# Random walks on groups and superlinear-divergent geodesics

KUNAL CHAWLA†, INHYEOK CHOI‡, VIVIAN HE§ and KASRA RAFI§

†*Department of Mathematics, Princeton University, Princeton, NJ, USA*
*(e-mail: kc7106@princeton.edu)*
‡*June E Huh Center for Mathematical Challenges of KIAS, Seoul, South Korea*
*(e-mail: inhyeokchoi48@gmail.com)*
§*Department of Mathematics, University of Toronto, Toronto, ON, Canada*
*(e-mail: vivian.he@mail.utoronto.ca, rafi@math.toronto.edu)*

*Abstract.* In this paper, we study random walks on groups that contain superlinear-divergent
geodesics, in the line of thoughts of Goldsborough and Sisto. The existence of a
superlinear-divergent geodesic is a quasi-isometry invariant which allows us to execute
Gouëzel's pivoting technique. We develop the theory of superlinear divergence and
establish a central limit theorem for random walks on these groups.

1. *Introduction*

Classical limit laws in probability theory concern the asymptotic behaviour of the random
variable (RV)

$$Z_n = \xi_1 + \xi_2 + \cdots + \xi_n.$$

for independent and identically distributed (i.i.d.) random variables $\xi_1, \xi_2, \ldots$ on $\mathbb{R}$. As
a non-commuting counterpart, Bellman, Furstenberg and Kesten initiated the study of
random walks on a matrix group $G$ [**Bel54**, **Kes59**, **FK60**, **Fur63**]. Given a probability
measure $\mu$ on $G$, the random walk generated by $\mu$ is a Markov chain on $G$ with transition
probabilities $p(x, y) := \mu(x^{-1}y)$. Our goal is to understand the $n$th step distribution

$$Z_n = g_1 \cdots g_n,$$

where $g_i$ are independent random variables distributed according to $\mu$.

There are several generalizations of Bellman, Furstenberg and Kesten's theory of non-commuting random walks: random walks on Lie groups (cf. [**BQ16**] and the references therein) and random conformal dynamics [**DK07**] to name a few. In geometric group theory, there is a strong analogy between rank-1 Lie groups and groups with a non-elementary action on a Gromov hyperbolic space $X$ [**MT18**]. Given a basepoint $o \in X$, the sample path $(Z_n o)_{n \geq 0}$ on $X$ tracks a geodesic and the displacement $d(o, Z_n o)$ at step $n$ grows like a sum of i.i.d. random variables with positive expectation. From this, one can derive a number of consequences, such as exponential bounds on the drift [**BMSS23**, **Gou22**], limit laws [**KM99**, **Bjö10**, **GS21**, **Gou17**, **Hor18**] and identification of the Poisson boundary [**MT18**, **Kai00**, **CFFT22**]. If the $G$-action on $X$ is compatible with the geometry of $G$ in a suitable sense, one can transfer these results on $X$ to $G$. One of the most successful results in this direction is due to Mathieu and Sisto [**MS20**], who proved a central limit theorem (CLT) for random walks on acylindrically hyperbolic groups. We refer the readers to [**BHS19**, **Osi16**] for examples of acylindrically hyperbolic groups and hierarchically hyperbolic groups.

Although the notion of acylindrical hyperbolicity captures a wide range of discrete groups, acylindrical hyperbolicity of a group is not known to be quasi-isometry (QI) invariant or even commensurability invariant. This is because there is no known natural way to transfer a group action through a quasi-isometry. To overcome this, the second author proposed a theory for random walks using a group-theoretic property that does not involve hyperbolic actions, namely, possessing a strongly contracting element [**Cho22**]. Nevertheless, this theory is still not invariant under quasi-isometry.

Meanwhile, certain hyperbolic-like properties are known to be quasi-isometry invariant, such as the existence of a Morse quasi-geodesic. Hence, one can expect that many consequences of hyperbolicity should hold under quasi-isometry invariant assumptions. To address this, Goldsborough and Sisto [**GS21**] developed a QI-invariant random walk theory for groups. Given a bijective quasi-isometry $f$ from a group $H$ to a group $G$, the pushforward of the random walk from $H$ to $G$ is not necessarily a random walk, but only an inhomogeneous Markov chain. Nonetheless, if one—equivalently both—groups are non-amenable, the induced Markov chain satisfies some sort of irreducibility, which the authors call *tameness*. At this moment, Goldsborough and Sisto require that $G$ acts on a hyperbolic space $X$ and contains what they call a 'superlinear-divergent' element $g$, that is, any path must spend a superlinear amount of time to deviate from the axis of $g$ (see §2 for the definition). Goldsborough and Sisto observed that along a random path arising from a tame Markov chain on $G$, some translates of the superlinear-divergent axis are aligned on $X$. Such alignment is also realized on the Cayley graph of $G$, and subsequently on $H$. As a consequence, they established a central limit theorem for random walks on $H$, which is only quasi-isometric to $G$.

In the setting of Goldsborough and Sisto, still, $G$ is required to possess an action on a hyperbolic space. Our purpose is to remove this assumption and establish a central limit theorem for groups satisfying a QI-invariant property, without referring to a hyperbolic space. Our main result describes the growth of the word norm $|Z_n|$ of a random element $Z_n$ picked by a simple random walk on $G$.

THEOREM A. *Let G be a finitely generated group with exponential growth, and suppose that G has a superlinear-divergent quasi-geodesic* $\gamma : \mathbb{Z} \to G$. *Let* $(Z_n)_{n \geq 1}$ *be a simple random walk on G. Then there exist constants* $\lambda, \sigma \geq 0$ *such that*

$$\frac{|Z_n| - \lambda n}{\sqrt{n}} \to \mathcal{N}(0, \sigma^2) \quad \text{in distribution.}$$

Note that we only assume existence of a superlinear-divergent quasi-geodesic, as opposed to a superlinear-divergent element. This makes our setting invariant under quasi-isometry; see Lemma 2.2. In addition, our proof only uses the classical theory of random walks and does not refer to tame Markov chains.

Theorem A does not exclude the possibility that the limiting Gaussian distribution is degenerate. To address this, Mathieu and Sisto give a sufficient criterion [**MS20**, Theorem 4.12] for positivity of the limiting variance. By using a variant of their criterion, we show that the constants $\lambda$ and $\sigma$ in Theorem A are strictly positive when $G$ is non-amenable, and the random walk under consideration is simple and symmetric (see §4.3).

This theorem applies to groups that are not flat but not of rank 1 either. For example, we can construct a superlinear-divergent element in any right-angled Coxeter group (RACG) that contains a periodic geodesic with geodesic divergence at least $r^3$.

PROPOSITION 1.1. *Let* $W_\Gamma$ *be a right-angled Coxeter group of thickness* $k \geq 2$. *Then, any Cayley graph of* $W_\Gamma$ *contains a periodic geodesic* $\sigma$ *which is* $(f, \theta)$-*divergent for some* $\theta > 0$ *and* $f(r) \gtrsim r^k$. *In particular, simple random walks on* $W_\Gamma$ *satisfy the central limit theorem.*

By $f(r) \gtrsim r^k$, we mean that $f(r) \geq cr^k$ for some sufficiently small $c > 0$ and $r$ sufficiently large. The proof of this proposition is Appendix A. Such RACGs can be produced following the method in [**BHS17**, **Lev22**] that shows there is an abundance of such groups. The recent paper of Behrstock, Çiçeksiz and Falgas-Ravry [**BCF24**, Theorem 1.2] provides a range of parameters for which the RACG on the corresponding Erdős–Renyi random graph $\Gamma$ has thickness exactly 2 asymptotically almost surely.

Lastly, let us mention the relationship between superlinear divergence and the strongly contracting property, which is a core ingredient of the second author's previous work [**Cho22**]. In general, a superlinear-divergent axis need not be strongly contracting and *vice versa*. Hence, the present theory and the theory in [**Cho22**] are logically independent. We elaborate this independence in Appendix B.

1.1. *Outline of the paper.* Our main idea is to bring Gouëzel's recent theory of pivotal time construction for random walks [**Gou22**]. Here, the key ingredient is a local-to-global principle for alignments between quasigeodesics. Lacking Gromov hyperbolicity of the ambient group, we establish such a principle among sufficiently long superlinear-divergent geodesics (Proposition 3.3). For this purpose, in §2, we continue to develop the theory of superlinear-divergent sets after Goldsborough and Sisto [**GS21**]. In §3, we discuss alignment of superlinear-divergent geodesics. In §4, we estimate the probability for alignment to happen during a random walk. This yields a deviation inequality (Lemma 4.7) that leads to the desired central limit theorem.

## 2. *Superlinear-divergence*

For this section, let $X$ be a geodesic metric space. For points $x, y \in X$, we will use the notation $[x, y]$ to mean an arbitrary geodesic between $x$ and $y$ (which may not be unique in general). For a quasi-geodesic $\alpha$ with points $x, y \in \alpha$, we use $[x, y]|_\alpha$ to denote the quasi-geodesic segment from $x$ to $y$ along $\alpha$. Throughout, all paths are continuous maps from an interval into $X$.

We adopt the definition in [**GS21**]. For a set $Z \subseteq X$ and constants $A, B > 0$, we say a map $\pi = \pi_Z : X \to Z$ is an $(A, B)$-coarsely Lipschitz projection if

$$\text{for all } x, y \in X, \quad d(\pi(x), \pi(y)) \leq Ad(x, y) + B$$

and

$$\text{for all } z \in Z, \quad d(\pi(z), z) \leq B.$$

We say that a map $\pi$ is coarsely Lipschitz if it is $(A, B)$-coarsely Lipschitz for some $A, B > 0$. Note that a coarsely Lipschitz projection is comparable to the closest point projection: for any $x \in X$, we have

$$
\begin{aligned}
d(x, \pi(x)) &\leq \inf_{z \in Z}(d(x, z) + d(z, \pi(z)) + d(\pi(z), \pi(x))) \\
&\leq \inf_{z \in Z}(d(x, z) + B + (Ad(x, z) + B)) \\
&\leq (A + 1)d(x, Z) + 2B. \tag{2.1}
\end{aligned}
$$

We say that a function $f : \mathbb{R}_+ \to \mathbb{R}_+$ is *superlinear* if it is convex, increasing and

$$\lim_{x \to \infty} \frac{f(x)}{x} = \infty.$$

*Definition 2.1.* (Cf. [**GS21**, Definition 3.1]) Let $Z$ be a closed subset of a geodesic metric space $X$, let $\theta > 0$ and let $f : \mathbb{R}_+ \to \mathbb{R}_+$ be superlinear. We say that $Z$ is $(f, \theta)$-*divergent* if there exists a coarsely Lipschitz projection $\pi_Z : X \to Z$ such that for any $R > 0$ and any path $p$ outside of the $R$-neighbourhood of $Z$, if the endpoints $p_-$ and $p_+$ of the path $p$ satisfy

$$d(\pi_Z(p_-), \pi_Z(p_+)) > \theta,$$

then the length of $p$ is at least $f(R)$.

We say that $Z$ is *superlinear divergent* if it is $(f, \theta)$-divergent for some constant $\theta > 0$, for some coarsely Lipschitz projection $\pi_Z$ and for a superlinear function $f : \mathbb{R}_+ \to \mathbb{R}_+$.

The following lemma shows that the existence of a superlinear-divergent quasi-geodesic in a group $G$ is a quasi-isometry invariance.

LEMMA 2.2. *Let $X$ and $Y$ be geodesic metric spaces where $X$ contains a superlinear-divergent subset $Z$, and let $\phi : X \to Y$ be a quasi-isometry. Then, $\phi(Z)$ is also superlinear divergent.*

*Proof.* Let $Z \subset X$ be $(f, \theta)$-divergent with a coarsely Lipschitz projection $\pi_Z$ (albeit with potentially different constants). Let $\phi : X \to Y$ be a $(q, Q)$-quasi-isometry.

Then $\pi_Z$ pushes forward to a coarsely Lipschitz projection $\pi_{\phi(Z)} = \phi \circ \pi_Z \circ \phi^{-1}$. Here, $\phi^{-1}$ is a quasi-inverse to $\phi$, that is to say, a map $\phi^{-1} : Y \to X$ such that $\sup_{x \in X} d_X(x, \phi^{-1}(\phi(x))) < \infty$.

Note that the image under $\phi^{-1}$ of a continuous path in $Y$ may not be a continuous path in $X$. However, by the taming quasi-geodesics lemma [**BH99**, Lemma III.H.1.11], we can find a continuous path within the $(q + Q)$-neighbourhood of $\phi^{-1}(p)$ with the same endpoints.

Fix $R > 0$. Suppose $p$ is a path in $Y$ outside of an $R$-neighbourhood of $\phi(Z)$, and suppose the endpoints $p_-$ and $p_+$ satisfy

$$d(\pi_{\phi(Z)}(p_-), \pi_{\phi(Z)}(p_+)) > \theta',$$

where $\theta' = q\theta + Q$. Then, let $p'$ be a continuous path in the $(q + Q)$-neighbourhood of $\phi^{-1}(p)$ with endpoints $p'_- \in \phi^{-1}(p_-)$ and $p'_+ \in \phi^{-1}(p_+)$. It follows that $p'$ is outside of the $(R/q - q - 2Q)$-neighbourhood of $Z$. Moreover, the endpoints have projections bounded by

$$d_Z(\pi_Z(p'_-), \pi_Z(p'_+)) > \theta.$$

Superlinear divergence of $Z$ lets us conclude that

$$l_X(p') > f\left(\frac{R}{q} - q - 2Q\right),$$

so $l_Y(p) > g(d)$, where

$$g(x) = \frac{1}{q} f\left(\frac{x}{q} - q - 2Q\right) - Q$$

is a superlinear function. $\qquad\square$

COROLLARY 2.3. *Suppose a finitely generated group $G$ contains a superlinear-divergent bi-infinite quasi-geodesic $\gamma : \mathbb{R} \to G$. Let $H$ be a finitely generated group quasi-isometric to $G$. Then, $H$ also contains a superlinear-divergent bi-infinite quasi-geodesic.*

We now establish basic consequences of superlinear divergence of a geodesic. In part, superlinear-divergent geodesics are 'constricting' (in the sense of [**ACT15**, **Sis18**]) up to a logarithmic error. This will be formulated more precisely in Lemma 2.6.

LEMMA 2.4. *For each superlinear function $f$ and positive constants $A, B, K, \theta, q, Q$, there exists a constant $K_0 > 1$ such that the following holds.*

*Let $Z$ be an $(f, \theta)$-divergent subset of $X$ with respect to an $(A, B)$-coarsely Lipschitz projection $\pi_Z$. Let $M > 0$ be a positive constant and let $\alpha : [0, M] \to X$ be a geodesic in $X$ such that*

$$d(\pi_Z\alpha(0), \pi_Z\alpha(M)) \geq \theta \quad and \quad d(\alpha(0), Z) > K_0.$$

*Then there exists $t \in [0, M]$ such that*

$$d(\pi_Z\alpha(0), \pi_Z\alpha(t)) \leq \theta + B,$$

*and either*

$$d(\alpha(t), Z) \geq K \cdot d(\alpha(0), Z) \quad or \quad d(\alpha(t), Z) \leq \frac{1}{K} \cdot d(\alpha(0), Z).$$

*Proof.* Let $A$, $B$ be the coarsely Lipschitz constants of $\pi_Z$. Choose $K' > 1$ large enough such that for all $t > K'$,

$$\frac{f(t)}{t} \geq K(K + 5B + \theta + 1)(A + 1).$$

Let

$$\tau := \inf\{t \in [0, M] : d(\pi_Z\alpha(0), \pi_Z\alpha(t)) \geq \theta\}.$$

By the $(A, B)$-coarse Lipschitzness of $\pi_Z$, we have

$$d(\pi_Z\alpha(0), \pi_Z\alpha(t)) \leq \theta + B$$

for all $t \in [0, \tau]$. We now take $K_0 = K'K$. For convenience, let $d_t := d(\alpha(t), Z)$ for each $t$. The desired conclusion holds if $d_t \leq d_0/K = K'$ for some $t \in [0, \tau]$; suppose not. Under this assumption, we show that $d_\tau > Kd_0$. By superlinear divergence of $Z$,

$$l(\alpha([0, \tau])) \geq f\left(\frac{d_0}{K}\right) \geq K(K+5B+\theta+1)(A+1) \cdot \left(\frac{d_0}{K}\right) \geq (K+5B+\theta+1)(A+1)d_0.$$

Using inequality (2.1) and the fact that $\alpha$ is a geodesic, we observe that

$$l(\alpha([0, \tau])) \leq d(\alpha(0), \pi_Z(\alpha(0))) + d(\pi_Z(\alpha(0)), \pi_Z(\alpha(\tau))) + d(\pi_Z(\alpha(\tau)), \alpha(\tau))$$
$$\leq ((A+1)d_0 + 2B) + (\theta + B) + [(A+1)d_\tau + 2B].$$

Combining these, we have

$$d_\tau \geq \frac{1}{A+1}[(K+5B+\theta+1)(A+1)d_0 - 5B - (A+1)d_0 - \theta]$$
$$\geq Kd_0 + (5B+\theta)\left(d_0 - \frac{1}{A+1}\right)$$
$$\geq Kd_0,$$

where the final inequality is due to $d_0 \geq K_0 = KK' > 1 > 1/(A+1)$.                    $\square$

The following lemma is a technical calculation that will be used in the proof of Lemma 2.6 to examine the behaviour of a sequence of points along a geodesic whose projections are making steady progress.

