CLT FOR RANGE OF RANDOM WALKS ON HYPERBOLIC GROUPS

MAKSYM CHAUDKHARI, KUNAL CHAWLA, CHRISTIAN GORSKI, EDUARDO SILVA

Abstract. We prove a central limit theorem for random walks on wreath products $A \wr H = \bigoplus_H F \wr H$, where A is a non-trivial group and H is a non-elementary acylindrically hyperbolic group, for adapted step distributions with a finite second moment and bounded lamp range. Additionally, we show that if H is a non-elementary hyperbolic group, then the central limit theorem holds for all adapted probability measures with a finite second moment.

1. Introduction

(1) Talk about random walks on groups; trying to prove limit laws in general
(2) Make a list of things known about lamplighters and why they are relevant
(3) Say that in this paper we concentrate on the CLT

Here are some things that we should cite... I am missing many more but this is a good start. [BQ16b, BQ16a, Bjö10, Cho23a, EZ22, GTT22, Gil08, Gou17, Hor18, Bar22, Led01, MSŠ23, Sal01, SS87].

1.1. Main results.

Consider the switch-walk-switch word length $|\cdot|$ on $A \wr H$.

Theorem 1.1. *Let A be a non-trivial group and H a non-elementary hyperbolic group. Consider a probability measure μ on $A \wr H$ such that μ_H is non-elementary and has a finite second moment, and such that μ has bounded lamp range. Denote by $\{w_n\}_{n\geq 0}$ the μ-random walk on $A \wr H$, and let $C = \lim_{n\to\infty} \frac{\mathbb{E}(|w_n|)}{n}$ be the drift of the μ-random walk on $A \wr B$. Then the sequence of normalized random variables $\frac{|w_n| - Cn}{\sqrt{n}}$, $n \geq 1$, converges in law to a non-degenerate gaussian law.*

What are the moment hypotheses for acylindrically hyperbolic base groups? Finite support? exponential tails?

We prove a central limit theorem for the range of random walks with a finite second moment on hyperbolic groups.

Theorem 1.2. *Let H be a non-elementary hyperbolic group and let μ be a non-elementary probability measure on H with a finite second moment. Let C be the probability that the μ-random walk on H starting at e_H never returns to e_H. Then the sequence of normalized random variables $\frac{|R_n| - Cn}{\sqrt{n}}$, $n \geq 1$, converges in law to a non-degenerate gaussian law.*

2. Preliminaries

2.1. Hyperbolic groups.

Say basic things about hyperbolicity; explain pivots

2.2. Lamplighter groups.

We consider the wreath products $A \wr H$, where A is a finite non-trivial group and H is a finitely generated group. Let S_H be a finite and symmetric generating set of H. Then we consider the *switch-walk-switch* S_{sws} generating set of $A \wr H$, given by

$$S_{\text{sws}} \coloneqq \left\{ (\delta_a, 0)(\mathbf{0}, s)(\delta_{a'}, 0) \middle| a, a' \in A \text{ and } s \in S_H \right\}.$$

Theorem 2.1 ([Par92, Theorem 1.2])**.** *For any $g = (f, x) \in A \wr H$, the word length of g with respect to the standard generating set is*

$$|g| = \mathrm{TSP}(e_H, x, \mathrm{supp}(f)).$$

2.3. Random walks on groups.

(1) Recall basic concepts of random walks on groups.

Date: January 2024.

2.4. Defective adapted cocycles and the central limit theorem.

(1) Introduce all necessary results and definitions of Mathieu-Sisto.
(2) Introduce pivots and the results of Gouëzel that we will use.

We will use the following general criterion of Mathieu-Sisto.

Definition 1 (Defective adapted cocycle)**.**

Theorem 2.2 ([MS20])**.** *Suppose that Q_n is a defective adapted cocycle with defect*

$$\Phi_{m,n} := Q_n - (Q_m + w_m Q_{m-n}).$$

Suppose that there exists $C > 0$ such that

$$\sup_{m,n\geq 1} \mathbb{E}[|\Phi_{m,n}|^2] \leq C.$$

Then the CLT holds for Q_n.

