \text{ for all } c, c\text{'} \iff F_c(r) \le CF_{c'}(r) \text{ for all } r. \tag{2.2}$$

Note that under these two hypotheses, $w(c) = w(c')$ is equivalent to $F_c \simeq F_{c'}$. In this case, except for notational convenience, it is obviously somewhat redundant to use both $\mathfrak{w}$ and $\mathfrak{F}$ since they contain more or less the same information.

Definition 2.7. Referring to the setting and notation introduced above, assume that the weight-function system $\mathfrak{F}$ and the weight system $\mathfrak{w}$ satisfy (2.1) and (2.2). For any $j = 1,\dots,j_*$, let $\mathbf{F}_j$ be a function such that for any commutator c with $w(c) = \bar{w}_j$, we have

$$\mathbf{F}_j \simeq F_c.$$

(The function $\mathbf{F}_j$ corresponding to commutators c with $w(c) = \bar{w}_j$ should not be confused $F_i = F_{s_i}$).

In the following definition, given a finite tuple Σ of elements of a nilpotent group G, we let $\Omega(\Sigma)$ be the set of all finite words with formal letters in $\Sigma \cup \Sigma^{-1}$. For $\omega \in \Omega(\Sigma)$, we write $\pi(\omega)$ to denote the corresponding element of G. For $\omega \in \Omega(\Sigma)$ and $\sigma \in \Sigma$, let $\deg_\sigma(\omega)$ is the number of occurrences of $\sigma^{\pm 1}$ in ω.

Definition 2.8. Let G be a nilpotent group generated by $S = (s_1, \dots, s_k)$. Let $\mathfrak{w}, \mathfrak{F}$ be a weight system and associated weight function system on a generating k-tuple S which satisfy (2.1) and (2.2). For any tuple Σ of elements in $\mathfrak{C}(S)$, set

$$F_\Sigma = F_c$$

where $w(c) = \min\{w(\sigma)\colon \sigma \in \Sigma\}$. For $g \neq e$, set

$$\begin{aligned}\|g\|_{\Sigma,\mathfrak{F}} = \min\{r \geq 1\colon & \text{there exists } \omega \in \Omega(\Sigma) \text{ such that} \\ & g = \pi(\omega) \text{ and} \\ & \deg_c(\omega) \leq F_c \circ F_\Sigma^{-1}(r) \text{ for all } c \in \Sigma\}.\end{aligned}$$

By convention, $\|e\|_{\Sigma,\mathfrak{F}} = 0$. Set also

$$Q(\Sigma, \mathfrak{F}, r) = \{g \in G\colon F_\Sigma^{-1}(\|g\|_{\Sigma,\mathfrak{F}}) \leq r\}.$$

Further, when S and $\mathfrak{w}, \mathfrak{F}$ are fixed, set

$$\|g\|_{\mathrm{com}} = \|g\|_{\mathfrak{F},\mathrm{com}} = \|g\|_{\mathfrak{C}(S),\mathfrak{F}}, \quad \|g\|_{\mathrm{gen}} = \|g\|_{\mathfrak{F},\mathrm{gen}} = \|g\|_{S,\mathfrak{F}}$$

and

$$Q_{\mathrm{com}}(r) = Q(\mathfrak{C}(S), \mathfrak{F}, r), \quad Q_{\mathrm{gen}}(r) = Q(S, \mathfrak{F}, r).$$

Note that $F_S = F_{\mathfrak{C}(S)}$.

According to this definition an element $g \in G$ has norm $\|g\|_{S,\mathfrak{F}}$ less or equal to r if and only if we can find a word ω in the generators and their inverses that represents $g \in G$ and satisfies the size condition

$$\max_{1 \leq i \leq k} \{F_S(F_i^{-1}(\deg_{s_i}(\omega)))\} \leq r.$$

Remark 2.9. If Σ generates G then $\|\cdot\|_{\Sigma,\mathfrak{F}}$ is a quasi-norm on G (see 5.1 below for a precise definition). It is a norm on G (i.e., satisfies the triangle inequality) if each of the functions $\{F_c \circ F_\Sigma^{-1}, c \in \Sigma\}$, defined on $[1, \infty)$ can be extended to a convex function on $[0, \infty)$ that vanishes at 0.

Example 2.4. The simplest example is when the weight system $\mathfrak{w}$ is one dimensional, generated by $w(s_i) = w_i \in [2, \infty)$, and the associated weight function system $\mathfrak{F}$ is generated by $F_i(r) = r^{w_i}$. In this case, it will sometimes be convenient to write $\|\cdot\|_{S,\mathfrak{w}}$ for $\|\cdot\|_{S,\mathfrak{F}}$ (resp. $\|\cdot\|_{\Sigma,\mathfrak{w}}$ for $\|\cdot\|_{\Sigma,\mathfrak{F}}$).

Example 2.5. For further illustration, consider the groups $\mathbb{Z}^3$ equipped with its natural generating 3-tuple $S = (s_i)_1^3$ and the discrete Heisenberg group (see Example 1.1) equipped with the generating 3-tuple $S = (s_1 = X, s_2 = Y, s_3 = Z)$ where X is the matrix with $x = 1, y = z = 0$ and Y, Z are defined similarly. Set $F_1(r) = r^{3/2}$, $F_2(r) = r^2 \log(e + r)$, $F_3(r) = r^\gamma$, $\gamma > 3/2$, and let $\mathfrak{F}$ be the associated weight-function system (we let the reader define the natural 2-dimensional weight system $\mathfrak{w}$ that is compatible with $\mathfrak{F}$).

On $\mathbb{Z}^3$, it is clear from the definition that

$$\|(x, y, z)\|_{\mathfrak{F},\mathrm{gen}} \simeq \max\left\{|x|, \frac{|y|^{3/4}}{\log(e + |y|)^{3/4}}, |z|^{3/(2\gamma)}\right\}.$$

On the Heisenberg group, it is not immediately obvious how to compute the $\|\cdot\|_{\mathfrak{F},\mathrm{gen}}$-norm of the element

$$g_{x,y,z} = \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix}.$$

Theorem 2.10 below (and the fact that the matrix representation of $g_{x,y,z}$ is unique) leads to the conclusion that

$$\|g_{x,y,z}\|_{\mathfrak{F},\mathrm{gen}} \simeq \max\left\{|x|, \frac{|y|^{3/4}}{\log(e + |y|)^{3/4}}, |z|^{3/(2\gamma)}\right\} \quad \text{if } \gamma > 7/2$$

and

$$\|g_{x,y,z}\|_{\mathfrak{F},\mathrm{gen}} \simeq \max\left\{|x|, \frac{|y|^{3/4}}{\log(e + |y|)^{3/4}}, \frac{|z|^{3/7}}{[\log(e + |z|)]^{3/7}}\right\} \quad \text{if } 3/2 \le \gamma \le 7/2.$$

One can check (without much trouble) that $\|\cdot\|_{\mathfrak{F},\mathrm{gen}}$ satisfies the triangle inequality in this case (on either $\mathbb{Z}$ or the Heisenberg group). We shall see that this choice of weight-function system is relevant to the study of the probability measure μ on G such that

$$\mu(s_i^n) \text{ is proportional to } \frac{1}{1 + |n| F_i^{-1}(|n|)}, \quad n \in \mathbb{Z}.$$

We will use this example to illustrate some of our main results in the rest of the paper.

The following theorem contains some of the key geometric results we will need to study the walk driven by measures of the type $\mu_{S,a}$. We say that a finite subset B of a finitely generated abelian group A is free if B is a basis of a free $\mathbb{Z}$-submodule of A.

Theorem 2.10 ($\mathfrak{w}$-F-adapted coordinates). *Let G be a nilpotent group equipped with a generating k-tuple $S = (s_1, \dots, s_k)$. Let $\mathfrak{w}$, $\mathfrak{F}$ be weight and weight-function systems on S satisfying (2.1) and (2.2). Let $R_j^{\mathfrak{w}}$ be as in Definition 1.6.*

Let $\Sigma = (c_1, \dots, c_t)$ be a tuple of formal commutators in $\mathfrak{C}(S)$ with non-decreasing weights $w(c_1) \preceq \cdots \preceq w(c_t)$. Let m_j, $j = 0, \dots, j_*$ be defined by

$$\{c_i : w(c_i) = \bar{w}_j\} = \{c_i : m_{j-1} < i \leq m_j\}.$$

Assume that (the image of) $\{c_i : w(c_i) = \bar{w}_j\}$ generates $G_j^{\mathfrak{w}}$ modulo $G_{j+1}^{\mathfrak{w}}$ and that $\{c_i : m_{j-1} < i \leq m_{j-1} + R_j^{\mathfrak{w}}\}$ is free in $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$. Then the following properties hold.

- *There exists a constant $C = C(G, S, \mathfrak{F})$ such that for any $r \geq 1$, if $g \in G$ can be expressed as a word ω over $\mathfrak{C}(S)$ with $\deg_c(\omega) \leq F_c(r)$ for all $c \in \mathfrak{C}(S)$ then g can be expressed in the form*

$$g = \prod_{i=1}^{t} c_i^{x_i} \text{ with } |x_i| \leq C \times \begin{cases} \mathbf{F}_j(r) & \text{if } m_{j-1} + 1 \leq i \leq R_j^{\mathfrak{w}}, \\ 1 & \text{if } R_j^{\mathfrak{w}} + 1 \leq i \leq m_j. \end{cases}$$

- *There exist an integer $p = p(G, S, \mathfrak{F})$, a constant $C = C(G, S, \mathfrak{F})$ and a sequence $(i_1, \dots, i_p) \in \{1, \dots, k\}^p$ such that if g can be expressed as a word ω over $\mathfrak{C}(S)$ with $\deg_c(\omega) \leq F_c(r)$ for some $r \geq 1$ and all $c \in \mathfrak{C}(S)$ then g can be expressed in the form*

$$g = \prod_{j=1}^{p} s_{i_j}^{x_j} \text{ with } |x_j| \leq C F_{i_j}(r).$$

This important theorem will be proved in the last section of this article. See also Theorem A.22 for an additional improvement of the the last statement of Theorem 2.10. Note that in the decomposition $g = \prod_{j=1}^{p} s_{i_j}^{x_j}$, the sequence $(i_j)_1^p$ is independent of the group element g.

The proof of the following simple corollary is omitted.

Corollary 2.11. *Referring to Definition 2.8, the quasi-norms* $\|\cdot\|_{\mathrm{com}}$ *and* $\|\cdot\|_{\mathrm{gen}}$ *defined on* G *satisfy*

$$\|\cdot\|_{\mathrm{gen}} \simeq \|\cdot\|_{\mathrm{com}} \quad \text{over } G.$$

Further, referring to the t*-tuple* $\Sigma = (c_1, \ldots, c_t)$ *of Theorem 2.10, we have*

$$F_\Sigma^{-1}(\|\cdot\|_{\Sigma,\mathfrak{F}}) \simeq F_S^{-1}(\|\cdot\|_{\mathrm{com}}) \quad \text{over } G.$$

Remark 2.12. In the case when the generators s_i are given equal weight-functions, i.e., $F_i = F_j$, $1 \le i \le j \le k$, the quasi-norms $\|\cdot\|_{S,\mathfrak{F}}$, $\|\cdot\|_{\Sigma,\mathfrak{F}}$ and $\|\cdot\|_{\mathfrak{C}(S),\mathfrak{F}}$ are all comparable to the usual word-norm $|\cdot|_S$.

2.2. Norm equivalences. In this section, we briefly discuss how changing weight functions affect the quasi-norms $\|\cdot\|_{\mathrm{com}}$ and $\|\cdot\|_{\mathrm{gen}}$ introduced in Definition 2.8.

Definition 2.13. Let G be a countable nilpotent group equipped with a generating k-tuple $S = (s_1, \ldots, s_k)$ and a (possibly multidimensional) weight system $\mathfrak{w}$ as above. For each $g \in G$, let

$$j_{\mathfrak{w}}(g) = \max\{j : \text{there exists } u \in \mathbb{N} \text{ such that } g^u \in G_j^{\mathfrak{w}}\}.$$

Let $\mathrm{core}(\mathfrak{w}, S)$ be the sub-sequence of S obtained by keeping only those s_i such that $w(s_i) = \bar{w}_{j_{\mathfrak{w}}(s)}$.

By construction, we always have $w(s) \le \bar{w}_{j_{\mathfrak{w}}(s)}$. Those generators $s \in S$ with $w(s) < \bar{w}_{j(s)}$ are, in some sense, inefficient. The following proposition makes this precise and motivates this definition.

Proposition 2.14. *Any formal commutator* $c \in \mathfrak{C}(S)$ *whose image in* G *is free in* $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$ *must only use letters in* $\mathrm{core}(\mathfrak{w}, S)$*. In particular, referring to the sequence of commutators* $c_1, \ldots, c_t$ *in Theorem 2.10, any formal commutator* c_i *with* $i \in m_{j-1}+1, \ldots, m_{j-1}+R_j^{\mathfrak{w}}$ *must only use letters in* $\mathrm{core}(\mathfrak{w}, S)$.

Proof. Assume that the image of c is in the torsion free part of $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$ and involves $s \notin \mathrm{core}(S)$, say $c = [c', [s, c'']]$. Then there exists $u \in \mathbb{N}$ such that $s^u \in G_{j(s)}^{\mathfrak{w}}$ with $\bar{w}_{j(s)} > w(s)$ (where we write $j(s) = j_{\mathfrak{w}}(s)$). From the linearity of brackets, we have

$$c^u \equiv [c', [s^u, c'']] \mod G_{j+1}^{\mathfrak{w}}$$

while $[c', [s^u, c'']] \in G_{j+1}^{\mathfrak{w}}$ since $s^u \in G_{j(s)}^{\mathfrak{w}}$ with $\bar{w}_{j(s)} > w(s)$. Therefore

$$c^u \equiv 0 \mod G_{j+1}^{\mathfrak{w}}.$$

This contradicts the assumption that c is free in $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$. The proposition follows. □

Definition 2.15. Let G be a countable nilpotent group equipped with a generating k-tuple $S = (s_1, \dots, s_k)$ and a (possibly multidimensional) weight system $\mathfrak{w}$ as above. Let $\Sigma = (c_1, \dots, c_t)$ be a sequence of formal commutators as in Theorem 2.10. Let $\operatorname{core}(\mathfrak{w}, S, \Sigma)$ be the sub-sequence of S of those letters s_δ that appear in the build-sequence of one or more of the formal commutators $c_i \in \Sigma$ with $i \in \cup_{j=1}^{q+1}\{m_{j-1}+1, \dots, m_{j-1}+R_j^{\mathfrak{w}}\}$.

Remark 2.16. Proposition 2.14 shows that, for any sequence Σ of formal commutators as in Theorem 2.10, we have

$$\operatorname{core}(\mathfrak{w}, S, \Sigma) \subset \operatorname{core}(\mathfrak{w}, S).$$

In what follows, given two tuples $S = \{s_1, \dots, s_k), \Theta = (\theta_1, \dots, \theta_\kappa)$ of elements of G (possibly of different length k, κ), we write $S \subset \Theta$ if there is a one to one map $J : \{1, \dots, k\} \to \{1, \dots, \kappa\}$ such that $s_{J(i)} = \theta_i$ in G. This applies, for instance, to the "inclusion" $\operatorname{core}(\mathfrak{w}, S, \Sigma) \subset \operatorname{core}(\mathfrak{w}, S)$ in the previous remark. Abusing notation, we will sometimes use the same letter s to denote an element of S and the associated element in Θ.

Proposition 2.17. *Referring to the setting and notation of Theorem 2.10, for each $g \in G$ either G is a torsion element and $\|g^n\|_{\mathrm{com}} \simeq 1$ for all n or*

$$\|g^n\|_{\mathrm{com}} \simeq F_S \circ \mathbf{F}_j^{-1}(n) \quad \textit{where } j = j_{\mathfrak{w}}(g). \tag{2.3}$$

Proof. The upper bound is very easy. Let κ be such that $g^\kappa \in G_j^{\mathfrak{w}}$, $j = j_{\mathfrak{w}}(g)$. Since g^κ is in $G_j^{\mathfrak{w}}$ it can be written as word ω using formal commutators of weight at least $\bar{w}_j$. Hence, $g^{\kappa n}$ can be written as a word ω_n, namely, ω repeated n times. Obviously, if $w(c) \geq \bar{w}_j$, $\deg_c(\omega_n) \leq \deg_c(\omega)n$. By definition, this implies $\|g^{\kappa n}\|_{\mathrm{com}} \leq CF_S \circ \mathbf{F}_j^{-1}(n)$. The estimate $\|g^n\|_{\mathrm{com}} \leq C' F_S \circ \mathbf{F}_j^{-1}(n)$ easily follows.

The lower bound is more involved. Using Theorem 2.10, it suffices to show that any writing of $g^{\kappa n}$ as a product

$$g^{\kappa n} = \prod_1^t c_i^{x_i} \quad \text{with } |x_i| \leq C \text{ for } i \in \bigcup_h \{m_{h-1} + R_h^{\mathfrak{w}} + 1, \dots, m_h\} \tag{2.4}$$

must have $\max_{i \in \{m_{j-1}+1, \dots, m_{j-1}+R_j^{\mathfrak{w}}\}}\{|x_i|\} \geq cn$. First, we claim that there exists a constant T (independent of g but depending on the structure of G, S, the weight system $\mathfrak{w}$ and the constant C appearing in the above displayed equation) such that for any n and any writing of $g^{\kappa n}$ as above we have

$$|x_i| \leq T \quad \text{for all } i \leq m_{h-1}, h \leq j. \tag{2.5}$$

The proof is by induction on $h \le j$. There is nothing to prove for $h = 1$. Assume that $h + 1 \le j$ and that we have proved that $|x_i| \le T$ for all $i \le m_{h-1}$. Since $g^\kappa, g^{\kappa n} \in G_h^{\mathfrak{w}}$, the product $\sigma = \prod_1^{m_{h-1}} c_i^{x_i}$ is in $G_h^{\mathfrak{w}}$. Since $|x_i| \le T$, $i \le m_{h-1}$, $\sigma = \prod_{i > m_{h-1}} c_i^{z_i}$ with $|z_i| \le T'$ where T' depends only on $G, S, \mathfrak{w}, T$ but not on g, n. Computing in $G_h^{\mathfrak{w}}$ modulo $G_{h+1}^{\mathfrak{w}}$, we have

$$g^{\kappa n} = \prod_{m_{h-1}+1}^{m_h} c_i^{x_i + z_i} = e \mod G_{h+1}^{\mathfrak{w}}.$$

The last equality holds because $g^{\kappa n} \in G_h^{\mathfrak{w}}$ and $h + 1 \le j$. Since

$$\{c_{m_{h-1}+1}, \dots, c_{m_{h-1}+R_h^{\mathfrak{w}}}\}$$

is free in $G_h^{\mathfrak{w}}/G_{h+1}^{\mathfrak{w}}$ and $\sup_i |z_i| \le T'$, $\sup\{|x_i| : m_{h-1} + R_h^{\mathfrak{w}} + 1 \le i \le m_h\} \le C$, there is a constant T'' depending only on $G, S, \mathfrak{w}, C$ and T' such that $|x_i| \le T''$ for $i \in \{m_{h-1} + 1, \dots, m_{h-1} + R_h^{\mathfrak{w}}\}$. This proves (2.5).

On the one hand, since j is the largest integer such that $g^u \in G_j^{\mathfrak{w}}$ for some u, it follows that for any n we can write

$$g^{\kappa n} = \prod_{i=m_{j-1}+1}^{m_j} c_i^{y_i} \mod G_{j+1}^{\mathfrak{w}} \quad \text{with} \sum_{i=m_{j-1}+1}^{m_{j-1}+R_j^{\mathfrak{w}}} |y_i| \ge cn$$

and

$$\max\{|y_i| : m_{j-1} + R_j^{\mathfrak{w}} + 1 \le i \le m_j\} \le C'.$$

On the other hand, since any writing of $g^{\kappa n}$ as in (2.4) satisfies (2.5), the same reasoning as in the induction step for (2.5) gives

$$g^{\kappa n} = \prod_{m_{j-1}+1}^{m_j} c_i^{y_i - x_i - z_i} = e \mod G_{j+1}^{\mathfrak{w}}$$

with $|z_i| \le T$. Since $\{c_i : m_{j-1} + 1 \le i \le m_{j-1} + R_j^{\mathfrak{w}}\}$ is free, the facts that

$$\sum_{i=m_{j-1}+1}^{m_{j-1}+R_j^{\mathfrak{w}}} |y_i| \ge cn, \quad \max\{|y_i| : m_{j-1} + R_j^{\mathfrak{w}} + 1 \le i \le m_j\} \le C'$$

and $|z_i| \le T$ together imply that

$$\sum_{i=m_{j-1}+1}^{m_{j-1}+R_j^{\mathfrak{w}}} |x_i| \ge c'n.$$

Hence, $\|g^{\kappa n}\|_{\mathrm{com}} \simeq F_S \circ \mathbf{F}_j^{-1}(n)$. □

Theorem 2.18. *Let G be a countable nilpotent group equipped with two generating tuples S, S' and associated multidimensional weight systems $\mathfrak{w}, \mathfrak{w}'$ as well as weight function systems $\mathfrak{F}, \mathfrak{F}'$ satisfying (2.1) and (2.2). By definition, F_S and $F'_{S'}$ are the weight functions associated with the smallest weights in $\mathfrak{w}$ and $\mathfrak{w}'$, respectively. Let $\Sigma = (c_1, \ldots, c_t)$ be a sequence of formal commutators as in Theorem 2.10 applied to $(S, \mathfrak{w}, \mathfrak{F})$.*

(1) *Assume that $S' \supset \mathrm{core}(\mathfrak{w}, S, \Sigma)$ and $F'_s \geq F_s$ for all $s \in \mathrm{core}(\mathfrak{w}, S, \Sigma)$. Then*

$$(F'_{S'})^{-1}(\|g\|_{S',\mathfrak{F}'}) \leq C F_S^{-1}(\|g\|_{S,\mathfrak{F}}), \quad \textit{for all } g \in G.$$

(2) *Assume that, for all $s \in S'$, $F'_s \leq \mathbf{F}_{j_{\mathfrak{w}}(s)}$. Then*

$$(F'_{S'})^{-1}(\|g\|_{S',\mathfrak{F}'}) \geq c F_S^{-1}(\|g\|_{S,\mathfrak{F}}), \quad \textit{for all } g \in G.$$

Proof. To prove the first statement, referring to the notation used in Theorem 2.10, Set

$$I_1 = \cup_j \{m_{j-1}+1, \ldots, m_{j-1}+R_j^{\mathfrak{w}}\}, \quad I_2 = \{1, \ldots, t\} \setminus I_1$$

and recall that any any $g \in G$ can be written as

$$g = \prod_1^t c_i^{x_i}, \quad |x_i| \leq C \begin{cases} F_{c_i}(F_S^{-1}(\|g\|_{\mathrm{com}})) & \text{if } i \in I_1, \\ 1 & \text{if } i \in I_2. \end{cases}$$

By hypothesis, $F'_{c_i} \geq F_{c_i}$ for $i \in I_1$. Further, each c_i, $i \in I_2$, is a product of elements in S'. Hence, we obtain an expression for g as a word ω on formal commutators on S' with

$$\deg_c(\omega) \leq C F'_c(F_S^{-1}(\|g\|_{\mathrm{com}})).$$

This proves that $(F'_{S'})^{-1}(\|g\|_{S',\mathfrak{F}'}) \leq C F_S^{-1}(\|g\|_{S,\mathfrak{F}})$ as desired.