LEMMA 2.5. *Let $\pi_Z : X \to Z$ be an $(A, B)$-coarsely Lipschitz projection onto a subset $Z$ of $X$ and let $K > 0$. Suppose that points $x, z \in X$ and a point $y \in [x, z]$ satisfy*

$$d(\pi_Z(x), \pi_Z(y)) < K,$$
$$d(\pi_Z(y), \pi_Z(z)) < K \text{ and}$$
$$d(x, Z), \ d(z, Z) \leq \frac{1}{8(A+1)}d(y, Z).$$

*Then, we have $d(y, Z) \leq 2K + 16B$.*

*Proof.* Suppose in contrast that $d(y, Z) > 2K + 16B$. First, the assumption together with inequality (2.1) tells us that

$$(A + 1)d(x, Z) + 2B \leq \tfrac{1}{8}d(y, Z) + 2B \leq \tfrac{1}{4}d(y, Z).$$

This forces

$$d(x, y) \geq d(y, \pi_Z(x)) - d(x, \pi_Z(x)) \geq d(y, Z) - (A + 1)d(x, Z) - 2B \geq \frac{3}{4}d(y, Z)$$

and similarly $d(y, z) \geq \frac{3}{4}d(y, Z)$, which leads to

$$d(x, z) = d(x, y) + d(y, z) \geq \tfrac{3}{2}d(y, Z). \tag{2.2}$$

Meanwhile, note that

$$
\begin{aligned}
d(x, z) & \\
&\leq d(x, \pi_Z(x)) + d(\pi_Z(x), \pi_Z(z)) + d(\pi_Z(z), z) && (\because \text{triangleinequality}) \\
&\leq (A + 1)d(x, Z) + d(\pi_Z(x), \pi_Z(z)) + (A + 1)d(z, Z) + 4B && (\because \text{inequality (2.1)}) \\
&\leq \tfrac{1}{4}d(y, Z) + 4B + 2K && (\because \text{theassumption}) \\
&\leq \tfrac{1}{2}d(y, Z) + 2K. && (\because \text{theassumption})
\end{aligned}
\tag{2.3}
$$

Combining equations (2.2) and (2.3), we have

$$\tfrac{3}{2}d(y, Z) \leq \tfrac{1}{2}d(y, Z) + 2K,$$

which contradicts the assumption. $\qquad\square$

The following is the main lemma.

LEMMA 2.6. *Let $Z$ be an $(f, \theta)$-divergent subset of $X$. Then, for any $\delta > 0$, there exists $K_1 = K_1(\delta, f, \theta) > 0$ such that the following holds. For any $x, y \in X$, if*

$$d(\pi_Z(x), \pi_Z(y)) > \delta(\log d(x, Z) + \log d(y, Z)) + K_1,$$

*then there exist a subsegment $[p_x, p_y]$ of $[x, y]$ and points $q_x, q_y \in Z$ such that:*
(1) $d(p_x, q_x), d(p_y, q_y) < K_1$;
(2) $d(q_x, \pi_Z(x)) \leq \delta \log d(x, Z) + K_1$;
(3) $d(q_y, \pi_Z(y)) \leq \delta \log d(y, Z) + K_1$;
(4) *the segment $[p_x, p_y]$ is in the $K_1$-neighbourhood of $Z$.*

Roughly speaking, parts (1), (2) and (3) state that the geodesic $[x, y]$ will approach the $K_1$-neighbourhood of $Z$ exponentially (with respect to the progress made by its projection along $Z$ from both sides), and part (4) states that it stays near $Z$ in the middle (see Figure 1).

*Proof.* Let $(A, B)$ be the coarsely Lipschitz constants for $\pi_Z$. Let $K' = 8(A + 1) + \exp((\theta + B)/\delta)$, let $K'' = K_0(K') + 2K + 16B$ where $K_0(K')$ is as in Lemma 2.4 and let

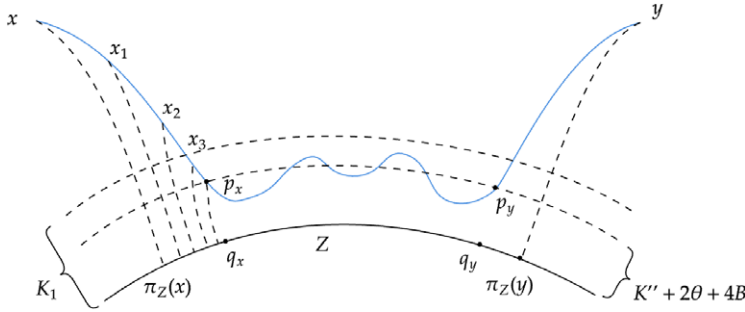$$K_1 = (2A + 3)(K'' + 2\theta + 4B) + 5B + \theta + \log K'.$$

FIGURE 1. A geodesics whose endpoints project sufficiently far apart onto a superlinear-divergent set $Z$ must enter and exit a small neighbourhood of $Z$ near the projections.

If $[x, y]$ entirely lies in the $(K'' + 2\theta + 4B)$-neighbourhood of $Z$, we can take $p_x = x$, $p_y = y$, $q_x = \pi_Z(x)$ and $q_y = \pi_Z(y)$.

If not, we analyse the subsegments of $[x, y]$ outside of the $(K'' + 2\theta + 4B)$- neighbourhood of $Z$. Fix an arbitrary connected component $[x', y']$ of

$$[x, y] \setminus N_{K''+2\theta+4B}(Z) := \{p \in [x, y] : d(p, Z) \geq K'' + 2\theta + 4B\}.$$

We will take a sequence of points $\{x_i\}_{i=0,1,2,...}$ on $[x', y']$, associated with a sequence of real numbers $\{r_i := d(x_i, Z)\}_{i=0,...,M}$ for some $M \in \mathbb{N}$ (Figure 1). We construct the sequence recursively. Start by choosing $x_0 := x'$, then recursively choose $x_{i+1} \in [x_i, y']$ such that

$$d(\pi_Z(x_i), \pi_Z(x_{i+1})) \leq \theta + B$$

and either

$$r_{i+1} \geq K' r_i \quad \text{or} \quad r_{i+1} \leq r_i / K'. \tag{2.4}$$

Such $x_{i+1}$ must exist when $d(\pi_Z(x_i), \pi_Z(y')) \geq \theta + B$, due to Lemma 2.4. We stop the process at step $M$ when $d(\pi_Z(x_M), \pi_Z(y')) < \theta + B$. By inequality (2.4), for each $i$, we have

$$d(x_i, x_{i+1}) \geq |d(x_i, Z) - d(x_{i+1}, Z)| \geq (K' - 1) \min(r_i, r_{i+1}) \geq (K' - 1)(K'' + 2\theta + 4B).$$

This implies that $M$ is finite and, in particular, $M \leq d(x, y)/(K' - 1)(K'' + 2\theta + 4B)$.

We first observe that, by Lemma 2.5, for any $i$, we cannot simultaneously have

$$r_i \geq K' r_{i-1} \quad \text{and} \quad r_i \geq K' r_{i+1}.$$

Hence, the only possibilities for the sequence are either:
(i)    $r_i$ keeps decreasing;
(ii)   $r_i$ keeps increasing; or
(iii)  $r_i$ decreases at first and then keeps increasing.
We will apply this observation in two cases depending on the endpoints of $[x', y']$.

*Case 1.* One (or both) of the endpoints is $x$ or $y$. We will first discuss the case $x' = x$. Note that the sequence $(r_i)_i$ terminates at step $M$ that satisfies

$$d(\pi_Z(x_M), \pi_Z(y')) \leq \theta + B. \tag{2.5}$$

Let $r_j := \min_{i=0,\dots,M} r_i$. Then considering scenarios (i)–(iii) discussed above, we have

$$\frac{r_M}{r_j} \cdot \frac{r_0}{r_j} \geq (K')^M. \tag{2.6}$$

Now, inequality (2.5) and our choices of $x_i$ values imply that

$$M \geq \frac{1}{\theta + B} d(\pi_Z(x), \pi_Z(y')) - 1. \tag{2.7}$$

Combining inequalities (2.6) and (2.7), we have

$$\log d(y', Z) + \log d(x, Z) - 2\log r_j \geq d(\pi_Z(x), \pi_Z(y')) \cdot \frac{\log K'}{\theta + B} - \log K'.$$

This implies that

$$d(\pi_Z(x), \pi_Z(y')) \leq \delta(\log d(x, Z) + \log d(y', Z))) + \log K'.$$

Considering the assumption of the lemma, we conclude that $y'$ cannot equal $y$. Hence, $[x, y']$ is not the entire $[x, y]$ and $d(y', Z) = K'' + 2\theta + 4B$ holds.

Now, at step $M$, we have either $r_M \geq K' r_{M-1}$ or $r_M \leq r_{M-1}/K'$. In the former case, note that

$$r_M \geq K' r_{M-1} \geq K' \cdot (K'' + 2\theta + 4B).$$

This implies that $r_M \geq K' d(y', Z)$ as well. Now, Lemma 2.5 asserts that $d(x_M, Z) \leq 2K + 16B \leq K'' + 2\theta + 4B$, which is a contradiction to the fact that $x_M \in [x', y']$ is $(K'' + 2\theta + 4B)$ away from $Z$. Hence, we are bound to the case that $r_M \leq r_{M-1}/K'$, which means scenario (i) out of the three scenarios mentioned before.

In particular, $(r_i)_i$ is decreasing for $i = 0, \dots, M$, and we have

$$(K')^M \leq r_0/r_M,$$

which implies

$$M \leq \frac{\log r_0 - \log r_M}{\log K'}.$$

This means

$$d(\pi_Z(x_M), \pi_Z(x)) \leq (\theta + B)M \leq \delta \log d(x, Z).$$

Now, recall that $y' \in N_{K''+2\theta+4B}(Z)$. Let $p_x = y'$ and take $q_x \in Z$ such that

$$d(p_x, q_x) \leq K'' + 2\theta + 4B.$$

This choice of $p_x$ and $q_x$ guarantees that

$$\begin{aligned} d(\pi_Z(p_x), q_x) &\leq d(\pi_Z(p_x), \pi_Z(q_x)) + d(\pi_Z(q_x), q_x) \\ &\leq (Ad(p_x, q_x) + B) + B \\ &\leq K_1 \end{aligned}$$

and consequently

$$d(q_x, \pi_Z(x)) \leq \delta \log d(x, Z) + K_1.$$

An analogous argument applies to the connected component $[x', y]$ of $[x, y] \setminus N_{K''+2\theta+4B}(Z)$ that contains $y$. In this case, we conclude that $x' \neq x$ and $d(x', Z) = K'' + 2\theta + 4B$. Moreover, if we set $p_y = x'$ and let $q_y \in Z$ be such that $d(p_y, q_y) = K'' + 2\theta + 4B$, then we conclude

$$d(p_y, q_y) \leq K_1 \quad \text{and} \quad d(q_y, \pi_Z(y)) \leq \delta \log d(y, Z) + K_1.$$

*Case 2.* The endpoints $x'$ and $y'$ both belong to the closure of $N_{K''+2\theta+4B}(Z)$. These are segments between our choice of $p_x$ and $p_y$. We show that they are within the $K_1$-neighbourhood of $Z$.

In this case,

$$d(x', Z) = d(y', Z) = K'' + 2\theta + 4B \leq d(p, Z) \quad \text{(for all } p \in [x', y']).$$

Observe that $r_i$ cannot decrease at first since $[x', y']$ lies outside the $(K'' + 2\theta + 4B)$-neighbourhood of $Z$. However, $r_i$ also cannot keep increasing, because $d(x', Z) = d(y', Z)$. So the process must stop at the very beginning, that is,

$$M = 0 \quad \text{and} \quad d(\pi_Z(x'), \pi_Z(y')) \leq \theta + B.$$

Then, we have

$$\begin{aligned}
d(x', y') &\leq d(x', \pi_Z(x')) + (\theta + B) + d(\pi_Z(y'), y') \\
&\leq ((A+1)d(x', Z) + 2B) + (\theta + B) + ((A+1)d(y', Z) + 2B) \\
&= 2(A+1)(K'' + 2\theta + 4B) + 4B + (\theta + B).
\end{aligned}$$

The second inequality is due to inequality (2.1). From this, we deduce that $[x', y']$ lies in the $K_1$-neighborhood of $Z$. $\square$

The next lemma helps us strengthen Lemma 2.6 to a statement about Hausdorff distance.

LEMMA 2.7. *Let $K, M, M'$ be positive constants, and $\alpha : [0, M] \to X$ and $\beta : [0, M'] \to X$ be $(q, Q)$-quasi-geodesics. Suppose that $\alpha$ is contained in a $K$-neighbourhood of $\beta$ and*

$$d(\alpha(0), \beta(m)) < K, \quad d(\alpha(M), \beta(n)) < K$$

*hold for some $0 \leq m < n \leq M'$. Then, we have*

$$d_{\mathrm{Haus}}(\alpha, \beta|_{[m,n]}) \leq 2(q^6 + q^4 + q^2 + 1)(Q + K).$$

*Proof.* Let us define a map $h$ from $[0, M]$ to $[0, M']$. For each $t \in [0, M]$, let $h(t) \in [0, M']$ be such that $d(\alpha(t), \beta(h(t))) \leq K$. This map is well defined, and is a $(q^2, 2qK + 2qQ)$-quasi-isometric embedding of $[0, M]$ into $\mathbb{R}$. Indeed, note that

$$\begin{aligned}
|h(t) - h(t')| &\leq q\,d(\beta(h(t)), \beta(h(t'))) + qQ \\
&\leq q\,d(\alpha(t), \alpha(t')) + 2qK + qQ \\
&\leq q^2|t - t'| + 2qK + 2qQ
\end{aligned}$$

and

$$|t - t'| \leq q d(\alpha(t), \alpha(t')) + q Q$$
$$\leq q d(\beta(h(t)), \beta(h(t'))) + 2q K + q Q$$
$$\leq q^2 |h(t) - h(t')| + 2q K + 2q Q.$$

From the very definition, it is clear that $\alpha$ and $\beta(h([0, M]))$ are within Hausdorff distance $K$. Next, as $h$ is a $(q^2, 2qK + 2qQ)$-QI-embedding of $[0, M]$, its image $h([0, M])$ is a $(2qK + 2qQ)$-connected subset of $\mathbb{R}$. Now, let $\tau \in [0, M]$ be such that $h(\tau) = \inf_{0 \leq t \leq M} h(t)$. Then, $h(\tau) \leq m < n = h(M)$ holds. Since $h([0, M])$ is $(2qK + 2qQ)$-connected, there exists $\tau_0 \in [\tau, M]$ such that $|h(\tau_0) - m| \leq 2qK + 2qQ$. Now, we have

$$|m - h(\tau)| = |h(0) - h(\tau)| \leq q^2 |\tau - 0| + 2q K + 2q Q$$
$$\leq q^2 |\tau_0 - 0| + 2q K + 2q Q$$
$$\leq q^2 (q^2 |h(\tau_0) - h(0)| + 2q K + 2q Q) + 2q K + 2q Q$$
$$\leq (q^4 + q^2 + 1)(2q K + 2q Q).$$

Similarly, we have $\sup_{0 \leq t \leq M} h(t) < n + (q^4 + q^2 + 1)(2qK + 2qQ)$. In other words, $h([0, M])$ is contained in

$$[m - 2(q^5 + q^3 + q)(Q + K), \quad n + 2(q^5 + q^3 + q)(Q + K)].$$

In particular, $h([0, M])$ and $[m, n]$ are within Hausdorff distance $2(q^5 + q^3 + q)(Q + K)$. By applying $\beta$, we deduce that $\beta(h([0, M]))$ and $\beta|_{[m,n]}$ are within Hausdorff distance $2(q^6 + q^4 + q^2)(Q + K) + Q$. Combining all these, we conclude that

$$d_{\text{Haus}}(\alpha, \beta|_{[m,n]}) \leq 2(q^6 + q^4 + q^2)(Q + K) + Q + K. \qquad \square$$

COROLLARY 2.8. *In the setting of Lemma* 2.6, *assume that $Z$ is a $(q, Q)$-quasi-geodesic. Then for some constant $K_2$ depending on $f, \theta, q, Q, \delta$,*

$$d_{\text{Haus}}([p_x, p_y], [q_x, q_y]|_Z) \leq K_2.$$

As another corollary of Lemma 2.6, we can replace a superlinear-divergent quasi-geodesic on $X$ with a superlinear-divergent geodesic.

COROLLARY 2.9. *Let $\gamma$ be a bi-infinite $(f, \theta)$-divergent quasigeodesic on a proper space $X$. Then there exists a bi-infinite $(f', \theta')$-divergent geodesic $\gamma'$ such that $d_{\text{Haus}}(\gamma, \gamma')$ is finite. Specifically, $f'(x) = f(x - C)$, $\theta' = \theta + 2C$, where $C$ is the Hausdorff distance between $\gamma$ and $\gamma'$.*

*Proof.* Let $\gamma : \mathbb{Z} \to X$ be an $(f, \theta)$-divergent $(q, Q)$-quasigeodesic on $X$. Let $K_1$ be the constant given by Lemma 2.6 for $Z = \gamma$ and $\delta = 0$. For each sufficiently large $n$, we note that

$$d(\pi_\gamma(\gamma(n)), \pi_\gamma(\gamma(-n))) \geq d(\gamma(n), \gamma(-n)) - 2B > \frac{2n}{q} - Q - 2B > K_1.$$

Lemma 2.6 tells us that there exists a subsegment $[p_{-n}, p_n]$ of $[\gamma(-n), \gamma(n)]$ and $j_{-n}, j_n \in \mathbb{Z}$ such that

$$d(p_{-n}, \gamma(j_{-n})) \le K_1, \quad d(p_n, \gamma(j_n)) \le K_1,$$

$$d(\gamma(j_{-n}), \gamma(-n)) \le d(\gamma(j_{-n}), \pi_\gamma(\gamma(-n))) + d(\pi_\gamma(\gamma(-n)), \gamma(-n)) \le K_1 + B,$$

$$d(\gamma(j_n), \gamma(n)) \le d(\gamma(j_n), \pi_\gamma(\gamma(n))) + d(\pi_\gamma(\gamma(n)), \gamma(n)) \le K_1 + B,$$

and such that $[p_{-n}, p_n] \subseteq N_{K_1}(\gamma)$. By Lemma 2.7, $[p_{-n}, p_n]$ and $\gamma([j_{-n}, j_n])$ are within Hausdorff distance $(q^5 + q^3 + q)(2K_1 + 2qQ) + Q + K_1$. For simplicity, let $C = (q^5 + q^3 + q)(2K_1 + 2qQ) + Q + K_1$. Note also that

$$j_{-n} < -n + q(K_1 + B) + Q < 0 < n - q(K_1 + B) - Q < j_n$$

for large enough $n$. In conclusion, $[p_{-n}, p_n]$ contains a point $p$ that is $C$-close to $\gamma(0)$. Moreover, the distance

$$d(\gamma(0), p_n) > d(\gamma(0), \gamma(j_n)) - 2K_1 - B$$

grows linearly, and likewise so does $d(\gamma(0), p_{-n})$. Using the properness of $X$ and Arzelà–Ascoli theorem, we conclude that the sequence $\{[p_{-n}, p_n]\}_{n>1}$ converges to a bi-infinite geodesic $\gamma'$, within a $K_1$-neighbourhood of $\gamma$. By Lemma 2.7 again, we have $d_{\mathrm{Haus}}(\gamma, \gamma') \le C$.