In this section we explain a generalization of [MS20, Theorems 4.1 & 4.2] (see Theorem 2.2)

Theorem 2.3. *Suppose that Q_n is a defective adapted cocycle with defect*

$$\Phi_{m,n} := Q_n - (Q_m + Z_m Q_{m-n})$$

and suppose that for some fixed polynomial p and $N_0 \in \mathbb{N}$ we have that

$$\mathbb{E}[|\Phi_{m,n}|^2] \leq p(\log(n))$$

whenever $m, n-m \geq N_0$. Then a CLT holds for Q_n.

Is this the formulation we want?

Here are the places in the proof of Mathieu-Sisto's CLT where the deviation inequality is used:

(1) In Theorem 4.4 one obtains $V(Q_n) \leq n\Big(p(4\mathbb{E}(Q_1^2) + 16\log(n))\Big)$
(2) Then in Lemma 4.5 it is cited a result of Hammersley [Ham62, Theorem 2]. This should be replaced by an inequality of the form $a_{n+m} \leq a_n + a_m + bp(log(n+m))$.

Lemma 2.4. *Let $\{a_n\}_{n\geq 0}$ be a sequence of non-negative real numbers. Suppose that there exists $b \geq 0$ and a polynomial $p : \mathbb{R}_+ \to \mathbb{R}_+$ such that*

$$a_{n+m} \leq a_n + a_m + b\sqrt{p(log(a_n + a_m))}, \text{ for each } m, n \geq 0.$$

Then the limit $\lim_{n\to\infty} \frac{a_n}{n}$ exists.

(3) The deviation inequality is used in multiple occasions during the proof of Lemma 4.6. One should do the appropriate modifications.

3. The CLT for lamplighter random walks on acylindrically hyperbolic groups

Theorem 3.1. *Let $G = \mathbb{Z}/2\mathbb{Z} \wr H$ be a wreath product over a Gromov-hyperbolic group H. Let $\tilde{Z}_n$ be a random walk on G with step distribution which is finitely supported and [what is the term for these types of generators? where everything is step or step+light]. switch-walk-switch? Then $d(o, \tilde{Z}_n)$ satisfies a CLT.*

(Maybe adopt the convention that $\tilde{Z}_n$ is the random walk on the wreath product G, while Z_n is its projection to the hyperbolic base group H. I think that the notation we've been using is that $R_n \subset H$ is the configuration of lights at time n, i.e. $\tilde{Z}_n = (R_n, Z_n)$. Since this looks like the range of the random walk, maybe we want to change notation, but this is the notation I'm going to use as I write the lemmas we need for now.)

For $R \subset H$, $x, y \in H$, denote by $TSP(x, R, y)$ a solution to the traveling salesman problem, that is, the edge path in the Cayley graph of H which starts at x, ends at y, visits each vertex in R, and has minimal length subject to this constraints. Note that for our choice of generators for G, we have that $d((\emptyset, 1), (R, x)) = |TSP(1, R, x)|$.

Lemma 3.2 (Deterministic TSP comparison)**.** *When we write this out properly, we'll figure out what all the proper polynomials are, and the proper definition of the "middle region", etc. There are fixed polynomials p_1, p_2, p_3, p_4 such that the following hold. Let C_1, C_2, and C_3 be constants. Let $(R_m, x_m), (R_n, x_n) \in G$ and define $R_{m,n} \subset H$ by $(R_{m,n}, x_m^{-1}x_n) := (R_m, x_m)^{-1}(R_n, x_n)$. Suppose that*

(1) *$R_m \cup x_m R_{m,n} \subset \mathcal{N}(\xi, C_1 p_1(\log n))$, that is, every vertex of $R_m \cup x_m R_{m,n}$ is within distance $C_1 p_1(\log n)$ of ξ, where ξ is a geodesic from 1 to x_n.*
(2) *Setting $r := d(o, x_m)$, we have $R_m \subset B(o, r + C_2 p_2(\log n))$ and $x_m R_{m,n} \subset B(o, r - C_2 p_2(\log n))^c$. (Not exactly clear yet what will be the best way to define the "middle/left/right" of the path yet, but I'm using this definition at least for now since it probably will come most easily out of the Hoeffding speed concentration estimates.*
(3) *Set*
$$M := B(o, r + C_2 p_2(\log n)) \cap B(o, r - C_2 p_2(\log n))^c.$$
Then we have that
$$|M \cap (R_m \cup x_m R_{n,m})| \leq C_3 p_3(\log n).$$

Then we have that
$$\begin{aligned}&|TSP(1, R_n, x_n)| \geq \\ &|TSP(1, R_m, x_m)| + |TSP(1, R_{m,n}, x_m^{-1} x_n)| - (C_1 + C_2 + C_3) p_4(\log n).\end{aligned}$$

Depending on the analysis, could maybe get some different expression in the C_i. If we have something of this form (or equivalently, a bound which is poly(log) times $\max(C_1, C_2, C_3)$), then to get the final step it will suffice to show the each C_i is square-integrable (as a random variable), but if other powers of the C_i become involved, might have to show stronger integrability.

