Electron. Commun. Probab. **28** (2023), article no. 36, 1–12.
https://doi.org/10.1214/23-ECP546
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Effective drift estimates for random walks on graph products

Kunal Chawla*

Abstract

We find uniform lower bounds on the drift for a large family of random walks on graph products, of the form $\mathbb{P}(|Z_n| \le \kappa n) \le e^{-\kappa n}$ for $\kappa > 0$. This includes the simple random walk for a right-angled Artin group with a sparse defining graph. This is done by extending an argument of Gouëzel, along with the combinatorial notion of a piling introduced by Crisp, Godelle, and Wiest. We do not use any moment conditions, instead considering random walks which alternate between one measure uniformly distributed on vertex groups, and another measure over which we make no assumptions.

Keywords: right-angled Artin group; random walk; drift; hyperbolic group; graph product; pivoting.
MSC2020 subject classifications: 60J10; 60B15.
Submitted to ECP on June 11, 2022, final version accepted on September 2, 2023.

1 Introduction

Suppose that G is a group acting on a metric space X, and that G is equipped with a probability measure μ. If $g_1, g_2, \dots$ are i.i.d. G-valued random variables with distribution μ, one can construct a random walk on X by picking a basepoint $o \in X$ and letting

$$Z_n \cdot o = g_1 \dots g_n \cdot o.$$

Often considered in the literature is qualitative long-term behaviour of Z_n. Furstenberg showed that random walks on semi-simple Lie groups converge almost surely to a point on a natural boundary at infinity [Fur63]. Kaimanovich identified the Poisson boundary for a general class of groups with hyperbolic properties [Kai00]. Karlsson and Margulis showed that certain random walks on Busemann non-positively curved spaces sublinearly track a geodesic [KM99], and Tiozzo exhibited a general condition to ensure sublinear tracking [Tio15]. Benoist and Quint [BQ16] exhibited a central limit theorem for random walks with finite variance on Gromov hyperbolic groups. Maher and Tiozzo showed that a non-elementary random walk on a (not necessarily proper) hyperbolic space converges to the boundary [MT18]. Nevo and Sageev identified the Poisson boundary for groups acting on CAT(0) cube complexes [NS13]. Most of these results rely on geometric assumptions about the group, usually some sort of negative curvature condition, as well as moment or entropy assumptions on μ.

*University of Toronto, Canada. E-mail: kunal.chawla@mail.utoronto.ca

In the recent literature are inquiries into large deviations principles for random walks on hyperbolic spaces. Let X be a Gromov hyperbolic G-space with a basepoint o. Maher and Tiozzo showed that if μ has finite support, then $\mathbb{P}(d(Z_n o, o) \leq \kappa n)$ decays exponentially for some κ. This was upgraded to an exponential moment condition by [Sun20]. Later, in [BMSS20] it was shown that this statement holds for all κ up to the rate of escape

$$\ell := \lim_{n\to\infty} \frac{\mathbb{E}\left[d(Z_n o, o)\right]}{n}.$$

Recently, Gouëzel [Gou23] has shown, with a clever geometric argument, that all moment assumptions can be removed. This argument does not rely on boundary theory, and is entirely quantitative. The idea is that one can decompose a sample path into segments which go in one direction, and 'pivots' where the random walk might travel in one of many directions. By hyperbolicity, in most directions the sample path will move further away from the basepoint. One can bound the number of pivotal points from below by a sum of i.i.d. random variables with positive expectation, and deduce linear progress with exponential decay. This type of argument has recently been used to explore genericity of pseudo-Anosovs [Cho21b], prove limit laws [Cho22], and identify the Poisson boundary for groups with WPD actions on hyperbolic spaces [CFFT22]. Similar arguments have appeared in the literature before (e.g. [BMSS20] or the notion of "persistent segments" in [MT18]). Essentially, this argument allows one to show that many statistics of random walk behave like a sum of i.i.d. random variables. Aoun and Sert show in [AS22], among other results, that the distance traveled by random walks on hyperbolic spaces admit subgaussian concentration bounds. In addition, Corso has recently [Cor21a, Cor21b] exhibited large deviations principles for free products of finitely generated groups and for relatively hyperbolic groups.

We apply this technique to give effective estimates for the drift of certain random walks on graph products. Furstenberg exhibited an integral formula for the drift, however this is requires knowledge about the harmonic measure and so is not amenable to computation. Furstenberg also used non-amenability to show positivity of drift, which follows by examining the spectrum of the averaging operator. This approach can be modified to get lower bounds for the drift in terms of spectral analysis of the averaging operator (see [Vir80, Nev03] or [AS22, Proposition 6.8]). In this paper, we consider a class of random walks on graph products of groups acting on their Cayley graphs, which are not usually hyperbolic. Let Γ be a graph, with vertex set V and edge set E. To each vertex $v \in V$ we associate a group G_v with the (not necessarily finite) presentation $G_v = \langle S_v | R_v \rangle$. Here S_v is a generating set for G_v, and R_v is a collection of relations. The graph product, denoted by $G = G(\Gamma)$, is the group defined by

$$G(\Gamma) = \langle \sqcup_{v\in V} S_v | \sqcup_{v\in V} R_v \sqcup_{(v,w)\in E} [S_v, S_w] \rangle.$$

In other words, two vertex groups G_v and G_w commute if and only if v and w are adjacent. For example, if the graph is a clique, then G is the direct product $G_1 \times \cdots \times G_n$. If the graph has no edges, then G is the free product $G_1 * \cdots * G_n$. If the graph is a path with 3 vertices and $G_1 = G_2 = G_3 = \mathbb{Z}$, then $G = F_2 \times \mathbb{Z}$, as shown in the following schematic:

$$\mathbb{Z} \;\text{——}\; \mathbb{Z} \;\text{——}\; \mathbb{Z}$$

Figure 1: A representation of $F_2 \times \mathbb{Z}$ as a graph product of three copies of $\mathbb{Z}$.