To prove the second statement, apply Theorem 2.10 to $(S', \mathfrak{w}', \mathfrak{F}')$ to write any $g \in G$ as a product

$$g = \prod_1^p (s'_{i_j})^{x_j} \quad \text{with } |x_j| \leq F'_{s'_{i_j}} \circ (F'_{S'})^{-1}(\|g\|_{S',\mathfrak{F}'})$$

where $s'_{i,j} \in S'$ (note that the sequence (i_j) and the integer p are fixed and independent of g). By Proposition 2.17 and the hypothesis $\mathbf{F}_{j_{\mathfrak{w}}(s)} \geq F'_s$ for all $s \in S'$, we obtain that $F_S^{-1}(\|g\|_{S,\mathfrak{F}}) \leq C(F'_{S'})^{-1}(\|g\|_{S',\mathfrak{F}'})$ as desired. □

Corollary 2.19. *Let G be a countable nilpotent group equipped with two generating tuple S, S' and associated multidimensional weight systems $\mathfrak{w}, \mathfrak{w}'$ with function systems $\mathfrak{F}, \mathfrak{F}'$ satisfying* (2.1) *and* (2.2)*. Let $\Sigma = (c_1, \dots, c_t)$ be a sequence of formal commutators as in Theorem* 2.10 *applied to $(S, \mathfrak{w}, \mathfrak{F})$. Assume that there exists $C \in (0, \infty)$ such that the following two conditions are satisfied*:

(i) $\operatorname{core}(\mathfrak{w}, S, \Sigma) \subset S'$ *and* $CF'_s \geq F_s$ *for all* $s \in \operatorname{core}(\mathfrak{w}, S, \Sigma)$;

(ii) $F'_s \leq C\mathbf{F}_{j_{\mathfrak{w}}(s)}$ *for all* $s \in S'$.

Then

$$(F'_{S'})^{-1}(\|g\|_{S',\mathfrak{F}'}) \simeq F_S^{-1}(\|g\|_{S,\mathfrak{F}}) \quad \text{for all } g \in G.$$

In particular,

$$\#Q(S', \mathfrak{F}', r) \simeq \#Q(S, \mathfrak{F}, r) \quad \text{for all } r > 0.$$

Example 2.6 (continuation of Example 2.5). Consider the discrete Heisenberg group as in Example 2.5 equipped with the generating 3-tuple

$$S = (s_1 = X, s_2 = Y, s_3 = Z) \quad \text{and} \quad S' = (s'_i = X, s'_2 = Y).$$

Set

$$F_1(r) = F'_1(r) = r^{3/2},$$
$$F_2(r) = F'_2(r) = r^2 \log(e + r),$$
$$F_3(r) = r^{\gamma}, \quad \gamma > 3/2,$$

and let $\mathfrak{F}, \mathfrak{F}'$ be the associated weight-function systems. The natural 2 dimensional weight systems $\mathfrak{w}, \mathfrak{w}'$ are generated by

$$w_1 = w'_1 = (3/2, 0),$$
$$w_2 = w'_2 = (2, 1),$$
$$w_3 = (\gamma, 0).$$

The first observation is that $\operatorname{core}(\mathfrak{w}, S) = (s_1, s_2, s_3)$ if $\gamma > 7/2$ and $\operatorname{core}(\mathfrak{w}, S) = (s_1, s_2)$ if $3/2 < \gamma \leq 7/2$. It follows that,

$$\|g\|_{S',\mathfrak{F}'} \simeq \|g\|_{S,\mathfrak{F}} \quad \text{for all } g \in G,$$

if $\gamma \in (3/2, 7/2]$ whereas these norms are not equivalent if $\gamma > 7/2$.

3. Volume estimates

This section gathers some of the main results we will need regrading volume estimates for the balls $Q(S, \mathfrak{F}, r)$ introduced in Definition 2.8. It also addresses the question of how changes in the weight-function system affect these volume estimates.

We start with a general and very flexible result which admits a rather simple proof. In this theorem, the weight-function system $\mathfrak{F}$ is not necessarily tightly related to the weight system $\mathfrak{w}$. The proof of this theorem will be given in the last section of this paper.

Theorem 3.1. *Let $\mathfrak{w}$ be a multidimensional weight system as in Section 2.1. Assume that we are given weight functions F_i, $1 \le i \le k$ satisfying (2.1). Let $\Sigma = (c_1, \dots, c_s)$ be a s-tuple of formal commutators on $\{s_i^{\pm 1} : 1 \le i \le k\}$. Assume that, for any h, the family $\{c_i : w(c_i) = \bar{w}_h\}$ projects to a free family in the abelian group $G_h^{\mathfrak{w}}/G_{h+1}^{\mathfrak{w}}$. Then there exist an integer $M = M_\Sigma$ and a sequence $(i_1, \dots, i_M) \in \{1, \dots, k\}^M$, depending on Σ such that for any $r > 0$ there exists a subset $K_\Sigma(r) \subset G$ satisfying the following two properties*:

(1) $\#K_\Sigma(r) \ge \prod_{i=1}^{s}(2F_{c_i}(r) + 1)$

(2) $g \in K_\Sigma(r) \implies g = \prod_{j=1}^{M} s_{i_j}^{x_j}, \; |x_j| \le F_{i_j}(r)$.

Further, every s_{i_j}, $1 \le j \le M$, belongs to the build-sequence of at least one $c_h \in \Sigma$.

Theorem 3.1 is very useful for comparing the volume growth associated with different "weight-function systems". See the proof of Theorem 3.4 below.

Next we state and prove sharp volume estimates related to Theorem 2.10.

Theorem 3.2. *Referring the setting and notation of Theorem 2.10, we have*

$$\#Q(\mathfrak{C}(S), \mathfrak{F}, r) \simeq \#Q(\Sigma, \mathfrak{F}, r) \simeq \#Q(S, \mathfrak{F}, r) \simeq \prod_{j=1}^{j_*} \mathbf{F}_j(r)^{R_j^{\mathfrak{w}}}.$$

Remark 3.3. Assume that the weight system $\mathfrak{w}$ is unidimensional, generated by $(w_i)_1^k \in (0, \infty)^k$, and the weight-functions F_i are power functions $F_i(r) = r^{\mathfrak{w}_i}$, $i = 1, \dots, k$. Then

$$Q(S, \mathfrak{F}, r) \simeq r^{D(S, \mathfrak{w})}$$

with $D(S, \mathfrak{w})$ as in Definition 1.7.

Proof. The equivalences

$$\#Q(\mathfrak{C}(S), \mathfrak{F}, r) \simeq \#Q(\Sigma, \mathfrak{F}, r) \simeq \#Q(S, \mathfrak{F}, r)$$

and the upper bound

$$\#Q(\Sigma, \mathfrak{F}, r) \le C \prod_{j=1}^{j_*} \mathbf{F}_j(r)^{R_j^{\mathfrak{w}}}$$

follows immediately from Theorem 2.10 and inspection.

The lower bound

$$\#Q(\Sigma, \mathfrak{F}, r) \ge c \prod_{j=1}^{j_*} \mathbf{F}_j(r)^{R_j^{\mathfrak{w}}}$$

requires an additional argument. Note that $Q(\Sigma, \mathfrak{F}, r)$ contains the image in G of

$$\prod_{j=1}^{j_*} \prod_{i=m_{j-1}+1}^{m_{j-1}+R_j} c_i^{x_i}, \quad |x_i| \le F_{c_i}(r).$$

Further, it is not hard to check that

$$\prod_j \prod_{i=m_{j-1}+1}^{m_{j-1}+R_j} c_i^{x_i} = \prod_j \prod_{i=m_{j-1}+1}^{m_{j-1}+R_j} c_i^{y_i}$$

implies

$$x_i = y_i, \quad i \in \bigcup_{j=1}^{j_*} \{m_{j-1}+1, \dots, m_{j-1}+R_j\}.$$

The desired lower bound follows. □

Theorem 3.4. *Let G be a countable nilpotent group equipped with two generating tuples S, S' and associated multidimensional weight systems $\mathfrak{w}, \mathfrak{w}'$ as well as weight function systems $\mathfrak{F}, \mathfrak{F}'$ satisfying* (2.1) *and* 2.2. *Let $\Sigma = (c_1, \dots, c_t)$ be a sequence of formal commutators as in Theorem* 2.10 *applied to $(S, \mathfrak{w}, \mathfrak{F})$. Assume that $S' \supset \mathrm{core}(\mathfrak{w}, S, \Sigma)$ and that*

$$F'_s \ge F_s \quad \textit{for all } s \in \mathrm{core}(\mathfrak{w}, S, \Sigma).$$

Then

$$\#Q(S', \mathfrak{F}', r) \simeq \prod_{j=1}^{j_*(\mathfrak{w}')} \mathbf{F}'_j(r)^{R_j^{\mathfrak{w}'}} \ge \#Q(S, \mathfrak{F}, r) \simeq \prod_{j=1}^{j_*(\mathfrak{w})} \mathbf{F}_j(r)^{R_j^{\mathfrak{w}}}.$$

Assume further that there exists $\sigma \in S'$ such that $F'_\sigma \geq \mathbf{F}_{j_{\mathfrak{w}}(\sigma)}$. Then

$$\#Q(S', \mathfrak{F}', r) \geq c\Big(\frac{F'_\sigma(r)}{\mathbf{F}_{j_{\mathfrak{w}}(\sigma)}(r)}\Big)\#Q(S, \mathfrak{F}, r).$$

Proof. Since $\mathrm{core}(\mathfrak{w}, S, \Sigma) \subset S'$ it follows that, for any $c_i \in \Sigma$, F'_{c_i} is well defined as the product of F'_s with $s \in \mathrm{core}(\mathfrak{w}, S, \Sigma) \subset S'$. Use the collection of commutators $c_i, i \in \{m_{j-1}+1, \dots, m_{j-1}+R_j^{\mathfrak{w}}\}$, $j = 1, \dots, j_*$ in Theorem 2.10 with the weight system $\mathfrak{w}$ and weight-function system $\mathfrak{F}'$. For each r, Theorem 3.1 provides a set $K(r) \in G$ such that

$$\#K(r) \geq \prod_{j=1}^{j_*(\mathfrak{w})} \prod_{i=m_{j-1}+1}^{m_{j-1}+R_i^{\mathfrak{w}}} F'_{c_i}(r) \tag{3.1}$$

and, by Theorem 2.10 , Theorem 3.1 and the definition of $\mathrm{core}(\mathfrak{w}, S, \Sigma)$,

$$K(r) \subset \{g \in G : \|g\|_{S', \mathfrak{F}'} \leq F'_{S'}(r)\}.$$

By Theorem 3.2, it follows that, for all r,

$$\#K(r) \leq \#Q(S', \mathfrak{F}', r).$$

By hypothesis, $F'_s \geq F_s$ if $s \in \mathrm{core}(\mathfrak{w}, S, \Sigma)$. Hence we have $F'_{c_i} \geq F_{c_i}$ (i.e., $w'(c_i) \geq w(c_i)$). By (3.1) and Theorem 3.2, this implies

$$\#K(r) \geq c \prod_{j=1}^{j_*(\mathfrak{w})} \mathbf{F}_j^{R_j^{\mathfrak{w}}}.$$

This proves the first statement.

Suppose that there exists $\sigma \in S'$ such that $w'(s) > \bar{w}_{j_{\mathfrak{w}}(\sigma)}$. Set $j_0 = j_{\mathfrak{w}}(\sigma)$. In the sequence of commutators $c_1, \dots, c_t$ used above, consider the the free family

$$\{c_i : i \in \{m_{j_0-1}+1, \dots, m_{j_0-1}+R_{j_0}^{\mathfrak{w}}\}\} \quad \text{in } G_{j_0}^{\mathfrak{w}}/G_{j_0+1}^{\mathfrak{w}}.$$

By hypothesis, there exists an integer u such that $\sigma^u \in G_{j_0}^{\mathfrak{w}}$ is free in $G_{j_0}^{\mathfrak{w}}/G_{j_0+1}^{\mathfrak{w}}$. Since a maximal free subset of $\{\sigma^u\} \cup \{c_i : i \in \{m_{j_0-1}+1, \dots, m_{j_0-1}+R_{j_0}^{\mathfrak{w}}\}\}$ in $G_{j_0}^{\mathfrak{w}}/G_{j_0+1}^{\mathfrak{w}}$ containing σ^u must contain $R_{j_0}^{\mathfrak{w}}$ elements, we can replace one of the c_i, say c_{i_*} by σ^u so that the $R_{j_0}^{\mathfrak{w}}$-tuple so obtained is free in $G_{j_0}^{\mathfrak{w}}/G_{j_0+1}^{\mathfrak{w}}$. Let $b_i = c_i$ if $i \neq i_*$, $b_{i_*} = \sigma^u$, $\widetilde{F}^i = F'_{c_i}$ if $i \neq i_*$, $\widetilde{F}^{i_*}(r) = F'_\sigma(r/|u|)$, and apply Theorem A.4. The desired result follows. □

4. Random walk upper bounds

This section is devoted to obtaining upper bounds on the return probability of a large collection of random walks including those driven by the measures $\mu_{S,a}$. Generalizing one of the approaches developed in [28] for simple random walks, we will make use of appropriate volume growth estimates and of the notion of pseudo-Poincaré inequality.

4.1. Pseudo-Poincaré inequality. Let G be a group generated by a finite symmetric set A. Then it holds that for any finitely supported function f on G,

$$\|f_g - f\|_2^2 \le C_A |g|_A^2 \mathcal{E}_A(f,f) \tag{4.1}$$

where

$$\mathcal{E}_A(f,f) = \frac{1}{2|A|} \sum_{x\in G, y\in A} |f(xy) - f(x)|^2.$$

This expression is the Dirichlet form associated with the simple random walk based on A. Inequality (4.1) captures a fundamental universal property of Cayley graphs. In [28], it is proved that this simple property implies interesting upper-bounds on $u_A^{(2n)}(e)$ in terms of the volume growth function V_A.

The main result of this section is a pseudo-Poincaré inequality adapted to probability measure of the form

$$\mu(g) = k^{-1} \sum_{j=1}^{k} \sum_{n\in\mathbb{Z}} \mu_i(n) \mathbf{1}_{s_i^n}(g). \tag{4.2}$$

where $(s_1,\dots,s_k)$ is a generating k-tuple in G and the μ_i's are probability measures on $\mathbb{Z}$ with truncated second moment

$$\mathcal{G}_i(n) := \sum_{|m|\le n} m^2 \mu_i(n) \tag{4.3}$$

satisfying

$$\mathcal{G}_i(n) \ge c n^{2-\tilde{\alpha}_i} L_i(n), \quad \tilde{\alpha}_i \in (0,2], \tag{4.4}$$

for some slowly positive varying functions L_i, $1 \le i \le k$. Under these circumstances, we let F_i denote the inverse function of $n \mapsto n^{\tilde{\alpha}_i}/L_i(n)$. The function F_i is a regularly varying function of positive index $1/\tilde{\alpha}_i \in [2,\infty)$. In addition, we assume that the μ_i's are essentially decreasing in the sense that there is a constant C_1 such that

$$\mu_i(n) \le C_1 \mu_i(m), \quad \text{for all } i = 1,\dots,k, 0 \le m \le n, \quad . \tag{4.5}$$

Example 4.1. The measure $\mu_{S,a}$ with $a = (\alpha_i)_1^k \in (0,\infty)^k$ satisfies

$$\mathcal{G}_i(n) \simeq \begin{cases} n^{2-\alpha_i} & \text{if } \alpha_i \in (0,2), \\ \log n & \text{if } \alpha_i = 2, \\ 1 & \text{if } \alpha_i > 2. \end{cases}$$

Hence, in this case, we have $\tilde{\alpha}_i = \min\{\alpha_i, 2\}$ and $L_i = 1$ except if $\alpha_i = 2$ in which case $L_i(n) = \log n$.

We will make use of the following general result (which is essentially well-known). We let $\mathcal{C}_c(G)$ be the set of all finitely supported function on G and set $f_g(x) = f(xg)$.

Theorem 4.1. *Let G be a finitely generated group. Let μ be a symmetric probability measure on G. Assume that for each $r \geq 1$ there is a subset $K(r)$ of G such that*

$$\|f_g - f\|_2^2 \leq C_0\, r\mathcal{E}_\mu(f,f), \quad \textit{for all } g \in K(r). \tag{4.6}$$

and

$$\#K(r) \geq v(r), \quad \textit{for all } r \geq 1. \tag{4.7}$$

where v is increasing and regularly varying of positive index (see Example 2.3). Let ψ be the right-continuous inverse of v. Then there is a function $\Psi \simeq \psi$ such that the Nash inequality

$$\|f\|_2^2 \leq \Psi(\|f\|_1^2/\|f\|_2^2)\mathcal{E}_\mu(f,f), \quad \textit{for all } f \in \ell^1(G). \tag{4.8}$$

is satisfied. Moreover

$$\mu^{(2n)}(e) \leq C_1\eta(n) \tag{4.9}$$

where η is defined implicitly by

$$\tau = \int_1^{1/\eta(\tau)} \Psi(s)\frac{ds}{s}, \quad \tau > 0.$$

Proof. Assuming (4.6) and $\#K(r) \geq v(r)$, the Nash inequality (4.8) easily follows from writing

$$\begin{aligned} \|f\|_2 &\leq \|f - f_{K(r)}\|_2 + \|f_{K(r)}\|_2 \\ &\leq (C_0 r\mathcal{E}_\mu(f,f))^{1/2} + v(r)^{-1/2}\|f\|_1 \end{aligned}$$

and optimizing in r. Here $f_{K(r)}(x)$ is the average of f over $K(r)$ so that, obviously,

$$\|f_{K(r)}\|_2 \leq (\#K(r))^{-1/2}\|f\|_1$$

and (4.11) gives

$$\|f - f_{K(r)}\|_2 \leq (C_0 r \mathcal{E}_\mu(f, f))^{1/2} \quad \text{with } C_0 = CMk.$$

The return probability estimate (4.9) is a well-known consequence of (4.8). See [7, 18]. □

Remark 4.2. In this theorem, the parametrization of the set $K(r)$ is chosen so that r appears on the right-hand side of (4.6) instead of r^2.