Proof. The outline of the proof is as follows. To simplify the notation, let $D = p_1(\log n)$. Notice that by quasi-convexity there is an absolute constant δ such that for any two points $x, y \in N(\xi, D)$ any geodesic between x and y will be contained in $N(\xi, D + \delta)$.

(1) We will split $N(\xi, D)$ into three regions - the initial part I close to the origin, the central region C and the terminal part close to x_n. More precisely, we let $I := N(\xi, D) \cap B(o, r - C_2 p_2(\log n)$, $C := N(\xi, D) \cap B(o, r + C_2 p_2(\log n)) \cap B(o, r - C_2 p_2(\log n))^c$ and $T = N(\xi, D) \cap B(o, r + C_2 p_2(\log n))^c$. Polynomial p_2 will be determined later.
(2) Pick any solution S to $TSP(1, R_n, x_n)$. We are going to show that it is possible to add a collection of segments of total length bounded by a fixed polynomial in $\log_n$ to $\overline{I \cap S}$ to get a path starting at 1, visiting every point in R_m and ending at x_m. Here $\overline{I \cap S}$ stands for the union of $I \cap S$ together with all geodesic segments of S that join two points in $R_n \cap I$ but possibly go outside of I. Then by the definition of TSP this path would have the length bounded below by $|TSP(1, R_m, x_m)|$.

   Similar argument will be applied to $\overline{S \cap T}$ and $|TSP(1, R_{m,n}, x_m^{-1} x_n)|$. Since $\overline{I \cap S}$ and $\overline{T \cap S}$ have no geodesic segments of S in common, we have
$$\begin{aligned}&|TSP(1, R_n, x_n)| = |S| \geq |\overline{I \cap S}| + |\overline{T \cap S}| \geq \\ &|TSP(1, R_m, x_m)| + |TSP(1, R_{m,n}, x_m^{-1} x_n)| - P(\log n)\end{aligned}$$

□

Since $\Psi_{n,m}$ is bounded above by $3n$, we should be able to reduce the problem to deterministic case using known deviation estimates and a simple union bound, so most likely we will not need the strongest versions of the next lemmas.

Lemma 3.3. *Define the random variable $\mathcal{C}_{1,m,n}$ to be the minimal value of C_1 such that (1) in Lemma 3.2 holds for $\tilde{Z}_m = (R_m, x_m)$ and $\tilde{Z}_n = (R_n, x_n)$, i.e.*
$$\mathcal{C}_{1,m,n} := \inf\{C_1 \geq 0 : d(r, \xi) \leq C_1 p_1(\log n) \text{ for all } r \in R_m \cup Z_m R_{n,m}\},$$
where ξ is a geodesic from 1 to $x_n = Z_n$. Then
$$\limsup_{m, n-m \to \infty} \mathbb{E}[\mathcal{C}_{1,m,n}^2] < \infty.$$

The proof of this should be via a tail bound which should be contained in "Tracking Rates of Random Walks" by Sisto.

Since $\Psi_{n,m}$ is bounded above by $3n$, we should be able to reduce the problem to deterministic case using known deviation estimates and a simple union bound, so most likely we will not need the strongest versions of the next lemmas.