Graph products need not be finitely generated or hyperbolic, for example if each vertex group is an infinite direct sum of copies of $\mathbb{Z}$. If each G_i is a copy of $\mathbb{Z}$, then G

is the right-angled Artin group (RAAG) on the graph in question. The graph product interpolates between the direct product and free product, where a sparse graph means that G is closer to a free product. Graph products were introduced by Green and are simple examples of non-hyperbolic groups with certain hyperbolic properites, as well as nice algorithmic properties [Gre90].

Given a graph Γ, let D be the number of vertices and B the maximum size of a 1-neighbourhood of a clique. For example, if Γ is a sufficiently long cycle, then $B = 4$. We say that Γ has *small cliques* if $D > 5B$. Moreover, we say that a probability measure on a graph product G is *alternating* if it is of the form $\mu * \nu$ where $\mu(G_v \setminus \{e\}) = 1/D$ for any $v \in V$, and ν is any probability measure on G. For example, if μ is the measure driving the simple random walk on a right-angled Artin group, then $\mu * \delta_e$ is alternating, where e is the identity element. This is because each vertex group has the same number of generators. The significance of picking an alternating measure is explained in our proof sketch below.

Also, given an element $g \in G$, let the *syllable length* $|g|$ be the minimum length of a representation $g = g_1 \cdots g_n$ where g_i, g_{i+1} are elements of distinct vertex groups for all $1 \leq i < n$.

In this article, we prove the following:

Theorem 1.1. *Let Γ be a graph with small cliques and let $G(\Gamma)$ be a graph product with vertex groups $G_1, \dots, G_D$. Then there exists an effective constant $\kappa = \kappa(\Gamma) > 0$ such that for any random walk (Z_n) driven by an alternating measure, we have*

$$\mathbb{P}(|Z_n| \leq \kappa n) \leq e^{-\kappa n} \qquad \text{for any } n \in \mathbb{N}.$$

We will see in Lemma 4.1 that κ can be effectively computed in terms of B and D. For example, if Γ is a D-cycle for $D > 20$ then $B = 4$, and $\kappa \geq 0.3$. In this example, we will see in proposition 4.3 that $\kappa \to 1$ as $D \to \infty$. This is asymptotically sharp, as seen by examining the simple random walk on a right-angled Artin group.

In the case where G is a Right-angled Artin group, this theorem gives a quantitative sense in which these RAAGs are closer to being a free group than a free abelian group. Indeed, if $d : G \times G \to \mathbb{N}$ is a word distance induced by a generating set where each generator lies inside a vertex group, then $d(e, g) \geq |g|$, so we can draw conclusions about the drift of a random walk on a RAAG. The use of an alternating random walk $\mu * \nu$ is notable because the measure ν is allowed to have arbitrarily fat tails. All of the regularity comes from the μ. This choice of κ is then uniform over a large class of random walks. Instead of relying on moment conditions to control backtracking, we combine some ideas from Gouëzel's argument with a combinatorial tool. In particular, we extend the notion of a 'piling' from [CGW08] to define pivotal points in the graph product setting.

Proof Sketch. To prove our main theorem, we write $g_i = s_i w_i$ where $s_i \sim \mu$ and $w_i \sim \nu$. We condition on the w_i's and keep the randomness coming from the s_i's. To each pair w_{n-1}, w_n ending and starting in words coming from certain vertex groups, we can find a s_n from a vertex far away on the graph, so that $|Z_n| > |Z_{n-1}|$. As the graph is sparse, there are many such choices for s_n uniformly over any realization of w_n's. If w_{n+1} causes large cancellations in $Z_n w_{n+1}$, we argue there are many other choices for s_n for which w_{n+1} does not cause cancellations. Hence we can bound the syllable length $|Z_n|$ from below by a sum of n i.i.d. copies of some random variable U with positive expectation. □

The paper is organized as follows: in Section 2 we define pilings and introduce the notion of a terminal and initial clique for elements of graph products. In Section 3 we describe the notion of pivotal points, inspired by Gouëzel. Finally, in Section 4, we prove our main theorem and state a formula for our drift estimate.

2 Pilings

The notion of a piling was introduced by Crisp, Godelle, and Wiest in [CGW08] to give a normal form for right-angled Artin groups. They used pilings to solve the conjugacy problem in this setting.

Since right-angled Artin groups interpolate between free groups and free abelian groups, one looks for a way to quantify how close a RAAG is to either extreme. One way to do this is to explore when a word in a RAAG locally looks like a word in a free product. Consider for example the group $\mathbb{Z}^2 * \mathbb{Z} = \langle a, b, c | [a, b] \rangle$. Then the word aba^{-1} can be shortened, whereas aca^{-1} cannot. Now consider a word of the form $acbscba^{-1}$ for some s chosen randomly from $\{a^{\pm}, b^{\pm}, c^{\pm}\}$. We want to estimate the probability with which the word can be shortened. This is the role of pilings in our argument.

We start off by extending the definition of pilings to graph products. We explain how to produce a piling for a word in $\sqcup_{i=1}^{n} G_i$, then show that this is independent of the choice of word representative. This will produce a well-defined piling for an element of G.

Let $\mathcal{A} = \sqcup_{i=1}^{D} G_i \setminus \{e\}$, and let $\mathcal{B} = \mathcal{A} \sqcup \{0\}$. Let $\mathcal{A}^\star$ (resp. $\mathcal{B}^\star$) the set of finite words in the alphabet $\mathcal{A}$ (resp. $\mathcal{B}$). We denote as ϵ the empty word. We now define a piling map $\Pi : \mathcal{A}^* \to (\mathcal{B}^*)^D$.