It remains to declare a coarsely Lipschitz projection $\pi_{\gamma'}$ onto $\gamma'$ and show that $\gamma'$ is $(f', \theta')$-divergent with respect to $\pi_{\gamma'}$. Since $d_{\mathrm{Haus}}(\gamma, \gamma') \le C$, we can define $\pi_{\gamma'}(z)$ to be a point on $\gamma'$ such that

$$d(\pi_{\gamma'}(z), \pi_\gamma(z)) < C.$$

Any path $p$ outside of the $R$-neighbourhood of $\gamma'$ is outside of the $(R - C)$-neighbourhood of $\gamma$. Moreover, if the endpoints $p_-$ and $p_+$ of $p$ satisfy that

$$d(\pi_{\gamma'}(p_-), \pi_{\gamma'}(p_+)) > \theta + 2C,$$

then by the construction of $\pi_{\gamma'}$,

$$d(\pi_\gamma(p_-), \pi_\gamma(p_+)) > \theta.$$

Superlinear divergence of $\gamma$ implies that the length of $p$ is at least $f(R - 2C)$. This concludes the proof. $\qquad\square$

2.1. *Convention.* From now on, we fix a finitely generated group $G$ with exponential growth which contains a superlinear-divergent bi-infinite geodesic $\gamma : \mathbb{R} \to G$: this is a QI-invariant property thanks to Corollaries 2.3 and 2.9.

3. *Alignment*
In this section, we define the alignment of sequences of (subsegments of) superlinear-divergent geodesics. The key lemma is Proposition 3.3, which promotes alignment between consecutive pairs to global alignment of a sequence.
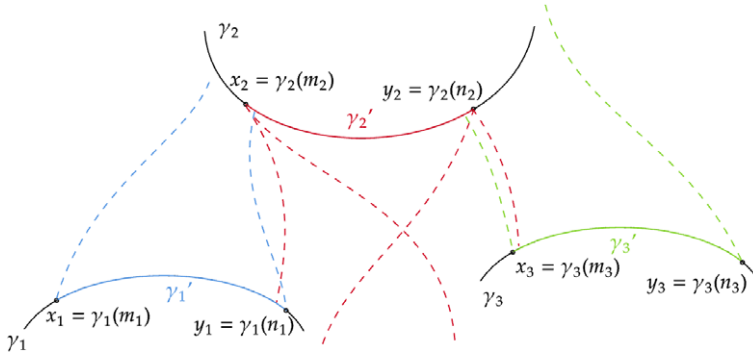
FIGURE 2. The segments satisfy $(\gamma_1', \gamma_2')$ is $(\eta\epsilon \log n)$-aligned and $(\gamma_2', \gamma_3')$ is $(2\eta\epsilon \log n)$-aligned.

*Definition 3.1.* Given paths $\gamma_1, \ldots, \gamma_N : \mathbb{Z} \to G$ equipped with $(A, B)$-coarse Lipschitz projections $\pi_{\gamma_1}, \ldots, \pi_{\gamma_N}$, integers $m_i \leq n_i$ and subpaths $\gamma_i' := \gamma_i([m_i, n_i])$, we say that $(\gamma_1', \ldots, \gamma_N')$ is $K$-aligned if:

(1)  $\pi_{\gamma_i}(\gamma_{i-1}')$ lies in $\gamma_i((-\infty, m_i + K])$; and

(2)  $\pi_{\gamma_i}(\gamma_{i+1}')$ lies in $\gamma_i([n_i - K, +\infty))$.

Note that $\gamma_i$ can be a single point. We will construct linkage words using $K$-aligned paths, starting with the following lemma.

LEMMA 3.2. *Given a superlinear function f, positive constants $\theta, A, B$ and $0 < \epsilon$, $\eta < 0.1$, there exists a constant $K_3 = K_3(f, \theta, A, B, \epsilon, \eta)$ such that the following holds.*

*For $i = 1, 2$, let $\gamma_i$ be an $(f, \theta)$-divergent geodesic with respect to an $(A, B)$-coarsely Lipschitz projection $\pi_{\gamma_i} : X \to \gamma_i$ and let $\gamma_i' = \gamma_i([m_i, n_i])$ be a subpath of $\gamma_i$. Let $z \in X$ and let $D > K_3$ be a constant such that:*

(1)  $\operatorname{diam}(\gamma_1' \cup \gamma_2' \cup z) \leq D$;

(2)  $|n_2 - m_2| \geq \epsilon \log D$;

(3)  $(\gamma_1', \gamma_2')$ *is $(\eta\epsilon \log D)$-aligned and $(\gamma_2', z)$ is $(2\eta\epsilon \log D)$-aligned.*

*Then $(\gamma_1', z)$ is $(2\eta\epsilon \log D)$-aligned (see Figure 2).*

*Proof.* We will assume that $D$ is much larger than the constants $K_1$ and $K_2$ that appear in the argument. For $i = 1, 2$, denote $x_i = \gamma(m_i)$ and $y_i = \gamma(n_i)$. Suppose for contradiction that $\pi_{\gamma_1}(z)$ lies in $\gamma_1((-\infty, n_1 - 2\eta\epsilon \log D))$ as in Figure 3. This implies that

$$d(\pi_{\gamma_1}(y_2), \pi_{\gamma_1}(z)) \geq \eta\epsilon \log D > \frac{\eta\epsilon}{3}(\log d(y_2, \gamma_1) + \log d(z, \gamma_1)) + K_1,$$

where $K_1$ is the constant given in Lemma 2.6 taking $\delta = \eta\epsilon/3$. By Lemma 2.6, there exist a subsegment $[p_1, p_2]$ of $[z, y_2]$ and time parameters $s, t$ of $\gamma_1$ such that $d(p_1, \gamma_1(s)) < K_1$, $d(p_2, \gamma_1(t)) < K_1$ and

$$d(\gamma_1(s), \pi_{\gamma_1}(z)) < \frac{\eta\epsilon}{3} \log d(z, \gamma_1) + K_1,$$

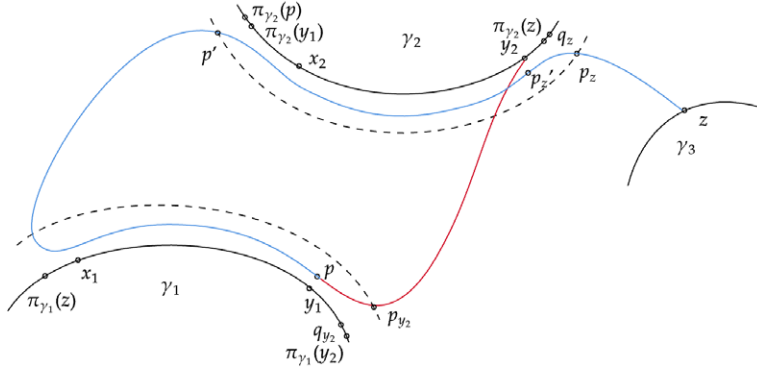$$d(\gamma_1(t), \pi_{\gamma_1}(y_2)) < \frac{\eta\epsilon}{3} \log d(y_2, \gamma_1) + K_1.$$

FIGURE 3. If $\pi_{\gamma_1}(z)$ lies in $\gamma_1((-\infty, n_1 - 2\eta\epsilon\log(n)))$, then the geodesic $[y_2, z]$ would fellow travel $\gamma_1'$ then $\gamma_2'$, causing a contradiction.

In particular, we have

$$
\begin{aligned}
s &< \gamma_1^{-1}(\pi_{\gamma_1}(z)) + \left(\frac{\eta\epsilon}{3}\log d(y_2, \gamma_1) + K_1\right) \\
&\leq n_1 - \frac{5}{3}\eta\epsilon\log D.
\end{aligned}
$$

A similar calculation shows that $t > n_1 - \frac{4}{3}\eta\epsilon\log D$. Now let $K_2$ be the constant in Corollary 2.8 so that $\gamma_1([s, t])$ and $[p_1, p_2]$ are within Hausdorff distance $K_2$ of each other. In particular, for $p' := \gamma_1(n_1 - 1.5\eta\epsilon\log D) \in \gamma_1([s, t])$, we have a point $p \in [p_1, p_2] \subseteq [z, y_2]$ such that $d(p, p') \leq K_2$.

Let us now investigate the relationship between $[p, y_2]$ and $\gamma_2'$. First, the coarse Lipschitzness of $\pi_{\gamma_2}$ tells us that

$$
\begin{aligned}
d(\pi_{\gamma_2}(p), \pi_{\gamma_2}(y_2)) &\geq d(\pi_{\gamma_2}(p'), y_2) - d(\pi_{\gamma_2}(p'), \pi_{\gamma_2}(p)) - d(y_2, \pi_{\gamma_2}(y_2)) \\
&\geq d(\pi_{\gamma_2}(p'), y_2) - AK_2 - 2B.
\end{aligned}
$$

Since $\pi_{\gamma_2}(p') \in \pi_{\gamma_2}(\gamma_1')$ is contained in $\gamma_2((-\infty, m_2 + \eta\epsilon\log D))$, we deduce that

$$
d(\pi_{\gamma_2}(p), \pi_{\gamma_2}(y_2)) > (n_2 - m_2) - \eta\epsilon\log D - AK_2 - 2B > \frac{\eta\epsilon}{3}(\log d(p, \gamma_2)) + K_1.
$$

Again, by Lemma 2.6, there exist a subsegment $[p_1', p_2'] \subseteq [p, z]$ and time parameters $s', t'$ of $\gamma_2$ with $d(p_1', \gamma_2(s')) < K_1, d(p_2', \gamma_2(t')) < K_1$ and

$$
\begin{aligned}
d(\gamma_2(s'), \pi_{\gamma_2}(p)) &< \frac{\eta\epsilon}{3q}\log d(p, \gamma_2) + K_1, \\
d(\gamma_2(t'), \pi_{\gamma_2}(y_2)) &< K_1.
\end{aligned}
$$

This means that

$$
\begin{aligned}
d(p_1', p_2') &\geq d(\pi_{\gamma_2}(p), \pi_{\gamma_2}(y_2)) - d(\gamma_2(t'), \pi_{\gamma_2}(y_2)) - d(\gamma_2(s'), \pi_{\gamma_2}(p)) \\
&\quad - d(p_1', \gamma_2(s')) - d(p_2', \gamma_2(t')) \\
&\geq \frac{(n_2 - m_2) - \eta\epsilon \log D}{q} - Q - \frac{1}{3}\eta\epsilon \log D - 4K_1 \\
&\geq \frac{2}{3}\eta\epsilon \log D.
\end{aligned}
$$

Hence, $[y_2, p_2']$ is longer than $\frac{2}{3}\eta\epsilon \log D$. However,

$$
d(p_2', y_2) \leq d(p_2', \gamma_2(t')) + d(\gamma_2(t'), \pi_{\gamma_2}(y_2)) + d(y_2, \pi_{\gamma_2}(y_2)) \leq 2K_1 + B.
$$

This is a contradiction for sufficiently large $D$. □

PROPOSITION 3.3. *Let $f$ be a superlinear function, $\theta, A, B > 0$, $0 < \epsilon, \eta < 0.1$ and let $K_3$ be the constant given in Lemma 3.2. Let $x, y \in X$, and for $i = 1, \ldots, N$, $\gamma_i$ be an $(f, \theta)$-divergent geodesic with respect to an $(A, B)$-coarse-Lipschitz projection and let $\gamma_i' = \gamma_i([m_i, n_i])$ be a subpath of $\gamma_i$. Let $D > K_3$ be a constant such that:*
(1)   $\operatorname{diam}(\gamma_1' \cup \cdots \cup \gamma_N') \leq D$;
(2)   $|n_i - m_i| \geq \epsilon \log D$ *for each $i$; and*
(3)   $(x, \gamma_1', \ldots, \gamma_N', y)$ *is $(\eta\epsilon \log D)$-aligned.*
*Then for each $i$, $(x, \gamma_i', y)$ is $(2\eta\epsilon \log D)$-aligned.*

*Proof.* This follows inductively from Lemma 3.2. Fixing $i < j$, we show that

$$
\pi_{\gamma_i}(\gamma_j') \in \gamma_i((-\infty, m_i + 2\eta\epsilon \log D]).
$$

If $i = j - 1$, immediately by assumption, we have

$$
\pi_{\gamma_i}(\gamma_j') \in \gamma_i((-\infty, m_i + \eta\epsilon \log D]) \subset \gamma_i((-\infty, m_i + 2\eta\epsilon \log D]).
$$

Now assuming

$$
\pi_{\gamma_{i+1}}(\gamma_j) \in \gamma_{i+1}((-\infty, m_{i+1} + 2\eta\epsilon \log D]),
$$

since $(\gamma_i, \gamma_{i+1})$ are $(\eta\epsilon \log D)$-aligned, the triple $(\gamma_i, \gamma_{i+1}, \gamma_j)$ satisfies the assumptions in Lemma 3.2. We conclude that

$$
\pi_{\gamma_i}(\gamma_j) \in \gamma_i((-\infty, m_i + 2\eta\epsilon \log D]).
$$

Applying the same argument to $\pi_{\gamma_j}(\gamma_i)$, $\pi_{\gamma_j}(\gamma_k)$ and $\pi_{\gamma_k}(\gamma_j)$ shows that

$$
\begin{aligned}
\pi_{\gamma_j}(\gamma_i) &\in \gamma_j([n_j - 2\eta\epsilon \log D, +\infty)), \\
\pi_{\gamma_j}(\gamma_k) &\in \gamma_j((-\infty, m_j + 2\eta\epsilon \log D]) \quad \text{and} \\
\pi_{\gamma_k}(\gamma_j) &\in \gamma_k([n_k - 2\eta\epsilon \log D, +\infty)). \qquad \square
\end{aligned}
$$

LEMMA 3.4. *Given a superlinear function $f$, positive constants $\theta, A, B$ and $0 < \epsilon$, $\eta < 0.1$, there exist constants $K_4 = K_4(f, \theta, A, B, \epsilon, \eta)$ and $C = C(A)$ such that the following holds.*

*Let $\alpha$ and $\beta$ be $(f, \theta)$-divergent geodesics with respect to $(A, B)$-coarsely Lipschitz projections. Let $\alpha'$ and $\beta'$ be their subsegments with beginning points $x_1$ and $x_2$, respectively, such that:*

(1)     *$D := \operatorname{diam}(\alpha' \cup \beta') \geq K_4$;*

(2)     *$\operatorname{diam}(\alpha') \geq \epsilon \log D$; and*

(3)     *$(\alpha', x_2)$ and $(x_1, \beta')$ are $\eta\epsilon \log D$-aligned.*

*Then, $(\alpha', \beta')$ is $(C\eta\epsilon \log D)$-aligned.*

*Proof.* Let $\alpha' = \alpha([m_1, n_1])$ and $\beta' = \beta([m_2, n_2])$. Denote $x_1 = \alpha(m_1)$, $y_1 = \alpha(n_1)$, $x_2 = \beta(m_2)$, $y_2 = \beta(n_2)$. Let $C' = 16(A + 1) + 1$ and $C = (C')^2 + 2$. We first show that

$$\pi_\beta(\alpha') \subset \beta((-\infty, m_2 + C'\epsilon \log D]) \subset \beta((-\infty, m_2 + C\epsilon \log D]).$$

Suppose in contrast that for some point $a \in \alpha'$, the projection

$$\pi_\beta(a) \in \beta([m_2 + C'\epsilon \log D, +\infty)).$$

Then, we have

$$d(\pi_\beta(x_1), \pi_\beta(a)) \geq (C' - 1)\eta\epsilon \log D$$
$$\geq \frac{1}{16A}(C' - 1)\eta\epsilon(\log d(x_1, \beta) + \log d(a, \beta)) + K_1,$$

where $K_1 > 0$ is the constant as in Lemma 2.6 taking $\delta = 1/16A(C' - 1)\eta\epsilon$. Then, there exists a subsegment $[p_{x_1}, p_a]|_\alpha \subset [x_1, a]|_\alpha \subset [x_1, y_1]|_\alpha$ and points $q_{x_1}, q_a$ on $\beta$ such that

$$d(p_{x_1}, q_{x_1}), d(p_a, q_a) < K_1,$$
$$d(q_{x_1}, \pi_\beta(x_1)) \leq \left( \frac{1}{16A}(C' - 1)\eta\epsilon \right) \log d(x_1, \beta) + K_1,$$
$$d(q_a, \pi_\beta(a)) \leq \left( \frac{1}{16A}(C' - 1)\eta\epsilon \right) \log d(a, \beta) + K_1.$$

Then, by Corollary 2.8, there is a point $p'_{x_1} \in [p_{x_1}, p_a]|_\alpha$ close to $x_2$. The point $p'_{x_1}$ is chosen to be $p_{x_1}$ if $q_{x_1} \in \beta((m_2, \infty))$, or the point where the Hausdorff distance $K_2$ is attained if $q_{x_1} \in \beta((-\infty, m_2])$. The distance is bounded by

$$d(x_2, p'_{x_1}) \leq \max \left( \left( \frac{1}{16A}(C' - 1)\eta\epsilon \right) \log d(x_1, \beta) + 2K_1, K_2 \right)$$
$$\leq \left( \frac{1}{16A}(C' - 1) \right)\eta\epsilon \log D + O(1)$$
$$= \left( \frac{A + 1}{A} \right)\eta\epsilon \log D + O(1),$$

where $K_2$ is the constant in Corollary 2.8 and $O(1)$ is the implied constant. Projecting to $\alpha$ gives that

$$d(\pi_\alpha(x_2), p'_{x_1}) \leq d(\pi_\alpha(x_2), \pi_\alpha(p'_{x_1})) + B \leq (A + 1)q\eta\epsilon \log D + O(1).$$

However, since $(\alpha', x_2)$ is $(\eta\epsilon \log D)$-aligned,

$$
\begin{aligned}
d(\pi_\alpha(x_2), p'_{x_1}) &\geq d(y_1, p'_{x_1}) - \eta\epsilon \log D \\
&\geq d(p_a, p'_{x_1}) - \eta\epsilon \log D \\
&\geq d(x_2, \pi_\beta(a)) - d(x_2, p'_{x_1}) - d(\pi_\beta(a), p_a) - \eta\epsilon \log D \\
&\geq \left(\frac{1}{q}C'\eta\epsilon \log D\right) - 2\left(\frac{C'-1}{16A}\eta\epsilon \log D\right) - \eta\epsilon \log D - O(1) \\
&\geq (14(A+1) - 1)\eta\epsilon \log D - O(1)
\end{aligned}
$$

contradicting the previous inequality when $D$ is sufficiently large.