Lemma 3.4. *Define the random variable $\mathcal{C}_{2,m,n}$ to be the minimal value of C_2 such that (2) in Lemma 3.2 holds for $\tilde{Z}_m = (R_m, x_m)$ and $\tilde{Z}_n = (R_n, x_n)$. Then*

$$\limsup_{m,n-m\to\infty} \mathbb{E}[\mathcal{C}_{2,m,n}^2] < \infty.$$

It seems likely enough that in the end the bound on C_2 will come simultaneously with the bound on either C_1 or C_3, since it should come from either tracking or speed. In fact, since C_2 is a parameter that we use to define the "middle" segment M, we will want to make it bigger *than the minimum possible. Maybe in the final argument we will just take $C_2 = 6C_1$ or $6C_3$ or something like that and this lemma won't exist.*

Lemma 3.5. *Define the random variable $\mathcal{C}_{3,m,n}$ to be the minimal value of C_3 such that (3) in Lemma 3.2 holds for $\tilde{Z}_m = (R_m, x_m)$ and $\tilde{Z}_n = (R_n, x_n)$. That is,*

$$\mathcal{C}_{3,m,n} := \frac{|M \cap (R_m \cup x_m R_{m,n})|}{p_3(\log n)}.$$

Then

$$\limsup_{m,n-m\to\infty} \mathbb{E}[\mathcal{C}_{3,m,n}^2] < \infty.$$

I expect this to follow from the concentration bounds for speed in "Random walks in hyperbolic spaces..." by Aoun and Sert. Probably not quite as immediate as bound on tail of $\mathcal{C}_1$, but I think should go through.

Proof of Theorem 3.1 given Lemmas. Again, this is assuming that the conclusion of the deterministic lemma has the form I wrote; write out that lemma carefully and adjust the proof here/lemmas above as needed. Define a defective adapted cocycle by

$$Q_n := |TSP(1, R_n, Z_n)|$$

where $(R_n, Z_n) = \tilde{Z}_n$. By construction of the $\mathcal{C}_i$ and by Lemma 3.2, we have that the defect satisfies

$$|\Psi_{m,n}| \le (\mathcal{C}_{1,m,n} + \mathcal{C}_{2,m,n} + \mathcal{C}_{3,m,n}) p_4(\log n).$$

By Lemmas 3.3, 3.4, and 3.5, there is some constant C (independent of m and n) and some $N_0 \in \mathbb{N}$ such that whenever $m, m-n \ge N_0$ we have that

$$\mathbb{E}[\mathcal{C}_{i,m,n}^2] \le C$$

for $i = 1, 2, 3$. Cauchy-Schwarz then tells us that for some C', whenever $m, n-m \ge N_0$ we have

$$\mathbb{E}[|\Psi_{m,n}|^2] \le C' p_4(\log n)^2.$$

Then applying Theorem 2.3 gives our result. □

4. CLT for the lamplighter over a hyperbolic group (using pivots)

4.1. **Pivots.** Let us consider a non-elementary hyperbolic group H, and let us fix a finite generating set S_H. Let $\delta \ge 0$ be the hyperbolicity constant of $\mathrm{Cay}(H, S_H)$, and let us denote by $d_H : H \times H \to \mathbb{Z}_{\ge 0}$ the word metric on H with respect to S_H.

Definition 2. Given a path $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_k)$ on $\mathrm{Cay}(H, S_H)$ and $g \in H$, let $\pi_\gamma(g)$ be the set of elements of γ that minimize the word metric to g. That is, we define

$$\pi_\gamma(g) := \{\gamma_i \in \gamma \mid d_S(\gamma_i, g) \le d_S(\gamma_j, g) \text{ for all } j = 1, \dots, k\}. \tag{1}$$

We now introduce the definition of pivots that we will use in the proof of Theorem 3.1. We refer to [Gou22, Section 4A] for details.

Definition 3. Let $C, D > 0$, $L \ge 20C + 100\delta + 1$, and $N \in \mathbb{Z}_{\ge 1}$. Let $\mathbf{w} = \{w_n\}_{n\ge 0} \in (A \wr H)^{\mathbb{N}}$ be a sample path of the μ-random walk on $A \wr H$. Denote by w_n^H the projection to H of w_n, for each $n \ge 0$. A time instant $m \ge 1$ is a (C, D, L, N)-*pivot* for the sample path $\mathbf{w}$ if the following three conditions hold.

(1) $d_H\left(w_m^H, w_{m+N}^H\right) \ge L$.

Let γ be an arbitrary geodesic path in $\mathrm{Cay}(H, S_H)$ that connects w_m^H to w_{m+N}^H. Then

(2)
$$d_H\left(\pi_\gamma\left(w_k^H\right), w_m^H\right) \leq C, \text{ for all } k \in \{0, 1, \ldots, m\},$$
(3) for all $m \leq k \leq m+N$ we have $d_H\left(w_k^H, \gamma\right) \leq D$, and
(4) for all $k \geq m+N$, we have $d_H\left(\pi_\gamma\left(w_k^H\right), w_{m+N}^H\right) \leq C$.