Theorem 4.3. *Let G be a finitely generated nilpotent group equipped with a generating k-tuple $(s_1, \ldots, s_k)$. Let μ be as in (4.2) with $(\tilde{\alpha}_i)_1^k$, L_i and F_i be as in (4.4). Assume that (4.5) holds. Assume that there exists an integer M and a sequence $(i_j)_1^M \in \{1, \ldots, k\}^M$ such that for each $r \geq 1$ there is a subset $K(r)$ of G with the property that*

$$g \in K(r) \implies g = \prod_1^M s_{i_j}^{x_j}, \quad \textit{with } |x_j| \leq F_{i_j}(r). \tag{4.10}$$

Then there exists a constant $C = C(\mu)$ such that

$$\|f_g - f\|_2^2 \leq CM^2\, r\mathcal{E}_\mu(f, f), \quad \textit{for all } g \in K(r). \tag{4.11}$$

Proof. Because we assume (4.10), the proof boils down to a collection of one dimensional inequalities, one for each of the measures μ_i on $\mathbb{Z}$ that appear in the definition (4.2) of μ. Indeed, Lemma 4.4 stated below shows that there exists a constant C such that, for each $i \in \{1, \ldots, k\}$ and $y \in \mathbb{Z}$ with $|y| \leq F_i(r)$ we have

$$\|f_{s_i^y} - f\|_2^2 \leq C\, r\, \mathcal{E}_\mu(f, f) \tag{4.12}$$

for any finitely supported function f on G. Together, (4.10) and (4.12) imply (4.11). Since there exists a constant C such that, for all $i \in \{1, \ldots, k\}$,

$$|y| \leq F_i(r) \text{ implies } \mathcal{G}_i(|y|)^{-1}|y|^2 \leq Cr,$$

the claim (4.12) follows from Lemma 4.4. □

Lemma 4.4. *Let ν be a symmetric probability measure on $\mathbb{Z}$ satisfying the following condition:*

$$\text{there exists } C_1 \text{such that } \nu(n) \le C_1 \nu(m) \text{ for all } 0 \le m \le n.$$

Let G be a finitely generated group equipped with a distinguished element s. Set

$$\mathcal{E}_{s,\nu}(f, f) = \frac{1}{2} \sum_{x \in G, z \in \mathbb{Z}} |f(xs^z) - f(x)|^2 \nu(z)$$

and

$$\mathcal{G}_\nu(m) = \sum_{|n| \le m} |n|^2 \nu(n).$$

(i) *For any finitely supported function f on G we have*

$$\|f_{s^y} - f\|_2^2 \le C_\nu (\mathcal{G}_\nu(|y|))^{-1} |y|^2 \mathcal{E}_{s,\nu}(f, f) \quad \text{for all } y \in \mathbb{Z}.$$

(ii) *Further, for any two finitely supported functions f, g and all $x \in G$, $y \in \mathbb{Z}$,*

$$|f * g(xs^y) - f * g(x)|^2 \le C_\nu (\mathcal{G}_\nu(|y|))^{-1} \, |y|^2 \, \mathcal{E}_{s,\nu}(f, f) \, \|g\|_2^2 .$$

Proof of (i). For any pair of integers $0 < m \le n$, write

$$n = a_m m + b_m$$

with $0 \le b_m < m$ and

$$\begin{aligned}
\|f - f_{s^n}\|_2^2 &= \sum_{x \in G} (f(xs^n) - f(x))^2 \\
&\le 2 \sum_{x \in G} (f(xs^{a_m m}) - f(x))^2 + 2 \sum_{x \in G} (f(xs^{b_m}) - f(x))^2 \\
&\le 2a_m^2 \sum_{x \in G} (f(xs^m) - f(x))^2 + 2 \sum_{x \in G} (f(xs^{b_m}) - f(x))^2.
\end{aligned}$$

This yields

$$\begin{aligned}
\|f - f_{s^n}\|_2^2 \Big(\sum_{m=1}^{n} m^2 \nu(m) \Big) &\le 2 \sum_{x \in G} \sum_{m=1}^{n} (f(xs^m) - f(x))^2 a_m^2 m^2 \nu(m) \\
&\quad + 2 \sum_{x \in G} \sum_{m=1}^{n} (f(xs^{b_m}) - f(x))^2 m^2 \nu(m).
\end{aligned}$$

Next, observe that

$$\begin{aligned}\sum_{x\in G}\sum_{m=1}^{n}(f(xs^m)-f(x))^2(a_m m)^2\nu(m)\\ \le n^2\sum_{x\in G}\sum_{m=1}^{n}(f(xs^m)-f(x))^2\nu(m)\\ \le n^2\mathcal{E}_{s,\nu}(f,f).\end{aligned}$$

Further, using the hypothesis that ν is essentially decreasing, i.e., $\nu(m)\le C_1\nu(b)$ is $0\le b\le m$, write

$$\begin{aligned}\sum_{x\in G}\sum_{m=1}^{n}(f(xs^{b_m})-f(x))^2m^2\nu(m)\\ =\sum_{x\in G}\sum_{b=1}^{n/2}\sum_{\substack{m|n-b\\ b<m\le n}}(f(xs^b)-f(x))^2m^2\nu(m)\\ \le C_1\sum_{x\in G}\sum_{b=1}^{n/2}\Big(\sum_{\substack{m|n-b\\ b<m\le n}}m^2\Big)(f(xs^b)-f(x))^2\nu(b).\end{aligned}$$

As

$$\sum_{\substack{m|n-b\\ b<m\le n}}m^2\le\Big(\sum_{1}^{\infty}i^{-2}\Big)n^2,$$

we obtain

$$\sum_{x\in G}\sum_{m=1}^{n}(f(xs^{b_m})-f(x))^2m^2\nu(m)\le C_2n^2\mathcal{E}_{s,\nu}(f,f).$$

It follows that, for both $n>0$ and $n<0$,

$$\|f-f_{s^n}\|_2^2\Big(\sum_{0<m\le|n|}m^2\nu(m)\Big)\le 2(1+C_2)n^2\mathcal{E}_{s,\nu}(f,f). \qquad \square$$

Proof of (ii). By Cauchy–Schwarz

$$\begin{aligned}|f*g(xs^y)-f*g(x)| &= \left|\sum_{z\in G}(f(z^{-1}xs^y)-f(z^{-1}x))g(z)\right| \\ &\le \Big(\sum_{z\in G}(f(z^{-1}xs^y)-f(z^{-1}x))^2\Big)^{\frac12}\Big(\sum_{z\in G}|g(z)|^2\Big)^{\frac12} \\ &= \|f-f_{s^y}\|_2\,\|g\|_2.\end{aligned}$$

Applying part (i) to $\|f-f_{s^y}\|_2$ yields the desired inequality. □

Remark 4.5. When $G=\mathbb{Z}$, Lemma 4.4 provides an interesting and new pseudo-Poincaré inequality for probability measure ν satisfying (4.5) (i.e., which are essentially decreasing) in terms of the truncated second moment $\mathcal{G}_\nu$. Namely, assuming (4.5), we have

$$\sum_{x\in\mathbb{Z}}|f(x+y)-f(x)|^2\le C_\nu\frac{|y|^2}{\mathcal{G}_\nu(|y|)}\mathcal{E}_\nu(f,f)$$

where

$$\mathcal{E}_\nu(f,f)=\frac12\sum_{x,z\in\mathbb{Z}}|f(x+z)-f(x)|^2\nu(z).$$

Together with the trivial fact that $\#\{y\colon |y|\le r\}=2r+1$, this pseudo-Poincaré inequality and Theorem 4.1 provide a sharp Nash inequality satisfied by $\mathcal{E}_\nu$.

4.2. Assorted return probability upper bounds. This section describes direct applications of Theorem 3.1 together with Theorems 4.1–4.3. We use the notation introduced in Sections 1.5 and 2.1.

Theorem 4.6. *Let G be a finitely generated nilpotent group equipped with a generating k-tuple $(s_1,\dots,s_k)$ and a k-tuple of positive reals $a=(\alpha_1,\dots,\alpha_k)$. Let $\mathfrak{w}$ be the weight system which assigns weight $w_i=1/\tilde{\alpha}_i$ to s_i where $\tilde{\alpha}_i=\min\{2,\alpha_i\}$. Then*

$$\mu^{(n)}_{S,a}(e)\le C_{S,a}n^{-D(S,\mathfrak{w})}$$

where $D(S,\mathfrak{w})=\sum_h \bar{w}_h\,\mathrm{rank}(G_h^{\mathfrak{w}}/G_{h+1}^{\mathfrak{w}})$.

Proof. By Theorem 3.1, for each $r\ge 1$ we can find a subset $K(r)$ of G such that $\#K(r)\ge r^{D(S,\mathfrak{w})}$ and $g\in K(r)$ implies $g=\prod_1^M s_{i_j}^{x_j}$ with $|x_i|\le r^{w(s_{i_j})}$. The result then follows from Theorems 4.1-4.3 □

Remark 4.7. If all the α_i's are in $(0,2)$ or, more generally, if $R_h^{\mathfrak{w}} > 0$ implies $\bar{w}_h > 1/2$, the upper bound given in Theorem 4.6 is sharp. Indeed, we will prove a matching lower bound in the next section.

If all the α_i's are greater than 2 the measure $\mu_{S,a}$ has finite second moment and $D(S,\mathfrak{w}) = \frac{1}{2}\sum h \operatorname{rank}(G_h/G_{h+1})$. In this case the upper bound of Theorem 4.6 is also sharp. It coincides with the bound provided by Corollary 1.12.

We conjecture that this upper bound is sharp when $\alpha_i \neq 2$ for all $i \in \{1,\dots,k\}$ but we have not been able to prove this conjecture when there exists i, j such that $\alpha_i < 2$ and $\alpha_j > 2$.

The next result shows that Theorem 4.6 is not always sharp when some of the α_i's are equal to 2.

Theorem 4.8. *Let G be a finitely generated nilpotent group equipped with a generating k-tuple $(s_1,\dots,s_k)$ and a k-tuple of positive reals $a = (\alpha_1,\dots,\alpha_k) \in (0,\infty]^k$. Let $\mathfrak{w} = \mathfrak{w}(a)$ be the two-dimensional weight system which assigns weight $w_i = (v_{i,1}, v_{i,2})$ to s_i where*

$$v_{i,1} = \frac{1}{\tilde{\alpha}_i}, \quad \tilde{\alpha}_i = \min\{2,\alpha_i\}$$

and

$$v_{i,2} = 0 \textit{ unless } \alpha_i = 2 \textit{ in which case } v_{i,2} = 1/2.$$

Then

$$\mu_{S,a}^{(n)}(e) \le C_{S,a} n^{-D_1(S,\mathfrak{w})}[\log(e+n)]^{-D_2(S,\mathfrak{w})}$$

where

$$D_i(S,\mathfrak{w}) = \sum_h \bar{v}_{h,i} \operatorname{rank}(G_h^{\mathfrak{w}}/G_{h+1}^{\mathfrak{w}}), \quad \bar{w}_h = (\bar{v}_{h,1}, \bar{v}_{h,2}).$$

Proof. The proof is the same as for Theorem 4.6 but uses a refined weight system and the associated weight function system $\mathfrak{F}(a)$ where the function F_c associated to a commutator of weight $v(c) = (v_1,v_2)$ is $F_c(r) = r^{v_1}[\log(e+r)]^{v_2}$. □

Remark 4.9. Referring to Theorem 4.8, let Σ be a sequence of formal commutators as in Theorem 2.10 applied to $S, \mathfrak{w}, \mathfrak{F}(a)$. Assume that for any i such that $s_i \in \operatorname{core}(\mathfrak{w}, S, \Sigma)$, we have $\alpha_i = 2$. Then $D_1(S,\mathfrak{w}) = D_2(S,\mathfrak{w}) = D(G)/2$ and

$$\mu_{S,a}^{(n)}(e) \le C_{S,a}[n\log n]^{-D(G)/2}.$$

Example 4.2. Let G be the group of 4 by 4 unipotent upper-triangular matrices

$$G = \left\{ \begin{pmatrix} 1 & x_{1,2} & x_{1,3} & x_{1,4} \\ 0 & 1 & x_{2,3} & x_{2,4} \\ 0 & 0 & 1 & x_{3,4} \\ 0 & 0 & 0 & 1 \end{pmatrix} : x_{i,j} \in \mathbb{Z} \right\}.$$

With obvious notation, let $X_{i,j}$ be the matrix in G with a 1 in position i, j and all other non-diagonal entries equal to 0. Consider the generating 4-tuple

$$S = (s_1 = X_{1,2}, s_2 = X_{2,3}, s_3 = X_{3,4}, s_4 = X_{1,4}).$$

The non-trivial brackets are

$$[X_{1,2}, X_{2,3}] = X_{1,3}, [X_{2,3}, X_{3,4}] = X_{2,4}, [X_{1,2}, X_{2,4}] = [X_{1,3}, X_{3,4}] = X_{1,4}.$$

Let $a = (1, 2, 5, 1/3)$. The 2-dimensional weight system $\mathfrak{w}$ is generated by

$$w(s_1) = (1, 0), \quad w(s_2) = \left(\frac{1}{2}, \frac{1}{2}\right),$$
$$w(s_3) = \left(\frac{1}{2}, 0\right), \quad w(s_4) = (3, 0).$$

This implies

$$w([X_{1,2}, X_{2,3}]) = \left(\frac{3}{2}, \frac{1}{2}\right), \qquad w([X_{2,3}, X_{3,4}]) = \left(1, \frac{1}{2}\right),$$
$$w([X_{1,2}, [X_{2,3}, X_{3,4}]]) = \left(2, \frac{1}{2}\right), \quad w([[X_{1,2}, X_{2,3}], X_{3,4}]) = \left(2, \frac{1}{2}\right).$$

Ignoring (as we may) the weight values that would obviously lead to trivial quotients $G_h^{\mathfrak{w}}/G_{h+1}^{\mathfrak{w}}$, we have

$$\bar{w}_1 = \left(\frac{1}{2}, 0\right), \quad \bar{w}_2 = \left(\frac{1}{2}, \frac{1}{2}\right), \quad \bar{w}_3 = (1, 0), \quad \bar{w}_4 = \left(1, \frac{1}{2}\right),$$
$$\bar{w}_5 = \left(\frac{3}{2}, \frac{1}{2}\right), \quad \bar{w}_6 = \left(2, \frac{1}{2}\right), \quad \bar{w}_7 = (3, 0).$$

Next we compute the groups $G_i^{\mathfrak{w}}$. We have

$$\begin{aligned} G_7^{\mathfrak{w}} = G_6^{\mathfrak{w}} = \langle X_{1,4} \rangle \subset G_5^{\mathfrak{w}} &= \langle X_{1,4}, X_{1,3} \rangle \\ \subset G_4^{\mathfrak{w}} &= \langle X_{1,4}, X_{1,3}, X_{2,4} \rangle \\ \subset G_3^{\mathfrak{w}} &= \langle X_{1,4}, X_{1,3}, X_{2,4}, X_{1,2} \rangle \\ \subset G_2^{\mathfrak{w}} &= \langle X_{1,4}, X_{1,3}, X_{2,4}, X_{1,2}, X_{2,3} \rangle \\ \subset G_1^{\mathfrak{w}} &= \langle X_{1,4}, X_{1,3}, X_{2,4}, X_{1,2}, X_{2,3}, X_{3,4} \rangle = G. \end{aligned}$$

This gives

$$D_1(S,\mathfrak{w}) = \frac{1}{2} + \frac{1}{2} + 1 + 1 + \frac{3}{2} + 3 = \frac{15}{2}$$

and

$$D_2(S,\mathfrak{w}) = 0 + \frac{1}{2} + 0 + \frac{1}{2} + \frac{1}{2} + 0 = \frac{3}{2}.$$

We believe that the associated upper bound $\mu_{S,a}^{(n)}(e) \le Cn^{-15/2}[\log n]^{-3/2}$ is sharp but, at this writing, we are not able to obtain a matching lower bound.

As a corollary of Theorem 4.8, we can prove Theorem 1.13. The bracket length $\ell(g)$ of an element of G is defined just before Theorem 1.13.

Corollary 4.10. *Referring to Theorem 4.8, assume that S and a are such that there exists $i \in \{1,\dots,k\}$ with the property that*

$$(\alpha_i, \ell(s_i)) = (2,1) \quad \textit{or} \quad \alpha_i \ell(s_i) < 2.$$

Then

$$\lim_{n\to\infty} n^{D(G)/2}\mu_{S,a}^{(n)}(e) = 0 \tag{4.13}$$

where $D(G) = \sum j \ \operatorname{rank}(G_j/G_{j+1})$ where G_j is the lower central series of G.

Proof. Pick i_0 among those $i \in \{1,\dots,k\}$ such that $(\alpha_i, \ell(s_i)) = (2,1)$ or $\alpha_i \ell(s_i) < 2$ so that α_{i_0} is smallest possible. Let $\mathfrak{w}' = \mathfrak{w}(a)$ be the 2-dimensional weight system introduced in Theorem 4.8 and let $\mathfrak{F}' = \mathfrak{F}(a)$ be the weight function system appearing in the proof of Theorem 4.8. Let $\mathfrak{w}$ be the weight system that assigns weight $(1/2, 0)$ to every $s_i \in S$ with weight function $F_{s_i} = (1+r)^{\frac{1}{2}}$.

If $\alpha_{i_0} < 2/\ell(s_{i_0})$ then by Theorem 3.4 shows that

$$D_1(S,\mathfrak{w}') > D(S,\mathfrak{w}) = D(G)/2.$$

If $\alpha_{i_0} = 2$ then we must have $\ell(s_{i_0}) = 1$. This time, it follows that

$$D_2(S,\mathfrak{w}') \ge 1/2 > D_2(S,\mathfrak{w}) = 0.$$

In both case, Theorem 4.8 show that $\mu_{S,a}^{(n)}(e) = o(n^{-D(G)/2})$ as desired. □

The next statement illustrates the use of a weight system $\mathfrak{w}$ and weight-functions system $\mathfrak{F}$ that are not tightly connected to each other (including cases when the weight functions F_c cannot be order in a useful way).

Theorem 4.11. *Let G be a finitely generated nilpotent group equipped with a generating k-tuple $(s_1, \dots, s_k)$. Assume that μ is a probability measure on G of the form (4.2) with*

$$\mu_i(n) = \kappa_i (1 + |n|)^{-\alpha_i - 1} \ell_i(|n|), \quad 1 \le i \le k,$$

where each ℓ_i is a positive slowly varying function satisfying $\ell_i(t^b) \simeq \ell_i(t)$ for all $b > 0$ and $\alpha_i \in (0, 2)$. Let $\mathfrak{w}$ be the power weight system associated with $a = (\alpha_1, \dots, \alpha_k)$ by setting $w_i = 1/\alpha_i$. Let $(c_i)_1^t$ be a t-tuple of formal commutators such that for each h, the family $\{c_i : w(c_i) = \bar{w}_h\}$ projects to a linearly independent family in $G_h^{\mathfrak{w}} / G_{h+1}^{\mathfrak{w}}$. Let $(s_{i_j}^{\pm 1})_{j=1}^N$ be the list of all the letters (with multiplicity) used in the build-words for the commutators c_i, $1 \le i \le t$. Then

$$\mu^{(n)}(e) \le C n^{-D(S,\mathfrak{w})} L(n)^{-1}$$

where

$$D(S, \mathfrak{w}) = \sum_h \bar{w}_h \operatorname{rank}(G_h^{\mathfrak{w}} / G_{h+1}^{\mathfrak{w}}) \quad \text{and} \quad L(n) = \prod_1^N \ell_{i_j}(n)^{1/\alpha_{i_j}}.$$

Note that this theorem does not offer one but many upper bounds. For each n, one can choose the commutator sequence $(c_i)_1^t$ so as to maximize the size of the resulting $L(n)$.

Example 4.3. Consider the Heisenberg group

$$G = \left\{ \begin{pmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{pmatrix} : x, y, z \in \mathbb{Z} \right\},$$

with generating 3-tuple $S = (X, Y, Z)$ where X is the matrix with

$$x = 1, \quad y = z = 0,$$

and Y, Z a defined similarly. Let $a = (\alpha_1, \alpha_2, \alpha_3) \in (0, 2)$ and let $\ell_1 \equiv 1, \ell_2, \ell_3$ be slowly varying functions such that $\ell_2 \le \ell_3$ if and only if $n \in \cup_k [n_{2k}, n_{2k+1}]$ for some increasing sequence n_k tending to infinity. We also assume that ℓ_2, ℓ_3 satisfy $\ell_i(t^b) \simeq \ell_i(t)$ for all $b > 0$. Applying Theorem 4.11, we obtain:

- if $\frac{1}{\alpha_3} < \frac{1}{\alpha_1} + \frac{1}{\alpha_3}$ then we have

$$\mu^{(n)}(e) \le C n^{-2(\frac{1}{\alpha_1}+\frac{1}{\alpha_2})} \ell_2(n)^{-\frac{2}{\alpha_2}};$$

- if $\frac{1}{\alpha_3} > \frac{1}{\alpha_1} + \frac{1}{\alpha_3}$ then we have

$$\mu^{(n)}(e) \le C n^{-\sum_1^3 \frac{1}{\alpha_i}} \ell_2(n)^{-\frac{1}{\alpha_2}} \ell_3(n)^{-\frac{1}{\alpha_3}};$$

- finally, if $\frac{1}{\alpha_3} = \frac{1}{\alpha_1} + \frac{1}{\alpha_3}$, we have

$$\mu^{(n)}(e) \le C n^{-\frac{2}{\alpha_3}} \begin{cases} \ell_2(n)^{-\frac{2}{\alpha_2}} & \text{if } n \in \cup_k [n_{2k-1}, n_{2k}], \\ \ell_2(n)^{-\frac{1}{\alpha_2}} \ell_3(n)^{\frac{1}{\alpha_3}} & \text{if } n \in \cup_k [n_{2k}, n_{2k+1}]. \end{cases}$$

Example 4.4 (continuation of Example 2.5, 2.6). Consider again the Heisenberg group with $S = (s_1 = X, s_2 = Y, s_3 = Z)$. Set

$$F_1(r) = r^{3/2},$$

$$F_2(r) = r^2 \log(e + r),$$

$$F_3(r) = r^\gamma,$$

with $\gamma > 3/2$. Let μ be the probability measure which assigns to s_i^n, $i = 1, 2, 3$, $n \in \mathbb{Z}$ a probability proportional to

$$\frac{1}{(1 + |n| F_i^{-1}(|n|))}.$$

Namely,

$$\mu(g) = \frac{1}{3} \sum_{i=1}^{3} \sum_{n \in \mathbb{Z}} \mu_i(n) \mathbf{1}_{s_i^n}(g), \quad \mu_i(n) = \frac{c}{1 + |n| F_i^{-1}(|n|)}.$$

Referring to the notation (4.3)(4.4), we have

- $\mathcal{G}_1(n) \simeq (1+n)^{2-(2/3)}$,
  $\tilde{\alpha}_1 = 2/3$,
  $L_1 \equiv 1$,

- $\mathcal{G}_2(n) \simeq (1+n)^{2-(1/2)} [\log(e+n)]^{-1/2}$,
  $\tilde{\alpha}_2 = 1/2$,
  $L_2(n) \simeq [\log(e+n)]^{-1/2}$,

- $\mathcal{G}_3(n) \simeq (1+n)^{2-1/\gamma}$,
  $\tilde{\alpha}_3 = 1/\gamma$,
  $L_3 \equiv 1$.

Apply Theorem 4.11 with $\alpha_i = \tilde{\alpha}_i, \ell_i = L_i$. If $\gamma \in (3/2, 7/2]$, use the sequence of formal commutators $(c_1 = s_1, c_2 = s_2, c_3 = [s_1, s_2])$. If $\gamma > 7/2$, use the sequence of formal commutators $(c_1 = s_1, c_2 = s_2, c_3 = s_3)$ instead. This gives

$$\mu^{(n)}(e) \leq C \begin{cases} (1+n)^{-7}[\log(e+n)]^{-2} & \text{if } \gamma \in (3/2, 7/2], \\ (1+n)^{-(7/2)-\gamma}[\log(e+n)]^{-1} & \text{if } \gamma > 7/2. \end{cases}$$

Below, we will prove a matching lower bound.

5. Norm-radial measures and return probability lower bounds

The aim of this section is to provide lower bounds for the return probability for the random walk driven by the measure $\mu_{S,a}$ on a nilpotent group G, that is, lower bounds on $\mu_{S,a}^{(n)}(e)$. These lower bounds are obtained via comparison with appropriate norm-radial measures.

5.1. Norm-radial measures. A (*proper*) *norm* $\|\cdot\|$ on a countable group G is a function

$$g \longmapsto \|g\| \in [0, \infty)$$

such that

- $\|g\| = 0$ if and only if $g = e$,
- $\#\{\|g\| \leq r\}$ is finite for all $r > 0$,
- $\|g\| = \|g^{-1}\|$, and
- $\|g_1 g_2\| \leq \|g_1\| + \|g_2\|$.