We now show that

$$
\pi_\alpha(\beta') \subset \alpha((n_1 - C\epsilon \log D, \infty)).
$$

Suppose in contrast that for some point $b \in \beta'$, the projection

$$
\pi_\alpha(b) \in \alpha((-\infty, n_1 - C\epsilon \log D)).
$$

We will discuss in two cases. If

$$
\pi_\alpha(b) \in \alpha((m_1, n_1 - C\epsilon \log D)) \subset \alpha',
$$

then the previous calculation shows that

$$
\pi_\beta(\pi_\alpha(b)) \in \beta((-\infty, m_2 + C'\eta\epsilon \log D]).
$$

This shows that $(\pi_\alpha(b), [x_2, b]|_\beta, )$ and $(b, [\pi_\alpha(b), y_1]|\alpha)$ are $(C'\eta\epsilon \log D)$-aligned. Moreover, $\mathrm{diam}([x_2, b]|_\beta \cup [\pi_\alpha(b), y_1]|_\alpha) < D$. So the exact same calculation as before shows that

$$
\pi_\alpha([x_2, b]|_\beta) \subset \alpha((-\infty, \pi_\alpha(b) + C'^2\epsilon \log D)) \subset \alpha((-\infty, n_1 - 2\epsilon \log D)).
$$

This contradicts that

$$
\pi_\alpha(x_2) \in \alpha((n_1 - \eta\epsilon \log D, \infty)).
$$

The remainder case is when $\pi_\alpha(b) \in \alpha((-\infty, m_1))$. We will show that this is impossible assuming $\eta < \min(1/(q + 2q^2), (A+2)/(A+q+2))$ and $\alpha'$ is long. In this case,

$$
\begin{aligned}
d(\pi_\alpha(b), \pi_\alpha(x_2)) &\geq \frac{1}{q}(1 - \eta)\epsilon \log D \\
&\geq \frac{1}{(2+A)}\eta\epsilon \log d(b, \alpha) + \log d(x_2, \alpha) - K_1,
\end{aligned}
$$

where $K_1$ is the constant in Lemma 2.6 choosing $\delta = 1/(2+A)\eta\epsilon$. Then, by Lemma 2.6, there are points $p_{x_2}, p_b \in [x_2, b]$ such that

$$
d(p_{x_2}, y_1) \leq \frac{1}{2+A}\eta\epsilon \log D + K_1 \quad \text{and}
$$

$$
d(p_b, x_1) \leq \frac{1}{2+A}\eta\epsilon \log D + K_1.
$$

Then,

$$d(\pi_\beta(x_1), p_b) \leq \frac{A}{2+A}\eta\epsilon \log D + O(1).$$

However, $\pi_\beta(x_1) \in \beta((-\infty, m_2 + \eta\epsilon \log D])$ implies that

$$\begin{aligned}
d(\pi_\beta(x_1), p_b) &\geq d(x_2, p_b) - \eta\epsilon \log D \\
&\geq d(p_{x_2}, p_b) - \eta\epsilon \log D \\
&\geq d(x_1, y_1) - d(x_1, p_b) - d(y_1, p_{x_2})\eta\epsilon \log D \\
&\geq \epsilon \log D - \frac{2}{2+A}\eta\epsilon \log D - \eta\epsilon \log D - O(1) \\
&> \frac{A}{2+A}\eta\epsilon \log n + O(1).
\end{aligned}$$

The last step is due to $\eta < 1/3$. This is a contradiction. □

We now construct linkage words. These play the role of Schottky sets in [**BMSS23**, **Gou22**]. We use the notation $\mathcal{B}(g, R) := \{h \in G : d(g, h) \leq R\}$ to mean the ball of radius $R$ around $g$, and $\mathcal{S}(g, R) := \{h \in G : d(g, h) = R\}$ to mean the sphere of radius $R$ around $g$.

LEMMA 3.5. *Let $\gamma : \mathbb{R} \to G$ be an $(f, \theta)$-divergent quasi-geodesic and let $\epsilon > 0$. For $K$ sufficiently large, the following holds. For each $m \in \mathbb{Z}$, there exists a subset $S \subseteq G$ with 100 elements such that for each pair of distinct elements $a, b \in S$, we have:*
(1) *$|a|, |b| = K$ and $|ba^{-1}|, |a^{-1}b| \geq 0.5K$;*
(2) *$\pi_\gamma(\gamma(0)a^{-1})$ and $\pi_\gamma(\gamma(0)a^{-1}b) \in \mathcal{B}(\gamma(0), \epsilon K)$; and*
(3) *$\pi_\gamma(\gamma(m)a)$ and $\pi_\gamma(\gamma(m)ab^{-1}) \in \mathcal{B}(\gamma(m), \epsilon K)$.*

*Proof.* Let $K_1 = K_1(0.1\epsilon, f, \theta)$ be as in Lemma 2.6.

Let $\lambda > 1$ be the growth rate of $G$. For $n$ large enough, we have

$$\lambda^n \leq \#\mathcal{S}(\mathrm{id}, n) \leq \lambda^{(1+0.1\epsilon)n}.$$

We consider the sets

$$\begin{aligned}
O_1 &:= \{g \in \mathcal{S}(\mathrm{id}, K) : d(\gamma(0), \pi_\gamma(\gamma(0)g)) \geq 0.5\epsilon K\}, \\
O_2 &:= \{g \in \mathcal{S}(\mathrm{id}, K) : d(\gamma(m), \pi_\gamma(\gamma(m)g)) \geq 0.5\epsilon K\}.
\end{aligned}$$

We will argue that both of these sets are much smaller than $\mathcal{S}(\mathrm{id}, K)$ and use a certain subset of $\mathcal{S}(\mathrm{id}, K) \setminus (O_1 \cup O_2)$ to construct our set $S$.

To show that $O_1, O_2$ are relatively small, let us now consider a word $a$ with $|a| = K$ and $d(\pi_\gamma(\gamma(0)a^{-1}), \gamma(0)) \geq 0.5\epsilon K$. Then, since (assuming $K$ is sufficiently large)

$$d(\pi_\gamma(\gamma(0)a^{-1}), \pi_\gamma(\gamma(0))) \geq 0.5\epsilon K - B \geq K_1 + 0.1\epsilon \log B + 0.1\epsilon \log K,$$
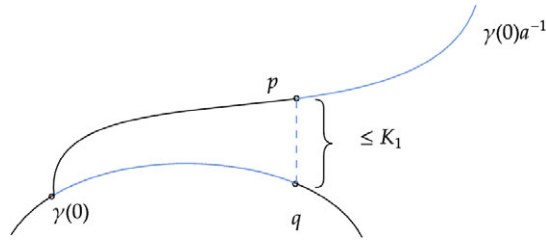
FIGURE 4. Decomposing $a^{-1}$ as a concatenation of well-controlled paths.

Lemma 2.6 asserts that there exist $p \in [\gamma(0), \gamma(0)a^{-1}]$ and $q \in \gamma$ such that $d(p, q) \leq K_1$ and $d(p, \pi_\gamma(\gamma(0)a^{-1})) \leq \log |a| + K_1$. In this case, we have

$$d(p, \gamma(0)a^{-1}) = d(\gamma(0), \gamma(0)a^{-1}) - d(\gamma(0), p)$$
$$\leq |a| - d(\gamma(0), \pi_\gamma(\gamma(0)a^{-1})) + d(p, \pi_\gamma(\gamma(0)a^{-1}))$$
$$\leq K - 0.5\epsilon K + \log K + K_1.$$

In summary,

$$a^{-1} = (\gamma(0)^{-1}q) \cdot (q^{-1}p) \cdot (p^{-1}\gamma(0)a^{-1})$$

where, as in Figure 4:

- $\gamma(0)^{-1}q = \gamma(0)^{-1}\gamma(k)$ for some $k$ between $-2qK - Q$ and $2qK + Q$;
- $|q^{-1}p| \leq K_1$; and
- $|p^{-1}\gamma(0)a^{-1}| \leq (1 - 0.5\epsilon)K + \log(1.5K) + K_1$.

For large enough $K$, the number of such elements is at most

$$5QK \cdot \lambda^{(1+0.1\epsilon)(1-0.4\epsilon)K} \leq 5QK\lambda^{(1-0.3\epsilon)K}.$$

Hence, the cardinality of

$$\mathcal{A} := \{(g_1, \ldots, g_{100}) \in \mathcal{S}(\mathrm{id}, K)^{100} : g_i \in O_1 \text{ for some } i \in [1, 100]\}$$

is at most $100 \cdot (\#S_0)^{99} \cdot 3QK\lambda^{(1+0.3)\epsilon K}$—we pick some index $i$ which satisfies the given condition and draw the rest of the elements from $\mathcal{S}(\mathrm{id}, K)$. This is exponentially small compared with $(\#\mathcal{S}(\mathrm{id}, K))^{100}$.

By a similar logic,

$$\mathcal{B} := \{(g_1, \ldots, g_{100}) \in \mathcal{S}(\mathrm{id}, K)^{100} : g_i \in O_2 \text{ for some } i \in [1, 100]\}$$

is exponentially small compared with $\mathcal{S}(\mathrm{id}, K)^{100}$.

Finally, we observe that for each $h \in \mathcal{S}(\mathrm{id}, K)$, there are at most

$$\#\mathcal{B}(h, 0.5K) \leq \lambda^{(1+\epsilon)0.5K}$$

elements $g$ such that $|g^{-1}h| \leq 0.5K$.

Hence, we deduce that the cardinality of

$$\mathcal{C} := \{(g_1, \ldots, g_{100}) \in S_0^{100} : d(g_i, g_j) \leq 0.5K \text{ for some } i \neq j\}$$

is at most $100 \cdot 99 \cdot 2 \cdot \lambda^{0.6K} \cdot (\#\mathcal{S}(\mathrm{id}, K)^{99}$, which is exponentially small compared with $(\#\mathcal{S}(\mathrm{id}, K))^{100}$. Given these estimates, we conclude that for sufficiently large $K$,

$$\mathcal{S}(\mathrm{id}, K)^{100} \setminus (\mathcal{A} \cup \mathcal{B} \cup \mathcal{C})$$

is non-empty.

Letting $(g_1, \ldots, g_{100})$ be one of its elements, we claim that the choice $S = \{g_i, i = 1, \ldots, 100\}$ satisfies the conditions of the lemma.

Note in particular that $g_i^{-1} g_j \neq \mathrm{id}$ since its norm is at least $0.5K$. We observe that:

(1)  $g_i$ are all distinct;
(2)  $|g_i| = K$ for all $1 \leq i \leq 100$;
(3)  $|g_i g_j^{-1}|, |g_i^{-1} g_j| \geq 0.5K$ for each $i \neq j$;
(4)  $\pi_\gamma(\gamma(0)g_i^{-1}) \in \mathcal{B}(\gamma(0), 0.5\epsilon K)$  and  $\pi_\gamma(\gamma(m)g_i) \in \mathcal{B}(\gamma(m), 0.5\epsilon K)$  for each $1 \leq i \leq 100$.

It remains to show that $d(\gamma(0), \pi_\gamma(\gamma(0)g_i^{-1}g_j)) < \epsilon K$ for each $i \neq j$. Suppose not; then for large enough $K$, we have

$$
\begin{aligned}
d(\pi_\gamma(\gamma(0)g_i^{-1}), \pi_\gamma(\gamma(0)g_i^{-1}g_j)) &\geq \epsilon K - 0.5\epsilon K \\
&> 2\epsilon \log K \\
&> \epsilon \log |g_i| + \epsilon \log(|g_i| + |g_j|) + K_1' \\
&\geq \epsilon \log d(\gamma, \gamma(0)g_i^{-1}) + \epsilon \log d(\gamma, \gamma(0)g_i^{-1}g_j) + K_1',
\end{aligned}
$$

where $K_1' = K_1(\epsilon, f, \theta)$ defined as in Lemma 2.6. Then by Lemma 2.6, there exists $p \in [\gamma(0)g_i^{-1}, \gamma(0)g_i^{-1}g_j]$ such that

$$d(p, \pi_\gamma(\gamma(0)g_i^{-1})) < \epsilon \log d(\gamma, \gamma(0)g_i^{-1}) + 2K_1' \leq \epsilon \log K + 2K_1'$$

and $d(\gamma(0), p) < 0.6\epsilon K$. Here, we have

$$d(p, \gamma(0)g_i^{-1}) \geq d(\gamma(0), \gamma(0)g_i^{-1}) - d(\gamma(0), p) \geq K - 0.6\epsilon K$$

and

$$
\begin{aligned}
d(\gamma(0), \gamma(0)g_i^{-1}g_j) &\leq d(\gamma(0), p) + d(p, \gamma(0)g_i^{-1}(g_j) \\
&\leq 0.6\epsilon K + [d(\gamma(0)g_i^{-1}, \gamma(0)g_i^{-1}g_j) - d(\gamma(0)g_i^{-1}, p)] \\
&\leq 0.6\epsilon K + 0.6\epsilon K.
\end{aligned}
$$

However, this contradicts $|g_i^{-1}g_j| \geq 0.5K$.                                                      $\square$

Given a translate of $\gamma$, we can naturally define the projection

$$\pi_{g\gamma}(x) := g\pi_\gamma(g^{-1}x).$$

Since $G$ acts by isometries, this is an $(A, B)$-coarse Lipschitz projection so long as $\pi_\gamma$ is as well. The following lemma describes projections between translates of superlinear-divergent quasi-geodesics.
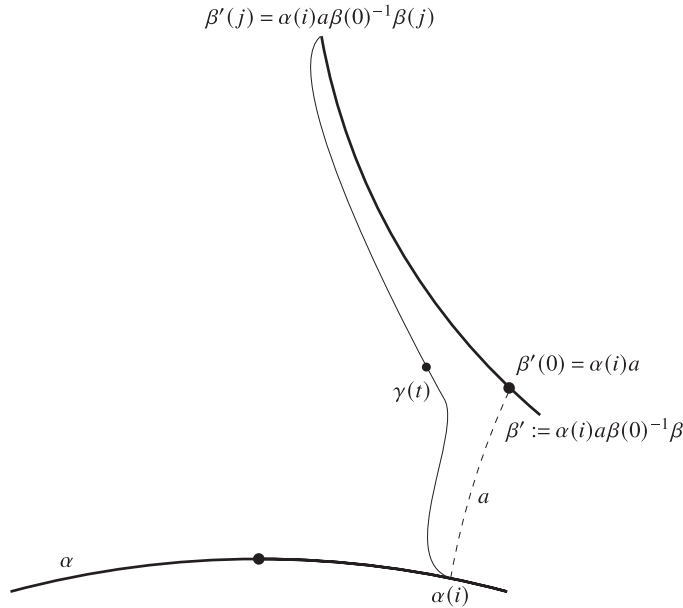
FIGURE 5. The geodesic $[\alpha(i), \beta'(j)]$ comes near $\beta'(0)$.

LEMMA 3.6. *Let $\alpha$ and $\beta$ be $(f, \theta)$-divergent quasi-geodesics equipped with $(A, B)$-coarse Lipschitz projections and let $0 < \epsilon < 1/10(A+1)$. Then, there exists $K_6 > 0$ such that the following holds. Suppose $a \in G$ and $i \in \mathbb{Z}$ satisfy that:*

(i)  $|a| > K_6$;

(ii)  $\pi_\beta(\beta(0)a^{-1}) \in \mathcal{B}(\beta(0), \epsilon|a|)$*; and*

(iii)  $\pi_\alpha(\alpha(i)a) \in \mathcal{B}(\alpha(i), \epsilon|a|)$.