The following lemma will be our main tool.

Lemma 4.1. *For any $C, D, \delta > 0$, and any $L \geq 20C + 100\delta + 1$, there exists $N, R > 0$ large such that*
$$\sup_{i \geq 1} \mathbb{P}\left(\exists m \in [i, i+k] \text{ such that } m \text{ is an } (C, D, L, N)\text{-pivot}\right) \geq 1 - Re^{-k/R}.$$

Sketch of proof. This is proven in proposition 4.11 in Gouezel's paper, where for Gouezel's definition of pivots, conditions 1,2, and 4 are met. To see why Gouezel's proof implies the lemma we state, we observe that for Gouezel's definition of pivots, the increments $w_n^{-1}w_{n+N}$ are drawn from some explicit finite set of isometries $S \subset H$ that receive positive support from μ_H^N. For this finite set of isometries, we can pick some $D > 0$ large enough so that μ^N gives positive mass to each of the sets $\pi_H^{-1}(s) = \{(\varphi, s), \text{supp}\varphi \subset B_D([e, s])\}$. Then tracing through the rest of Gouezel's proof we have the estimate required. □

4.2. **TSP structure along pivots.**

Proposition 4.2. *Suppose that we are looking at a sample path $\{w_n\}_{n\geq 0}$ and that we have a pivoting time m. Then the group element $w_n = (f_n, x_n)$ satisfies the following. The support of f_n can be decomposed as a disjoint union*
$$\text{supp}(f_n) = P_1' \cup P_2' \cup P_3',$$
that satisfies the following properties. Let us denote $P_1 = P_1' \cup \{w_{m+u}^H\}$, $P_2 = P_2' \cup \{w_{m+u}^H, w_{m+u+N}^H\}$ and $P_3 = P_3' \cup \{w_{m+u+N}^H\}$. Let γ be an arbitrary geodesic path from w_m^H to w_{m+N}^H on $\text{Cay}(H, S_H)$*. Then we have*

(1) *for all $g \in P_1$, we have $d_H\left(\pi_\gamma(g), w_m^H\right) \leq C$,*
(2) *for all $g \in P_2$ we have $d_H(g, \gamma) \leq D$, and*
(3) *for all $g \in P_3$, we have $d_D\left(\pi_\gamma(g), w_{m+N}^H\right) \leq C$.*

Definition 4. Let $g = (f, x) \in A \wr H$ and suppose that $\text{supp}(f) = P_1 \cup P_2 \cup P_3$. Let η be a path on $\text{Cay}(A \wr H, S_{\text{sws}})$ that realizes $|g|_{S_{\text{sws}}}$. We define the associated *coding* of η as the word u in the alphabet $\{P_1, P_2, P_3\}$, such that $u_i = P_j$ if and only if at the i-th step of η, there is a lamp at a position in P_j which was modified for the first time.

We are going to abuse notation (in this draft) and not make a distinction between the elements visited by a path, and the coding in the alphabet $\{P_1, P_2, P_3\}$ associated with it.

Definition 5. Given a path η, let us call a *backtracking* a subpath of η that is of the form $P_1P_2^*P_3^\varepsilon P_2^*P_1^{\varepsilon'}P_2^*P_3$ for $\varepsilon, \varepsilon' \geq 1$. Here the $*$ symbolizes 0 or more occurrences.

Lemma 4.3. *Let η be a solution to the TSP for $|w_n|$. Then the coding of η does not have a subword of the form $P_1P_3^\varepsilon P_1^{\varepsilon'}P_3$, for $\varepsilon, \varepsilon' \geq 1$.*

Proof. Surgery, meaning that you glue together the excursions to P_1, and you glue together the excursions to P_3, and connect them with any path through P_2. This gives something even shorter than optimal since each gluing strictly reduces the length of the path. □

Corollary 4.4. *Let η be a solution to the TSP for $|w_n|$. Then the number of backtrackings of η is at most $|P_2|$.*

Proof. Every backtracking must contain at least one element of P_2. (Recall that the path only has an element in its coding if it has not been visited before). □