If the triangle inequality is replaced by the weaker property that there exists K such that

$$\|g_1 g_2\| \leq K(\|g_1\| + \|g_2\|),$$

we say that $\|\cdot\|$ is a *quasi-norm.*

The associated left-invariant distance is obtained by setting

$$d(g_1, g_2) = \|g_1^{-1} g_2\|.$$

A norm is κ-geodesic if for any element $g \in G$ there is a sequence $g_1, \dots, g_N$ with $N \leq \kappa\|g\|$ such that

$$\|g_i^{-1} g_{i+1}\| \leq \kappa.$$

A simple observation is that any two κ-geodesic proper norms $\|\cdot\|_1, \|\cdot\|_2$ are comparable in the sense that there is a constant $C \in (0, \infty)$ such that

$$C^{-1}\|g\|_1 \leq \|g\|_2 \leq C\|g\|_1.$$

The word-length norm associated to any finite symmetric generating set is a proper 1-geodesic norm. Most of the quasi-norms that we will consider below are not κ-geodesic. In general, they are not norms but only quasi-norms.

Theorem 5.1. *Let G be a finitely generated group. Let $\|\cdot\|$ be a norm on G such that, for some $d > 0$,*

$$V(r) = \#\{g : \|g\| \leq r\} \simeq r^D, \quad \text{for all } r \geq 1.$$

Fix $\gamma \in (0, 2)$ and set

$$\nu_\gamma(g) = \frac{C_\gamma}{(1 + \|g\|)^\gamma V(\|g\|)}, \quad C_\gamma^{-1} = \sum_g \frac{1}{(1 + \|g\|)^\gamma V(\|g\|)}.$$

Then we have

$$\nu_\gamma^{(n)}(e) \simeq n^{-D/\gamma} \quad \text{for all } n \in \mathbb{N}. \tag{5.1}$$

Remark 5.2. This is a subtle result in that, as stated, it depends very much on the fact that $\|\cdot\|$ is norm versus a quasi-norm. Indeed, the lower bound in (5.1) is false if $\gamma \geq 2$ and the only thing that prevents us to apply the result to $\|\cdot\|^\theta$ with $\theta > 1$ is that, in general, $\|\cdot\|^\theta$ is only a quasi-norm when $\theta > 1$. Note that, by Theorem 1.9, (5.1) holds true for any measure ν such that $\nu \simeq \nu_\gamma$.

Remark 5.3. Definition 2.8 provides a great variety of examples of norms to which Theorem 5.1 applies.

Proof. This result is a straightforward corollary of the much more general result described in [2, Theorem 0.3, Theorem 1.2]. The article [2] is itself part of a long sequence of works by a variety of authors on heat kernel bounds for continuous time jump processes. For instance, see [1, Theorem 1.5] which treats the case of graphs.

To any probability measure, say ν, on a group G, one can associate (a) the discrete time random walk whose law at time n is $\nu^{(n)}$ (here, and in what follows, we assume that the starting point is at the identity element e) and (b) the continuous time random walk whose jumps are taken according to ν but occur at exponentially distributed random time intervals. This description is equivalent to say that the law of this continuous time process is given at time t by

$$p_t^\nu = e^{-t} \sum_0^\infty \frac{t^n}{n!} \nu^{(n)}.$$

It is well known that, under the assumption that ν is symmetric, we have

$$\nu^{(2n)}(e) \simeq p_n^\nu(e). \tag{5.2}$$

See, e.g., [19, Sect. 3.2]. This estimate shows that (5.1) is equivalent to the analogous continuous time statement

$$p_t^{\nu_\gamma}(e) \simeq t^{-D/\gamma} \quad \text{for all } t \geq 1. \tag{5.3}$$

Here we also use the fact that, since $\nu_\gamma(e) > 0$, the control of $\nu_\gamma^{(n)}(e)$ at even times is enough to control odd times as well.

Next, we observe that [2, Theorem 1.4] applies to our context and gives the estimate

$$p_t^{\nu_\gamma}(x) \leq C_1 \min\left\{t^{-D/\gamma}, \frac{t}{\|x\|^{\gamma+D}}\right\}. \tag{5.4}$$

Indeed, all the key hypotheses of [2, Theorem 1.4] are trivially satisfied in our context with the exception of the Nash inequality denoted by (H2) in [2]. As explained in [2] (just after their Theorem 1.5), (H2) follows from the hypothesis $V(r) \simeq r^D$, $r \geq 1$, and the form of ν_γ (see [2, (1.12)]) by a simple argument recorded in [14].

In our context, it is clear that the continous jump process in question cannot escape to infinity in finite time. This property is called *stochastic completeness* and it allows us to use [2, Theorem 0.3] to obtain the lower bound

$$p_t^{\nu_\gamma}(e) \geq t^{-D/\gamma}, \quad t \geq 1.$$

Indeed, stochastic completness (i.e., the fact that $\sum_G p_t^{\nu_\gamma} = 1$) and [2, (0.12)] implies that, for a large enough constant C and all $t \geq 1$, we have

$$\sum_{x\colon \|x\| \leq C t^{1/\gamma}} p_t^{\nu_\gamma}(x) \geq 1/2.$$

Since $\max_{x \in G}\{p_t^{\nu_\gamma}(x)\} = p_t^{\nu_\gamma}(e)$, the desired lower bound follows the volume assumption $V(r) \simeq r^D$, $r \geq 1$. For a more self-contained treatment of Theorem 5.1 and generalizations connected to the present work, see [17, 24]. □

5.2. Comparisons between $\mu_{S,a}$ and radial measures. Let G be a countable group. Let $\|\cdot\|$ be a quasi-norm on G. Set

$$V(r) = \#\{g\colon \|g\| \leq r\}, \quad \text{for all } r \geq 1.$$

Let

$$\phi\colon [0,\infty) \longrightarrow (0,\infty)$$

be continuous. Consider the following hypotheses:

$$\text{there exists } C \text{ such that } V(2r) \leq C V(r) \text{for all } r \geq 0; \tag{5.5}$$

$$\text{there exists } C \text{ such that } \phi(t) \leq C\phi(\lambda t) \text{for all } \lambda \in (1/2, 2),\ t \in (0,\infty); \tag{5.6}$$

and

$$\sum_g \frac{1}{\phi(\|g\|)V(\|g\|)} < \infty. \tag{5.7}$$

Lemma 5.4. *Assume (5.5)–(5.7). For each $n \in \mathbb{Z}$, let $g_n \in G$ and $\Lambda_n \subset G$ be such that*

(1) $g \in \Lambda_n \implies \|g\| \leq C\|g_n\|$;

(2) $V(\|g_n\|) \leq C n \#\Lambda_n$;

(3) $\#\{n\colon g \in \Lambda_n\} \leq C$ *and* $\#\{n\colon g \in g_n^{-1}\Lambda_n\} \leq C$ *for all* $g \in G$.

Then there is a constant C_1 such that

$$\sum_{n\in\mathbb{Z}} \sum_{x\in G} \frac{|f(xg_n) - f(x)|^2}{(1+n)\phi(\|g_n\|)} \leq C_1 \sum_{x,g\in G} \frac{|f(xg) - f(x)|^2}{\phi(\|g\|)V(\|g\|)}.$$

Proof. Using (2), (1), and (3) successively, write

$$\begin{aligned}
&\sum_n \sum_x \frac{|f(xg_n)-f(x)|^2}{(1+n)\phi(\|g_n\|)} \\
&\le C \sum_n \sum_x \frac{|f(xg_n)-f(x)|^2 \#\Lambda_n}{\phi(\|g_n\|)V(\|g_n\|)} \\
&\le 2C \sum_n \sum_{g\in\Lambda_n} \sum_x (|f(xg_n)-f(xg)|^2 + |f(xg)-f(x)|^2) \frac{1}{\phi(\|g_n\|)V(\|g_n\|)} \\
&\le C' \sum_n \sum_{g\in\Lambda_n} \sum_x \Big(\frac{|f(xg^{-1}g_n)-f(x)|^2}{\phi(\|g^{-1}g_n\|)V(\|g^{-1}g_n\|)} + \frac{|f(xg)-f(x)|^2}{\phi(\|g\|)V(\|g\|)} \Big) \\
&\le C'' \sum_{x,g} \frac{|f(xg)-f(x)|^2}{\phi(\|g\|)V(\|g\|)}. \qquad \square
\end{aligned}$$

Remark 5.5. Note that under the hypotheses of Lemma 5.4, we have

$$\sum \frac{1}{(1+n)\phi(\|g_n\|)} < \infty.$$

The next lemma will allow us to apply Lemma 5.4 in the context of Theorem 2.10. Assume that G is a nilpotent group generated by the k -tuple $(s_1,\dots,s_k)$. In addition, we are given a weight system $\mathfrak{w}$ and weight functions F_c such that (2.1) and (2.2) holds. Observe that for any commutators c, c', we have

$$F_{c'} \circ F_c^{-1}(r_1+r_2) \simeq F_{c'} \circ F_c^{-1}(r_1) + F_{c'} \circ F_c^{-1}(r_2), \quad \text{for all } r_1, r_2 \ge 1. \tag{5.8}$$

Indeed, it follows from our hypotheses that $F_{c'} \circ F_c^{-1}$ is an increasing doubling function.

Lemma 5.6. *Referring to the setting of Theorem 2.10, fix* $h \in \{1,\dots,q\}$, $i \in \{m_{h-1}+1,\dots m_{h-1}+R_h\}$ *and an integer* u. *For each* $n \in \mathbb{Z}$, *let* $z_n \in G^{\mathfrak{w}}_{h+1}$ *with* $\|z_n\|_{\mathfrak{F},\mathrm{com}} \le F_{c_1} \circ F_{c_i}^{-1}(n)$. *Set*

$$g_n = \pi(c_i^{un})z_n \in G$$

and

$$\Lambda_n = \Big\{ g = \pi\Big(\prod_1^q \prod_{m_{h-1}+1}^{m_{h-1}+R_h} c_j^{x_j} \Big) : |x_j| \le F_{c_j} \circ F_{c_i}^{-1}(n),\ x_i = \Big\lfloor \frac{un}{2} \Big\rfloor \Big\}.$$

Then (g_n) *and* (Λ_n) *satisfy the hypotheses* 1, 2, *and* 3 *of Lemma* 5.4.

Proof. By Proposition 2.17 and Theorem 2.10, $\|g_n\|_{\mathfrak{F},\mathrm{com}} \simeq F_{c_1} \circ F_{c_i}^{-1}(n)$ and $g \in \Lambda_n$ implies

$$\|g\|_{\mathfrak{F},\mathrm{com}} \le CF_{c_1} \circ F_{c_i}^{-1}(n),$$

so, Property 1 in Lemma 5.4 is satisfied. Property 2 also follows from Theorem 2.10 and the proof of Theorem 3.2.

Suppose that $g \in \Lambda_n \cap \Lambda_m$. Then, computing modulo $G_{h+1}^{\mathfrak{w}}$ and using the fact that $[G_h^{\mathfrak{w}}, G_h^{\mathfrak{w}}] \subset G_{h+1}^{\mathfrak{w}}$ we obtain that $\lfloor un/2 \rfloor = \lfloor um/2 \rfloor$. Similarly, $g \in g_n^{-1}\Lambda_n \cap g_m^{-1}\Lambda_m$ implies $n + \lfloor un/2 \rfloor = m + \lfloor um/2 \rfloor$. In both cases we must have $|n - m| \le 1$. This shows that Property 3 of Lemma 5.4 is satisfied. □

The main result of this section is the following theorem.

Theorem 5.7. *Let G be a nilpotent group with generating k-tuple $S = (s_1, \dots, s_k)$. Let $I_{\mathrm{tor}} = \{i \in \{1, \dots, k\} : s_i$ is torsion in $G\}$. Fix a weight system $\mathfrak{w}$ and a weight-function system $\mathfrak{F}$ such that (2.1) and (2.2) are satisfied. Let $\|\cdot\| = \|\cdot\|_{\mathfrak{F},\mathrm{com}}$ be the associated quasi-norm introduced in Definition 2.8. For each $i \in \{1, \dots, k\} \setminus I_{\mathrm{tor}}$, let*

$$h_i = j_{\mathfrak{w}}(s_i).$$

Let ϕ be such that (5.6) and (5.7) are satisfied.

Let μ be a probability measure on G of the form

$$\mu(g) = \frac{1}{k} \sum_{j=1}^{k} \sum_{n \in \mathbb{Z}} \mu_i(n) \mathbf{1}_{s_i^n}(g)$$

where μ_i is an arbitrary symmetric probability measure on $\mathbb{Z}$ if $i \in I_{\mathrm{tor}}$ and

$$\mu_i(n) = \frac{C_i}{(1+n)\phi(F_{c_1} \circ \mathbf{F}_{h_i}^{-1}(n))}, \quad C_i^{-1} = \sum_n \frac{1}{(1+n)\phi(F_{c_1} \circ \mathbf{F}_{h_i}^{-1}(n))},$$

for $i \in \{1, \dots, k\} \setminus I_{\mathrm{tor}}$. Then there exists C such that

$$\mathcal{E}_\mu(f, f) \le C\mathcal{E}_\nu(f, f)$$

where

$$\nu(g) = \frac{C_\phi}{\phi(\|g\|)V(\|g\|)}, \quad C_\phi^{-1} = \sum_g \frac{1}{\phi(\|g\|)V(\|g\|)}.$$

In particular, there are constants $c > 0$ and N such that

$$\mu^{(2n)}(e) \ge c\nu^{(2Nn)}(e).$$

Proof. Fix i and write $s = s_i$. By Definition 2.13, either s is a torsion element and $s^\kappa = e$ for some κ or $j_{\mathfrak{w}}(s) = h < \infty$. In the second case we can find κ such that

$$s^\kappa = \pi\Big(\prod_{m_{h-1}+1}^{m_{h-1}+\rho} c_i^{x_i}\Big)z, \quad x_{m_{h-1}+\rho} \neq 0, z \in G^{\mathfrak{w}}_{h+1}.$$

If s is torsion, it is very easy to see that $\mathcal{E}_{s,\mu_i}(f,f) \leq C(f,f)$. In the course of this proof, C denotes a generic constant that may change from line to line. If s is not torsion and

$$s^\kappa = \pi\Big(\prod_{m_{h-1}+1}^{m_{h-1}+\rho} c_i^{x_i}\Big)z, \quad x_{m_{h-1}+\rho} \neq 0, z \in G^{\mathfrak{w}}_{h+1},$$

set

$$F = F_{c_{m_{h-1}+1}}$$

(we have $F \simeq F_{c_j}$, $j \in \{m_{h-1}+1, m_h\}$). Then, for any n,

$$s^{\kappa n} = \pi\Big(\prod_{m_{h-1}+1}^{m_{h-1}+\rho} c_i^{x_i n}\Big)z_n \quad \text{with } \|z_n\| \leq C F_{c_1} \circ F^{-1}(|n|), z_n \in G^{\mathfrak{w}}_{h+1}.$$

Now, write

$$n = \kappa u_n + v_n,$$

with

$$|v_n| < \kappa$$

and

$$\sum_g |f(gs^n) - f(g)|^2 \leq 2\Big(\sum_g |f(gs^{\kappa u_n}) - f(g)|^2 + \sum_g |f(gs^{v_n}) - f(g)|^2\Big).$$

By Lemma 5.6 and Remark 5.5, the hypotheses of Theorem 5.7 imply that

$$\sum ((1+n)\phi(\|s^n\|))^{-1} < \infty.$$

Hence, it is is easy to check that

$$\sum_g \sum_n \frac{|f(gs^{v_n}) - f(g)|^2}{(1+n)\phi(\|s^n\|)} \leq C\mathcal{E}_\nu(f,f). \tag{5.9}$$

Consequently, it suffices to show that

$$\sum_g \sum_n \frac{|f(gs^{\kappa u_n}) - f(g)|^2}{(1+n)\phi(\|s^n\|)} \leq C\mathcal{E}_\nu(f,f).$$

We have

$$\|s^n\| \simeq \|s^{\kappa u_n}\| \simeq F_{c_1} \circ F^{-1}(\kappa u_n).$$

Hence

$$\sum_g \sum_n \frac{|f(gs^{\kappa u_n}) - f(g)|^2}{(1+n)\phi(\|s^n\|)} \leq C \sum_g \sum_\ell \frac{|f(gs^{\kappa \ell}) - f(g)|^2}{\ell \phi(F_{c_1} \circ F^{-1}(\ell))}. \tag{5.10}$$

Next, set

$$i_1 = m_{h-1} + 1, \quad i_2 = m_{h-1} + \rho$$

and write

$$\begin{aligned} &\sum_g \sum_\ell |f(gs^{\kappa\ell}) - f(g)|^2 \\ &\quad \leq \rho\Big(\sum_g \sum_\ell \sum_{i=i_1}^{i_2-1} |f(g\pi(c_i^{x_i \ell})) - f(g)|^2 \\ &\qquad\qquad + \sum_g \sum_\ell |f(g\pi(c_{i_2}^{x_{i_2}\ell}) z_\ell) - f(g)|^2 \Big). \end{aligned}$$

By Lemmas 5.4-5.6, for each $i = i_1, \ldots, i_2 - 1$, we have

$$\sum_g \sum_\ell \frac{|f(g\pi(c_i^{x_i \ell})) - f(g)|^2}{(1+\ell)\phi(\|\pi(c_i^{x_i \ell})\|)} \leq C\mathcal{E}_\nu(f, f)$$

and, since $z_\ell \in G^{\mathfrak{w}}_{h+1}$ and $\|z_\ell\| \leq CF_{c_1} \circ F^{-1}(\ell)$,

$$\sum_g \sum_\ell \frac{|f(g\pi(c_{i_2}^{x_{i_2}\ell}) z_\ell) - f(g)|^2}{(1+\ell)\phi(\|\pi(c_{i2}^{x_{i_2}\ell}) z_\ell\|)} \leq C\mathcal{E}_\nu(f, f).$$

Further, for each $i = i_1, \ldots, i_2$ with $x_i \neq 0$, we have

$$\|\pi(c_i^{x_i \ell})\| \simeq F_{c_1} \circ F^{-1}(\ell)$$

as well as

$$\|\pi(c_i^{x_{i_2}\ell}) z_\ell\| \simeq F_{c_1} \circ F^{-1}(\ell).$$

Hence (5.10) and the above estimates give

$$\sum_g \sum_n \frac{|f(gs^{\kappa u_n}) - f(g)|^2}{(1+n)\phi(\|s^n\|)} \leq C\mathcal{E}_\nu(f, f).$$

Together with (5.9), this gives

$$\sum_{g \in G} \sum_{n \in \mathbb{Z}} \frac{|f(gs^n) - f(g)|^2}{(1+n)\phi(\|s^n\|)} \leq C\mathcal{E}_\nu(f, f).$$

Since this holds true for each $s = s_i$, $i = 1, \ldots, k$, the desired result follows. □

5.3. Assorted corollaries: return probability lower bounds. In this section we use the comparison with norm-radial measures to obtain explicit lower estimates on $\mu_{S,a}^{(n)}(e)$. The simplest and most important result of this type is as follows.

Theorem 5.8. *Let G be a finitely generated nilpotent group equipped with a generating k-tuple $(s_1, \ldots, s_k)$ and a k-tuple of positive reals $a = (\alpha_1, \ldots, \alpha_k) \in (0,2)^k$. Let $\mathfrak{w}$ be the weight system which assigns weight $w_i = 1/\alpha_i$ to s_i. Then*

$$\mu_{S,a}^{(n)}(e) \geq c_{S,a} n^{-D(S,\mathfrak{w})}$$

where $D(S,\mathfrak{w}) = \sum_h \bar{w}_h \operatorname{rank}(G_h^{\mathfrak{w}}/G_{h+1}^{\mathfrak{w}})$.

Remark 5.9. This lower bound matches precisely the upper bound given by Theorem 4.6. Thus, as stated in Theorems 1.2-1.8, for any $a \in (0,2)^k$,

$$\mu_{S,a}^{(n)}(e) \simeq n^{-D(S,\mathfrak{w})}.$$

Note however that, in Theorems 1.2-1.8, the constraints on the α_i's is weaker. This more general case will be treated below.

Proof. Fix a sequence $\Sigma = (c_i)_1^t$ of commutators as in Theorem 2.10 and let $\|\cdot\|$ be the associated norm $\|\cdot\| = \|\cdot\|_{\Sigma,\mathfrak{F}}$ introduced in Definition 2.8 where the weight function system. is generated by $F_i(r) = r^{w_i}$. Note that, by Remark 2.9, $\|\cdot\|$ is indeed not only a quasi-norm but a norm. By hypothesis, $1/w(c_1) < 2$. Hence Theorem 5.1, together with Theorem 3.2, shows that the norm-radial measure

$$\nu(g) = \frac{C}{(1+\|g\|)^{1/w(c_1)} V(\|g\|)}$$

satisfies

$$\nu^{(n)}(e) \geq cn^{-w(c_1)D(S,\mathfrak{w})/w(c_1)} = cn^{-D(S,\mathfrak{w})}. \tag{5.11}$$

Theorem 5.7 produces a symmetric measure μ such that $\mathcal{E}_\mu \leq C\mathcal{E}_\nu$. This measure μ is given by

$$\mu(g) = \frac{1}{k}\sum_{j=1}^{k}\sum_{n\in\mathbb{Z}} \mu_i(n)\mathbf{1}_{s_i^n}(g)$$

where μ_i is an arbitrary symmetric probability measure on $\mathbb{Z}$ if $i \in I_{\text{tor}}$ and

$$\mu_i(n) = \frac{C_i}{(1+n)(1+F_{c_1}\circ \mathbf{F}_{h_i}^{-1}(n))^{1/w(c_1)}}$$

with

$$C_i^{-1} = \sum_n \frac{1}{(1+n)(1+F_{c_1} \circ \mathbf{F}_{h_i}^{-1}(n))^{1/w(c_1)}}$$

for $i \in \{1, \ldots, k\} \setminus I_{\mathrm{tor}}$. In the latter case, we have

$$\mathbf{F}_{h_i}(t) = t^{\bar{w}_{h_i}}$$

with $\bar{w}_{h_i} \geq w(s_i) = 1/\alpha_i$ and $F_{c_1}(t) = t^{w(c_1)}$. Hence

$$\mu_i(n) \simeq \frac{C_i}{(1+n)^{1+1/\bar{w}_{h_i}}} \geq \frac{C_i'}{(1+n)^{1+\alpha_i}}.$$

It follows that if we pick μ_i to be given by

$$\mu_i(n) = c_i(1+n)^{-(1+\alpha_i)} \quad \text{for } i \in I_{\mathrm{tor}},$$

and

$$\mu_i = c_i(1+n)^{1+1/\bar{w}_{h_i}} \quad \text{if } i \in I \setminus I_{\mathrm{tor}},$$

then we obtain a measure μ such that

$$\mathcal{E}_{\mu_{S,a}} \leq C\,\mathcal{E}_\mu \leq C'\mathcal{E}_\nu.$$

By Theorem 1.9, this implies that there are $c, N \in (0, \infty)$ such that

$$\mu_{S,a}^{(2n)}(e) \geq c\nu^{(2nN)}(e).$$

Thus the lower bound stated in Theorem 5.8 follows from (5.11). □

The following theorem extends the range of applicability of the previous result. In particular, the statement is different but equivalent to the statement recorded in Theorem 1.8. See also Theorem 5.13 below.