*Then, for each $j \in \mathbb{Z}$, $\pi_\alpha(\alpha(i)a\beta(0)^{-1}\beta(j))$ is within distance $\epsilon \log|j| + 2|a|$ from $\alpha(i)$.*

*Proof.* For simplicity, we denote and parametrize the translate of $\beta$

$$\beta'(j) = \alpha(i)a\beta(0)^{-1}\beta(j).$$

Let $\gamma : [0, M] \to G$ be a geodesic connecting $\alpha(i)$ and $\beta'(j)$, see Figure 5. The projection of $\alpha(i)$ onto $\beta'$ is near $\alpha(i)a$:

$$d(\alpha(i)a, \pi_{\beta'}(\alpha(i))) = d(\beta(0), \pi_\beta(\beta(0)a^{-1})) \le \epsilon|a|.$$

Then, there exists $t \in [0, M]$ such that $\gamma(t) \in \mathcal{B}(\alpha(i)a, 2\epsilon|a|)$. If $d(\beta'(j), \alpha(i)a) < 2\epsilon|a|$, simply take $t = M$ so that $\gamma(t) = \beta'(j)$. Additionally, if $d(\beta'(j), \alpha(i)a) \ge 2\epsilon|a|$, we obtain such $t$ by applying Lemma 2.6. Notice

$$
\begin{aligned}
&d(\pi_{\beta'}(\alpha(i)a), \pi_{\beta'}(\beta'(j)))\\
&\quad \ge d(\alpha(i)a, \beta'(j)) - d(\alpha(i)a, \pi_{\beta'}(\alpha(i)a)) - d(\pi_{\beta'}(\beta'(j)), \beta'(j))\\
&\quad \ge 2\epsilon|a| - \epsilon|a| - B\\
&\quad \ge \epsilon(\log d(\beta'(j), \beta') + \log d(\alpha(i), \beta')) + K_1,
\end{aligned}
$$

where $K_1$ is the constant from Lemma 2.6 taking $\delta = \epsilon$. The last inequality holds when $|a|$ is sufficiently large. Then Lemma 2.6 implies that for some $t$,

$$
\begin{aligned}
d(\gamma(t), \alpha(i)a) &\le d(\gamma(t), \pi_{\beta'}(\alpha(i))) + d(\pi_{\beta'}(\alpha(i)), \alpha(i)a) \\
&\le (\epsilon \log |a| + K_1) + \epsilon |a| \\
&\le 2\epsilon |a|
\end{aligned}
$$

for sufficiently large $|a|$. Note that

$$
t = d(\gamma(0), \gamma(t)) \le d(\alpha(i), \alpha(i)a) + 2\epsilon|a| \le \tfrac{3}{2}|a|.
$$

Now, if

$$
\begin{aligned}
d(\pi_\alpha(\gamma(0)), \pi_\alpha(\gamma(M))) &> \epsilon \log(|j| + |a|) \\
&\ge \epsilon(\log d(\gamma(0), \alpha) + \log d(\gamma(M), \alpha)) - K_1,
\end{aligned}
$$

where $K_1$ is the constant from Lemma 2.6 taking $\delta = \epsilon$, then there exists $\tau \in [0, M]$ such that

$$
d(\gamma(\tau), \pi_\alpha(\gamma(M))) \le \epsilon \log(|j| + |a|) + K_1
$$

and $\gamma|_{[0,\tau]}$ is contained in the $K_1$-neighbourhood of $\alpha$. Notice that $\tau$ cannot be larger than $t$, otherwise $\gamma(t)$ is $K_1$-close to $\alpha$; let $p \in \alpha$ be the point such that $d(\gamma(t), p) \le K_1$. Then, when $|a|$ is sufficiently large,

$$
\begin{aligned}
d(\alpha(i), \pi_\alpha(\alpha(i)a)) &\ge d(\alpha(i), p) - d(\pi_\alpha(\alpha(i)a), p) \\
&\ge d(\alpha(i)a, \alpha(i))) - d(\alpha(i)a, q) - (Ad(p, \alpha(i)a) + 2B) \\
&\ge |a| - (A+1)(2\epsilon|a| + K_1) - 2B > \epsilon|a|.
\end{aligned}
$$

This is a contradiction, so we must have $\tau \le t$. We then have

$$
\begin{aligned}
d(\pi_\alpha(\alpha(i)a\beta(0)^{-1}\beta(j)), \alpha(i)) &\le d(\pi_\alpha(\gamma(M)), \gamma(\tau)) + d(\gamma(\tau), \gamma(0)) \\
&\le \epsilon \log(|j| + |a|) + K_1 + \tfrac{3}{2}|a| \\
&\le \epsilon \log |j| + 2|a|. \qquad \square
\end{aligned}
$$

## 4. *Probabilistic part*

In this section, fixing a small enough $\epsilon > 0$, we study the situation where a random path (id $=: Z_0, Z_1, \ldots, Z_n$) is seen by a superlinear-divergent direction, or to be precise, where $(Z_i, \ldots, Z_{i+\epsilon \log n})$ is (a part of) an $(f, \theta)$-divergent quasigeodesic and

$$
(\mathrm{id}, (Z_i, Z_{i+1}, \ldots, Z_{i+\epsilon \log n}), Z_n)
$$

is $\epsilon^2$-aligned for some $i \ll n$. We will prove in Corollary 4.6 and Lemma 4.7 that this happens with overwhelming probability.

To make an analogy, consider a random path (id $=: Z_0, Z_1, \ldots, Z_n$) arising from a simple random walk on the Cayley graph of a free group $F_2 \simeq \langle a, b \rangle$. Here, we similarly expect that $Z_n = \mathrm{id}$ is not desirable and (id, $(Z_i, Z_{i+1}), Z_n$) is aligned for some $i \ll n$.

In fact, the alignment fails with exponentially decaying probability. A classical argument using martingales can be described as follows:

(1)  construct a 'score' that marks the progress made till step $i$;
(2)  prove that at each step $i$, it is more probable to earn a score rather than lose one;
(3)  sum up the difference at each step and use concentration inequalities to deduce an exponential bound.

Here, the score at step $i$ should be determined by information up to time $i$. Moreover, when the score grows, the recorded local progresses should also pile up. To realize these features on a general Cayley graph other than tree-like ones, we employ the concatenation lemma proven in §3.

### 4.1. *Combinatorial model.*

In the following, let $\gamma$ be an $(f, \theta)$-divergent geodesic on $G$ with $\gamma(0) = $ id and $\epsilon > 0$ be a small enough constant. Let us fix some constants:

* $K_3 = K_3(f, \theta, q, Q, A, B, \epsilon^3, \epsilon)$ be as in Lemma 3.2;
* $K$ is larger than $K_5 = K_5((1/10q)\epsilon^4)$ and the twice of $K_6 = K_6(0.1\epsilon^4)$ given by Lemmas 3.5 and 3.6, respectively;
* $N_0$ is a threshold such that

$$\epsilon^4 n > 10(K + K_3 + \log n)$$

for all $n > N_0$.

After multiple applications of our alignment lemmas, the exponent on $\epsilon$ will decrease, which is why we start with $\epsilon^4$.

Throughout this section, we will consider the following combinatorial model. Fix $w_0, w_1, \ldots \in G$. Now, given a sequence of 3-tuples $s_i = (a_i, b_i, c_i) \in S^3$, we consider a word of the form

$$W_k = w_0 \cdot a_1\gamma(\epsilon \log n)b_1\gamma(\epsilon \log n)c_1 \cdot w_1 \cdots a_k\gamma(\epsilon \log n)b_k\gamma(\epsilon \log n)c_k \cdot w_k.$$

To ease the notation, let us also define

$$V_k = W_{k-1}a_k, \quad U_k = W_{k-1}a_k\gamma(\epsilon \log n)b_k.$$

We also denote

$$s = (a_1, b_1, c_1, \ldots, a_k, b_k, c_k).$$

We will argue that for most choices of $s \in S^{3k}$, a certain subsequence of

$$(\text{id}, \ V_1\gamma|_{[0,\epsilon \log n]}, U_1\gamma|_{[0,\epsilon \log n]} \ldots, U_{k-1}\gamma|_{[0,\epsilon \log n]}, W_k)$$

is well aligned. In §4.2, we will derive from this a deviation inequality (Lemma 4.7) and deduce a central limit theorem.

To show well alignment, we argue analogously to [**BMSS23**, **Gou22**, **Cho22**], by keeping track of times in which the random walk may travel along different translates of $\gamma|_{[0,\epsilon \log n]}$ and arguing that, at most of these times, most directions of the random walk do not backtrack. To implement, we need the following Proposition 4.1. We remark that for the rest of the paper, whenever we discuss alignment of a sequence of points and geodesic segments, the only segments used are translates of $\gamma|_{[0,\epsilon \log n]}$.

PROPOSITION 4.1. *Let $g \in G$ and let $n$ be an integer greater than $N_0$ and $|g|$. Let $S$ be the subset of $\mathcal{S}(\mathrm{id}, K)$ described in Lemma 3.5 for $m = \epsilon \log n$. Then for any distinct $a, b \in S$, at least one of*

$$(\gamma|_{[0,\epsilon \log n]}, \gamma(\epsilon \log n)ag) \quad and \quad (\gamma|_{[0,\epsilon \log n]}, \gamma(\epsilon \log n)bg)$$

*is $\epsilon^4 \log n$-aligned. Likewise, at least one of*

$$(a^{-1}g, \gamma|_{[0,\epsilon \log n]}) \quad and \quad (b^{-1}g, \gamma|_{[0,\epsilon \log n]})$$

*is $\epsilon^4 \log n$-aligned.*

*Proof.* We prove the first claim only. Let $t \in \mathbb{Z}$ be such that $\gamma(t) = \pi_\gamma(\gamma(\epsilon \log n)ag)$. If $t$ is greater than $\epsilon(1 - \epsilon^3) \log n$, we deduce that $(\gamma|_{[0,\epsilon \log n]}, \gamma(\epsilon \log n)ag)$ is $\epsilon^4 \log n$-aligned as desired. Let us deal with the remaining case: we assume

$$t \in (-\infty, \epsilon \log n - \epsilon^4 \log n]. \tag{4.1}$$

Consider two translates of $\gamma$:

$$\gamma_1 = a^{-1}\gamma(\epsilon \log n)^{-1}\gamma \quad and \quad \gamma_2 = b^{-1}\gamma(\epsilon \log n)^{-1}\gamma,$$

and their subpaths

$$\gamma_1' := \gamma_1|_{[t,\epsilon \log n]} \quad and \quad \gamma_2' := \gamma_2|_{[0,\epsilon \log n]}.$$

Let $\bar{\gamma}_2'$ be the reversal of $\gamma_2'$.

By the definition of $t$, $(\gamma(\epsilon \log n)ag, \gamma|_{[t,\epsilon \log n]})$ is automatically 0-aligned or, equivalently, $(g, \gamma_1')$ is 0-aligned. Next, since $a$ and $b$ are chosen from $S$, the subset of $\mathcal{S}(\mathrm{id}, K)$ as described in Lemma 3.5, we have that

$$\pi_\gamma(\gamma(\epsilon \log n)ab^{-1}) \quad \text{is within} \quad \mathcal{B}(\gamma(\epsilon \log n), 0.1\epsilon^4|ab^{-1}|)$$

and

$$\pi_\gamma(\gamma(\epsilon \log n)ba^{-1}) \quad \text{is within} \quad \mathcal{B}(\gamma(\epsilon \log n), 0.1\epsilon^4|ab^{-1}|).$$

Moreover, we have

$$|ba^{-1}| \geq 0.5K \geq K_6.$$

By plugging in $\alpha = \gamma$ and $\beta = \bar{\gamma}$ (that is, $\beta(t) = \gamma(\epsilon \log n - t)$ for each $t \in \mathbb{Z}$), we can apply Lemma 3.6. The required assumptions are

$$\pi_\beta(\beta(0)(ba^{-1})^{-1}) = \pi_\gamma(\gamma(\epsilon \log n)ab^{-1})$$
$$\in \mathcal{B}(\gamma(\epsilon \log n), 0.1\epsilon^4|ab^{-1}|) = \mathcal{B}(\beta(0), 0.1\epsilon^4|ab^{-1}|)$$

and

$$\pi_\alpha(\alpha(\epsilon \log n), ba^{-1}) = \pi_\gamma(\gamma(\epsilon \log n), ba^{-1}) \in \mathcal{B}(\gamma(\epsilon \log n), 0.1\epsilon^4|ab^{-1}|).$$

As a result, for each $j \in \mathbb{Z}$, we have

$$d(\pi_\gamma(\gamma(\epsilon \log n)ab^{-1}\gamma(\epsilon \log n)^{-1}\gamma(j)), \gamma(\epsilon \log n)) \leq 0.1\epsilon^4 \log |j - \epsilon \log n| + 2|ab^{-1}|.$$

In other words, we have

$$\pi_{\gamma_1}(\gamma_2') \in \gamma_1((\epsilon \log n - 0.1\epsilon^4 \log(\epsilon \log n) - 2K, +\infty)).$$

Similarly, we deduce that

$$\pi_{\gamma_2}(\gamma_1') \in \gamma_2((\epsilon \log n - 0.1\epsilon^4 \log(\epsilon \log n) - 2K, +\infty)).$$

We conclude that $(\gamma_1', \bar{\gamma}_2')$ is $(0.1\epsilon^4 \log(\epsilon \log n) + 2K)$-aligned.

We now let $D = |g| + 2\epsilon \log n + 2K + K_3$; note that

$$D > \text{diam}(g^{-1} \cup \gamma_1' \cup \gamma_2').$$

Moreover, the lengths of $\gamma_1'$ and $\gamma_2'$ are at least $\epsilon^4 \log n$ and we have

$$\epsilon^4 \log n \geq \epsilon^3 \log D.$$

Finally, $(g, \gamma_1', \bar{\gamma}_2')$ is $(0.1\epsilon^4 \log m + 2K)$-aligned, and hence $0.2\epsilon^4 \log D$-aligned. Lemma 3.2 now tells us that $(g^{-1}, \bar{\gamma}_2')$ is $\epsilon^4 \log D$-aligned. This implies that $(\gamma|_{[0,m]}, \gamma(m)bg)$ is $\epsilon^4 \log n$-aligned, as desired. $\qquad\square$

Following Boulanger *et al* [**BMSS23**] and Gouëzel [**Gou22**], we define the *set of pivotal times* $P_k(s)$ inductively. We will suppress the notation $P_k := P_k(s)$ when it is unambiguous, and the remaining notation follows from the beginning of this section. First, set $P_0 = \emptyset$ and $z_0 = \text{id}$. Given $P_{k-1} \subseteq \{1, \ldots, k-1\}$ and $z_{k-1} \in G$, $P_k$ and $z_k$ are determined by the following criteria.

(a)  When $(z_{k-1}, V_k\gamma|_{[0,\epsilon \log n]}, U_k\gamma|_{[0,\epsilon \log n]}, W_k)$ is $\epsilon^3 \log n$-aligned, we set $P_k = P_{k-1} \cup \{k\}$ and $z_k = U_k$.

(b)  Otherwise, we find the maximal index $m \in P_{k-1}$ such that $(V_m\gamma|_{[0,\epsilon \log n]}, W_k)$ is $\epsilon^3 \log n$-aligned and let $P_k = P_{k-1} \cap \{1, \ldots, m-1\}$ (that is, we gather all pivotal times in $P_{k-1}$ *smaller than m*) and $z_k = V_m$. If such an $m$ does not exist, then we set $P_k = \emptyset$ and $z_k = \text{id}$.

Given input $w_0, w_1, \ldots, w_k \in G$ and $s \in S^{3k}$, this algorithm outputs a subset $P_k(s)$ of $\{1, \ldots, k\}$. Our first lemma tells us that $P_k(s)$ effectively records the alignment.

LEMMA 4.2.  *The following holds for all $n > N_0$.*

*Let $P_k = \{i(1) < \cdots < i(M)\}$ and suppose that $\epsilon \log(|w_0| + \cdots + |w_k| + k\epsilon \log n) \leq \log n$. Then, there exists $N \in \mathbb{N}$ and $g_1, \ldots, g_N = z_k$ such that $(V_{i(1)}, U_{i(1)}, \ldots, V_{i(M)}, U_{i(M)})$ is a subsequence of $(g_1, \ldots, g_N)$ and*

$$(\text{id}, g_1\gamma|_{[0,\epsilon \log n]}, \ldots, g_N\gamma|_{[0,\epsilon \log n]}, W_k)$$

*is $\epsilon^2 \log n$-aligned.*

*Proof.*  We induce on $k$. If we added a pivot, $P_k = P_{k-1} \cup \{k\}$, there are the following two cases.

(1) $P_{k-1} = \emptyset$. Then, $(\text{id}, V_k\gamma|_{[0,\epsilon \log n]}, U_k\gamma|_{[0,\epsilon \log n]}, W_k)$ is $(\epsilon^3 \log n)$-aligned, with $z_k = U_k$, as desired.

(2) $P_{k-1} = \{i(1) < \cdots < i(M-1)\}$ is non-empty. Then, there exist $g_1, \ldots, g_N$ such that $(V_{i(1)}, \ldots, V_{i(M-1)})$ is a subsequence of $(g_1, \ldots, g_N)$, $g_N = z_{k-1}$ and

$$(\text{id}, \ g_1\gamma|_{[0,\epsilon \log n]}, \ldots, g_N\gamma|_{[0,\epsilon \log n]}, W_{k-1})$$

is $\epsilon^2 \log n$-aligned. Moreover,

$$(z_{k-1}, V_k\gamma|_{[0,\epsilon \log n]}, U_k\gamma|_{[0,\epsilon \log n]}, W_k)$$

is $(\epsilon^3 \log n)$-aligned. Here, since $(z_{k-1}\gamma|_{[0,\epsilon \log n]}, W_{k-1})$ is $\epsilon^3 \log n$-aligned,

$$(z_{k-1}\gamma|_{[0,\epsilon \log n]}, W_{k-1}a_k) = (z_{k-1}\gamma|_{[0,\epsilon \log n]}, V_k)$$

is also $(\epsilon^3 \log n + AK + B)$-aligned. Now, Lemma 3.4 asserts that for large enough $n$, $(z_{k-1}\gamma|_{[0,\epsilon \log n]}, V_k\gamma|_{[0,\epsilon \log n]})$ is $\epsilon^2 \log n$-aligned. As a result,

$$(\text{id}, \ g_1\gamma|_{[0,\epsilon \log n]}, \ldots, g_N\gamma|_{[0,\epsilon \log n]}, \ V_k\gamma|_{[0,\epsilon \log n]}, \ U_k\gamma|_{[0,\epsilon \log n]}, W_k)$$

is $\epsilon^2 \log n$-aligned, with $z_k = U_k$.