Definition 2.1. *Let G be a graph product with vertex groups $G_1, \dots, G_D$. A* piling $\Pi(h)$ *for a word $h \in \mathcal{A}^\star$ is an ordered list of D words in the alphabet $\mathcal{B}$, defined inductively as follows:*

- *The piling $\Pi(\varepsilon)$ for the empty word is $(\varepsilon, \dots, \varepsilon)$.*
- *If $h = h' g_i$, where $g_i \in G_i \setminus \{e\}$, then*

  1. *If the ith word of $\Pi(h')$ is empty or ends in a 0, then the ith word of $\Pi(h)$ is given by appending g_i to the ith word of $\Pi(h')$, and a 0 to the jth word for every j such that vertices v_i and v_j are not adjacent.*
  2. *If the ith word of $\Pi(h')$ ends in an element g_i' of G_i, compute $g = g_i' g_i$. If g is a nontrivial element of G_i, then the ith word of $\Pi(h)$ is given by replacing the last letter g_i' with g. If $g_i' g_i$ is the identity, then the ith word of $\Pi(h)$ is given by erasing g_i' from the ith word of $\Pi(h')$, and all other words of $\Pi(h)$ are given by removing the final 0 on the jth words, where v_i and v_j are not adjacent.*

Example 1. Consider the group $\mathbb{Z}^2 * \mathbb{Z} = \langle a, b, c | [a, b] \rangle$. Then

- The piling for a is $(a, \epsilon, 0)$.
- The piling for ac is $(a0, 0, 0c)$.
- The piling for acb is $(a0, 0b, 0c0)$.
- The piling for $acba$ is $(a0a, 0b, 0c00)$.
- The piling for $acbaa^{-1}$ is $(a0, 0b, 0c0)$.
- Meanwhile, the piling for $acbaca^{-1}$ is $(a00a^{-1}, 0b0, 0c00c^{-1}0)$.

Observe that the piling for $acbaa^{-1}$ is equal to that of acb.

One can see that the last occurence of a c presents as a barrier to the cancellation $aa^{-1} \to e$. In the context of random walks, we will use pilings to argue that there are many such barriers with high probability. To make this rigorous, we must verify that the choice of piling of a group element does not depend on its word representation.

Proposition 2.2. *The piling map $\Pi : \mathcal{A}^\star \to (\mathcal{B}^\star)^D$ induces a well-defined map $G \to (\mathcal{B}^\star)^D$.*

Proof. We need to show that for any two words $h_1, h_2 \in \mathcal{A}^*$ that represent the same group element, the pilings $\Pi(h_1)$ and $\Pi(h_2)$ are equal. Observe that for any $h, h' \in \mathcal{A}^*$ the following holds:

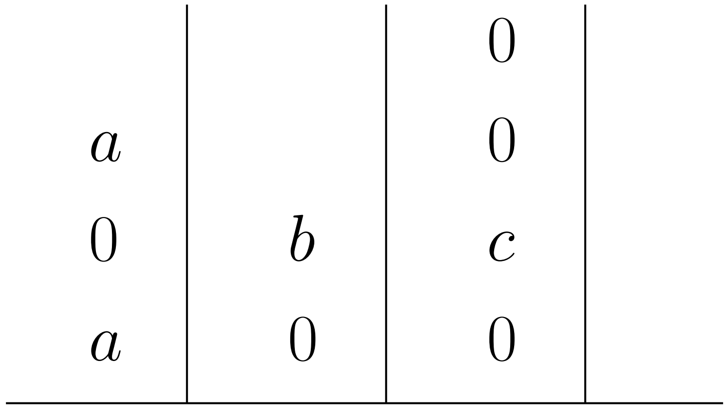

Figure 2: A piling for $acba$.

1. If s_i and s_j are elements of the adjacent groups G_i and G_j, then $\Pi(hs_is_jh') = \Pi(hs_js_ih')$.
2. For any $s_i \in G_i$, we have $\Pi(hs_is_i^{-1}h') = \Pi(hh')$.
3. If s_i and s_i' are element of the same vertex group, with $s_is_i' = s_i''$ nontrivial, then $\Pi(hs_is_i'h') = \Pi(hs_i''h')$. If s_i'' is trivial, then $\Pi(hs_is_i'h') = \Pi(hh')$.

Now let $h_1 = g_1g_2...g_m$ and $h_2 = g_1'...g_n'$ be two words in $\mathcal{A}^*$ which represent the same group element in G. Since both words represent the same group element, then we can obtain h_2 from h_1 by some sequence of the following moves:

1. If g_i and g_{i+1} are elements of adjacent vertex groups, replace $...g_ig_{i+1}...$ with $...g_{i+1}g_i...$.
2. If g_i, g_{i+1} are from the same vertex group, then compute $g_i'' = g_ig_{i+1}$. If g_i'' is trivial, replace $...g_ig_{i+1}...$ with $...\varepsilon...$ If g_i'' is nontrivial, replace $...g_ig_{i+1}...$ with $...g_i''...$.
3. If $g_i = g'g''$ for two nontrivial elements in the vertex group G_i, then replace $...g_i...$ with $...g'g''...$.

All these moves leave the piling unchanged. □

From now on, when we refer to a piling, we are referring to the induced map $G \to (\mathcal{B}^*)^D$. In an abuse of notation, we also denote this map by Π.

Remark 2.3. The notion of piling resembles that of a “pruning” in [HM95]. A pruning of an element g in a graph product is a representation of g as a product of elements in vertex groups, which is minimal with respect to syllable length and with respect to a ShortLex ordering induced by an ordering of the vertices. Like pilings, it can be used to produce a normal form for an element of a graph product [HM95, Proposition 3.1], and the proof of this is essentially the same as Proposition 2.2. However, a piling is a collection of words indexed by the vertices of a graph, while a pruning is one word.