Theorem 5.10. *Let G be a finitely generated nilpotent group equipped with a generating k-tuple $(s_1, \ldots, s_k)$ and a k-tuple of positive reals $a = (\alpha_1, \ldots, \alpha_k) \in (0, \infty]^k$. Set $\tilde{\alpha}_i = \min\{\alpha_i, 2\}$. Let $\mathfrak{w}$ be the weight system which assigns weight $1/\tilde{\alpha}_i$ to $s_i \in S$. Let Σ be a sequence of formal commutators as in Theorem 2.10. Assume that $w(s) > 1/2$ for all $s \in \mathrm{core}(\mathfrak{w}, S, \Sigma)$. Then*

$$\mu_{S,a}^{(n)}(e) \simeq n^{-D(S,\mathfrak{w})}.$$

Proof. The upper bound follows from Theorem 4.6. The lower bound is more subtle. Consider any $s \in S$ such that $w(s) = 1/2$ (i.e., $s = s_i$ with $\alpha_i \geq 2$). Observe that $1/2$ is the lowest possible value for weights in $\mathfrak{w}$ and that the hypothesis that $w > 1/2$ on core$(\mathfrak{w}, S, \Sigma)$ implies that $G_1^{\mathfrak{w}}/G_2^{\mathfrak{w}}$ is a torsion group. In particular, this implies that $\bar{w}_{j_{\mathfrak{w}}(s)} > 1/2 = w(s)$. By Corollary 2.19, the weight system $\mathfrak{w}'$ generated by

$$w'(s) = \begin{cases} w(s) & \text{if } w(s) \neq 1/2, \\ \bar{w}_2 & \text{if } w(s) = 1/2, \end{cases}$$

is such that $w(s) \leq w'(s) \leq \bar{w}_{j_{\mathfrak{w}}(s)}$ for all $s \in S$ and $w'(s) > 1/2$ for all $s \in S$. Now, Theorem 5.7 gives the comparison $\mathcal{E}_{\mu_{S,a}} \leq C\mathcal{E}_\nu$ with

$$\nu(g) \simeq \frac{1}{(1 + \|g\|_{\Sigma,\mathfrak{w}})^{1/w_\Sigma} V_{\Sigma,\mathfrak{w}}(\|g\|_{\Sigma,\mathfrak{w}})}.$$

However, since the minimum weight value w_Σ may be equal to $1/2$, we cannot apply Theorem 5.1 directly. We proceed as follows. By the definition of w' and Corollary 2.19, we have

$$\|g\|_{\Sigma,\mathfrak{w}}^{1/w_\Sigma} \simeq \|g\|_{S,\mathfrak{w}'}^{1/w'_S} \quad \text{for all } g \in G.$$

Note that this implies that

$$V_{\Sigma,\mathfrak{w}}(\|g\|_{\Sigma,\mathfrak{w}}) = \#\{g' \in G : \|g'\|_{\Sigma,\mathfrak{w}} \leq \|g\|_{\Sigma,\mathfrak{w}}\} \simeq V_{S,\mathfrak{w}'}(\|g\|_{S,\mathfrak{w}'}).$$

Hence we have

$$\mathcal{E}_\nu \simeq \mathcal{E}_{\nu'}$$

where

$$\nu'(g) \simeq \frac{1}{(1 + \|g\|_{S,\mathfrak{w}'})^{1/w'_S} V_{S,\mathfrak{w}'}(\|g\|_{S,\mathfrak{w}'})}.$$

Now, since by construction $w'_S > 1/2$, we can apply Theorem 5.1 which gives

$$(\nu')^{(n)}(e) \simeq n^{-D(S,\mathfrak{w}')} = n^{-D(S,\mathfrak{w})}.$$

Also, we have

$$\mathcal{E}_{\mu_{S,a}} \leq C\mathcal{E}_\nu \simeq \mathcal{E}_{\nu'}.$$

Hence

$$\mu_{S,a}^{(n)}(e) \geq cn^{-D(S,\mathfrak{w})}.$$

This ends the proof of Theorem 5.10. □

Our next results provides a comparison between the behaviors of two measures $\mu_{S,a}$ and $\mu_{S',a'}$. Compare to Corollary 1.12 and Theorem 1.13 which treats comparison with $\mu_{S',a'}$ when $a' = (\alpha'_i)_1^{k'} \in (2,\infty]^{k'}$, a case that is excluded in Theorem 5.11.

Theorem 5.11. *Let G be a finitely generated nilpotent group equipped with a generating k-tuple $(s_1,\dots,s_k)$ and a k-tuple of positive reals $a = (\alpha_1,\dots,\alpha_k) \in (0,\infty]^k$. Set $\tilde{\alpha}_i = \min\{\alpha_i, 2\}$. Let $\mathfrak{w}$ be the weight system which assigns weight $1/\tilde{\alpha}_i$ to $s_i \in S$. Fix another weight system $\mathfrak{w}' = (w'_1,\dots,w'_k)$ with minimal weight $w'_S > 1/2$. Let Σ be a sequence of formal commutators as in Theorem 2.10 for $(S,\mathfrak{w}')$. Assume that $w(s) \geq w'(s)$ for all $s \in \mathrm{core}(\mathfrak{w}', S, \Sigma)$. Then*

$$\mu_{S,a}^{(n)}(e) = o(n^{-D(S,\mathfrak{w}')})$$

if and only if there exists $s \in S$ such that $w(s) > \bar{w}'_{j_{\mathfrak{w}'}(s)}$.

Proof. Apply Theorem 4.6 and Theorem 5.10 together with Corollary 2.19 and Theorem 3.4. □

Theorem 5.12. *Let G be a finitely generated nilpotent group equipped with a generating k-tuple $(s_1,\dots,s_k)$ and a k-tuple of positive reals $a = (\alpha_1,\dots,\alpha_k) \in (0,\infty]^k$. Set $\tilde{\alpha}_i = \min\{\alpha_i, 2\}$. Let $\mathfrak{w}$ be the weight system which assigns weight $w_i = 1/\tilde{\alpha}_i$ to s_i. Then there exists $A \geq 0$ such that*

$$\mu_{S,a}^{(n)}(e) \geq c_{S,a} n^{-D(S,\mathfrak{w})}[\log n]^{-A}.$$

Further, let Σ be as in Theorem 2.10 applied to $(S,\mathfrak{w})$ and assume that $\alpha_i = 2$ for all $i \in \{1,\dots,k\}$ such that $s_i \in \mathrm{core}(S,\mathfrak{w},\Sigma)$. Then

$$\mu_{S,a}^{(n)}(e) \simeq [n \log n]^{-D(G)/2}.$$

Proof. The proof of the general lower bound is essentially the same as for Theorem 5.8, except that we cannot rule out the possibility that $w(c_1) = 1/2$. If $w(c_1) > 1/2$ then the previous proof applies and we obtain $\mu_{S,a}^{(n)}(e) \geq cn^{-D(S,\mathfrak{w})}$ which is better than the statement we need to prove. If $w(c_1) = 1/2$ then we have a comparison

$$\mathcal{E}_{\mu_{S,a}} \leq C\mathcal{E}_\nu \tag{5.12}$$

with

$$\nu(g) = \frac{C}{(1+\|g\|)^2 V(\|g\|)}.$$

To conclude, we need a lower bound on $\nu^{(n)}(e)$. This turns out to be rather subtle and difficult question in the present generality. In [24, Theorem 5.6] we show that there exists $A \geq 0$ such that

$$\nu^{(n)}(e) \geq cn^{-D(S,\mathfrak{w})}[\log n]^{-A}. \tag{5.13}$$

This proves the desired lower bound on $\mu^{(n)}_{S,a}(e)$.

When $\alpha_i = 2$ for all $i \in \operatorname{core}(S, \mathfrak{w}, \Sigma)$, it follows that

$$D(S,\mathfrak{w}) = D(G)/2 \quad \text{and} \quad \|g\| \simeq |g|_S$$

where $|g|_S$ denotes the usual word-length of g over the symmetric generating set $\{s_i^{\pm 1} : 1 \leq i \leq k\}$. Theorem 4.8 provides the upper bound

$$\mu^{(n)}_{S,a}(e) \leq C[n \log n]^{-D(G)/2}.$$

For the lower bound, by the Dirichlet form inequality (5.12), it suffices to bound $\nu^{(n)}(e)$ from below. Using the fact that $\|g\| \simeq |g|_S$, we prove in [24, Theorem 5.5] that, in this special case, (5.13) holds with $A = D(G)/2$. This provides the desired matching lower bounds

$$\mu^{(n)}_{S,a}(e) \geq c[n \log n]^{-D(G)/2}. \qquad \square$$

Theorem 5.13. *Let G be a finitely generated nilpotent group equipped with a generating k-tuple $(s_1,\dots,s_k)$ and a k-tuple of positive reals $a = (\alpha_1,\dots,\alpha_k) \in (0,\infty]^k$. Set $\tilde{\alpha}_i = \min\{\alpha_i, 2\}$ and $w_i = 1/\tilde{\alpha}_i$. Let $\mathfrak{w}$ be the associated weight system. Let Σ be as in Theorem 2.10 applied to $(S,\mathfrak{w})$. Let*

$$\Theta = (\theta_1 = s_{i_1},\dots,\theta_\kappa = s_{i,\kappa}) = \operatorname{core}(S,\mathfrak{w},\Sigma).$$

Let H be the subgroup of G generated by Θ. Set $b = (\beta_1 = \alpha_{i_1},\cdots,\beta_\kappa = \alpha_{i_\kappa})$, $\tilde{\beta}_i = \tilde{\alpha}_{i_j}$, $v(\theta_i) = w(s_{i_j})$. Let $\mathfrak{v}$ be the weight system associated to v on (H,Θ), respectively. Then

$$D(\Theta,\mathfrak{v}) = D(S,\mathfrak{w}).$$

In particular, letting e_H, e_G be the identity elements in H and G, respectively, we have:

- *if $\alpha_i \in (0,2)$ for all i such that $s_i \in \operatorname{core}(S,\mathfrak{w},\Sigma)$ then*

$$\mu^{(n)}_{S,a}(e_G) \simeq \mu^{(n)}_{\Theta,b}(e_H) \simeq n^{-D(\Theta,\mathfrak{v})}.$$

- *if $\alpha_i = 2$ for all i such that $s_i \in \operatorname{core}(S,\mathfrak{w},\Sigma)$ then*

$$\mu^{(n)}_{S,a}(e_G) \simeq \mu^{(n)}_{\Theta,b}(e_H) \simeq [n \log n]^{-D(H)/2}.$$

Remark 5.14. One can easily prove that H is a subgroup of finite index in G. It is also easy to prove by the direct comparison techniques of [19] that

$$\mu_{S,a}^{(2Kn)}(e_G) \le C\mu_{\Theta,b}^{(2n)}(e_H), \quad \text{for all } n \in \mathbb{N},$$

for some integer K and constant C and for each $a = (\alpha_1, \dots, \alpha_k)$. The converse inequality seems significantly harder to prove although we conjecture it does hold true.

Proof. First we observe that $D(\Theta, \mathfrak{v}) \le D(S, \mathfrak{w})$. Indeed, this follows immediately from the obvious fact that

$$\{g \in H : \|g\|_{\Theta,\mathfrak{v}}^{1/v_\Theta} \le r\} \subset \{g \in G : \|g\|_{S,\mathfrak{w}}^{1/w_S} \le r\}.$$

To prove that $D(\Theta, \mathfrak{v}) \ge D(S, \mathfrak{w})$, it is convenient to introduce the generating k-tuple $S^* = (s_i^*)_1^k$ of H such that $s_{i,j}^* = s_{i_j}$ if $s_{i_j} = \theta_j \in \Theta$, and $s_{i_j}^* = e$ otherwise. Both S and S^* are equipped with the weight system $\mathfrak{w}$. Obviously, the non-decreasing sequence of subgroups $(H_j^{\mathfrak{w}})$ is a trivial refinement of the sequence $(H_j^{\mathfrak{v}})$ in the sense that the two sequences differ only by insertion of some repetitions. For instance, A, B, C may become A, A, B, B, B, B, C. It follows that $D(\Theta, \mathfrak{v}) = D(S^*, \mathfrak{w})$. The notational advantage is that the weight system $\mathfrak{w}$ with increasing weight-value sequence $\bar{w}_j$ is now shared by S and S^*. We wish to prove that

$$\operatorname{rank}(H_j^{\mathfrak{w}}/H_{j+1}^{\mathfrak{w}}) \ge \operatorname{rank}(G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}).$$

The (torsion free) rank of an abelian group can be computed as the cardinality of a maximal free subset. Set $R = R_j^{\mathfrak{w}}$ be the torsion free rank of $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$. Let $(c_{m_{j-1}+1}, \dots, c_{m_{j-1}+R})$ be the formal commutators given by Theorem 2.10 which form a maximal free subset of $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$. By definition of $\operatorname{core}(S, \mathfrak{w}, \Sigma)$, the images of these formal commutators in G belong to H. In fact, they clearly belong to $H_j^{\mathfrak{w}} \subset G_j^{\mathfrak{w}}$. Now, we also have $H_{j+1}^{\mathfrak{w}} \subset G_{j+1}^{\mathfrak{w}}$. Assume that

$$\prod_{m_{j-1}+1}^{m_{j-1}+R} c_i^{x_i} = e$$

in $H_j^{\mathfrak{w}}/H_{j+1}^{\mathfrak{w}}$. Then, a fortiori, this product is trivial in

$$H_j^{\mathfrak{w}} G_{j+1}^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}} \simeq H_j^{\mathfrak{w}}/(H_j^{\mathfrak{w}} \cap G_{j+1}^{\mathfrak{w}})$$

since $(H_j^{\mathfrak{w}} \cap G_{j+1}^{\mathfrak{w}}) \subset H_{j+1}^{\mathfrak{w}}$. In particular, this product must be trivial in $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$. This implies that $x_i = 0$ for all i so that $H_j^{\mathfrak{w}}/H_{j+1}^{\mathfrak{w}}$ admits a free subset of size R. It follows that $\operatorname{rank}(H_j^{\mathfrak{w}}/H_{j+1}^{\mathfrak{w}}) \ge R$ as desired. □

To state the final result of this section, we need some preparation. Consider the class of measure μ of the form (4.2) with

$$\mu_i(n) = \kappa_i(1+|n|)^{-\alpha_i-1}\ell_i(|n|), \quad 1 \le i \le k, \tag{5.14}$$

where each ℓ_i is a positive slowly varying function satisfying $\ell_i(t^b) \simeq \ell_i(t)$ for all $b > 0$ and $\alpha_i \in (0,2)$ (regarding the notions of slowly and regularly varying functions, see Example 2.3 or consult [6]). Consider the weight-function system $\mathfrak{F}$ generated by letting F_i be the inverse function of $r \mapsto r^{\alpha_i}/\ell_i(r)$. Note that F_i is regularly varying of order $1/\alpha_i$ and that $F_i(r) \simeq [r\ell_i(r)]^{1/\alpha_i}$, $r \ge 1$, $i = 1,\dots,k$. We make the fundamental assumption that the functions F_i have the property that for any $1 \le i, j \le k$, either $F_i(r) \le CF_j(r)$ of $F_j(r) \le CF_i(r)$. For instance, this is clearly the case if all α_i are distinct. Without loss of generality, we can assume that there exists a multidimensional weight system $\mathfrak{w}$, say of dimension d, with

$$w_i = (v_i^1,\dots,v_i^d), \;\; v_i^1 = 1/\alpha_i, \;\; 1 \le i \le k,$$

and such that $\mathfrak{w}$ and $\mathfrak{F}$ are compatible in the sense that (2.1)-(2.2) hold true. Separately, consider also the one-dimensional weight system $\mathfrak{v}$ generated by $v_i = 1/\alpha_i$, $1 \le i \le k$. Note that one can check that

$$D(S,\mathfrak{v}) = \sum_j \bar{v}_j R_j^{\mathfrak{v}} = \sum_j \bar{v}_i^1 R_j^{\mathfrak{w}}$$

where, by definition, $\bar{w}_j = (\bar{v}_j^1,\dots,\bar{v}_j^d)$. Fix $\alpha_0 \in (0,2)$ such that

$$\alpha_0 > \max\{\alpha_i : 1 \le i \le k\}$$

and $\alpha_0/\alpha_i \notin \mathbb{N}$, $i = 1,\dots,k$. Observe that there are convex functions $K_i \ge 0$, $i = 0,\dots,k$, such that $K_i(0) = 0$ and

$$F_i(r^{\alpha_0}) \simeq K_i(r), \quad r \ge 1. \tag{5.15}$$

Indeed, $r \mapsto F_i(r^{\alpha_0})$ is regularly varying of index α_0/α_i with $1 < \alpha_0/\alpha_i \notin \mathbb{N}$. By [6, Theorems 1.8.2-1.8.3] there are smooth positive convex functions $\tilde{K}_i$ such that $\tilde{K}_i(r) \sim F_i(r^{\alpha_0})$. If $\tilde{K}_i(0) > 0$, it is easy to construct a convex function $K_i : [0,\infty) \to [0,\infty)$ such that $K_i \simeq \tilde{K}_i$ on $[1,\infty)$ and $K_i(0) = 0$.

Theorem 5.15. *Let G be a finitely generated nilpotent group equipped with a generating k-tuple $(s_1,\dots,s_k)$. Assume that μ is a probability measure on G of the form* **(4.2)** *with μ_i as in* **(5.14)***. Let ℓ_i, F_i, $\mathfrak{F},\mathfrak{w},\mathfrak{v}$ be as described above. Let $(c_i)_1^t$ be a t-tuple of formal commutators as in* ***Theorem* 2.10** *applied to $G,S,\mathfrak{w},\mathfrak{F}$. Let $(s_{i_j}^{\pm 1})_{j=1}^N$ be the list of all the letters (repeated according to multiplicity) used in the build-words for the commutators c_i with $i\in\bigcup_j\{m_{j-1}+1,\dots,m_{j-1}+R_j^{\mathfrak{w}}\}$. Then*

$$\mu^{(n)}(e)\simeq n^{-D(S,\mathfrak{v})}L(n)^{-1}$$

where

$$L(n)=\prod_1^N \ell_{i_j}(n)^{1/\alpha_{i_j}}.$$

Proof. The upper bounds follows immediately from Theorem 4.11. For the lower bound, it is technically convenient to adjoint to S the dummy generator $s_0=e$ with associated weight function $F_0(r)=r^{1/\alpha_0}$. Let $\mathfrak{W}_0,\mathfrak{F}_0$ we the weight systems induced by $S_0=(e,s_1,\dots,s_k)$, $F_0,F_1,\dots,F_k$.

Apply Theorem 5.7 to $G,S,\mathfrak{w}_0,\mathfrak{F}_0$ to obtain that $\mathcal{E}_\mu\le C\mathcal{E}_\nu$ where

$$\nu(g)\simeq\frac{1}{(1+\|g\|_{\mathfrak{F}_0,\mathrm{com}})^{\alpha_0}V_{\mathfrak{F}_0,\mathrm{com}}(\|g\|_{\mathfrak{F}_0,\mathrm{com}})}$$

with $V_{\mathfrak{F}_0,\mathrm{com}}(r)=\#\{g\in G:\|g\|_{\mathfrak{F}_0,\mathrm{com}}\le r\}$. By construction,

$$\nu(g)\simeq\frac{1}{(1+\|g\|)^{\alpha_0}V(\|g\|)}$$

where $\|\cdot\|$ is the norm $\|\cdot\|_{\mathfrak{K},\mathrm{com}}$ based on the convex function $K_i\simeq F_i(r^{\alpha_0})$ provided by (5.15) and V denotes the associated volume function. Indeed, by construction we have $\|\cdot\|\simeq\|\cdot\|_{\mathfrak{F}_0,\mathrm{com}}^{\alpha_0}$. As $\|\cdot\|$ is a norm, an extension of Theorem 5.1 obtained in [24] and which allows volume growth of regular variation with positive index gives

$$\nu^{(n)}(e)\simeq\frac{1}{V(n)}\simeq\frac{1}{V_{\mathfrak{F}_0,\mathrm{com}}(n^{1/\alpha_0})}\simeq\frac{1}{\#Q(S_0,\mathfrak{F}_0,n)}\simeq\frac{1}{\#Q(S,\mathfrak{F},n)}.$$

Using the notation introduced in Theorem 5.15, we have

$$\#Q(S,\mathfrak{F},r)\simeq n^{D(S,\mathfrak{v})}L(n)$$

which yields the desired result. □

5.4. Near diagonal lower bounds. In this section we use Lemma 4.4(ii) to turn the sharp *on diagonal lower bounds* of the previous section into *near diagonal lower bounds*. The key tool is the following lemma.