Now, suppose we backtracked: $P_k = P_{k-1} \cap \{1, \ldots, m-1\}$ for some $m \in P_{k-1}$. Letting $M = \#P_{k-1}$, so that $\#P_k = M + 1$, our induction hypothesis tells us that there exist $g_1, \ldots, g_N$ such that $(V_{i(1)}, U_{i(1)}, \ldots, V_{i(M+1)}, U_{i(M+1)})$ is a subsequence of $(g_1, \ldots, g_N)$ and

$$(\text{id}, \ g_1\gamma|_{[0,\epsilon \log n]}, \ldots, g_N\gamma|_{[0,\epsilon \log n]}, W_{k-1})$$

is $\epsilon^2 \log n$-aligned. Moreover, we have that $(V_{i(M+1)}\gamma|_{[0,\epsilon \log n]}, W_k)$ is $\epsilon^3 \log n$-aligned by the criterion. It follows that

$$(\text{id}, \ g_1\gamma|_{[0,\epsilon \log n]}, \ldots, V_{i(M+1)}\gamma|_{[0,\epsilon \log n]}, W_k)$$

is $\epsilon^2 \log n$-aligned, with $z_k = V_m = V_{i(M+1)}$, as desired. □

Next, we have the following lemma.

LEMMA 4.3. *Let us fix $a_1, b_1, c_1, \ldots, a_k, b_k, c_k$ and draw $a_{k+1}, b_{k+1}, c_{k+1}$ in $S^3$ according to the uniform measure. For $n \in \mathbb{N}$ sufficiently large, the probability that $\#P_{k+1} = \#P_k + 1$ is at least $9/10$.*

*Proof.* We need to choose $a_{k+1}, b_{k+1}, c_{k+1}$ in $S^3$ such that

$$(z_k, \ V_{k+1}\gamma|_{[0,\epsilon \log n]}, \ U_{k+1}\gamma|_{[0,\epsilon \log n]}, W_{k+1})$$

is $\epsilon^3 \log n$-aligned. By Proposition 4.1, there are at least 99 choices of $a_{k+1}$ such that

$$(z_k, \ V_{k+1}\gamma|_{[0,\epsilon \log n]})$$

is $\epsilon^3 \log n$-aligned.

Likewise, there are at least 98 choices of $b_{k+1}$ such that both

$$(V_k, \ U_{k+1}\gamma|_{[0,\epsilon \log n]}) \quad \text{and} \quad (V_{k+1}\gamma|_{[0,\epsilon \log n]}, \ U_{k+1})$$

are $\epsilon^4 \log n$-aligned. From Lemma 3.4, for sufficiently large $n$, this tells us there are at least 98 choices of $b_{k+1}$ such that $(V_{k+1}\gamma|_{[0,\epsilon \log n]}, \ U_{k+1}\gamma|_{[0,\epsilon \log n]})$ is $\epsilon^4 \log n$-aligned.

Finally, there are at least 99 choices of $c_{k+1}$ such that $(U_{k+1}\gamma|_{[0,\epsilon \log n]}, W_{k+1})$ is $\epsilon^3 \log n$-aligned.

We are done as $99/100 \cdot 98/100 \cdot 99/100 > 9/10$. $\qquad\square$

Given a sequence $s = (a_i, b_i, c_i)_{i=1}^k$, we say that another sequence $s' = (a_i', b_i', c_i')_{i=1}^k$ is *pivoted from s* if they have the same pivotal times, $(a_l, c_l) = (a_l', c_l')$ for all $l = 1, \ldots, k$, and $b_l = b_l'$ for all $l$ except for $l \in P_k(s)$. We observe that being pivoted is an equivalence relation. The following lemma shows that these equivalence classes are large.

LEMMA 4.4. *Given $s = (a_i, b_i, c_i)_{i=1}^k$ and a pivotal time $\ell \in P_k(s)$, construct a new sequence $s'$ by replacing $b_\ell$ with another $b_\ell' \in S$ such that*

$$(z_{\ell-1}, \ V_\ell\gamma|_{[0,\epsilon \log n]}, \ U_\ell\gamma|_{[0,\epsilon \log n]}, \ W_\ell)$$

*is $\epsilon^3 \log n$-aligned. Then, $s'$ is pivoted from s.*

*Proof.* We need to show that both sequences $s$ and $s'$ have the same set of pivotal times. Before time $\ell$, the sequences are identical, so that $P_j(s) = P_j(s')$ for $j < \ell$. By our choice of $b_\ell'$, we know that the time $\ell$ is added as a pivot, and so $z_\ell' = U_\ell'$. Now we induce on $j > \ell$: suppose that all pivotal times in $P_{j-1}(s)$ are still in $P_{j-1}(s')$.

To determine $P_j(s)$, either we added a new pivotal time $j$ or we backtracked. In the former case, we have that $(z_{j-1}, V_j\gamma|_{[0,\epsilon \log n]}, U_j\gamma|_{[0,\epsilon \log n]}, W_j)$ is $\epsilon^3 \log n$-aligned. Since $G$ acts on itself by isometries, this happens if and only if the sequence

$$(z_\ell'(z_\ell^{-1})z_{j-1}, z_\ell'(z_\ell^{-1})V_j\gamma|_{[0,\epsilon \log n]}, z_\ell'(z_\ell^{-1})U_j\gamma|_{[0,\epsilon \log n]}, z_\ell'(z_\ell^{-1})W_j)$$

is $\epsilon^3 \log n$-aligned. However, this is the same as requiring that

$$(z_{j-1}', \ V_j'\gamma|_{[0,\epsilon \log n]}, \ U_j'\gamma|_{[0,\epsilon \log n]}, \ W_j')$$

is $\epsilon^3 \log n$-aligned, so that $j \in P_j(s')$.

In the latter case, we found the maximum $M$ such that $(V_M\gamma|_{[0,\epsilon \log n]}, W_k)$ is $\epsilon^3 \log n$-aligned. Since $\ell \in P_k(s)$, we know that $M > \ell$. Hence, this is the same as requiring that

$$(z_\ell'(z_\ell^{-1})V_M\gamma|_{[0,\epsilon \log n]}, z_\ell'(z_\ell^{-1})W_k) = (V_M'\gamma|_{[0,\epsilon \log n]}, W_k')$$

is $\epsilon^3 \log n$-aligned. Therefore, $j \in P_k(s')$. $\qquad\square$

Now fixing $w_i$ terms, we regard $W_k$ as a random variable depending on the choice of

$$(a_1, b_1, c_1, \ldots, a_k, b_k, c_k),$$

which are distributed according to the uniform measure on $S^{3k}$.

Fixing a choice $s = (a_1, \ldots, c_k)$, let $\mathcal{E}_k(s)$ be the set of choices $s'$ that are pivoted from $s$. Since being pivoted is an equivalence relation, the collection of $\mathcal{E}_k(s)$ terms partitions the space of sequences $S^{3k}$. We claim that most of these equivalence classes are large: at pivotal times $\ell \in P_k$, one can replace $b_\ell$ with one of many other $b_\ell'$ terms while remaining pivoted.

LEMMA 4.5. *Let $s = (a_1, b_1, c_1, \ldots, a_k, b_k, c_k)$. Draw $(a_{k+1}, b_{k+1}, c_{k+1})$ according to the uniform measure on $S^3$. Then, for all $j \geq 0$,*

$$\mathbb{P}(\#P_{k+1}(s', s_{k+1}) < \#P_k(s') - j \mid (s', s_{k+1}) \in \mathcal{E}_k(s) \times S^3) \leq (\tfrac{1}{10})^{j+1}.$$

We remark that the conditional measure $\mathbb{P}(\cdot|\mathcal{E}_k(s) \times S^3)$ on $S^{3(k+1)}$ is the same as the uniform measure on $\mathcal{E}_k(s) \times S^3 \subset S^{3(k+1)}$, because $\mathbb{P}(\cdot)$ is the uniform measure on a finite set.

*Proof.* We induce on $j \geq 0$. The $j = 0$ case is Lemma 4.3. We prove it for $j = 1$. Suppose that we made some choice of $s_{k+1} := (a_{k+1}, b_{k+1}, c_{k+1})$ that led to backtracking. We must show that for such an $s_{k+1}$,

$$\mathbb{P}(\#P_{n+1}(s', s_{k+1}) < \#P_n(s') - 1 \mid s' \in \mathcal{E}_k(s)) \leq \tfrac{1}{10}.$$

To this end, we examine the final pivot $s_\ell$. By Lemma 4.4, we can replace $b_\ell$ with any distinct $b'_\ell \in S$ such that

$$(z_{\ell-1}, \ V_\ell \gamma|_{[0, \epsilon \log n]}, \ U_\ell \gamma|_{[0, \epsilon \log n]}, \ W_\ell)$$

is $\epsilon^4 \log n$-aligned. There are at least 98 choices of such a $b'_\ell$, by Proposition 4.1.

Likewise, there are at least 98 choices of $b'_\ell \neq b_\ell$ such that $(U_\ell \gamma_{[0, \epsilon \log n]}, W_k)$ is $\epsilon^4 \log n$-aligned. From Lemma 3.2, we know that

$$(V_\ell \gamma|_{[0, \epsilon \log n]}, \ W_k)$$

is $\epsilon^3 \log n$-aligned. For this choice of $s'$, we have $P_{k+1}(s') = P_k(s') \cap \{0, \ldots, \ell - 1\}$. In particular, $\#P_{k+1} = \#P_k - 1$. Hence,

$$\mathbb{P}(\#P_{k+1} < \#P_k - 1 \mid \mathcal{E}_k(s), s_{k+1}) \leq (\tfrac{4}{100}) < (\tfrac{1}{10}).$$

To handle the induction step for $j \geq 2$, the same argument works, except we condition not only on $s_{k+1}$ but also on the final $j$ pivotal increments which resulted in backtracking. $\square$

COROLLARY 4.6. *Let $P_k$ be the set of pivotal times. Then we have*

$$\mathbb{P}(\#P_k \leq k/2) < (1/10)^k.$$

### 4.2. *Random walks.*

Recall that $G$ contains an $(f, \theta)$-superdivergent $(q, Q)$-quasi-geodesic $\gamma : \mathbb{Z} \to G$ with $\gamma(0) = \text{id}$.

Let $\mu$ be a probability measure on $G$ whose support generates $G$ as a semigroup. Passing to a convolution power if necessary, assume that $\mu(a) > 0$ for all $a$ in our finite generating set $A \subset G$. Let $(Z_n)_{n \geq 1}$ be the simple random walk generated by $\mu$ and let $\alpha \in (0, 1)$. We can define

$$p = \min\{\mu(a), a \in A\},$$

$$\epsilon = \frac{\alpha/100}{\log(1/p)},$$

so that $p^{\epsilon \log n} = n^{-\alpha/100}$. Then for any path $\eta$ of length $100\epsilon \log n$ and any $k \in \mathbb{Z}$, we have

$$\mathbb{P}((g_{k+1}, \ldots, g_{k+100\epsilon \log n}) = \eta) \geq n^{-\alpha}.$$

Also recall that for any three points $o, x, y \in G$, we can define the Gromov product, given by

$$(x, y)_o = \tfrac{1}{2}(d(o, x) + d(o, y) - d(x, y)).$$

We now have the following lemma.

LEMMA 4.7. *For any $0 < \alpha < 1$, there exists $K > 0$ such that for each $n \in \mathbb{N}$ and $x \in \mathcal{B}(\mathrm{id}, 2n)$, we have*

$$\mathbb{P}[(x, Z_n)_{\mathrm{id}} \geq n^{3\alpha})] \leq K e^{-n^\alpha/K}.$$

*Proof.* First, we would like to find a nice decomposition of our random walk, which will allow us to analyse the sample paths using our combinatorial model in §4.1.

Let $\overline{\lambda}_i$ be i.i.d. distributed according to the uniform measure on the subset $S' \subset G^5$ defined by

$$S' := \{(a, \gamma', b, \gamma', c) : a, b, c \in S, \gamma' = \gamma(\epsilon \log n)\}.$$

Then, the evaluation $\lambda_i = a \cdot \gamma' \cdot b \cdot \gamma' \cdot c$ is distributed according to the measure $\mu_S * \gamma' * \mu_S * \gamma' * \mu_S$, where $\mu_S$ is uniform over $S$.

Let $N = 3K + 2\epsilon \log n$. By our choice of $p$, for each $a, b, c \in S$, we have $\mu^{*N}(a\gamma'b\gamma'c) \geq p^N$. Then, we can decompose

$$\mu^{*N} = 10^6 p^N (\mu_S * \gamma' * \mu_S * \gamma' * \mu_S) + (1 - 10^6 p^N)\nu$$

for some probability measure $\nu$.

Now, we consider the following coin-toss model. Let $\rho_i$ be independent 0-1 valued random variables, each with probability $10^6 \cdot p^N$ of being equal to 1. Also, let $\xi_i$ be i.i.d. distributed according to $\nu$. We set

$$g_i = \begin{cases} \lambda_i & \text{if } \rho_i = 1, \\ \xi_i & \text{otherwise.} \end{cases}$$

Then, $(g_1 \cdots g_n)_n$ has the same distribution as $(Z_{Nn})$, because each $g_i$ is distributed according to $\mu^{*N}$.

Hoeffding's inequality tells us that

$$\mathbb{P}\left( \sum_{i=1}^{n^{3\alpha}} \rho_i \geq 0.5n^{3\alpha} \cdot n^{-\alpha} \right) \geq 1 - 2 \exp\left( -\frac{2(0.5n^{2\alpha})^2}{n^{3\alpha}} \right) \geq 1 - 2 \exp(-0.5n^\alpha).$$

After tossing away an event of probability at most $2 \exp(-0.5n^\alpha)$, we assume $\sum_{i=1}^{n^{3\alpha}} \rho_i \geq 0.5n^{2\alpha}$.

To apply the analysis of our combinatorial model, we condition on the values of $\rho_i, \xi_i$ and only keep the randomness coming from the $\eta_i$ terms. Let

$$i(1) < i(2) < \cdots < i(M)$$

be the indices in $[1, n^{3\alpha}]$, where $\rho_i = 1$. Then, we can write

$$x^{-1} \cdot Z_n$$
$$= w_0 \cdot a_1 \gamma(\epsilon \log n) b_1 \gamma(\epsilon \log n) c_1 \cdot w_1 \cdots a_M \gamma(\epsilon \log n) b_M \gamma(\epsilon \log n) c_M \cdot w_M,$$

where

$$w_0 = x^{-1} g_1 \cdots g_{N(i(1)-1)-1},$$
$$w_1 = g_{Ni(1)+1} \cdots g_{N(i(2)-1)-1},$$
$$\vdots$$
$$w_M = g_{Ni(M)+1} \cdots g_n$$

and $a_i, b_i, c_i$ are i.i.d.s distributed according to the uniform measure on $S$. As in the previous section, we set $s = (a_1, b_1, c_1, \ldots, a_M, b_M, c_M)$. By Lemma 4.6, the set of pivots $P_M(s)$ is non-empty with probability at least $1 - (1/10)^M \geq 1 - (1/10)^{0.5n^{2\alpha}}$. By Proposition 3.3, for any pivotal time $i \in P_M(s)$, we have

$$(\mathrm{id}, x^{-1} Z_{N(i-1)} \gamma|_{\epsilon \log n}, x^{-1} Z_n) = (\mathrm{id}, (x^{-1} Z_{N(i-1)}, \ldots, x^{-1} Z_{N(i-1)+\epsilon \log n}), x^{-1} Z_n)$$

is $\epsilon \log n$-aligned. Lemma 2.6 implies that $[\mathrm{id}, x^{-1} Z_n]$ passes through the $K_1$- neighbourhood of $(x^{-1} Z_{N(i-1)}, \ldots, x^{-1} Z_N)$. In other words, $[x, Z_n]$ passes through the $(Ni + K_0)$-neighbourhood of id, which is within the $n^{3\alpha}$-neighbourhood of id when $n$ is large. $\qquad \square$

COROLLARY 4.8. *For any $\alpha > 0$, there exists $K'$ such that for each $0 \leq m \leq n$, we have*

$$\mathbb{E}[(\mathrm{id}, Z_n)^2_{Z_m}] \leq n^{6\alpha} + K e^{-n^\alpha/K} \cdot n \leq n^{6\alpha} + K'.$$

The following lemma states that our deviation inequality (Corollary 4.8) implies a rate of convergence in Fekete's subadditivity lemma.

LEMMA 4.9. *Let*

$$\lambda := \lim_{n \to \infty} \frac{1}{n} \mathbb{E}[d(\mathrm{id}, Z_n)].$$

*Then,*

$$\lambda - \frac{1}{n} \mathbb{E}[d(\mathrm{id}, Z_n)] = o\left(\frac{1}{\sqrt{n}}\right).$$

*Proof.* Note that by the definition of the Gromov product, we have

$$\mathbb{E}[d(\mathrm{id}, Z_{n2^k})] = \sum_{i=1}^{2^k} \mathbb{E}[d(Z_{n(i-1)}, Z_{ni})] - 2 \sum_{i=1}^{k} \sum_{t=1}^{2^{k-i}} \mathbb{E}[(Z_{n2^i(t-1)}), Z_{n2^i t})_{Z_{n(2^i t-1)}}].$$

Also, by Corollary 4.8,

$$\mathbb{E}[(Z_{n2^i(t-1)}), Z_{n2^i t})_{Z_{n(2^i t-1)}}] \leq 2(n2^{i-1})^{6\alpha} + K'$$

and we also know that $\mathbb{E}[d(Z_{n(i-1)}, Z_{ni})] = \mathbb{E}[d(\mathrm{id}, Z_n)]$ for any $i \in \mathbb{N}$. Hence, for any sufficiently small $\alpha > 0$, we have

$$\left| \frac{1}{2^k} \mathbb{E}[d(\mathrm{id}, Z_{n2^k})] - \mathbb{E}[d(\mathrm{id}, Z_n)] \right| \leq \frac{2}{2^k} \sum_{i=1}^{k} 2^{k-i} \cdot (2(n2^{i-1})^{6\alpha} + K')$$

$$\lesssim n^{6\alpha} \sum_{i=1}^{k} 2^{-i/2}.$$

As $k \to \infty$, the quantity $2^{-k} \mathbb{E}[d(\mathrm{id}, Z_{n2^k})]$ converges to $L$. Picking $\alpha < 1/12$, we can send $k \to \infty$ and divide by $n$ to conclude. $\qquad\square$

We now prove the CLT (Theorem A). It is essentially the same argument as [**MS20**], but with a different deviation inequality as input.