Remark 2.4. Observe that the piling for $\Pi(g^{-1})$ is given by reversing all words in $\Pi(g)$ and swapping each g_i with g_i^{-1}. This can be proven by induction on the syllable length of g.

Definition 2.5. *Given an element $g \in G$, the terminal clique $\mathrm{term}(g)$ is the set of vertices v_i in the graph such that the ith word in the piling $\Pi(g)$ ends in a nontrivial element of G_i.*

We define the *initial clique* $\mathrm{init}(g)$ to be the set of vertices v_i such that the ith word of $\Pi(g)$ starts with a nontrivial element of g_i. The terminal and initial cliques of the trivial element are the empty set. Observe that $\mathrm{init}(g) = \mathrm{term}(g^{-1})$, which we leverage in the proof of the following lemma:

Lemma 2.6. *Let g, h be elements in the graph product G such that $\mathrm{term}(g) \cap \mathrm{init}(h) = \varnothing$. Then $\Pi(gh) = \Pi(g)\Pi(h)$.*

Proof. We induct on the syllable length of h.

In our base case, h is just an element from one vertex group G_i, so that $\mathrm{init}(h) = \{v_i\}$. As this set is disjoint from $\mathrm{term}(g)$, we know that the piling for g ends with a 0 in the ith word, or the ith word is empty. In either case, the piling $\Pi(gh)$ is given by concatenating $\Pi(g), \Pi(h)$.

Now suppose that the syllable length for h is equal to $n > 1$, and let $h = w_1 \dots w_n$ be a representation of minimal syllable length. Let v be the vertex corresponding to w_1.

We claim that the initial clique of $w_1^{-1}h$ is disjoint from $\{v\}$. If w_1 is the only letter w_i lying in the vertex group G_v, the claim holds. Now suppose that w_1 and w_k lie in the same vertex group, then we can find some i strictly between 1 and k such that v and v_i are not adjacent. If such an i did not exist, we could apply our commutation relations to see that $w_1w_2...w_k...w_n = (w_1w_k)w_2...w_{k-1}w_{k+1}...w_n$ as group elements, contradicting minimality of our representation. Thus the piling for h, in some coordinate, begins with $w_1 0$. As a result, the piling for $w_1^{-1}h$ begins with a 0 in the coordinate corresponding to w_1, so that the initial clique of $w_1^{-1}h$ is disjoint from $\{v\}$.

Applying our induction hypothesis, since the syllable length of $w_1^{-1}h$ is smaller than that of h, we have

$$\begin{aligned}\Pi(g)\Pi(h) &= \Pi(g)\Pi(w_1w_1^{-1}h)\\ &= \Pi(g)\Pi(w_1)\Pi(w_1^{-1}h)\\ &= \Pi(gw_1)\Pi(w_1^{-1}h)\\ &= \Pi(gh)\end{aligned}$$

□

As a result of this lemma, we can read off the syllable length of an element $g \in G$ from its piling.

Lemma 2.7. *Let $g \in G$. The syllable length of g is equal to the number of nonzero characters in the piling $\Pi(g)$.*

In other terms, for each of the D words in $\Pi(g)$, we count the length of the word minus the number of occurences of the character 0, and add this up over all D words.

Proof. Like before, we induct on the syllable length of g. When $|g| = 1$, then $g \in G_{v_i}$, so that the piling $\Pi(g)$ consists of the character g in the ith string, along with some number of 0's. Now let $|g| = n$, and let $g' \in G_v$ for some vertex group G_v, such that $|gg'| = n + 1$. Let $g = g_1...g_n$ be a representation of g where each g_i, g_{i+1} lie in different vertex groups. Every group element h of syllable length $n + 1$ can be obtained in this fashion: write $h = h_1...h_{n+1}$ where h_i, h_{i+1} are elements of distinct vertex groups, then pick $g = h_1...h_n$ and $g' = h_{n+1}$.

Since $gg' = g_1...g_ng'$ has syllable length $n + 1$, then this representation of gg' must be of minimal length. Then it must be the case that $\mathrm{term}(g) \cap \mathrm{init}(g') = \varnothing$, because otherwise we could apply our commutation relations to get a representation of shorter length. Therefore $\Pi(gg') = \Pi(g)\Pi(g')$ and so the number of nonzero characters in $\Pi(gg')$ is equal to $n + 1$. □

3 Pivotal times

Our plan is to control the behaviour of a sample path by considering times where it goes in independent directions. In the free group, there are points where the sample path lies in some subtree forever after some time n. Intuitively, we should be able to pivot a sample path about such a point, and get another sample path with the same drift. We plan to show that there are many such points, and that for each point there are many directions in which the sample path moves further away from the identity.

Recall that the paths for our walks are of the form $Z_n = s_1 w_1 \dots s_n w_n$ where $s_i \sim \mu$ and $w_i \sim \nu$ for $i = 1, \dots, n$.