Lemma 5.16. *Let G be a finitely generated nilpotent group equipped with a generating k-tuple $(s_1,\dots,s_k)$ and a k-tuple of positive reals $a=(\alpha_1,\dots,\alpha_k)\in(0,\infty]^k$. Let $\mathfrak{w}=\mathfrak{w}(a)$ be the two-dimensional weight system which assigns weight $w_i=(v_{i,1},v_{i,2})$ to s_i where*

$$v_{i,1}=\frac{1}{\tilde{\alpha}_i},\quad \tilde{\alpha}_i=\min\{2,\alpha_i\}$$

and

$$v_{i,2}=0 \text{ unless } \alpha_i=2 \text{ in which case } v_{i,2}=1/2.$$

Let $\mathfrak{F}$ be the associated weight function system generated by

$$F_i(r)=r^{v_{i,1}}[\log(1+r)]^{v_{i,2}},\quad 1\le i\le k.$$

Then

$$|\mu_{S,a}^{(2n+m)}(xg)-\mu_{S,a}^{(2n+m)}(x)|\le C(F_S^{-1}(\|g\|_{\Sigma,\mathfrak{F}})/m)^{1/2}\mu_{S,a}^{(2n)}(e).$$

Proof. By Theorem 2.10, there is an integer $p=p(G,S,\mathfrak{w})$ such that any g with $F_S^{-1}(\|y\|_{S,\mathfrak{F}})=r$ can be expressed as

$$g=\prod_{j=1}^{p}s_{i_j}^{x_j}\quad \text{with } |x_j|\le CF_{i_j}(r).$$

Write

$$\mu_{S,a}^{(2n+m)}=\mu_{S,a}^{(n+m)}*\mu_{S,a}^{(n)}$$

and, for each step $s_{i_j}^{x_j}$, apply Lemma 4.4(ii) to obtain

$$\begin{aligned}&|\mu_{S,a}^{(2n+m)}(zs_{i_j}^{x_j})-\mu_{S,a}^{(2n+m)}(z)|\\&\quad\le C\mathcal{G}_{i_j}(|x_j|)^{-1/2}|x_j|\mathcal{E}_{\mu_{S,a}}(\mu_{S,a}^{(n+m)},\mu_{S,a}^{(n+m)})^{1/2}\|\mu_{S,a}^{(n)}\|_2\\&\quad\le Cr^{1/2}\mathcal{E}_{\mu_{S,a}}(\mu_{S,a}^{(n+m)},\mu_{S,a}^{(n+m)})^{1/2}\|\mu_{S,a}^{(n)}\|_2.\end{aligned}$$

Here, according to Lemma 4.4, $\mathcal{G}_i(r)=r^{2-\tilde{\alpha}_i}$ if $v_i,2=0$ and $\mathcal{G}_i(r)=\log(1+r)$ if $v_{i,2}=1/2$ (i.e., if $\alpha_i=2$). Hence, $s^2/\mathcal{G}_i(s)\simeq F_i^{-1}(s)$, which gives the last inequality.

By [13, Lemma 3.2], we also have

$$\mathcal{E}_{\mu_{S,a}}(\mu_{S,a}^{(n+m)},\mu_{S,a}^{(n+m)})^{1/2}\le Cm^{-1/2}\|\mu_{S,a}^{(n)}\|_2=Cm^{-1/2}\mu_{S,a}^{(2n)}(e)^{1/2}.$$

This gives the desired inequality. □

Theorem 5.17. *Let G be a finitely generated nilpotent group equipped with a generating k-tuple $(s_1,\dots,s_k)$ and a k-tuple of positive reals $a=(\alpha_1,\dots,\alpha_k)\in(0,\infty]^k$. Set $\tilde{\alpha}_i=\min\{\alpha_i,2\}$. Let $\mathfrak{w}$ be the weight system which assigns weight $1/\tilde{\alpha}_i$ to $s_i\in S$. Let Σ be a sequence of formal commutators as in Theorem 2.10. Assume that $w(s)>1/2$ for all $s\in\operatorname{core}(\mathfrak{w},S,\Sigma)$. Then, there exists $\epsilon>0$ such that, uniformly over the region $\{x\in G\colon \|x\|_{S,\mathfrak{w}}\le F_S(\epsilon n)\}$, we have*

$$\mu_{S,a}^{(n)}(x)\simeq n^{-D(S,\mathfrak{w})}.$$

Proof. Theorem 5.10 gives $\mu_{S,a}^{(n)}(e)\simeq n^{-D(S,\mathfrak{w})}$. This, together with Lemma 5.16, yields the desired lower bound. □

Theorem 5.18. *Let G be a finitely generated nilpotent group equipped with a generating k-tuple $(s_1,\dots,s_k)$ and a k-tuple of positive reals $a=(\alpha_1,\dots,\alpha_k)\in(0,\infty]^k$. Set $\tilde{\alpha}_i=\min\{\alpha_i,2\}$. Let $\tilde{\mathfrak{w}}$ be the weight system which assigns weight $\tilde{w}_i=1/\tilde{\alpha}_i$ to s_i. Let Σ be as in Theorem 2.10 applied to $(S,\tilde{\mathfrak{w}})$ and assume that $\alpha_i=2$ for all $i\in\{1,\dots,k\}$ such that $s_i\in\operatorname{core}(S,\tilde{\mathfrak{w}},\Sigma)$. Then there exists $\epsilon>0$ such that, uniformly over the region*

$$\{x\in G\colon |x|_S^2[\log|x|_S]^{-1}\le\epsilon n\},$$

we have

$$\mu_{S,a}^{(n)}(x)\simeq[n\log n]^{-D(G)/2}.$$

Proof. By Theorem 5.12, we have $\mu_{S,a}^{(n)}(e)\simeq[n\log n]^{-D(G)/2}$. Let $\mathfrak{w},\mathfrak{F}$ be the two dimensional weight system and weight function system introduced above in Lemma 5.16. It follows from Theorems 2.10-A.22 and Corollary 2.19 that $F_S^{-1}(\|\cdot\|_{S,\mathfrak{F}})\simeq|\cdot|_S^2/\log|\cdot|_S$. The result follows. □

Appendix A. Approximate coordinate systems

This appendix contains the proofs of the key results stated in Sections 2.1–3, namely, Theorems 2.10–3.1. Throughout this section, G is a finitely generated nilpotent group equipped with a generating k-tuple $(s_1,\dots,s_k)$. Formal commutators refer to commutators on the alphabet $\{s_i^{\pm1}\colon 1\le i\le k\}$.

A.1. Proof of Theorem 3.1 and assorted results. Theorem 3.1 is one of the keys to the random walk upper bounds of Section 4. It can be understood as providing a volume lower bound for the volume of certain balls together with some additional "structural information" on the balls in question.

Fix a weight system $\mathfrak{w}$ and weight functions F_c as in Theorem 3.1. Let $G_h^{\mathfrak{w}}$ be the associated descending normal series in G. By construction, $G_h^{\mathfrak{w}}$ is normal in G and, for all p, q, j such that $\bar{w}_p + \bar{w}_q \geq \bar{w}_j$, we have (see Section 1.3)

$$[G_p^{\mathfrak{w}}, G_q^{\mathfrak{w}}] \subset G_j^{\mathfrak{w}}.$$

It follows that the commutators map

$$G_p^{\mathfrak{w}} \times G_q^{\mathfrak{w}} : (u, v) \longmapsto [u, v] \in G_j^{\mathfrak{w}}$$

induces a group homomorphism

$$G_p^{\mathfrak{w}}/G_{p+1}^{\mathfrak{w}} \otimes G_q^{\mathfrak{w}}/G_{q+1}^{\mathfrak{w}} \longrightarrow G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}.$$

This yields the following lemma.

Lemma A.1 (Similar to [3, Lemma 3]). *Let c be a formal commutator of weight $\bar{w}_j$ and let g_c be its image in G. There is an integer $\ell = \ell(c) \leq 8^j$ and a sequence $(i_1, \dots, i_\ell) \in \{1, \dots, k\}^\ell$ such that, for any $r \geq 1$ and $n \in \mathbb{Z}$ satisfying $|n| \leq F_c(r)$, we have*

$$g_c^n = s_{i_1}^{n_1} s_{i_2}^{n_2} \cdots s_{i_\ell}^{n_\ell} \mod G_{j+1}^{\mathfrak{w}}$$

for some $n_{i_j} \in \mathbb{Z}$ with $|n_j| \leq F_{s_{i_j}}(r)$.

Proof. The proof is by induction on j. For $j = 1$, c must have length 1 and $g_c^n = s_i^n$ for some $i \in \{1, \dots, k\}$. Assume the result holds true for all $h < j$ and let c be a commutator of weight $\bar{w}_j$. Either c has length 1 and the result is trivial or $c = [u, v]$ where u, v are commutators of weights $\bar{w}_p, \bar{w}_p, \bar{w}_p + \bar{w}_q = \bar{w}_j$. Since $F_c = F_u F_v$, for all $|n| \leq F_c(r)$ we can write $n = ab + d$ with $|a|, |d| \leq F_u(r)$, $0 \leq d \leq F_v(r)$. Then

$$g_c^n = [u, v]^{ab}[u, v]^d = [u^a, v^b][u^d, v] \mod G_{j+1}^{\mathfrak{w}}.$$

The desired result follows from the induction hypothesis. □

Definition A.2. Given c, $\ell = \ell(c)$ and $(i_1, \dots, i_\ell)$ as in Lemma A.1, for any $\mathbf{x} = (x_1, \dots, x_\ell) \in \mathbb{Z}^\ell$, set

$$\mathbf{g}_c(\mathbf{x}) = \mathbf{g}_c(x_1, \dots, x_\ell) = s_{i_1}^{x_1} s_{i_2}^{x_2} \cdots s_{i_\ell}^{x_\ell} \in G.$$

Set

$$F_j^c = F_{s_{i_j}} = F_{i_j}, \quad 1 \leq j \leq \ell.$$

By Lemma A.1, if $w(c) = \bar{w}_j$ and $|n| \le F_c(r)$ then

$$g_c^n = \mathbf{g}_c(\mathbf{n}(c)) \mod G_{j+1}^{\mathfrak{w}}$$

for some $\mathbf{n}(c) = (n_1(c), \ldots, n_\ell(c))$ with $|n_j(c)| \le F_{s_{i_j}}(r) = F_j^c(r)$.

Theorem A.3. *Let $c_1, \ldots c_t$ be a sequence of formal commutators with non-decreasing $\mathfrak{w}$-weights and such that, for each h, the image in $G_h^{\mathfrak{w}}/G_{h+1}^{\mathfrak{w}}$ of the family $\{c_i : w(c_i) = \bar{w}_h\}$ is a linearly independent family. Set*

$$K(r) = \left\{ g \in G : g = \prod_{i=1}^{t} \mathbf{g}_{c_i}(\mathbf{x}_i),\ \mathbf{x}_i = (x_1^i, \ldots, x_{\ell(c_i)}^i) \in \mathbb{Z}^{\ell(c_i)}, |x_j^i| \le F_j^{c_i}(r) \right\}.$$

Then

$$\# K(r) \ge \prod_1^t (2F_{c_i}(r) + 1) \ge \prod_{i=1}^{t} \prod_{j=1}^{\ell(c_i)} F_j^c(r).$$

Proof. For each $(y_i)_1^t \in \mathbb{Z}^t$ with $|y_i| \le F_{c_i}(r)$, let $\mathbf{y}_i = (y_j^i)_1^{\ell(c_i)}$, $1 \le i \le t$, be such that

$$g_{c_i}^{y_i} = \mathbf{g}_{c_i}(\mathbf{y}_i) \mod G_{j+1}^{\mathfrak{w}}, \quad w(c_i) = \bar{w}_j,\ 1 \le i \le t.$$

Such a $(\mathbf{y}^i)_1^t$ is given by Lemma A.1. Assume that two sequences $(y_i)_1^t$ and $(\tilde{y}_i)_1^t$ are such that $\prod_{i=1}^t \mathbf{g}_{c_i}(\mathbf{y}_i) = \prod_{i=1}^t \mathbf{g}_{c_i}(\tilde{\mathbf{y}}_i)$. Then by projecting on $G_1^{\mathfrak{w}}/G_2^{\mathfrak{w}}$ and using the assumed linear independence of the collection of the c_i's with $w(c_i) = \bar{w}_1$ in $G_1^{\mathfrak{w}}/G_2^{\mathfrak{w}}$ and the fact that $g_{c_i}^{y_i} = \mathbf{g}_{c_i}(\mathbf{y}^i)$ in $G_1^{\mathfrak{w}}/G_2^{\mathfrak{w}}$,we find that $y_i = \tilde{y}_i$ for those i with $w(c_i) = \bar{w}_1$. This implies that $\mathbf{y}_1 = \tilde{\mathbf{y}}_1$. Proceeding further up in the weight filtration shows that we must have $\mathbf{y}_i = \tilde{\mathbf{y}}_i$ for all $1 \le i \le t$. This shows that there are at least $\prod_1^t (2F_{c_i}(r) + 1)$ distinct elements in $K(r)$ which is the desired result. □

Theorem A.4. *Fix a weight system* $\mathfrak{w}$ *and weight functions* F_c *as in Theorem 3.1. Let* $b_1, \dots b_t$ *be a sequence of elements in* G. *Assume the following hypotheses.*

(1) *For each* $i = 1, \dots, t$, *there exists an integer* $h(i)$ *such that* $b_i \in G^{\mathfrak{w}}_{h(i)}$ *and* b_i *is torsion free in* $G^{\mathfrak{w}}_{h(i)}/G^{\mathfrak{w}}_{h(i)+1}$. *Further, for each* h, *the system* $\{b_i : h(i) = h\}$ *is free in* $G^{\mathfrak{w}}_{h(i)}/G^{\mathfrak{w}}_{h(i)+1}$.

(2) *For each* $i = 1, \dots, t$, *there exists and increasing function* $\widetilde{F}^i$, *a positive integer* $\ell(i)$ *and a sequence* $j^i_1, \dots, j^i_{\ell(i)}$ *such that, for any* $r > 0$ *and any integer* n *with* $|n| \leq \widetilde{F}^i(r)$, *there exists* $\mathbf{n}^i = (n^i_1, \dots, n^i_{\ell(i)})$ *with* $|n^i_q| \leq F_{j^i_q}(r)$ *satisfying*

$$b_i^n = \prod_{q=1}^{\ell(i)} s_{j^i_q}^{n^i_q} \mod G^{\mathfrak{w}}_{h(i)+1}.$$

For $\mathbf{x} = (x_1, \dots, x_{\ell(i)}) \in \mathbb{Z}^{\ell(i)}$, *set*

$$\mathbf{b}_i(\mathbf{x}) = \prod_{q=1}^{\ell(i)} s_{j^i_q}^{x_q} \in G$$

and

$$K(r) = \left\{ g \in G : g = \prod_{i=1}^{t} \mathbf{b}_i(\mathbf{x}_i),\ \mathbf{x}_i = (x^i_1, \dots, x^i_{\ell(i)}) \in \mathbb{Z}^{\ell(i)}, |x^i_q| \leq F_{j^i_q}(r) \right\}.$$

Then

$$\#K(r) \geq \prod_{1}^{t} (2\widetilde{F}_i(r) + 1).$$

Proof. This a straightforward generalization of Theorem A.3. Instead of considering commutators and their natural weight function F_c, we consider arbitrary group elements b with associated weight function $\widetilde{F}$ with the property that b is free in $G^{\mathfrak{w}}_h/G^{\mathfrak{w}}_{h+1}$, for some u, h, and b^n, $|n| \leq \widetilde{F}(r)$, can be express modulo $G^{\mathfrak{w}}_{h+1}$ as a fixed product of powers of generators with properly controlled exponents. The proof is essentially the same as that of Theorem A.3. Namely, for each $(y_i)_1^t \in \mathbb{Z}^t$ with $|y_i| \leq \widetilde{F}^i(r)$, let $\mathbf{y}_i = (y^i_j)_1^{\ell(i)}$, $1 \leq i \leq t$, be such that

$$b_i^{u_i y_i} = \mathbf{b}_i(\mathbf{y}_i) \mod G^{\mathfrak{w}}_{h(i)+1}, \quad 1 \leq i \leq t.$$

Such a $(\mathbf{y}^i)_1^t$ exists by hypothesis. Assume that two sequences $(y_i)_1^t$ and $(\tilde{y}_i)_1^t$ are such that $\prod_{i=1}^t \mathbf{b}^i(\mathbf{y}_i) = \prod_{i=1}^t \mathbf{b}^i(\tilde{\mathbf{y}}_i)$. Then by projecting on $G_1^{\mathfrak{w}}/G_2^{\mathfrak{w}}$ and using the assumed freeness of the collection of the b_i's with $h(i) = 1$ in $G_1^{\mathfrak{w}}/G_2^{\mathfrak{w}}$ and the fact that $b_i^{u_i y_i} = \mathbf{b}^i(\mathbf{y}^i)$ in $G_1^{\mathfrak{w}}/G_2^{\mathfrak{w}}$,we find that $y_i = \tilde{y}_i$ for those i with $h(i) = 1$. This implies $\mathbf{y}_1 = \tilde{\mathbf{y}}_1$. Proceeding further up in the weight filtration shows that we must have $y_i = \tilde{y}_i$ for all $1 \leq i \leq t$. This shows that there are at least $\prod_1^t(2\widetilde{F}^i(r) + 1)$ distinct elements in $K(r)$, as desired. □

Remark A.5. Theorem A.4 allows for much more freedom than Theorem A.3. This freedom is used in the proof of Theorem 3.4.

A.2. Commutator collection on free nilpotent groups. We prove the following weak version of Theorem 2.10.

Theorem A.6. *Referring to the setting and notation of Theorem 2.10, assume that (2.1) and (2.2) hold true. Then there exist an integer $t = t(G, S, \mathfrak{w})$, a constant $C = C(G, S, \mathfrak{w}) \geq 1$, and a sequence Σ of commutators (depending on $G, S, \mathfrak{w}$)*

$$c_1, \dots, c_t \text{ with non-decreasing weights } w(c_1) \preceq \cdots \preceq w(c_t)$$

such that

(i) *For any $r > 0$, if $g \in G$ can be expressed as a word ω over $\mathfrak{C}(S)^{\pm 1}$ with $\deg_c(\omega) \leq F_c(r)$ for all $c \in \mathfrak{C}(S)$ then g can be expressed in the form*

$$g = \prod_{i=1}^t c_i^{x_i} \text{ with } |x_i| \leq F_{c_i}(Cr) \quad \text{for all } i \in \{1, \dots, t\}.$$

(ii) *There exist an integer*

$$p = p(G, S, \mathfrak{w})$$

and a p-tuple $(i_j)_1^p \in \{1, \dots, k\}^p$ (also depending on $(G, S, \mathfrak{w})$ such that, if g can be expressed as a word ω over $\{c_i^{\pm 1} : 1 \leq i \leq t\}$ with $\deg_{c_i}(\omega) \leq F_{c_i}(r)$ for some $r > 0$ then g can be expressed in the form

$$g = \prod_{j=1}^p s_{i_j}^{x_j} \quad \text{with } |x_j| \leq F_{i_j}(Cr).$$

Remark A.7. It must be the case that, for any j, the image of $\{c_i : w(c_i) = \bar{w}_j\}$ in $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$ generates $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$. The key difference with Theorem 2.10 is that Theorem A.6 does not identify a maximal subset of $\{c_i : w(c_i) = \bar{w}_j\}$ that is free in $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$.

The proof of Theorem A.6 requires a number of steps. The first observation is that it is enough to prove Theorem A.6 in the case of the free nilpotent group $N(k,\ell)$ on k generators $s_1,\dots,s_k$ and of nilpotency class ℓ. Indeed, once Theorem A.6 is proved on $N(k,\ell)$, the same statement holds on any nilpotent G of nilpotency class ℓ equipped with a generating k-tuple S via the canonical projection from $N(k,\ell)$ to G (by definition, the canonical projection is the group homeomorphism from $N(k,\ell)$ onto G which sends the canonical k generators of $N(k,\ell)$ to the given k generators of G).

Notation A.8. For the rest of this section, we assume that $G = N(k,\ell)$ is the free nilpotent group $N(k,\ell)$ equipped with its canonical generating set $S = (s_1,\dots,s_k)$ and the multidimensional weight-system $\mathfrak{w}$ generated by the $(w_1,\dots,w_k)$. Without loss of generality, we assume that the commutator set $\mathfrak{C}(S)$ is equipped with a total order $\prec$ such that the function

$$w\colon \mathfrak{C}(S) \ni c \longmapsto w(c) \in (0,\infty)\times\mathbb{R}^{d-1}$$

associated with the given weight system $\mathfrak{w}$ is non-decreasing. Hence, $c \prec c'$ implies $w(c) \preceq w(c')$. In addition, we let $\mathfrak{F}$ be a weight function system that is compatible with $\mathfrak{w}$ in the sense that (2.1) and (2.2) hold true.

Notation A.9. Recall that $\deg_c(\omega)$ denotes the number of occurrences of $c^{\pm 1}$ in the word ω over $\mathfrak{C}(S)$. Similarly, we define $\deg_c^*(\omega)$ to be the number of occurrences of c minus the number of occurrences of c^{-1} in a word over $\mathfrak{C}(S)$.

On $\mathfrak{C}(S)$, consider the map J such that $J(s_i^{\pm 1}) = s_i^{\mp 1}$ and $J([a,b]) = [b,a]$. Abusing notation, we also write $J(c) = c^{-1}$. Note that J^2 is the identity. Restrict J to $\mathfrak{C}^*(S) = \{c\colon J(c) \neq c\}$ (where $J(c) = c$ is understood as equality as formal commutator so that $J(s_i) \neq s_i$ and $J([a,b]) = [a,b]$ if and only if $a = b$). Let $\mathfrak{C}^*_+$ be the set of representative of $\mathfrak{C}^*(S)/J$ given by $c \in \mathfrak{C}^*_+(S)$ if and only if $c = s_i$ or $c = [a,b]$ with $a \succ b$.

It is convenient to enumerate all formal commutators in $\mathfrak{C}^*_+(S,\ell)$ and write

$$\mathfrak{C}^*_+(S,\ell) = \{c_1,\dots,c_t\}, \quad t = \#\mathfrak{C}^*_+(S,\ell).$$

Since ℓ is fixed throughout, we write

$$\mathfrak{C}^*_+(S) = \mathfrak{C}^*_+(S,\ell).$$

Note that, a priori, this list contains commutators that are trivial in $N(k,\ell)$. This does not matter although these formal commutators can be omitted if desired. Let us describe the basic collecting process on $N_{k,\ell}$.