Lemma 4.5. *Consider a sequence of points $\{w_n\}_n$ of H, that satisfies the decomposition of* $\text{supp}(f_n)$ *given by the three conditions of Proposition 4.2.*

Let T be the length of a solution to $\text{TSP}(w_0, w_n, \text{supp}(f_n)) = \text{TSP}\left(w_0^H, w_n^H, P_1 \cup P_2 \cup P_3\right) = |w_n|_{A\wr H}$*. Then there exists a path η that starts at w_0^H, finishes at w_n^H and visits all points in* $\text{supp}(f_n)$ *such that length$(\eta) \leq T + 100N(L + 2D)$, and such that, in the coding of η, all the elements of P_1 appear before any of the elements of P_3.*

Proof. Let us first consider η_0 the optimal solution to the TSP.

Lemma 4.3 implies Corollary 4.4 that the total number of backtrackings is the size of P_2.

Finally, the argument goes as follows: first do all excursions of η_0 on P_1, then visit all elements in P_2, and then do all excursion in P_3. In total we added at most $2D\times$(number of backtrackings)+(length of solution of TSP in P2 that visits all elements in P2). And the number of backtrackings is at most $|P_2|$ by the previous claim.

□

In other words, we are saying that trying to solve the problem by first visiting all elements of P_1, and then visiting all elements of P_3, and crossing the middle section only once, is at a bounded length of being optimal.

Lemma 4.6. *For any $N \in \mathbb{N}$ there exists some $C > 0$ such that the following holds. set $m \in \mathbb{N}$ be an integer and let U be the waiting time until the first pivot after time m. Then we have*

$$\sup_{m\geq 1} \mathbb{E}\left[\left(\sum_{i=m}^{m+U+N} |g_i|\right)^2\right] \leq C.$$

Proof. I'll go into the construction of pivots and explain why this is true. Maybe there's a simpler reasoning using only the exponential estimates on U.

Recall that the way Gouezel constructs pivots is as follows: if we let $S \subset H$ be our finite Schottky set, then we can decompose some convolution power μ_H^N as

$$\mu_H^N = \alpha\mu_S + (1-\alpha)\nu$$

for some positive $\alpha > 0$. Then we draw our increments as follows: let $\{\varepsilon_i\}_i$ be i.i.d. Bernoulli(α) random variables. If $\varepsilon_i = 1$, we draw $g_i' = s_i$ according to μ_S. Else we draw $g_i' = w_i$ according to ν. We observe that the sequence $\{g_1' ... g_k'\}_k$ has the same distribution as $\{g_1 ... g_k\}_k$ for $g_i \sim \mu_H^N$.

Now we denote the resampled random walk by $g_1' ... g_n' = w_1 ... w_{k_1} s_1 w_{k_1+1} ... w_{k_2} s_2 ...$, where the strings between $s_i's$ may be empty. Now each string $w_{k_{i-1}+1} ... w_{k_i} s_i$ is distributed according to $\nu^Z * \mu_S$, where Z is a geometric random variable with parameter α.

Now Gouezel tells us that, conditional on any realization of the increments drawn from ν, the number of μ_S increments ℓ until we see a pivot has an exponential tail. This implies that

$$\mathbb{E}\left[\left(\sum_{i=m}^{m+U+N} |g_i|\right)^2 |\{w_i\}_i\right] \leq \mathbb{E}\left[\left(L\ell + \sum_{i=0}^{\ell-1}\sum_{k=K_i}^{K_{i+1}-1} |w_i|\right)^2 |\{w_i\}_i\right].$$

Now we can integrate over the possible values of w_i and use independence in order to conclude that

$$\mathbb{E}\left[\left(\sum_{i=m}^{m+U+N} |g_i|\right)^2\right]$$

is bounded uniformly over m. □

4.3. **Proof of the CLT.** Let us define $\Phi_{n,m} = |w_n| - |w_m^{-1}w_n| - |w_m|$. Thanks to Theorem 2.2, it suffices to show that $\sup_{m,n\geq 1} \mathbb{E}(|\Phi_{n,m}|^2)$ is bounded.

We will do the proof for finitely supported μ.

We fix m and n. Let $m+u$ be the first instant after m that you see a pivot.