Definition 3.1. *Given a finite sequence $(s_1, w_1, \dots, s_n, w_n)$ of length $2n$ with all s_i's nontrivial elements of some vertex groups, and a set of times $P \subset \{1, \dots, n\}$, the sequence is* pivotal *with respect to P if*

- *For all $k \in P$, the vertex $\mathrm{init}(s_k)$ is not adjacent to $\mathrm{term}(Z_{k-1})$*
- *If $k < k'$ are subsequent elements of P and $w_k ... w_{k'-1}$ is nontrivial, then the vertex $\mathrm{term}(s_k)$ is not adjacent to $\mathrm{init}(w_k ... w_{k'-1})$. If $w_k ... w_{k'-1}$ is trivial, then $\mathrm{term}(s_k)$ is not adjacent to $\mathrm{init}(s_{k'})$.*
- *If k is the greatest element of P, then the terminal clique of s_k is not adjacent to $\mathrm{init}(w_k ... w_n)$.*

By "adjacent" we mean "adjacent to or contained in". Observe that there may be multiple choices of P with respect to which the sequence is pivotal. For any such choice P, letting $k_1, \dots, k_{\#P}$ be our collection of pivotal times, these three conditions along tell us that $\mathrm{term}(Z_{k-1} s_k) = \mathrm{term}(s_k)$, so applying Lemma 2.7 allows us to deduce

$$0 \leq |Z_{k_1 - 1}| < |Z_{k_1 + 1} s_{k_1}| \leq |Z_{k_2 - 1}| < |Z_{k_2 - 1} s_{k_2}| \leq \dots \leq |Z_{k_{\#P} - 1}| < |Z_{k_{\#P-1}} s_{k_{\#P}}| \leq |Z_n|.$$

We summarize this in the following:

Lemma 3.2. *If $(s_1, w_1, \dots, s_n, w_n)$ is pivotal with respect to $P \subset \{1, \dots, n\}$, then the syllable length $|Z_n|$ is at least $\#P$.*

To estimate the distance travelled by our random walk, it suffices to estimate the maximal size of a set of pivotal times. To do this, we argue that pivotal times are in a quantitative sense persistent: once the time k is pivotal for the random walk at the kth step, is it very likely to remain pivotal up to time n.

In the proof of the main theorem, we will condition on the group elements $w_1, \dots, w_n$ coming from the measure ν, and keep the randomness from the group elements $s_1, \dots, s_n$. Hence to understand our random walk, we should consider the sequence $(s_1, \dots, s_n)$ associated to some sample path.

For a fixed sequence $(w_1, \dots, w_n)$ and a subset $P \subset \{1, \dots, n\}$, denote by $\mathcal{E}(P)$ the collection of sequences $(s_1, \dots, s_n)$ such that $(s_1, w_1, \dots, s_n, w_n)$ is pivotal with respect to P. We argue that for any P, these collections are large.

Lemma 3.3. *Fix $(w_1, \dots, w_n)$ and $P \subset \{1, \dots, n\}$, and suppose that $(s_1, \dots, s_n) \in \mathcal{E}(P)$. Then for any $k \in P$, there exist cliques $C_1, C_2 \subset \Gamma$ such that if s_k is replaced by any nontrivial s'_k with $\mathrm{init}(s'_k)$ not adjacent to $C_1 \cup C_2$, then $(s_1, \dots, s'_k, \dots, s_n) \in \mathcal{E}(P)$.*

Proof. We set $C_1 = \mathrm{term}(Z_{k-1})$. If k is not the last pivotal time, choose $C_2 = \mathrm{init}(w_k ... w_{k'-1})$ if $w_k ... w_{k'-1}$ is not trivial, else $C_2 = \mathrm{init}(s_{k'})$ where k' is the next pivotal time. If k is the last pivotal time, pick $C_2 = \mathrm{init}(w_k ... w_n)$. We claim that these choices work.

Say for some $k \in P$, we replace s_k with some nontrivial s'_k satisfying $\mathrm{init}(s'_k) \notin C_1 \cup C_2$.

By construction, s'_k still satisfies the conditions imposed on s_k for k to be a pivotal time.

If $j \in P$ is a pivotal time strictly before k, we have $\mathrm{init}(s_j)$ is not adjacent to $\mathrm{term}(Z_{j-1})$. Let $j' \in P$ be the pivotal time immediately after j. If the intermediate word $w_j...w_{j'-1}$ is not trivial or $j' \neq k$, then s_j still satisfies the same conditions imposed for j to be a pivotal time. If the intermediate word is trivial and $j' = k$, then since $\mathrm{term}(Z_{k-1}) = \mathrm{term}(Z_{j-1}s_j) = \mathrm{term}(s_j)$, so by our choice of s'_k we know that $\mathrm{term}(s_j)$ is not adjacent to $\mathrm{init}(s'_k)$.

If $j > k$, then the second and third conditions in Definition 3.1 are immediately met, so it remains to check the first condition for any pivotal times $j > k$. First suppose that j is the pivotal time immediately after k. If the intermediate word $w_k...w_{j-1}$ is trivial, we have $Z_{j-1} = Z_{k-1}s'_k$, so that $\mathrm{init}(s_j)$ is not adjacent to $\mathrm{term}(Z_{j-1}) = \mathrm{term}(s'_k)$. If the intermediate word is not trivial, then the fact that $\mathrm{term}(Z_{k-1}s'_k) = \mathrm{term}(s'_k)$ and $\mathrm{term}(s'_k) \cap \mathrm{init}(w_k...w_{j-1}) = \varnothing$ implies $\mathrm{term}(Z_{k-1}s'_k...w_{j-1}) = \mathrm{term}(w_k...w_{j-1})$ and likewise for $\mathrm{term}(Z_{j-1})$. Since $\mathrm{init}(s_j)$ is not adjacent to $\mathrm{term}(Z_{j-1})$, we deduce that it is also not adjacent to $\mathrm{term}(Z_{k-1}s'_k...w_{j-1})$.