Commutator collecting algorithm

- Given a word $\omega = c_{i_1}^{\epsilon_{i_1}} c_{i_2}^{\epsilon_{i_2}} ... c_{i_m}^{\epsilon_{i_m}}$ in $\mathfrak{C}_+^*(S) \cup \mathfrak{C}_+^*(S)^{-1}$, first identify the commutator of lowest order with respect to $\prec$, say it is commutator c_{i_j}, mark all the contributions of c_{i_j} to ω from left to right in order: $\{y_1, ..., y_q\}$, $y_j \in \{c_{i_j}^{\pm 1}\}$.
- Starting with y_1, move $y_1, ..., y_q$ to the left one by one by successive commutation. Note that every time c_{i_j} jumps backward over a commutator c, the jump produces the sequence $...c_{i_j} c [c, c_{i_j}]...$. It follows that all commutators that are created in this process belong to $\mathfrak{C}_+^*(S)$ and have weight $\succeq 2w(c_{i_j}) \succ w(c_{i_j})$.
- After $y_1, ..., y_q$ have been moved to the left, we obtain a word $y_1 ... y_q \omega'$ with the same image as ω, and where ω' is a word in commutators $\succ c_{i_j}$.
- Apply the previous steps to ω', producing ω'' and continue until the process terminates after at most $\#\mathfrak{C}_+^*(S)$ steps.

This proves the following weak version of M. Hall basis theorem [11, Theorem 11.2.3] (in Hall's more sophisticated version, only the so called "basic" commutators are used and this results in a unique representation of any element of $N(k, \ell)$).

Proposition A.10. *Any element $g \in N(k, \ell)$ has a representation*

$$g = c_1^{x_1} c_2^{x_2} ... c_t^{x_t}, \quad x_i \in \mathbb{Z}.$$

Next we want to have some control over $\{x_i,\ 1 \le i \le t\}$. Let's start with a simple binomial counting lemma adapted from [11, p. 173] and [26]. We will use the following notation. For any two commutators $c_j \succ c_i$, let $C_{n-1}(i, j)$ be the sets of all commutators $c \in \mathfrak{C}_+^*(S)$ such that there exist $\epsilon_0, \ldots, \epsilon_n \in \{-1, 1\}$ such that $c_j^{\epsilon_n} = [\cdots[c^{\epsilon_0}, c_i^{\epsilon_1}], \ldots, c_i^{\epsilon_{n-1}}]$ (as formal commutators in $\mathfrak{C}(S)$).

Lemma A.11. *Consider a word ω in $\{c_j : c_j \succeq c_i\}^{\pm 1}$. Let $m = \deg_{c_i} \omega$, and let $\{y_1, ..., y_m\}$, $y_j \in \{c_i^{\pm 1}\}$, be the left to right contribution of c_i to ω. For $0 \le q \le m$, there is a word ω_q in $\{c_j : c_j \succeq c_i\}^{\pm 1}$ which starts with $y_1 ... y_q$, whose left to right contribution of $c_i^{\pm 1}$ is $y_1, \ldots, y_m$, and in which, for all $c_j \succ c_i$,*

$$\deg_{c_j}(\omega_q) \le \deg_{c_j}(\omega) + q \sum_{c \in C_1(i,j)} \deg_c(\omega) + \binom{q}{2} \sum_{c \in C_2(i,j)} \deg_c(\omega)$$
$$+ \cdots + \binom{q}{\ell} \sum_{c \in C_\ell(i,j)} \deg_c(\omega)$$

Further, if c' denotes the lowest commutator in ω with $c' \succ c_i$ then contributions of commutators c with $w(c) \prec w(c') + w(c_i)$ remain unchanged in ω_q.

Remark A.12. Note that, after we move all contributions of c_i to ω to the left, we obtain a word ω_m with same image as ω of the form

$$\omega_m = c_i^x \omega_m'$$

where $x = \deg^*_{c_i}(\omega)$, ω_m' is a word in $[\mathfrak{C}^*_+(S) \cap \{c \succ c_i\}]^{\pm 1}$, and in which the contributions of commutators c with $w(c) \prec w(c') + w(c_i)$ remain the same than in ω.

Proof. The proof is by induction on q. It holds trivially for $q = 0$. The induction hypothesis gives us a word ω_{q-1} with

$$\deg_{c_j}(\omega_{q-1}) \le \deg_{c_j}(\omega) + (q-1) \sum_{c \in C_1(i.j)} \deg_c(\omega) + \binom{q-1}{2} \sum_{c \in C_2(i,j)} \deg_c(\omega)$$
$$+ \cdots + \binom{q-1}{\ell} \sum_{c \in C_\ell(i,j)} \deg_c(\omega).$$

Now, we move y_q to the left as in the collecting process by successive commutations. To keep track of contribution of c_j, notice that a new contribution of c_j is produced only if y_q jumps over a commutator $c^{\pm 1}$ such that

$$[c^{\pm 1}, y_q] = c_j^{\pm 1}.$$

Further,

$$w([c^{\pm 1}, y_q]) = w(c) + w(c_i) \succeq w(c') + w(c_i).$$

Hence, c_j must satisfies

$$w(c_j) \succeq w(c') + w(c_i).$$

Therefore we eventually get a word ω_q in $[\mathfrak{C}^*_+(S) \cap \{c \succeq c_i\}]^{\pm 1}$ with $\pi(\omega_q) = \pi(\omega)$, in which the left to right contribution of c_i is the same as in ω, which starts with $y_1 ... y_q$, and such that

$$\deg_{c_j}(\omega_q) \le \deg_{c_j}(\omega_{q-1}) + \sum_{c \in C_1(i,j)} \deg_c(\omega_{q-1}).$$

Using the induction hypothesis on ω_{q-1} and the fact that all brackets of length at least $\ell + 1$ drop out,

$$\sum_{c \in C_1(i,j)} \deg_c(\omega_{q-1}) = \sum_{c = c_\alpha \in C_2(i,j)} \sum_{p=0}^{\ell} \binom{q-1}{p} \sum_{\tilde{c} \in C_p(i,\alpha)} \deg_{\tilde{c}}(\omega)$$
$$\le \sum_{p=1}^{\ell} \binom{q-1}{p-1} \sum_{\tilde{c} \in C_p(i,j)} \deg_{\tilde{c}}(\omega).$$

Hence, we have

$$\begin{aligned}\deg_{c_j}(\omega_q) &\le \deg_{c_j}(\omega_{q-1}) + \sum_{c\in C_2(i,j)} \deg_c(\omega_{q-1})\\ &\le \sum_{p=0}^{\ell}\left(\binom{q-1}{p}+\binom{q-1}{p-1}\right)\sum_{\tilde{c}\in C_p(i,j)} \deg_{\tilde{c}}(\omega)\\ &= \sum_{p=0}^{\ell}\binom{q}{p}\sum_{\tilde{c}\in C_p(i,j)} \deg_{\tilde{c}}(\omega).\end{aligned}$$ □

Lemma A.13. *There exists a constant $C > 0$ such that for any word ω in $[\mathfrak{C}^*_+(S)\cap\{c \succeq c_i\}]^{\pm 1}$ with $\deg_c \omega \le F_c(d)$ for all $c \succeq c_i$, there exists a word ω' in $[\mathfrak{C}^*_+(S)\cap\{c \succeq c_i\}]^{\pm 1}$ in collected form:*

$$\omega' = \prod_{j=i}^{t} c_j^{x_j}$$

such that $\pi(\omega') = \pi(\omega)$, $x_j = \deg^*_{c_j}\omega$ for those j such that $w(c_j) \prec 2w(c_i)$ and $|x_j| \le F_{c_j}(Cd)$ for all $i \le j \le t$.

Proof. The proof is by backward induction on i. For $i = t$, the statement holds trivially since commutators with $c \succeq c_t$ commute.

Suppose the assertion holds for $i+1$. As in the lemma, consider a word ω on $[\mathfrak{C}^*_+(S)\cap\{c \succeq c_i\}]^{\pm 1}$. Let $\{y_1, ..., y_q\}$ be the contribution of c_i to ω, $q = \deg_{c_i}\omega$. The previous lemma yields

$$\omega_q = y_1 ... y_q \omega'_q,$$

where ω'_q is a word in $[\mathfrak{C}^*_+(S)\cap\{c \succeq c_{i+1}\}]^{\pm 1}$. From the hypothesis on the degrees of ω,

$$\deg_{c_j}(\omega_k) \le \sum_{p=0}^{\ell}\binom{k}{p}\sum_{c\in C_p(i,j)} F_c(d)$$

By definition, if $c \in C_p(i,j)$ then

$$F_c F^p_{c_i} = F_{c_j}.$$

Further,

$$\#C_p(i,j) \le t = \#\mathfrak{C}_+^*(S)$$

and

$$q = \deg_{c_i} \omega \le F_{c_i}(d).$$

Therefore, we obtain

$$\begin{aligned}\deg_{c_j}(\omega_q) &\le t F_{c_j}(d)\Big(\sum_{p=0}^{\ell}\binom{q}{p} F_{c_i}(d)^{-p}\Big)\\ &\le t F_{c_j}(d)\Big(\sum_{p=0}^{\ell} q^p F_{c_i}(d)^{-p}\Big)\\ &\le t(1+\ell)F_{c_j}(d).\end{aligned}$$

By assumption (2.1), there exists a constant C_1 such that

$$t(1+\ell)F_c(d) \le F_c(C_1 d)$$

for all c and $d \ge 1$. □

Lemma A.13 with $i = 1$ proves Theorem A.6(i). Next we work on improving Theorem A.6(i) in the special case of the free nilpotent group $N(k,\ell)$. This improvement will be instrumental in proving Theorem A.6(ii). It is based on the following important Lemma.

Lemma A.14. *For each j, $N(k,\ell)^{\mathfrak{w}}_j / N(k,\ell)^{\mathfrak{w}}_{j+1}$ is a finitely generated free abelian group.*

Proof. The proof is by a backward induction on ℓ. If $\ell = 1$, $N(k,1)$ is the free abelian group on k generators and the desired result holds by inspection. Let $g \in N(k,\ell)^{\mathfrak{w}}_j$ such that $g \notin N(k,\ell)^{\mathfrak{w}}_{j+1}$. Let $N_\ell = N(k,\ell)_\ell$ be the center of $N(k,\ell)$ (i.e., the subgroup generated by commutators of length ℓ). Assume first that $g \in N(k,\ell)^{\mathfrak{w}}_{j+1} N_\ell$. Since

$$N(k,\ell)^{\mathfrak{w}}_{j+1} N_\ell / N(k,\ell)^{\mathfrak{w}}_{j+1} \simeq N_\ell / [N(k,\ell)^{\mathfrak{w}}_{j+1} \cap N_\ell],$$

and $N(k,\ell)^{\mathfrak{w}}_{j+1} \cap N_\ell$ is generated by the basic commutators of weight $\bar{w}_j$ and length ℓ, $N_\ell / [N(k,\ell)^{\mathfrak{w}}_{j+1} \cap N_\ell]$ is torsion free. It thus follows that g is not torsion in $N(k,\ell)^{\mathfrak{w}}_j / N(k,\ell)^{\mathfrak{w}}_{j+1}$.

Now, consider the case when $g \notin N(k,\ell)_j^{\mathfrak{w}} N_\ell$. Let g' be the projection of g in $N(K,\ell)/N_\ell = N(k,\ell-1)$. Clearly $g' \in N(k,\ell-1)_j^{\mathfrak{w}}$ and $g' \notin N(k,\ell-1)_{j+1}^{\mathfrak{w}}$ because the inverse image of $N(k,\ell-1)_{j+1}^{\mathfrak{w}}$ under this projection is $N(k,\ell)_{j+1}^{\mathfrak{w}} N_\ell$. Further,

$$N(k,\ell)_j^{\mathfrak{w}} N_\ell / N(k,\ell)_{j+1}^{\mathfrak{w}} N_\ell \simeq N(k,\ell-1)_j^{\mathfrak{w}} / N(k,\ell-1)_{j+1}^{\mathfrak{w}}.$$

By the induction hypothesis, g' is not torsion in $N(k,\ell-1)_j^{\mathfrak{w}}/N(k,\ell-1)_{j+1}^{\mathfrak{w}}$. It follows that g is not torsion in $N(k,\ell)_j^{\mathfrak{w}}/N(k,\ell)_{j+1}^{\mathfrak{w}}$. □

Next, let $(b_i)_1^\tau$ be a sequence of elements of $\mathfrak{C}_+^*(S)$ such that $\{b_i : w(b_i) = \bar{w}_j\}$ projects to a basis of $N(k,\ell)_j^{\mathfrak{w}}/N(k,\ell)_{j+1}^{\mathfrak{w}}$. Let $R_j^{\mathfrak{w}}$ be the rank of this torsion free abelian group and set $m'_j = \sum_1^j R_i^{\mathfrak{w}}$ so that $\tau = m'_{j_*}$. Set also $m_j = \max\{i : w(c_i) = \bar{w}_j\}$. Without loss of generality, we can assume that our ordering on $\mathfrak{C}_+^*(S)$ is such that

$$(b_i)_{m'_{j-1}+1}^{m'_j} = (c_j)_{m_{j-1}+1}^{m_{j-1}+R_j^{\mathfrak{w}}}.$$

Lemma A.15. *Referring o the above setup and notation, there exists a constant $C > 0$ such that for any word ω in $\{c_i : w(c_i) \succeq \bar{w}_h\}^{\pm 1}$ with $\deg_{c_j} \omega \le F_{c_j}(d)$ for all j, there is a word ω_h*

$$\omega_h = \prod_{j=m'_{h-1}+1}^{\tau} b_j^{x_j}$$

such that

$$\pi(\omega_h) = \pi(\omega)$$

and

$$|x_j| \le CF_{c_j}(Cd), \quad m'_{h-1}+1 \le j \le m'_h.$$

Proof. The proof is by backward induction on h. When $h = j_*$, $N(k,\ell)_{j_*}^{\mathfrak{w}}$ is abelian and this is just linear algebra.

For a word ω as in the lemma, Lemma A.13 gives a word

$$\omega' = \prod_{i \ge m_{h-1}+1} c_i^{x_i}, \quad |x_i| \le F_{c_i}(Cd)$$

with the same image as ω. Set

$$I_1(h) = \{m_{h-1}+1, \dots, m_{h-1}+R_h^{\mathfrak{w}}\},$$

$$I_2(h) = \{m_{h-1}+R_h^{\mathfrak{w}}+1, \dots, m_h\}.$$

For $i \in I_2(h)$, c_i has the same image than

$$\prod_{j\in I_1(h)} c_j^{z_{j,i}} v_i$$

with v_i a word in $\{c_p : w(c_p) \succeq \bar{w}_{h+1}\}^{\pm 1}$. Hence

$$\omega'' = \prod_{j\in I_1(h)} c_j^{x_j} \prod_{i\in I_2(h)} \Big(\prod_{j\in I_1(h)} c_j^{z_{i,j}} v_i \Big)^{x_i} \prod_{p>m_h} c_p^{x_p}$$

has the same image than ω. Applying Lemma A.13 to this word ω'' gives

$$\omega'_h = \prod_{j\in I_1(h)} c_j^{x_j + \sum_{i\in I_2 h} z_{i,j} x_i} \prod_{p>m_h} c_p^{x'_p}$$

with the same image than ω'' and $|x'_p| \le F_{c_p}(Cd)$ for $p > m_h$. Further, since $F_{c_i} \simeq F_{c_j} \simeq \mathbf{F}_h$, for $i \in I_1(h)$, $j \in I_2(h)$, we have

$$|x_j + \sum_{i\in I_2(h)} z_{i,j} x_i| \le F_{c_j}(Cd).$$

Applying the induction hypothesis to rewrite $\prod_{p>m_h} c_p^{x'_p}$ finishes the proof. □

Theorem A.16. *Assume that the free nilpotent group $N(k,\ell)$ is equipped with its canonical generating k-tuple $S = (s_1, \dots, s_k)$ and a weight system $\mathfrak{w}$ and weight-function system $\mathfrak{F}$ such that (2.1) and (2.2) hold true. Let b_i, $1 \le i \le \tau$, be a sequence of elements of $C_+^*(S)$ with $w(b_i) \preceq w(b_{i+1})$, $1 \le i \le \tau - 1$ and such that, for each j, $\{b_i : w(b_i) = \bar{w}_j\}$ is a basis of the free abelian group $N(k,\ell)_j^{\mathfrak{w}} / N(k,\ell)_{j+1}^{\mathfrak{w}}$. Then*

(i) *Any element $g \in N(k,\ell)$ can be expressed uniquely in the form*

$$g = \prod_{i=1}^{\tau} b_i^{x_i}, \quad x_i \in \mathbb{Z}, i \in \{1, \dots, \tau\}.$$

Further,

$$F_S^{-1}(\|g\|_{\mathfrak{C}(S),\mathfrak{F}}) \simeq \max_{1\le i\le \tau} \{F_{b_i}^{-1}(|x_i|)\}.$$

(ii) *There exist an integer p and $(i_j)_1^p \in \{1, \dots, k\}^p$ such that any $g \in N(k,\ell)$ with $\|g\|_{\mathfrak{C}(S),\mathfrak{F}} \le F_S(r)$, $r > 0$, can be expressed in the form*

$$g = \prod_{j=1}^{p} s_{i_j}^{y_j} \text{ with } |y_j| \le F_{i_j}(Cr),\ j \in \{1, \dots, p\}.$$

Remark A.17. This result is a strong version of Theorem 2.10 in the special case when $G = N(k,\ell)$.

Proof of (i). The first assertion follows from Lemma A.15. Uniqueness is clear if one considers the projections of g onto the successive free abelian groups $N(k,\ell)^{\mathfrak{w}}_j / N(k,\ell)^{\mathfrak{w}}_{j+1}$. □

The proof of the the second assertion requires some preparation. Given a commutator c with length $m \le \ell$, let $\sigma = \sigma_1 ... \sigma_m$ be the formal word on the alphabet S obtained from c by removing brackets and inverses. For $\vec{a} = (a_1, ..., a_\ell) \in \mathbb{Z}^\ell$, $\Theta(\vec{a}, c)$ is defined as the expression we get by substituting in c each σ_i by $\sigma_i^{a_i}$, while keeping all the brackets and signs unchanged. For example, if $c = [[s_{i_1}, s_{i_2}^{-1}], s_{i_3}^{-1}]$, and $\vec{a} = (a_1, a_2, a_3, 0, ..., 0)$, we have

$$\Theta(\vec{a}, c) = [[s_{i_1}^{a_1}, s_{i_2}^{-a_2}], s_{i_3}^{-a_3}].$$

Lemma A.18. *For a commutator c with length $m \le \ell$, let $\sigma = \sigma_1 ... \sigma_m$ be the formal word associated with it. Suppose $a_1, ..., a_m \in \mathbb{Z}$ are such that $|a_j| \le F_{\sigma_j}(d)$ for all $1 \le j \le m$, $d > 0$. Set $\vec{a} = (a_1, ..., a_m, 0, ..., 0) \in \mathbb{Z}^\ell$ and consider the element $u \in N(k,\ell)$ such that*

$$uc^{a_1 ... a_k} = \Theta(\vec{a}, c).$$

Then u can be represented by a word ω on $\{c_j : w(c_j) \succ w(c)\}^{\pm 1}$ with

$$\deg_{c_j}(\omega) \le F_{c_j}(Cd) \quad \textit{for all } c_j \textit{ with } w(c_j) \succ w(c).$$

Proof. The proof is by induction on the length m of the commutator c. When $m = 1$, the statement is trivial.

Suppose the statement is true for commutators of length $\le m-1$. Let c be a commutator with length m, say $c = [f_1, f_2]$, where f_1, f_2 are commutators of length $m_1, m_2 < m$. Write $\vec{a}_1 = (a_1, \dots, a_{m_1}, 0, \dots, 0)$ and $\vec{a}_2 = (a_{m_1+1}, \dots, a_{m_1+m_2}, 0, \dots, 0)$, then by definition

$$\Theta(\vec{a}, c) = [\Theta(\vec{a}_1, f_1), \Theta(\vec{a}_2, f_2)].$$

By the induction hypothesis,

$$\Theta(\vec{a}_1, f_1) = u_1 f_1^{a_1 ... a_{m_1}}, \quad \Theta(\vec{a}_2, f_2) = u_2 f_2^{a_{m_1+1} ... a_{m_1+m_2}}$$

where u_1 can be represented by a word ω_1 in commutators c_p with $w(c_p) \succ w(f_1)$ and $\deg_{c_p}(\omega) \le F_{c_p}(Cd)$. Similarly, u_2 can be represented by a word ω_2 in commutators c_p with $w(c_p) \succ w(f_2)$ and $\deg_{c_p}(\omega) \le F_{c_p}(Cd)$.