If $m+u+N > n$, then we use Lemma 4.6

$$\mathbb{E}|\Phi_{n,m}|^2 \leq \sup_{m\geq 1} \mathbb{E}\left[\left(\sum_{i=m}^{m+u} |g_i|\right)^2\right] \leq C.$$

Otherwise, $m+u+N \leq n$ and we do the following.

(1) The three conditions at the beginning of this subsection are satisfies.
(2) Our objective is to get a good upper bound for $\Phi_{n,m}$ in the inequality

$$|w_n| \geq |w_m| + |w_m^{-1}w_n| - \Phi_{n,m}.$$

(3) We first note that $||w_m| - |w_{m+u}||$ has a finite second moment. Indeed, this amount is controlled by the increments done during u steps, and we know the distribution of how large u can be. That is, we use Lemma 4.6 to justify this. The same is true for $\big||w_m^{-1}w_n| - |w_{m+u}^{-1}w_n|\big|$. Again, this follows from a triangular inequality and Lemma 4.6.
(4) From this, we just need a good upper bound for $\Phi_{n,m}$ in the inequality

$$|w_n| \geq |w_{m+u}| + |w_{m+u}^{-1}w_n| - \Phi_{n,m}.$$

(5) We note that $\big||w_{m+u+N}^{-1}w_n| - |w_{m+u}^{-1}w_n|\big|$ is a bounded constant (since it only depends on N), and in particular has a finite second moment.
(6) From this, we just need a good upper bound for $\Phi_{n,m}$ in the inequality

$$|w_n| \geq |w_{m+u}| + |w_{m+u+N}^{-1}w_n| - \Phi_{n,m}.$$

(7) We look at the TSP between time 0 and n, we use the path η from the previous lemma to get a path which is near optimal and crosses only once the neighborhood of γ.
(8) From this path we obtain near-optimal paths from $|w_{m+u}|$ and for $|w_{m+u+N}^{-1}w_n|$, by doing surgery near the endpoints of γ and possibly adding a constant bounded amount of length.
Indeed, we first take the path from the starting point to the last visit to P_1, and we connect it to w_{m+u}. This is at most $Optimal + L + 2D$. Similarly we look at the first time we enter P_3, and connect that to a path to w_{m+u+N}. This again adds at most $Optimal + L + 2D$.
(9) From this, we directly apply 2.2.

5. The CLT for the range of random walks on hyperbolic groups

We borrow the framework from [MS20] for proving a CLT - We observe the following trivial fact that whenever $1 \leq m \leq n$ and denoting $R_{m,n}$ for the range between times m and n we have

$$|R_n| = |R_m| + |R_{m,n}| - |R_m \cap R_{m,n}|.$$

In the language of [MS20], we say that $\{|R_n|\}_{n\geq 1}$ is a *defective adapted cocycle* with defect $\Phi_{m,n} := |R_m \cap R_{m,n}|$.

By theorem 4.2 in [MS20], to prove a CLT for the sequence $|R_n|$ it is enough to show a second-moment deviation inequality: that

$$\mathbb{E}[\Phi_{m,n}^2] \leq C.$$

For some C not depending on m, n. We instead prove a stronger version of the deviation inequality:

Proposition 5.1. *There exists $C > 0$ such that for any $1 \leq m \leq n$.*

$$\mathbb{P}(\Phi_{m,n} \geq k) \leq Ce^{-k/C},$$

Proof. Let $\hat{R}_n$ denote the range of the reversed random walk - that is, the random walk driving by $\hat{\mu}$.' If is enough to show that for any $n, n' \in \mathbb{N}$ we have

$$\mathbb{P}(\sup_{n,n'} |\hat{R}_n \cap R_n| \geq k) \leq Ce^{-k/C}.$$

This is an immediate consequence of lemma 5.3 of [Cho23b]. (maybe this is actually Lemma 4.9 of the arxiv version of [Cho23b]?) □

This concludes the proof of Theorem 1.2.

I have a question: do we know of ANY random walk that is transient but does not satisfy a CLT for range? I think this may be unknown

Whenever you have a positive density of cut times, I think you should be able to make a regeneration-type argument to prove a CLT for the range, so you need to look for transient random walks which travel sublinearly, maybe $\mathbb{Z}^3$ is a good candidate?

Yes indeed, on Z^3 the variance of the range is not like $\sqrt{n}$. However, with the appropriate normalization there is a CLT; this is proved here [JP70, JP71]. Maybe one can construct groups that interpolate between $\mathbb{Z}^3$ and something else to find a group that has different limit laws along different subsequences?