Since this argument only used the facts that $\mathrm{init}(s'_k)$ is not adjacent to $\mathrm{term}(Z_{k-1})$ or the appropriate choice of $\mathrm{init}(s_{k'}), \mathrm{init}(w_k...w_{k'-1}), \mathrm{init}(w_k...w_n)$, then the same argument shows that every pivotal time $j > k$ remains pivotal. ☐

4 Main argument

In this Section we prove Theorem 1.1. First we recall some notation. Let G be a graph product with D vertices and vertex groups $G_1, \dots, G_D$. Let B be the maximum size of the 1-neighbourhood $N_1(K)$ where $K \subset \Gamma$ ranges through all cliques. Let μ be a probability measure on G such that $\mu(G_i \setminus \{e\}) = \frac{1}{D}$. In other words, μ is equally likely to pick out a nontrivial element of any group. Let ν be an arbitrary probability measure on G and consider the random walk driven by $\mu * \nu$.

We want to show that there exists some $\kappa > 0$ such that

$$\mathbb{P}(|Z_n| \leq \kappa n) \leq e^{-\kappa n},$$

where $|Z_n|$ denotes the syllable length of Z_n. Write $Z_n = s_1 w_1 ... s_n w_n$ where $s_i \sim \mu$ and $w_i \sim \nu$. We will condition on the w_n's and keep the randomness coming from the s_n's. Then we will find a working κ that is independent of our conditioning. To this end, we will use our assumptions on the graph to show that the drift is bounded from below by a sum of i.i.d. variables with positive expectation.

For any $n \in \mathbb{N}$, a sequence $(w_1, \dots, w_n)$, and a random choice of $s_1, \dots, s_n$ drawn i.i.d. according to μ, let A_n be the random variable that is the maximum cardinality of a subset $P \subset \{1, \dots, n\}$ with respect to which $(s_1, \dots, w_n)$ is pivotal. In the following lemma we argue that the small cliques assumption implies that there are many pivotal times.

Lemma 4.1. *Suppose that $D > 5B$. Fix $\{w_i\}_{1 \leq i \leq n+1}$. Also let U be an integer-valued random variable, independent of $(s_1, \dots, s_{n+1})$ and with distribution*

$$\mathbb{P}(U = 1) = \frac{D - 2B}{D},$$

and

$$\mathbb{P}(U \leq -j) = \frac{2B}{D} \cdot \left(\frac{B}{D - 2B} \right)^j$$

for all $j \geq 0$. Then A_{n+1} stochastically dominates $A_n + U$ in the sense that

$$\mathbb{P}(A_{n+1} \geq i) \geq \mathbb{P}(A_n + U \geq i) \qquad \text{for all } i.$$

Proof. Fix a subset $P = (k_1, \dots, k_q) \subset \{1, \dots, n\}$ and condition on the event that $(s_1, \dots, s_n) \in \mathcal{E}(P)$.

For any $(s_1, \ldots, s_n) \in \mathcal{E}(P)$, there are at least $D - B$ vertices v_i which are not in or adjacent to $\operatorname{term}(Z_n)$. Likewise, we know that there are at most B vertices contained in or adjacent to $\operatorname{init}(w_{n+1})$, therefore there are at least $D - B$ choices for $G_{i_{n+1}}$ such that s_{n+1} is not in a vertex group adjacent to that of w_{n+1}. Hence there are at least $D - 2B$ choices for $G_{i_{n+1}}$ such that $(s_1, \ldots, w_{n+1})$ is pivotal with respect to $P \cup \{n+1\}$. As our probability distribution μ is uniform over our set of groups then

$$\mathbb{P}(A_{n+1} \geq A_n + 1 | \mathcal{E}(P)) \geq \frac{D - 2B}{D}.$$

Now fix $j > 0$ and consider the event that $A_{n+1} \leq A_n - j$. This first requires that s_{n+1} is not disjoint from one of $\operatorname{term}(Z_n)$ or is adjacent to $\operatorname{init}(w_{n+1})$, which happens with probability at most $\frac{2B}{D}$. For k_q to no longer be pivotal, this requires that $\operatorname{term} s_{k_q}$ is now adjacent to the initial clique of $w_{k_q} \ldots s_{n+1} w_{n+1}$, which happens with probability at most $\frac{B}{D-2B}$. Conditional on the event that $(s_1, \ldots, s_n) \in \mathcal{E}(P)$, the group elements s_{n+1} and s_{k_q} are independent. We remark that the conditioning is only necessary because we make reference to the final pivotal time k_q, and independence is because $k_q \neq n+1$ by definition. Hence

$$\mathbb{P}(A_{n+1} \leq A_n - 1 | \mathcal{E}(P)) \leq \frac{2B}{D} \cdot \frac{B}{D - 2B}.$$

For k_{q-1} to no longer be pivotal, this requires that $\operatorname{term}\left(s'_{k_{q-1}}\right)$ is now adjacent to $\operatorname{init}(w_{k_{q-1}} ... s_{n+1} w_{n+1})$. Conditional on the event that $(s_1, \ldots, s_n) \in \mathcal{E}(P)$, and a choice of s_{k_q}, s_{n+1} such that $(s_1, \ldots, w_{n+1})$ fails to be pivotal with respect to P, this has probability at most $\frac{B}{D-2B}$. For any choice of s_{k_q} such that the sequence $(s_1, \ldots, w_{n+1})$ is pivotal with respect to P, this has probability 0. Hence we have

$$\mathbb{P}(A_{n+1} \leq A_n - 2 | \mathcal{E}(P)) \leq \frac{2B}{D} \cdot \left(\frac{B}{D - 2B}\right)^2$$

Continuing in this fashion, we have

$$\mathbb{P}(A_{n+1} \leq A_n - j | \mathcal{E}(P)) \leq \frac{2B}{D} \cdot \left(\frac{B}{D - 2B}\right)^j$$

for all $j > 0$.

As this bound is uniform over conditioning on $\mathcal{E}(P)$, we have the conclusion of the lemma. □

We also make use of the following probabilistic lemma.