Suppose $w(f_1) = \bar{w}_{h_1}$, $w(f_2) = \bar{w}_{h_2}$, and $w([f_1, f_2]) = \bar{w}_h$. By the natural group homomorphism

$$N^{\mathfrak{w}}_{h_1}/N^{\mathfrak{w}}_{h_1+1} \otimes N^{\mathfrak{w}}_{h_2}/N^{\mathfrak{w}}_{h_2+1} \longrightarrow N^{\mathfrak{w}}_{h}/N^{\mathfrak{w}}_{h+1},$$

we have that

$$\begin{aligned}[\Theta(\vec{a}_1, f_1), \Theta(\vec{a}_2, f_2)] &\equiv [f_1^{a_1 \dots a_{m_1}}, f_2^{a_{m_1+1} \dots a_{m_1+m_2}}] \mod N^{\mathfrak{w}}_{h+1} \\ &\equiv [f_1, f_2]^{a_1 \dots a_{m_1+m_2}} \mod N^{\mathfrak{w}}_{h+1} \\ &\equiv c^{a_1 \dots a_m} \mod N^{\mathfrak{w}}_{h+1}.\end{aligned}$$

Therefore $u = \Theta(\vec{a}, c)c^{-a_1 \dots a_m} \in N^{\mathfrak{w}}_{h+1}$, and since

$$u = [u_1 f_1^{a_1 \dots a_{k_1}}, u_2 f_2^{a_{k_1+1} \dots a_{k_1+k_2}}]c^{-a_1 \dots a_k},$$

it can be represented by a word ω such that $\deg_{c_i} \omega \le 5F_{c_i}(Cd)$ for all i. Then by Theorem A.16(i), we have

$$u = \prod_{j : w(b_j) \succeq \bar{w}_h} b_j^{x_j}.$$

with $|x_j| \le F_{b_j}(C'd)$. ∎

Lemma A.19. *For any h, there exist constants $M_h > 0$ and $C_h > 0$ such that, for any $c \in \mathfrak{C}^*_+(S)$ with $w(c) \succeq \bar{w}_h$, there a integer $p = p(c)$ with $0 \le p \le M_h$ and a p-tuple $(i_1, \dots, i_p) \in \{1, \dots, k\}^p$, such that for any $x \in \mathbb{Z}$ with $|x| \le F_c(d)$, $d > 0$, we have*

$$c^x = s_{i_1}^{x_1} s_{i_2}^{x_2} \dots s_{i_p}^{x_p} \text{ with } x_j \in \mathbb{Z}, \quad |x_j| \le F_{i_j}(Cd), \ j = 1, \dots, p.$$

Proof. The proof is by backward induction on h. When $h = j_*$ and c is a commutator with $w(c) = \bar{w}_{j^*}$, let $\sigma = \sigma_1 \dots \sigma_m$, $\sigma_i \in \{s_1, \dots, s_k\}$ be the formal word associated with c (by forgetting brackets and inverses). Write

$$x = a_0 \prod_{1 \le j \le m} \lfloor F_{\sigma_j}(d) \rfloor + a_1 \prod_{2 \le j \le m} \lfloor F_{\sigma_j}(d) \rfloor + \dots + a_{m-1} \lfloor F_{\sigma_m}(d) \rfloor + a_m$$

with $a_j \in \mathbb{Z}$, $|a_0| \le C$ and $\left|a_j\right| \le F_{\sigma_j}(d)$. Write

$$\vec{a}_0 = (a_0 \lfloor F_{\sigma_1}(d) \rfloor, \lfloor F_{\sigma_2}(d) \rfloor ..., \lfloor F_{\sigma_m}(d) \rfloor),$$

$$\vec{a}_j = (\underbrace{1, ..., 1}_{j-1}, a_j, \lfloor F_{\sigma_{j+1}}(d) \rfloor, ..., \lfloor F_{\sigma_m}(d) \rfloor),$$

then

$$c^x \equiv \Theta(\vec{a}_1, c)...\Theta(\vec{a}_k, c) \mod N(k,\ell)^{\mathfrak{w}}_{j*+1}.$$

Since $N(k,\ell)^{\mathfrak{w}}_{j*+1} = \{e\}$, we actually have equality. Unraveling the brackets in $\Theta(\vec{a}_j, c)$ we get an expression in the powers of the generators satisfying the desired conditions.

Suppose the claim holds for $h+1$. Given a commutator c with $w(c) = \bar{w}_h$, let again $\sigma_1, \ldots \sigma_m$ (m depends on c) be the formal word on the generators associated with c. For $x \in \mathbb{Z}$, $|x| \le F_c(d)$, decompose x as above and use Lemma A.18 to write

$$c^x = u_0^{-1}\Theta(\vec{a}_0, c) \ldots u_m^{-1}\Theta(\vec{a}_m, c),$$

where $u_i \in N(k,\ell)^{\mathfrak{w}}_{h+1}$ can be represented by a word ω_i with $\deg_{c_j} \upsilon_i \le F_{c_j}(Cd)$ for all j. By Lemma A.15, u_i can also be represented in the form $\prod_{j \ge h+1} b_j^{y_{i,j}}$ with $|y_{i,j}| \le F_{b_j}(Cd)$. Applying the induction hypothesis to each terms of these products we can now write c^x in the desired form $c^x = s_{i_1}^{x_1} s_{i_2}^{x_2} ... s_{i_p}^{x_p}$. □

Proof of Assertion (ii) in Theorem A.16. By Theorem A.16(i), any $g \in N(k,\ell)$ with $\|g\|_{S,\mathfrak{F}} \le F_S^{-1}(r)$, $r > 0$, as a unique representation of the form $g = \prod_1^\tau b_j^{x_j}$ with $|x_j| \le F_{b_j}(Cr)$. Applying Lemma A.19 with $c = b_j, x = x_j$ for each $j = 1, \ldots, \tau$ produces a sequence $((i_n)_1^p$ (independent of g) and a sequence $(x'_n) \in \mathbb{Z}^p$ (depending on g) with $|x'_n| \le F_{s_{i_n}}(Cr)$ for all $n \in \{1, \ldots, p\}$ and such that

$$g = \prod_1^p s_{i_n}^{x'_n}.$$ □

A.3. End of the proof of Theorem 2.10. In order to finish the proof of Theorem 2.10 for a general finitely generated nilpotent group G, we simply need to improve upon Theorem A.6(i). Namely, Theorem A.6(i) provide a decomposition of any element g with $\|f\|_{\mathfrak{C}(S),\mathfrak{F}} \le F_S(r)$ in the form

$$g = \prod_1^t c_i^{x_i}, \quad |x_i| \le F_{c_i}(Cr).$$

Here $(c_i)_1^t$ is an enumeration of $\mathfrak{C}_+^*(S)$ so that $w(c_i) \preceq w(c_{i+1})$.

Now, let $(b_i)_1^\tau$ be a collection of formal commutators with $w(b_i) \preceq w(b_{i+1})$. For $j \in \{1, \dots, j_*\}$, let

$$m_j = \max\{i : w(b_i) = \bar{w}_j\}.$$

Clearly, $w(b_i) = \bar{w}_j$ if and only if $m_{j-1}+1 \le i \le m_j$. Recall that $R_j^{\mathfrak{w}}$ is the torsion free rank of the abelian group $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$. We make two natural assumptions on the sequence (b_i):

(A1) For each j, $\{b_i' : m_{j-1} < i \le m_j\}$ generates $G_j^{\mathfrak{w}}$ modulo $G_{j+1}^{\mathfrak{w}}$.

(A2) For each j, $\{b_i' : m_{j-1} < i \le m_{j-1} + R_j^{\mathfrak{w}}\}$ is free in $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$.

Note that, since $R_j^{\mathfrak{w}}$ is the torsion free rank of $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$, (A2) implies that (the image of) $\{b_i' : m_{j-1} < i \le m_{j-1} + R_j^{\mathfrak{w}}\}$ generates a subgroup of finite index in $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$.

Lemma A.20. *Referring to the notion introduce above, assume that $(b_i)_1^\tau$ satisfies* (A1). *Then there exists $C \in (0, \infty)$ such that, for any $h = 1, \dots, j_*$, any $g \in G$ that can be written in the form*

$$g = \prod_{i : w(c_i) \succeq \bar{w}_h} c_i^{x_i}, \quad |x_i| \le F_{c_i}(r),$$

can also be written in the from

$$g = \prod_{i : w(b_i) \succeq \bar{w}_h} b_i^{y_i}, \quad |x_i| \le F_{b_i}(Cr).$$

Proof. The proof is by backward induction on h and is similar to the proof of Lemma A.15. The details are omitted. ☐

Proposition A.21. *Assume that, for each j, the image of*

$$\{b_i : m_{j-1} + 1 \le i \le m_{j-1} + R_j\}$$

in $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$ generates a subgroup of finite index in $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$. Then there exists a constant $C > 0$ such that for any word ω in $\{b_i : w(b_i) \succeq \bar{w}_h\}^{\pm 1}$ with $\deg_{b_i} \omega \le F_{b_i}(r)$ for all i, there is a word ω' of the form

$$\omega' = \prod_{i=m_{h-1}+1}^{\tau} b_i^{x_i}$$

with

$$|x_i| \le \begin{cases} F_{b_i}(Cr) & \text{for } m_{j-1} + 1 \le i \le m_{j-1} + R_j^{\mathfrak{w}}, \\ C & \text{for } m_{j-1} + R_j^{\mathfrak{w}} + 1 \le i \le m_j \end{cases}$$

for $j \in \{h, \dots, j_*\}$ and such that $\pi(\omega') = \pi(\omega)$.

Proof. The proof is by backward induction on h. When $h = j_*$,$G^{\mathfrak{w}}_{j_*}$ is abelian and the desired result holds.

In general, let ω as in the proposition. By Lemmata A.13–A.20, we obtain a word

$$\omega_1 = \prod_{j=m_{h-1}+1}^{t} b_j^{x_j} \quad \text{with } |x_j| \leq F_{b_j}(Cr) \text{ for all } j \geq m_{h-1}+1$$

and such that

$$\pi(\omega) = \pi(\omega_1).$$

By hypothesis, the images of the commutators b_j, $m_{h-1}+1 \leq j \leq m_{h-1}+R^{\mathfrak{w}}_h$, generates a subgroup of finite index in $G^{\mathfrak{w}}_h/G^{\mathfrak{w}}_{h+1}$. Let N_h denote the index. Then for $m_{h-1}+R^{\mathfrak{w}}_h+1 \leq j \leq m_h$, there exists $a_1^{(j)}, \ldots, a_{R^{\mathfrak{w}}_h}^{(j)} \in \mathbb{Z}$ such that

$$b_j^{N_h} = b_{m_{h-1}+1}^{a_1^{(j)}} \ldots b_{m_{h-1}+R^{\mathfrak{w}}_h}^{a_{R^{\mathfrak{w}}_h}^{(j)}} \mod G^{\mathfrak{w}}_{h+1},$$

that is

$$\pi(b_j^{N_h}) = \pi(b_{m_{h-1}+1}^{a_1^{(j)}} \ldots b_{m_{h-1}+R^{\mathfrak{w}}_h}^{a_{R^{\mathfrak{w}}_h}^{(j)}} v_j),$$

where v_j is a word in $\{c_i : w(c) \succeq \bar{w}_{h+1}\}^{\pm 1}$. In

$$\omega_1 = \prod_{j=m_{h-1}+1}^{t} b_j^{x_j},$$

for each $j \in \{m_{h-1}+R^{\mathfrak{w}}_h+1, \ldots, m_h\}$, write

$$x_j = z_j N_h + y_j, \quad \text{with } 0 \leq y_j < N_h,$$

and replace $b_j^{N_h}$ by the word

$$\omega_j = b_{m_{h-1}+1}^{a_1^{(j)}} \ldots b_{m_{h-1}+R^{\mathfrak{w}}_h}^{a_{R^{\mathfrak{w}}_h}^{(j)}} v_j.$$

This produce a new word

$$\omega_1' = \prod_{j=m_{h-1}+1}^{m_{h-1}+R^{\mathfrak{w}}_h} b_j^{x_j} \cdot \prod_{j=m_{h-1}+1+R^{\mathfrak{w}}_h}^{m_h} \omega_j^{z_j} b_j^{y_j} \cdot \prod_{j=m_h+1}^{t} b_j^{x_j}$$

satisfying $\pi(\omega_1') = \pi(\omega_1)$. For $m_{h-1}+1 \leq j \leq m_{h-1}+R^{\mathfrak{w}}_h$,

$$\deg_{b_j} \omega_1' \leq |x_j| + \sum_{m_{h-1}+R^{\mathfrak{w}}_h+1 \leq i \leq m_h} |a^{(i)}_{j-m_{h-1}}|\,|x_i|,$$

By hypothesis, $\deg_{b_j} \omega \le F_{b_j}(Cd) \le \mathbf{F}_h(C_1 d)$ for all $m_{h-1}+1 \le j \le m_h$ and

$$\max\{|a_n^{(i)}| : m_{h-1} + R_h^{\mathfrak{w}} + 1 \le i \le m_h, 1 \le n \le R_h^{\mathfrak{w}}\} = C_h < \infty.$$

Hence, for $m_{h-1}+1 \le j \le m_{h-1} + R_h^{\mathfrak{w}}$, we obtain

$$\deg_{b_j} \omega_1' \le C_1(m_h - m_{h-1}) \mathbf{F}_h(Cd) \le \mathbf{F}_h(C_2 d).$$

For $m_{h-1} + R_h^{\mathfrak{w}} + 1 \le j \le m_h$, $\deg_{b_j} \omega \le N_h$. Finally, for any $c \in \{c_i : 1 \le i \le t\}$ with $w(c) \succ \bar{w}_h$, we have $F_c \succ \mathbf{F}_h$ and

$$\begin{aligned} \deg_c \omega_1' &\le \deg_c \omega_1 + \sum_{m_{h-1}+R_h^{\mathfrak{w}}+1 \le k \le m_h} |z_k| \deg_c v_k \\ &\le F_c(C_3 d). \end{aligned}$$

Applying Lemmata A.13–A.20 to ω_1', we obtain a word ω' with $\pi(\omega) = \pi(\omega')$ and

$$\omega_2 = \prod_{j=m_{h-1}+1}^{m_{h-1}+R_h^{\mathfrak{w}}} b_j^{\widetilde{x_j}} \prod_{j=m_{h-1}+1+R_h^{\mathfrak{w}}}^{m_h} b_j^{y_j} \prod_{j>m_h} b_j^{\widetilde{x_j}}$$

where $|\widetilde{x_j}| \le \mathbf{F}_h(C_1 d)$ for $m_{h-1}+1 \le j \le m_{h-1} + R_h^{\mathfrak{w}}$; $0 \le y_j < N_h$ for $m_{h-1} + R_h^{\mathfrak{w}} + 1 \le j \le m_h$, and $|\tilde{x}_j| \le F_{c_j}(C_2' d)$ for all $j > m_h$. Now, apply the induction hypothesis to $\prod_{j=m_h+1}^{t} b_j^{\widetilde{x_j}}$, to obtain the desired conclusion. □

We end with the following simple improvement of the last statement in Theorem 2.10. The proof is a simple combination of the previous proposition together with Lemma A.19.

Theorem A.22. *Let G be a nilpotent group equipped with a generating k-tuple $S = (s_1, \dots, s_k)$. Let $\mathfrak{w}$, $\mathfrak{F}$ be weight and weight-function systems on S satisfying (2.1) and (2.2). Let $\Sigma = (c_1, \dots, c_t)$ be a tuple of formal commutators in $\mathfrak{C}(S)$ with non-decreasing weights $w(c_1) \preceq \cdots \preceq w(c_t)$. Let m_j, $j = 0, \dots, j_*$ be defined by*

$$\{c_i : w(c_i) = \bar{w}_j\} = \{c_i : m_{j-1} < i \le m_j\}.$$

Assume that (the image of) $\{c_i : w(c_i) = \bar{w}_j\}$ generates $G_j^{\mathfrak{w}}$ modulo $G_{j+1}^{\mathfrak{w}}$ and that $\{c_i : m_{j-1} < i \le m_{j-1} + R_j^{\mathfrak{w}}\}$ is free in $G_j^{\mathfrak{w}}/G_{j+1}^{\mathfrak{w}}$.

There exist an integer $p = p(G, S, \mathfrak{F})$, *a constant* $C = C(G, S, \mathfrak{F})$ *and a sequence* $(i_1, \dots, i_p) \in \{1, \dots, k\}^p$ *such that if* g *can be expressed as a word* ω *over* $\mathfrak{C}(S)$ *with* $\deg_c(\omega) \leq F_c(r)$ *for some* $r \geq 1$ *and all* $c \in \mathfrak{C}(S)$ *then* g *can be expressed in the form*

$$g = \prod_{j=1}^{p} s_{i_j}^{x_j} \quad \textit{with } |x_j| \leq C \begin{cases} F_{i_j}(r) & \textit{if } s_{i_j} \in \mathrm{core}(S, \mathfrak{w}, \Sigma), \\ 1 & \textit{if } s_{i_j} \notin \mathrm{core}(S, \mathfrak{w}, \Sigma). \end{cases}$$

References

[1] M. Barlow, R. Bass, and T. Kumagai, Parabolic Harnack inequality and heat kernel estimates for random walks with long range jumps. *Math. Z.* **261** (2009), no. 2, 297–320. Zbl 1159.60021 MR 2457301

[2] M. T. Barlow, A. Grigor′yan, and T. Kumagai, Heat kernel upper bounds for jump processes and the first exit time. *J. Reine Angew. Math.* **626** (2009), 135–157. Zbl 1158.60039 MR 2492992

[3] H. Bass, The degree of polynomial growth of finitely generated nilpotent groups. *Proc. London Math. Soc.* (3) **25** (1972), 603–614. Zbl 0259.20045 MR 0379672

[4] A. Bendikov and L. Saloff-Coste, Random walks driven by low moment measures. *Ann. Probab.* **40** (2012), no. 6, 2539–2588. Zbl 1262.60005 MR 3050511

[5] A. Bendikov and L. Saloff-Coste, Random walks on groups and discrete subordination. *Math. Nachr.* **285** (2012), no. 5-6, 580–605. Zbl 1251.60004 MR 2902834

[6] N. H. Bingham, C. M. Goldie, and J. L. Teugels, *Regular variation.* Encyclopedia of Mathematics and its Applications, 27. Cambridge University Press, Cambridge, 1987. Zbl 0617.26001 MR 0898871

[7] Th. Coulhon, Ultracontractivity and Nash type inequalities. *J. Funct. Anal.* **141** (1996), no. 2, 510–539. Zbl 0887.58009 MR 1418518

[8] W. A. de Graaf, Classification of 6-dimensional nilpotent Lie algebras over fields of characteristic not 2. *J. Algebra* **309** (2007), no. 2, 640–653. Zbl 1137.17012 MR 2303198

[9] P. Griffin, Matrix normalized sums of independent identically distributed random vectors. *Ann. Probab.* **14** (1986), no. 1, 224–246. Zbl 0602.60031 MR 0815967

[10] P. Griffin, N. Jain, and W. Pruitt, Approximate local limit theorems for laws outside domains of attraction. *Ann. Probab.* **12** (1984), no. 1, 45–63. Zbl 0539.60022 MR 0723729

[11] M. Hall, Jr., *The theory of groups.* Reprinting of the 1968 edition. Chelsea Publishing Co., New York, 1976. Zbl 0354.20001 MR 0414669

[12] Ph. Hall, *The collected works of Philip Hall.* Compiled and with a preface by K. W. Gruenberg and J. E. Roseblade. With an obituary by Roseblade. Oxford Science Publications. The Clarendon Press, Oxford University Press, New York, 1988. Zbl 0644.20001 MR 0986732

[13] W. Hebisch and L. Saloff-Coste, Gaussian estimates for Markov chains and random walks on groups. *Ann. Probab.* **21** (1993), no. 2, 673–709. Zbl 0776.60086 MR 1217561

[14] J. Hu and T. Kumagai, Nash-type inequalities and heat kernels for non-local Dirichlet forms. *Kyushu J. Math.* **60** (2006), no. 2, 245–265. Zbl 1119.60068 MR 2268236

[15] H. Kesten, Full Banach mean values on countable groups. *Math. Scand.* **7** (1959), 146–156. Zbl 0092.26704 MR 112053

[16] H. Kesten, Symmetric random walks on groups. *Trans. Amer. Math. Soc.* **92** (1959), 336–354. Zbl 0092.33503 MR 0109367

[17] M. Murugan and L. Saloff-Coste, Transition probability estimates for symmetric jump processes. In preparation.

[18] Ch. Pittet and L. Saloff-Coste, A survey on the relationships between volume growth, isoperimetry, and the behavior of simple random walk on cayley graphs, with examples. Preprint 2000. www.math.cornell.edu/~lsc/surv.ps.gz

[19] Ch. Pittet and L. Saloff-Coste, On the stability of the behavior of random walks on groups. *J. Geom. Anal.* **10** (2000), no. 4, 713–737. Zbl 0988.60517 Zbl 0985.60043 MR 1817783

[20] L. Saloff-Coste, Sur la décroissance des puissance de convolution sur les groupes. *Bull. Sci. Math.* (2) **113** (1989), no. 1, 3–21. Zbl 0682.43004 MR 0984425

[21] L. Saloff-Coste, Sobolev inequalities and polynomial decay of convolution powers and random walks. In A. B. Cruzeiro and J.-C. Zambrini (eds.), *Stochastic analysis and applications.* (Lisbon, 1989.) Progr. Probab., 26, Birkhäuser Boston, Boston, MA, 1991, 176–189. Zbl 751.60067 MR 1168075

[22] L. Saloff-Coste Probability on groups: random walks and invariant diffusions. *Notices Amer. Math. Soc.* **48** (2001), no. 9, 968–977. Zbl 0987.60018 MR 1854532

[23] L. Saloff-Coste, Analysis on Riemannian co-compact covers. In A. Grigor′yan and S. T. Yau (eds.), *Surveys in differential geometry.* Vol. IX, Surv. Differ. Geom., 9, Int. Press, Somerville, MA, 2004, 351–384. Zbl 1082.31006 MR 2184990

[24] L. Saloff-Coste and T. Zheng, On some random walks driven by spread-out measures. Preprint 2013. arXiv:1309.6296 [math.PR]

[25] F. Spitzer, *Principles of random walk.* The University Series in Higher Mathematics. D. Van Nostrand, Princeton etc., 1964. Zbl 0119.34304 MR 0171290

[26] J. Tits, Appendix to: "Groups of polynomial growth and expanding maps" [*Inst. Hautes Études Sci. Publ. Math.* **53** (1981), 53–73] by M. Gromov. *Inst. Hautes Études Sci. Publ. Math.* **53** (1981), 74–78. Zbl 0474.20018 MR 0623535

[27] N. Th. Varopoulos, Convolution powers on locally compact groups. *Bull. Sci. Math.* (2) **111** (1987), no. 4, 333–342. Zbl 0626.22004 MR 0921558

[28] N. Th. Varopoulos, L. Saloff-Coste, and T. Coulhon, *Analysis and geometry on groups.* Cambridge Tracts in Mathematics, 100. Cambridge University Press, Cambridge, 1992. MR 1218884 Zbl 1179.22009

Received December 27, 2012

Laurent Saloff-Coste, Department of Mathematics, 320A/567 Malott Hall,
Cornell University, Ithaca, NY 14853-4201, USA

e-mail: lsc@math.cornell.edu

Tianyi Zheng, Department of Mathematics, Stanford University, 450 Serra Mall,
Building 380, Stanford, CA 94305-2125, USA

e-mail: tzheng2@stanford.edu