References

[Bar22] Corentin Le Bars. Central limit theorem on cat(0) spaces with contracting isometries, 2022. [Cited on page 1.]

[Bjö10] Michael Björklund. Central limit theorems for Gromov hyperbolic groups. *J. Theoret. Probab.*, 23(3):871–887, 2010. [Cited on page 1.]

[BQ16a] Y. Benoist and J.-F. Quint. Central limit theorem on hyperbolic groups. *Izv. Math.*, 80(1):3–23, 2016. [Cited on page 1.]

[BQ16b] Yves Benoist and Jean-François Quint. Central limit theorem for linear groups. *Ann. Probab.*, 44(2):1308–1340, 2016. [Cited on page 1.]

[Cho23a] Inhyeok Choi. Central limit theorem and geodesic tracking on hyperbolic spaces and Teichmüller spaces. *Adv. Math.*, 431:68, 2023. Id/No 109236. [Cited on page 1.]

[Cho23b] Inhyeok Choi. Random walks and contracting elements I: Deviation inequality and limit laws, 2023. [Cited on page 7.]

[EZ22] Anna Erschler and Tianyi Zheng. Law of large numbers for the drift of the two-dimensional wreath product. *Probab. Theory Related Fields*, 182(3-4):999–1033, 2022. [Cited on page 1.]

[Gil08] L. A. Gilch. Acceleration of lamplighter random walks. *Markov Process. Related Fields*, 14(4):465–486, 2008. [Cited on page 1.]

[Gou17] Sébastien Gouëzel. Analyticity of the entropy and the escape rate of random walks in hyperbolic groups. *Discrete Anal.*, 2017:37, 2017. Id/No 7. [Cited on page 1.]

[Gou22] Sébastien Gouëzel. Exponential bounds for random walks on hyperbolic spaces without moment conditions. *Tunis. J. Math.*, 4(4):635–671, 2022. [Cited on page 4.]

[GTT22] Ilya Gekhtman, Samuel J. Taylor, and Giulio Tiozzo. Central limit theorems for counting measures in coarse negative curvature. *Compos. Math.*, 158(10):1980–2013, 2022. [Cited on page 1.]

[Ham62] J. M. Hammersley. Generalization of the fundamental theorem on sub-additive functions. *Proc. Cambridge Philos. Soc.*, 58:235–238, 1962. [Cited on page 2.]

[Hor18] Camille Horbez. Central limit theorems for mapping class groups and Out(F_N). *Geom. Topol.*, 22(1):105–156, 2018. [Cited on page 1.]

[JP70] Naresh C. Jain and William E. Pruitt. The central limit theorem for the range of transient random walk. *Bull. Amer. Math. Soc.*, 76:758–759, 1970. [Cited on page 7.]

[JP71] Naresh C. Jain and William E. Pruitt. The range of transient random walk. *J. Analyse Math.*, 24:369–393, 1971. [Cited on page 7.]

[Led01] François Ledrappier. Some asymptotic properties of random walks on free groups. In *Topics in probability and Lie groups: boundary theory*, volume 28 of *CRM Proc. Lecture Notes*, pages 117–152. Amer. Math. Soc., Providence, RI, 2001. [Cited on page 1.]

[MS20] P. Mathieu and A. Sisto. Deviation inequalities for random walks. *Duke Math. J.*, 169(5):961–1036, 2020. [Cited on pages 2 and 7.]

[MSŠ23] Rudi Mrazović, Nikola Sandrić, and Stjepan Šebek. Capacity of the range of random walks on groups. *Kyoto J. Math.*, 63(4):783–805, 2023. [Cited on page 1.]

[Par92] Walter Parry. Growth series of some wreath products. *Trans. Amer. Math. Soc.*, 331(2):751–759, 1992. [Cited on page 1.]

[Sal01] François Salaün. Marche aléatoire sur un groupe libre: théorèmes limite conditionnellement à la sortie. *C. R. Acad. Sci. Paris Sér. I Math.*, 333(4):359–362, 2001. [Cited on page 1.]

[SS87] Stanley Sawyer and Tim Steger. The rate of escape for anisotropic random walks in a tree. *Probab. Theory Related Fields*, 76(2):207–230, 1987. [Cited on page 1.]