Lemma 4.2. *Let $U_1, \ldots, U_n$ be i.i.d. copies of a random variable U and let $t > 0$ be such that $\mathbb{E}[e^{-tU}] < 1$. Then for $\kappa < \frac{-\ln \mathbb{E}[e^{-tU}]}{1+t}$ we have*

$$\mathbb{P}(U_1 + \cdots + U_n \leq \kappa n) \leq e^{-\kappa n}.$$

Proof. By Markov's inequality and independence we have, for any $\kappa > 0$,

$$\mathbb{P}(U_1 + \cdots + U_n \leq \kappa n) \leq e^{t\kappa n} \left(\mathbb{E}\left[e^{-tU}\right]\right)^n.$$

If we pick

$$\kappa < \frac{-\ln \mathbb{E}[e^{-tU}]}{1+t},$$

then the right hand side is less than $e^{-\kappa n}$. □

Now we are ready to prove our main theorem.

Proof. We have

$$\mathbb{E}[U] = \frac{D-2B}{D} - \frac{2B}{D}\left(\frac{D-2B}{D-3B}\right).$$

If $D > 5B$, then $\mathbb{E}[U] > 0$. Then $\frac{d}{dt}|_{t=0^+}\mathbb{E}[e^{-tU}] < 0$, so there exists some $t > 0$ such that $\mathbb{E}[e^{-tU}] < 1$. Then there exists some positive κ with

$$\kappa < \frac{-\ln \mathbb{E}[e^{-tU}]}{1+t}.$$

Now let $U_1, \ldots, U_n$ be n i.i.d. copies of U. Condition on all of the steps $\{w_i\}_{i\geq 1}$. Iterating Lemma 4.1 we get that the random variable A_n, defined conditionally on $\{w_i\}_{i\geq 1}$, stochastically dominates $U_1 + \cdots + U_n$. Hence by our large deviations bound we have

$$\mathbb{P}(|Z_n| \leq \kappa n | \{w_i\}_{i\geq 1}) \leq \mathbb{P}(A_n \leq \kappa n | \{w_i\}_{i\geq 1}) \leq \mathbb{P}(U_1 + \cdots + U_n \leq \kappa n) \leq e^{-\kappa n}.$$

As this bound is uniform over conditioning, we are done. □

To conclude, we discuss asymptotic sharpness of our results in a certain family of graph products. Given a graph product of groups, one can estimate the drift for a random walk induced by an alternating measure as follows:

1. Verify that $D > 5B$.
2. Maximize the quantity
$$\frac{-\ln \mathbb{E}[e^{-tU}]}{1+t}$$
with the constraints $t > 0$, $\mathbb{E}[e^{-tU}] < 1$.

For an example, we consider the family of graphs which are cycles of length D. In this case we have $B = 4$, so that the theorem applies for $D > 20$. We numerically compute the lower bound on the drift afforded from Theorem 1.1 for $20 < D \leq 12000$, shown in Figure 2.

In the case where our alternating measure is $\mu * \delta_e$, where μ is the simple random walk on a RAAG, the drift is at most 1. With further assumptions on B and D we can derive sharp asymptotics for our drift estimate.

Proposition 4.3. *Suppose that $B \leq o(D)$ as $D \to \infty$. Let T be the set of positive t such that $\mathbb{E}[e^{-tU}] < 1$. Then*

$$\sup_{t\in T} \frac{-\ln \mathbb{E}\left[e^{-tU}\right]}{1+t} \to 1$$

as $D \to \infty$.

Proof. Since $-\ln$ is convex, by Jensen's inequality we have

$$\frac{-\ln \mathbb{E}\left[e^{-tU}\right]}{1+t} \leq \frac{\mathbb{E}\left[-\ln e^{-tU}\right]}{1+t} = \frac{t\mathbb{E}U}{1+t}.$$

Since $\mathbb{E}U > 0$, then this final term is increasing in t and goes to $\mathbb{E}U$ as $t \to \infty$. Since $B \leq o(D)$ then $\mathbb{E}U \to 1$ as $D \to \infty$. Therefore

$$\lim_{D\to\infty} \sup_{t\in T} \frac{-\ln \mathbb{E}\left[e^{-tU}\right]}{1+t} \leq 1.$$

Pick $t_D = \ln\left(\frac{D-2B}{2B}\right)$, so that t_D satisfies $\mathbb{E}\left[e^{-t_D U}\right] < 1$ for large enough D.
Then

$$\mathbb{E}\left[e^{-t_D U}\right] = e^{-t_D}\frac{D-2B}{D} + \frac{2B}{D}\frac{D-3B}{D-2B-e^{t_D}B}.$$

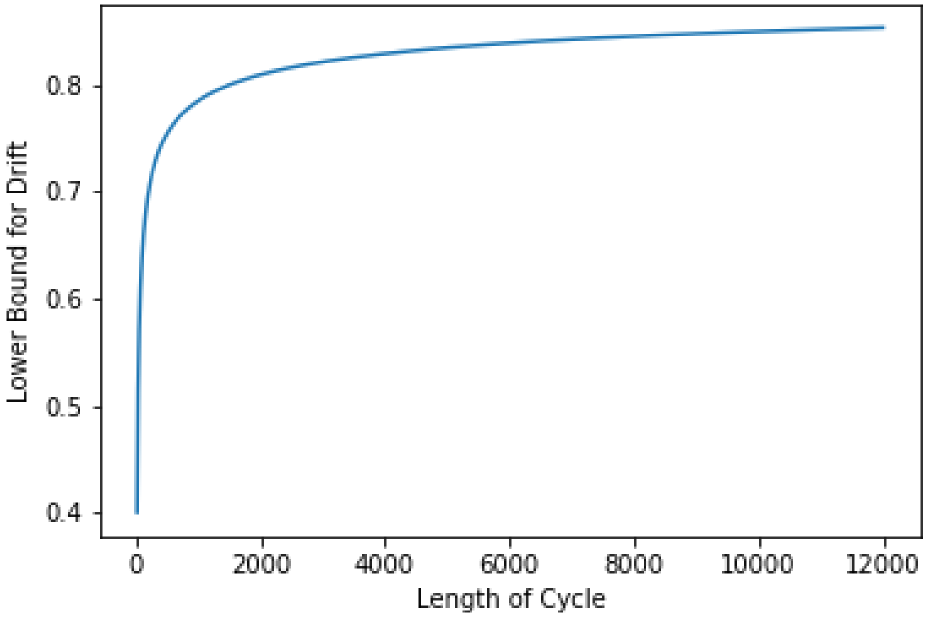

Figure 3: Our drift estimate for an alternating random walk on a graph product given by a D-cycle.