*Proof.* We claim that for any $\epsilon > 0$, there exists $N$ sufficiently large, such that the sequence

$$\frac{1}{\sqrt{Nk}}(d(\mathrm{id}, Z_{Nk}) - \mathbb{E}[d(\mathrm{id}, Z_{Nk})])$$

converges to a Gaussian distribution up to an error at most $\epsilon$ in the Lévy distance.

Indeed, the sequence

$$\left\{ \frac{1}{\sqrt{k}}(d(\mathrm{id}, Z_k) - \mathbb{E}[d(\mathrm{id}, Z_k)]) \right\}_{k>0}$$

is eventually $2\epsilon$-close to a distribution $X$ (in the Lévy distance) as long as its $N$-jump subsequence $\{1/\sqrt{Nk}(d(\mathrm{id}, Z_{Nk}) - \mathbb{E}[d(\mathrm{id}, Z_{Nk})])\}_{k>0}$ is eventually $\epsilon$ close to $X$ for some $N \in \mathbb{N}$. To see this, we note that the difference

$$\frac{1}{\sqrt{k}}(d(\mathrm{id}, Z_k) - \mathbb{E}[d(\mathrm{id}, Z_k)]) - \frac{1}{\sqrt{k+1}}(d(\mathrm{id}, Z_{k+1}) - \mathbb{E}[d(\mathrm{id}, Z_{k+1})])$$

is $O(k^{-1/2})$ with probability $1 - o_k(1)$ (in fact, since the step sizes are bounded, this holds with probability 1, but when we consider unbounded step distributions, only $1 - o_k(1)$ is necessary). Moreover, from Lemma 4.9, we know that

$$\mathbb{E}[d(\mathrm{id}, Z_{Nk})] = LNk + o(\sqrt{Nk}).$$

To show the claim, we first take a sequence

$$0 = j(0) < j(1) < \cdots < j(2^{\lfloor \log_2 k \rfloor}) = k$$

such that $j(t+1) - j(t) = 1$ or $2$ for each $t$. The easiest way is to keep halving the numbers, that is,

$$j(2^t k) := \left\lfloor \frac{j(2^t(k-1)) + j(2^t(k+1))}{2} \right\rfloor$$

for each $t$ and odd $k$. Let $T$ be the collection of $j(t)$ terms such that $j(t+1) - j(t) = 2$.

Then,

$$\frac{1}{\sqrt{Nk}}(d(\mathrm{id}, Z_{Nk}) - \mathbb{E}[d(\mathrm{id}, Z_{nk})]) = I_1 - I_2 - I_3 m,$$

where

$$I_1 = \sum_{i=1}^{k} \frac{1}{\sqrt{k}} \left[ \frac{d(Z_{N(i-1)}, Z_{Ni}) - \mathbb{E}[d(Z_{N(i-1)}, Z_{Ni})]}{\sqrt{N}} \right],$$

$$I_2 = \frac{2}{\sqrt{Nk}} \sum_{t \in T} ((Z_{Nj(t)}, Z_{N(j(t)+2)}) Z_{N(j(t)+1)} - \mathbb{E}[(Z_{Nj(t)}, Z_{N(j(t)+2)}) Z_{N(j(t)+1)}])$$

and

$$I_3 = \frac{2}{\sqrt{Nk}} \sum_{t=1}^{\lfloor \log_2 k - 1 \rfloor} \sum_{l=1}^{2^{\lfloor \log_2 k \rfloor - t - 1}} ((Z_{N2^t l}, Z_{N2^t(l+2)}) Z_{N2^t(l+1)} - \mathbb{E}[(Z_{N2^t l}, Z_{N2^t(l+2)}) Z_{N2^t(l+1)}]).$$

We claim that for sufficiently large $N \in \mathbb{N}$, $I_2$ and $I_3$ are small (in terms of the Lévy distance). Then the only non-negligible term $I_1$ is a sum of i.i.d random variables, normalized to converge to a Gaussian as $k \to \infty$.

The second summation $I_2$ is the sum of at most $k$ independent RVs whose variance is bounded by

$$\frac{4}{Nk} \cdot 3N^{6\alpha}.$$

Hence, the second summation has variance at most $12N^{6\alpha-1}$ and

$$\mathbb{P}(|I_2| \geq N^{-\alpha}) \leq 12N^{8\alpha-1}$$

by Chebyshev.

Now, for each $t$,

$$I_{3;t} := \frac{2}{\sqrt{Nk}} \sum_{l=1}^{2^{\lfloor \log_2 k \rfloor - t - 1}} ((Z_{N2^t l}, Z_{N2^t(l+2)}) Z_{N2^t(l+1)} - \mathbb{E}[(Z_{N2^t l}, Z_{N2^t(l+2)}) Z_{N2^t(l+1)}])$$

is the sum of at most $k/2^t$ independent RVs whose variance is bounded by $4/Nk \cdot 3(N2^t)^{6\alpha}$. This means that $I_{3;t}$ has variance at most $12N^{6\alpha-1} \cdot 2^{(6\alpha-1)t}$ and

$$\mathbb{P}(|I_{3;t}| \geq N^{-\alpha}2^{-\alpha t}) \leq 12N^{8\alpha-1}2^{(8\alpha-1)t}$$

by Chebyshev.

Summing them up, we have

$$\left| I_2 + \sum_t I_{3;t} \right| \leq N^{-\alpha} \sum_t 2^{-\alpha t}$$

outside a set of probability $N^{8\alpha-1} \sum_t 2^{(8\alpha-1)t}$. These are small, regardless of the range of $t$. More precisely, by setting $\alpha = 1/10$, we deduce that

$$|I_2 + I_3| = O(N^{-1/10})$$

outside a set of probability $O(N^{-1/10})$, ending the proof. $\qquad\square$

4.3. *Some comments on Theorem A.* We can also establish the CLT for random walks with finite $p$th moment for some $p > 2$. It suffices to show that Corollary 4.8 holds for such random walks.

For some $q > 0$, let $E$ be the event that $\sum_{i=1}^{n} |g_i|$ is at least $n^q$. We note the following inequality:

$$\mathbb{E}[n^{q(p-2)}\bigg(\sum_{i=1}^{n} |g_i|\bigg)^2 1_{\sum_{i=1}^{n} |g_i| \geq n^q}] \leq \mathbb{E}\bigg[\bigg(\sum_{i=1}^{n} |g_i|\bigg)^p\bigg]$$

$$\leq \mathbb{E}\bigg[(n \max_{i=1}^{n} |g_i|)^p\bigg]$$

$$\leq n^p \mathbb{E}\bigg[\sum_{i=1}^{n} |g_i|^p\bigg]$$

$$\leq n^{p+1} \mathbb{E}|g|^p.$$

This implies that $\mathbb{E}[(\sum_{i=1}^{n} |g_i|)^2 1_E] \leq Cn^{(p+1)-q(p-2)}$. By taking $q > (p+1)/(p-2)$, we can keep this bounded.

Now, on the event $E^c$, we argue as in Lemma 4.7. We remark that the only place we used the finite support assumption was to invoke Lemma 4.2. In particular, we needed

$$\epsilon \log(|w_0| + \cdots + |w_k| + k\epsilon \log n) \leq \log n,$$

where $w_i$. However, on the event $E^c$, this assumption is still met, replacing $\epsilon$ with $\epsilon/q$ if necessary. Then, we may still apply Lemma 4.2. Hence, we get

$$\mathbb{E}[(\mathrm{id}, Z_n)^2_{Z_m}] \leq n^{6\alpha} + Ke^{-n\alpha/K} \leq 2n^{6\alpha} + K'.$$

Given this estimate, we get the following theorem.

THEOREM 4.10. *Let $G$ be a finitely generated group with exponential growth and suppose that $G$ has a superlinear-divergent quasi-geodesic. Let $\mu$ be an admissible measure on $G$ with finite $p$-moment for some $p > 2$, and $(Z_n)_n$ be the random walk on $G$ generated by $\mu$. Then, there exist constants $\lambda, \sigma \geq 0$ such that*

$$\frac{|Z_n| - \lambda n}{\sqrt{n}} \to \mathcal{N}(0, \sigma^2).$$

We now discuss the positivity of the escape rate $\lambda$ and the asymptotic variance $\sigma$ in Theorem A. We first recall a classical result about non-amenable groups [**Gui80**, **Kes59**].

THEOREM 4.11. *For any simple symmetric random walk $(Z_n)_{n>0}$ on a non-amenable group, we have*

$$\lim_{n \to \infty} \frac{\mathbb{E}[|Z_n|]}{n} > 0.$$

Using this theorem and [**MS20**, Theorem 4.12], we observe the following.

THEOREM 4.12. *Let G be a finitely generated non-amenable group and suppose that G has a superlinear-divergent quasi-geodesic* $\gamma : \mathbb{Z} \to G$. *Let* $(Z_n)_{n \geq 1}$ *be a simple symmetric random walk on G. Then, there exist* $\lambda, \sigma > 0$ *such that*

$$\frac{|Z_n| - \lambda n}{\sqrt{n}} \to \mathcal{N}(0, \sigma^2).$$

*Proof.* First note that the escape rate $\lambda := \lim_n \mathbb{E}|Z_n|/n$ is positive by Theorem 4.11.

Let

$$\sigma_n^2 := Var\left(\frac{1}{\sqrt{n}}d(\mathrm{id}, Z_n)\right).$$

In the proof of Theorem A, we showed that for each $\epsilon > 0$, for sufficiently large $N$,

$$\left\{\frac{1}{\sqrt{k}}(d(\mathrm{id}, Z_k) - \mathbb{E}[d(\mathrm{id}, Z_k)])\right\}_{k \geq 1}$$

converges to a Gaussian RV $\mathcal{N}(0, \sigma_N^2)$ in distribution (by the classical CLT) up to an error of $O(N^{-1/10})$ in Lévy distance. From this, we deduce that $\sigma_N$ converges to a constant $\sigma \geq 0$ as $N$ tends to infinity. If we show that $\sigma > 0$, then the sequence of random variables $\{|Z_n| - \mathbb{E}|Z_n|/\sqrt{n}\}_{n \geq 1}$ converges in distribution to a non-degenerate Gaussian $\mathcal{N}(0, \sigma^2)$, as desired.

First, the proof of [**MS20**, Theorem 4.12] guarantees an $\epsilon > 0$ such that

$$\mathbb{E}((|Z_{2n}| - 2\lambda n)^2) + \mathbb{E}((|Z_{2n+2\lceil\sqrt{n}\rceil}| - 2\lambda n - 2\lambda\lceil\sqrt{n}\rceil)^2) > \epsilon n$$

holds for all large enough $n$. This follows from the fact that $\mu^{*2}$ puts non-zero weight on the identity element (since $\mu$ is symmetric), which implies that the even-step distributions $(Z_{2n})_{n>0}$ are {id}-consistent. Note that, even though [**MS20**, Theorem 4.12] is stated with the second moment deviation inequality assumption, the above claim only requires {id}-consistency. As a result, we conclude that

$$\limsup_n \frac{1}{n}\mathbb{E}((|Z_{2n}| - 2\lambda n)^2) > \epsilon.$$

Now, observe that

$$\frac{1}{n}\mathbb{E}((|Z_{2n}| - 2\lambda n)^2) = \frac{1}{n}\mathbb{E}((|Z_{2n}| - \mathbb{E}|Z_{2n}| + \mathbb{E}|Z_{2n}| - 2\lambda n)^2)$$

$$= \frac{1}{n}(\mathbb{E}((|Z_{2n}| - \mathbb{E}|Z_{2n}|)^2) + (\mathbb{E}|Z_{2n}| - 2\lambda n)^2),$$

and $(\mathbb{E}|Z_{2n}| - 2\lambda n)^2$ is $o(n)$ by Lemma 4.9. Combining these, we also deduce that

$$\limsup_n \sigma_n^2 = \limsup_n \frac{1}{n}\mathbb{E}((|Z_{2n}| - \mathbb{E}|Z_{2n}|)^2) > \epsilon$$

and $\sigma = \lim_n \sigma_n > 0$, as desired. $\qquad\square$

## A. *Appendix. Right-angled Coxeter groups*

Let $\Gamma = (V, E)$ be a finite simple graph. We can define the *right-angled Coxeter group* by the presentation

$$W_\Gamma = \langle v \in V | v^2, [v, w], (v, w) \in E \rangle.$$

In this appendix, we show the following.

LEMMA A.1. *Let $W_\Gamma$ be a right-angled Coxeter group of thickness $k \geq 2$. Then, any Cayley graph of $\Gamma$ contains a periodic geodesic $\sigma$ which is $(f, \theta)$-divergent for some $\theta > 0$ and $f(r) \gtrsim r^k$.*

We only need to slightly modify the proof of [**Lev22**, Theorem C]. They show that an RACG of thickness at least $k$ has divergence at least polynomial of degree $k + 1$. To do this, they construct a periodic geodesic $\gamma$ such that for any path $\kappa$ with endpoints on $\gamma$ and avoiding an $r$-neighbourhood of $\gamma$ term's midpoint, any segment of $\kappa$ with projection at least some constant has to have length at least $r^k$. By integrating, they get $r^{k+1}$. For completeness, we include the proof below.

*Proof.* Since the claim is quasi-isometry invariant, we work on the Davis complex $\Sigma_\Gamma$. We modify the proof of [**Lev22**, Theorem C], borrowing their notation and terminology. Take the word $w$ in the proof, so that $\sigma$ is a bi-infinite geodesic which is the axis of $w$, and set $p_i = w^i$. Since the Davis complex is a CAT(0) cube complex, the nearest point projection $\pi : \Sigma_\Gamma \to \sigma$ is well defined and 1-Lipschitz.

Let $\kappa : [0, t] \to \Sigma_\Gamma$ be a path whose projection has diameter at least $2|w|$, which is disjoint from the $|w|r$-neighbourhood around some $w^i$. As the projection of $\kappa$ has length at least $2|w|$, we can find some points $p_j$, $p_k$ such that

$$\pi(\kappa(0)) < p_j < p_k < \pi(\kappa(t))$$

in the orientation on $\sigma$. Here, $p_j, p_k = w^j, w^k$.

For the rest of the proof, we follow [**Lev22**]. For some $j \leq i < k$, let $H_i$ (respectively $K_i$) be the hyperplane dual to the edge of $\sigma$ which is adjacent to $p_i$ (respectively $p_{i+1}$) and is labelled by $s_0$ (respectively $s_n$). As hyperplanes separate $\Sigma_\Gamma$ and do not intersect geodesics twice, it follows that $H_i$ (respectively $K_i$) intersects $\kappa$. Let $e_i$ (respectively $f_i$) be the last (respectively first) edge of $\kappa$ dual to $H_i$ (respectively $K_i$). Let $\gamma_i$ (respectively $\eta_i$) be a minimal length geodesic in the carrier $N(H_i)$ (respectively $N(K_i)$) with starting point $p_i$ (respectively $p_{i+1}$) and endpoint on $e_i$ (respectively $f_i$).

Let $\alpha_i$ be the subpath of $\kappa$ between $\gamma_i \cap \kappa$ and $\eta_i \cap \kappa$. As $w$ is a $\Gamma$-complete word, no pair of hyperplanes dual to $\sigma$ intersect. By our choices, $\alpha_i \cap \alpha_j$ is either empty or a single vertex for all $i \neq j$. Let $D_i$ be the disk diagram with boundary path $\gamma_i \alpha_i \eta_i^{-1} \beta_i$, where $\beta_i$ has label $w^{-1}$. For each $0 \leq i \leq r - 2$, we observe the following.

(1)    The path $\gamma_i$ is reduced.
(2)    By Lemma 7.2, no $(k-1)$-fence connects $\gamma_i$ to $\eta_i^{-1}$ in any disk diagram with boundary path $\gamma_i \alpha_i \eta_i^{-1} \beta_i$.
(3)    The path $\alpha_i$ does not intersect the ball $B_{p_i}(|w|(r))$.

Thus, we can apply [**Lev22**, Theorem 6.2] to $D_i$ by setting, in that theorem,

$$\gamma = \gamma_i, \quad \alpha = \alpha_i, \quad \eta = \eta_i^{-1}, \quad \beta = \beta_i \quad \text{and} \quad L = k - 1 \, R = |w|(r - i).$$

We conclude that for $r$ large enough,

$$|\alpha_i| \geq C'(|w|(r)^k).$$

As $\alpha_i$ is a subsegment of $p$, we are done. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

## B. *Appendix. Superlinear-divergence and contracting properties*

Superlinear divergence is reminiscent of the weakly and strongly contracting properties of subsets of a metric space. As we shall see, superlinear divergence implies weakly contracting property, but it is logically independent with strongly contracting property.

The notion of a weakly contracting subset was first introduced in [**MM99**, Definition 2.2]. We recall the definition here.

*Definition B.1.* (Weakly contracting sets) For a subset $Z \subseteq X$ of a metric space $X$, constants $K > 0$ and $0 < \rho < 1$, the subset $Z$ is said to be $(\rho, K)$-*weakly contracting* if there exists a coarsely Lipschitz projection $\pi : X \to Z$ such that if $d(x, \pi(x)) \geq R$ and $d(x, y) \leq \rho d(x, \pi(x))$, then

$$\operatorname{diam}(\pi(x) \cup \pi(y)) \leq K.$$

It follows immediately from Lemma 2.4 that a superlinear-divergent subset is weakly contracting.