Therefore

$$\begin{aligned}\frac{-\ln \mathbb{E}\left[e^{-t_D U}\right]}{1+t_D} &= \frac{-\ln\left(\frac{2B}{D-2B}\frac{D-2B}{D}+\frac{2B}{D}\frac{D-3B}{D/2-B}\right)}{1+\ln\left(D/B\right))} \\ &\sim \frac{\ln\left(D/B\right))}{1+\ln\left(D/B\right))} \\ &\to 1,\end{aligned}$$

as $D\to\infty$, where "$f(D)\sim g(D)$ means $f(D)/g(D)\to 1$", and the second line comes from the fact that $B\le o(D)$. □

References

[AS22] Richard Aoun and Cagri Sert. Random walks on hyperbolic spaces: Concentration inequalities and probabilistic tits alternative. *Probability Theory and Related Fields*, 184(1-2):323–365, 2022. MR4498512

[BMSS20] Adrien Boulanger, Pierre Mathieu, Cagri Sert, and Alessandro Sisto. Large deviations for random walks on hyperbolic spaces. *arXiv preprint arXiv:2008.02709*, 2020. MR4079419

[BQ16] Yves Benoist and Jean-François Quint. Central limit theorem on hyperbolic groups. *Izvestiya: Mathematics*, 80(1):3–23, 2016. MR3462675

[CFFT22] Kunal Chawla, Behrang Forghani, Joshua Frisch, and Giulio Tiozzo. The Poisson boundary of hyperbolic groups without moment conditions. *arXiv preprint arXiv:2209.02114*, 2022.

[CGW08] John Crisp, Eddy Godelle, and Bert Wiest. The conjugacy problem in right-angled Artin groups and their subgroups. *arXiv preprint arXiv:0802.1771*, 2008. MR2546582

[Cho21a] Inhyeok Choi. Central limit theorem and geodesic tracking on hyperbolic spaces and Teichmüller spaces. *arXiv preprint arXiv:2106.13017*, 2021. MR4627690

[Cho21b] Inhyeok Choi. Pseudo-Anosovs are exponentially generic in mapping class groups. *arXiv preprint arXiv:2110.06678*, 2021.

[Cho22] Inhyeok Choi. Random walks and contracting elements I: Deviation inequality and limit laws. *arXiv preprint arXiv:2207.06597v2*, 2022.

[Cor21a] Emilio Corso. Large deviations for irreducible random walks on relatively hyperbolic groups. *arXiv preprint arXiv:2110.14592*, 2021.

[Cor21b] Emilio Corso. Large deviations for random walks on free products of finitely generated groups. *Electronic Journal of Probability*, 26:1–22, 2021. MR4343563

[Fur63] Harry Furstenberg. A Poisson formula for semi-simple Lie groups. *Annals of Mathematics*, pages 335–386, 1963. MR0146298

[Gou23] Sébastien Gouëzel. Exponential bounds for random walks on hyperbolic spaces without moment conditions. *Tunisian Journal of Mathematics*, 4(4):635–671, 2023. MR4533553

[Gre90] Elisabeth Ruth Green. *Graph products of groups*. PhD thesis, University of Leeds, 1990.

[HM95] Susan Hermiller and John Meier. Algorithms and geometry for graph products of groups, 1995. MR1314099

[Kai00] Vadim A Kaimanovich. The Poisson formula for groups with hyperbolic properties. *Annals of Mathematics*, pages 659–692, 2000. MR1815698

[KM99] Anders Karlsson and Gregory A Margulis. A multiplicative ergodic theorem and nonpositively curved spaces. *Communications in Mathematical Physics*, 208(1):107–123, 1999. MR1729880

[MT18] Joseph Maher and Giulio Tiozzo. Random walks on weakly hyperbolic groups. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2018(742):187–239, 2018. MR3849626

[Nev03] Amos Nevo. The spectral theory of amenable actions and invariants of discrete groups. *Geometriae Dedicata*, 100:187–218, 2003. MR2011122

[NS13] Amos Nevo and Michah Sageev. The Poisson boundary of cat(0) cube complex groups. *Groups, Geometry, and Dynamics*, 7(3):653–695, 2013. MR3095714

[Sun20] Matthew H Sunderland. Linear progress with exponential decay in weakly hyperbolic groups. *Groups, Geometry, and Dynamics*, 14(2):539–566, 2020. MR4118628

[Tio15] Giulio Tiozzo. Sublinear deviation between geodesics and sample paths. *Duke Mathematical Journal*, 164(3):511–539, 2015. MR3314479

[Vir80] AD Virtser. On products of random matrices and operators. *Theory of Probability & Its Applications*, 24(2):367–377, 1980.

Acknowledgments. The author would like to thank Giulio Tiozzo for many helpful conversations and comments. The author would also like to thank the referees for multiple helpful comments which greatly improved the exposition, and for pointing out helpful references. This work was completed under funding from an NSERC USRA.