The difference between the two notions can be intuitively understood as follows. Let $B$ be a ball of radius $R$ disjoint from a subset $Z$. If $Z$ is weakly contracting, the diameter of the coarsely Lipschitz projection of $B$ to $Z$ is at most logarithmic to $R$. This is analogous to Lemma 2.6, but strictly weaker. As $\delta$ goes to 0, Lemma 2.6 can be understood to say that, if $Z$ is assumed to be superlinear divergent, the diameter of the projection of $B$ to $Z$ is sub-logarithmic to $R$. This is the key difference between the two definitions; the arguments of this paper do not work for weakly contracting geodesics.

We now give two constructions that illustrate the logical independence between superlinear divergence and strongly contracting property. We first recall the definition of strongly contracting property.

*Definition B.2.* (Strongly contracting sets) For a subset $Z \subseteq X$ of a metric space $X$, we define the *closest point projection* of $x \in X$ to $Z$ by

$$\pi_Z(x) := \{a \in Z : d_X(x, a) = d_X(x, Z)\}.$$

For $K > 0$, the subset $Z$ is said to be *K-strongly contracting* if:

(1)   $\pi_Z(z) \neq \emptyset$ for all $z \in X$; and
(2)   for any geodesic $\eta$ such that $\eta \cap N_K(Z) = \emptyset$, we have $\mathrm{diam}(\pi_Z(\eta)) \leq K$.

LEMMA B.3. *There exists a finitely generated group G containing an element whose axis is strongly contracting but not superlinear divergent.*

*Proof.* Let $G$ be the group constructed by Gersten in [**Ger94**]:

$$G = \langle x, y, t \,|\, (\mathrm{txt}^{-1} = y, xy = yx \rangle.$$

The group $G$ naturally acts on the universal cover of its presentation complex, which is a CAT(0) cube complex. Recall that the presentation complex of $G$ is defined as follows: start with a single 0-cell, attach a 1-cell for each of the three generators $x$, $y$, $t$, and attach a 2-cell for each of the relations $[x, y]$ and $(\mathrm{txt}^{-1}y^{-1}$. Let $X$ be the universal cover of this complex, which Gersten shows is CAT(0) [**Ger94**, Proposition 3.1].

The induced combinatorial metric on $X$ is isometric to the word metric with respect to $\{x, y, t\}$.

Let $g = (\mathrm{tx}$ and $\gamma$ be a path connecting $(\ldots, \mathrm{id}, \mathrm{t}, (\mathrm{tx}, (\mathrm{txt}, ((\mathrm{tx})^2, \ldots)$. Then, $\gamma$ is a $g$-invariant geodesic, and $\gamma$ does not bound a flat half-plane (the cone angle of $\gamma$ at its each vertex is $3\pi/2$). Hence, $\gamma$ is rank-1 and we can conclude that $g$ is strongly contracting.

Meanwhile, by [**Ger94**, Theorem 4.3], $G$ has quadratic divergence. Given an appropriate action of $G$ on a hyperbolic space, we would be able to conclude from [**GS21**, Lemma 3.6] that $\gamma$ is not superlinear divergent. Since we do not assume a hyperbolic action, we instead present a modification of Goldborough and Sisto's argument.

Suppose that there exists an $(A, B)$-coarsely Lipschitz projection $\pi_\gamma : G \to \gamma$, a constant $\theta > 0$ and a superlinear function $f$ such that $\gamma$ is $(f, \theta)$-divergent with respect to $\pi_\gamma$. Up to a finite additive error, we may assume that $\pi_\gamma$ takes the values $\{(zx)^i : i \in \mathbb{Z}\}$.

Let $\epsilon = 1/2(A + 3)$ and let $n$ be a sufficiently large integer. We make the following claim.

CLAIM B.4. *If a point $p \in G \setminus B(\mathrm{id}, n)$ satisfies $d(p, \gamma) \leq \epsilon n$, then $\pi_\gamma(p) = (\mathrm{tx})^i$ for some $|i| > 0.5n$.*

*Proof of Claim B.4.* First, from $d(p, \gamma) \leq \epsilon n$ and the coarse Lipschitzness of $\pi_\gamma$, we deduce

$$d(p, \pi_\gamma(p)) \leq (A + 1)\epsilon n + 2B.$$

Hence, we have

$$d(\mathrm{id}, \pi_\gamma(p)) \geq d(\mathrm{id}, p) - d(p, \pi_\gamma(n)) \geq n - (A + 1)\epsilon n - 2B > 0.5n$$

and the claim follows.   □

Next, we let

$$a_n = (\mathrm{tx})^{(1-\epsilon)n} y^{-\lfloor \epsilon n \rfloor}, \quad b_n = (\mathrm{tx})^{-(1-\epsilon)n} y^{-\lfloor \epsilon n \rfloor}$$

and let $\eta$ be an arbitrary path in $G \setminus B(\mathrm{id}, n)$ connecting $a_n$ and $b_n$. Let $m, m' \in \mathbb{Z}$ be such that $\pi_\gamma(a_n) = (\mathrm{tx})^m$ and $\pi_\gamma(b_n) = ((\mathrm{tx})^{m'}$. We then have

$$d(((\mathrm{tx})^n, \pi_\gamma(a_n)) \le d((\mathrm{tx})^n, a_n) + d(a_n, \pi_\gamma(a_n)) \le (A+2)\epsilon n + 2B < 0.5n.$$

It follows that $m > n - 0.5n \ge 0.5n$. Similarly, we deduce $m' < -0.5n$.

We examine the two connected components of $\eta \setminus N_{\epsilon n}(\gamma)$ as well as $\eta \cap N_{\epsilon n}(\gamma)$. Each component of $\eta \cap N_{\epsilon n}(\gamma)$ attains values of $\pi_\gamma(\cdot)$ in

$$\{(\mathrm{tx})^i : i < -0.5n\} \quad \text{or} \quad \{(\mathrm{tx})^i : i > 0.5n\},$$

by Claim B.4, but not in both (by the coarse Lipschitzness of $\pi_\gamma$). Meanwhile, the endpoints of $\eta$ attain values of $\pi_\gamma(\cdot)$ in $\{(\mathrm{tx})^i : i < -0.5n\}$ and $\{(\mathrm{tx})^i : i > 0.5n\}$, respectively. As a result, there exists a subsegment $\eta'$ of $\eta$, as a component of $\eta \setminus N_{\epsilon n}(\gamma)$, such that

$$\pi_\gamma(\eta'^+) \in \{(\mathrm{tx})^i : i > 0.5n\} \quad \text{and} \quad \pi_\gamma(\eta'^-) \in \{(\mathrm{tx})^i : i < -0.5n\}.$$

It follows that the length of $\eta'$ is at least $(n/\theta) \cdot f(\epsilon n)$. Since $\eta$ is longer than $\eta'$, we deduce that an arbitrary path in $G \setminus B(\mathrm{id}, n)$ connecting $a_n, b_n \in B(\mathrm{id}, n)$ is longer than $(n/\theta) \cdot f(\epsilon n)$. When $n$ increases, this contradicts the quadratic divergence of $G$. Hence, we deduce that $\gamma$ is not superlinear divergent. $\qquad\square$

LEMMA B.5. *There exists a proper geodesic metric space $X$ that contains a superlinear-divergent geodesic $\gamma$ that is not strongly contracting.*

*Proof.* Let $X = \mathbb{H}^2$ and $\gamma$ be a bi-infinite geodesic $\gamma$ on $X$ with respect to the standard Poincaré metric $ds_0^2$. Let $o \in \gamma$ be a reference point on $\gamma$ and let $\mathrm{proj}_\gamma$ be the closest point projection onto $\gamma$ with respect to $ds_0^2$. For each $x \in X$, let $r$ be the (directed) distance from $x$ to $\gamma$ and let $\tau$ be the (directed) distance from $o$ to $\mathrm{proj}_\gamma(x)$. Since $(r, \tau)$ is an orthogonal parametrization of $X$, there exists a continuous coefficient $F_0$ such that

$$ds_0^2 = dr^2 + F_0(x)d\tau^2$$

holds at each point $x \in X$. We note that $F_0(x) \sim e^{\kappa r(x)}$ for some $\kappa > 0$ (due to the Gromov hyperbolicity of $(X, ds_0^2)$) and $F_0(x) \ge 1$.

We will now define a new metric $ds^2$ as follows. For each $i > 0$ and $j \in \mathbb{Z}$, let

$$I_{i,j} = \{(r, \tau) : r = 4^{2^i}, 2j + i \le \tau \le 2j + i + 1\}$$

and let

$$S := \bigcup_{i > 0, \, j \in \mathbb{Z}} I_{i,j}.$$

Let $\chi : \mathbb{R}^2 \to [0, 1]$ be a smooth function that takes value 0 on $S$ and 1 on $\mathbb{R}^2 \setminus N_{0.1}(S)$. We finally define

$$F(x) := F_0(x) \cdot \chi(r(x), \tau(x)) + (1 - \chi(r(x), \tau(x)))$$

and

$$ds^2 := dr^2 + F(x)d\tau^2.$$

First, $\mathrm{proj}_\gamma$ is still the closest point projection with respect to $ds^2$. Indeed, the shortest path from $x \in X$ to $\gamma$ is the one that does not change in the value of $\tau$. As a corollary, the $K$-neighbourhoods of $\gamma$ with respect to the two metrics coincide.

Let $i$ be a positive integer and let $x, y \in X$ be such that $r(x) = r(x) = 4^{2^{4i}}$ and $\tau(x) = 0$, $\tau(y) = 2i$. We first consider a path $\eta$ connecting $x$ to $y$ while passing through $N_K(\gamma)$. Then, the total length is at least $2 \cdot (4^{2^{4i}} - K)$. Next, we take a piecewise geodesic path $\eta'$ that goes like

$$(r, \tau) = (4^{2^{4i}}, 0) - (4^{2^{4i}}, 1) - (4^{2^{4i-1}}, 1) - (4^{2^{4i-1}}, 2) - \cdots$$
$$- (4^{2^{3i+1}}, i) - (4^{2^{3i}}, i) - (4^{2^{3i}}, i+1) - (4^{2^{3i+1}}, i+1) - \cdots - (4^{2^i}, 2i).$$

Then, the total length is $2(4^{2^{4i}} - 4^{2^{3i}}) + 2i$. Note also that $\eta'$ does not intersect $N_K(\gamma)$. We conclude that the geodesic connecting $x$ to $y$ does not touch $N_K(\gamma)$. Note also that the projection is larger than $2i$. By increasing $i$, we conclude that $\gamma$ is not $K$-strongly contracting for any $K > 0$.

Meanwhile, it is superlinear divergent. To see this, suppose a path $\eta$ lies in $X \setminus N_R(\gamma)$ and satisfies $\pi_\gamma(\eta) > 4$. Then, $\pi_\gamma(\eta)$ contains $[2k, 2k+2]$ for some integer $k$, and by restricting the path if necessary, we may assume $\pi_\gamma(\eta) = \gamma([2k, 2k+2])$.

If $r(\eta)$ ever takes two values among $\{4^{2^i} : i > 0\} \cap [R, +\infty)$, say $4^{2^m}$ and $4^{2^{m'}}$ for some $m < m'$, then the total variation of $r(\eta(t))$ is at least

$$4^{2^{m+1}} - 4^{2^m} = 4^{2^m}(4^{2^m} - 1) \geq R^2/2.$$

Consequently, we have $l(\eta) \geq 0.5R^2$.

If not, $r(\eta)$ takes at most one value $4^{2^i}$ among $\{4^{2^j} : j > 0\}$. If $i$ is even, then

$$F(\eta(t)) = F_0(\eta(t))$$

for $t$ such that $\tau(\eta(t)) \in [2k+1.1, 2k+1.9]$. Since

$$F_0(\eta(t)) \geq e^{\kappa r(\eta(t))} \geq e^{\kappa R},$$

we have

$$l(\eta) \geq \int F(\eta) \, d\tau(\eta) \geq e^{\kappa R} \times 0.8 = 0.8^{\kappa R}.$$

Similarly, we have $l(\eta) \geq 0.8 e^{\kappa R}$ when $i$ is odd. This concludes that $\gamma$ is superlinear divergent. $\qquad \square$

Finally, we remark that superlinear divergence is invariant under quasi-isometry, but the notion of strongly contracting is not. For example, let $X$ be the Cayley graph of a group $G$ equipped with the word metric associated to some finite generating set $\mathcal{S}$ and let $Z$ be a superlinear-divergent set in $X$. Then, changing the generating set changes the metric in $X$ by a quasi-isometry, and hence, $Z$ is still a superlinear-divergent set. However, if $\gamma$ is a strongly contracting geodesic in $X$, it may not be strongly contracting with respect to the new metric.

As an explicit example, it was shown in [SZ22, Theorem C] that each mapping class group admits a proper cobounded action on a metric space $X$ such that all pseudo-Anosov

elements have strongly contracting quasi-axes in *X*. To contrast, it was shown in [**RV21**, Theorem 1.4] that the mapping class group of the five-times punctured sphere can be equipped with a word metric such that the axis of a certain pseudo-Anosov map in the Cayley graph is not strongly contracting.

## References

[ACT15]   G. N. Arzhantseva, C. H. Cashen and J. Tao. Growth tight actions. *Pacific J. Math.* **278**(1) (2015), 1–49.

[BCF24]   J. Behrstock, R. A. Çiçeksiz and V. Falgas-Ravry. A threshold for relative hyperbolicity in random right-angled Coxeter groups. *Preprint*, 2024, arXiv:2407.12959.

[Bel54]   R. Bellman. Limit theorems for non-commutative operations. I. *Duke Math. J.* **21** (1954), 491–500.

[BH99]   M. R. Bridson and A. Haefliger. *Metric Spaces of Non-positive Curvature (Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], 319)*. Springer-Verlag, Berlin, 1999.

[BHS17]   J. Behrstock, M. F. Hagen and A. Sisto. Thickness, relative hyperbolicity, and randomness in Coxeter groups. *Algebr. Geom. Topol.* **17**(2) (2017), 705–740; with an appendix written jointly with P.-E. Caprace.

[BHS19]   J. Behrstock, M. Hagen and A. Sisto. Hierarchically hyperbolic spaces II: combination theorems and the distance formula. *Pacific J. Math.* **299**(2) (2019), 257–338.

[Bjö10]   M. Björklund. Central limit theorems for Gromov hyperbolic groups. *J. Theoret. Probab.* **23**(3) (2010), 871–887.

[BMSS23]   A. Bounlanger, P. Mathieu, Ç. Sert and A. Sisto. Large deviations for random walks on hyperbolic spaces. *Ann. Sci. Éc. Norm. Supér. (4)* **56**(3) (2023), 885–944.

[BQ16]   Y. Benoist and J.-F. Quint. *Random Walks on Reductive Groups (Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], 62)*. Springer, Cham, 2016.

[CFFT22]   K. Chawla, B. Forghani, J. Frisch and G. Tiozzo. The Poisson boundary of hyperbolic groups without moment conditions. *Preprint*, 2022, arXiv:2209.02114.

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

[DK07]   B. Deroin and V. Kleptsyn. Random conformal dynamical systems. *Geom. Funct. Anal.* **17**(4) (2007), 1043–1105.

[FK60]   H. Furstenberg and H. Kesten. Products of random matrices. *Ann. Math. Stat.* **31** (1960), 457–469.

[Fur63]   H. Furstenberg. Noncommuting random products. *Trans. Amer. Math. Soc.* **108** (1963), 377–428.

[Ger94]   S. M. Gersten. Quadratic divergence of geodesics in spaces. *Geom. Funct. Anal.* **4**(1) (1994), 37–51.

[Gou17]   S. Gouëzel. Analyticity of the entropy and the escape rate of random walks in hyperbolic groups. *Discrete Anal.* **7** (2017), 37.

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

[GS21]   A. Goldsborough and A. Sisto. Markov chains on hyperbolic-like groups and quasi-isometries. *Preprint*, 2021, arXiv:2111.09837.

[Gui80]   Y. Guivarc'h. Sur la loi des grands nombres et le rayon spectral d'une marche aléatoire. *Astérisque* **74**(3) (1980), 47–98.

[Hor18]   C. Horbez. Central limit theorems for mapping class groups and $Out(F_N)$. *Geom. Topol.* **22**(1) (2018), 105–156.

[Kai00]   V. A. Kaimanovich. The Poisson formula for groups with hyperbolic properties. *Ann. of Math. (2)* **152**(3) (2000), 659–692.

[Kes59]   H. Kesten. Symmetric random walks on groups. *Trans. Amer. Math. Soc.* **92** (1959), 336–354.

[KM99]   A. Karlsson and G. A. Margulis. A multiplicative ergodic theorem and nonpositively curved spaces. *Comm. Math. Phys.* **208**(1) (1999), 107–123.

[Lev22]   I. Levcovitz. Characterizing divergence and thickness in right-angled Coxeter groups. *J. Topol.* **15**(4) (2022), 2143–2173.

[MM99]   H. Masur and Y. Minsky. Geometry of the complex of curves I: Hyperbolicity. *Invent. Math.* **138** (1999), 103–149.

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

[MT18] J. Maher and G. Tiozzo. Random walks on weakly hyperbolic groups. *J. Reine Angew. Math.* **742** (2018), 187–239.

[Osi16] D. Osin. Acylindrically hyperbolic groups. *Trans. Amer. Math. Soc.* **368**(2) (2016), 851–888.

[RV21] K. Rafi and Y. Verberne. Geodesics in the mapping class group. *Algebr. Geom. Topol.* **21**(6) (2021), 2995–3017.

[Sis18] A. Sisto. Contracting elements and random walks. *J. Reine Angew. Math.* **742** (2018), 79–114.

[SZ22] A. Sisto and A. Zalloum. Morse subsets of injective spaces are strongly contracting. *Preprint*, 2022, arXiv:2208.13859.