Illinois Journal of Mathematics
Volume 58, Number 4, Winter 2014, Pages 939–980
S 0019-2082

# DIVERGENCE, THICK GROUPS, AND SHORT CONJUGATORS

JASON BEHRSTOCK AND CORNELIA DRUŢU

ABSTRACT. The notion of thickness, introduced in (*Math. Ann.* **344** (2009) 543–595), is one of the first tools developed to study the quasi-isometric behavior of weakly relatively hyperbolic groups. In this paper, we further this exploration through a relationship between thickness and the divergence of geodesics. We construct examples, for every positive integer $n$, of CAT(0) groups which are thick of order $n$ and with polynomial divergence of order $n+1$. With respect to thickness, these examples show the non-triviality at each level of the thickness hierarchy defined in (*Math. Ann.* **344** (2009) 543–595). With respect to divergence, our examples provide an answer to questions of Gromov (In *Geometric Group Theory* (1993) 1–295 Cambridge Univ. Press) and Gersten (*Geom. Funct. Anal.* **4** (1994) 633–647; *Geom. Funct. Anal.* **4** (1994) 37–51). The divergence questions were independently answered by Macura in (CAT(0) spaces with polynomial divergence of geodesics (2011) Preprint).

We also provide tools for obtaining both lower and upper bounds on the divergence of geodesics and spaces, and we prove an effective quadratic lower bound for Morse quasi-geodesics in CAT(0) spaces, generalizing results of Kapovich–Leeb and Bestvina–Fujiwara (*Geom. Funct. Anal.* **8** (1998) 841–852; *Geom. Funct. Anal.* **19** (2009) 11–40).

In the final section, we obtain linear and quadratic bounds on the length of the shortest conjugators for various families of

Received May 16, 2014; received in final form June 5, 2015.

The research of the first author was partially supported as an Alfred P. Sloan Fellow and by NSF grant #DMS-1006219. The research of the second author was supported in part by the EPSRC grant "Geometric and analytic aspects of infinite groups" and by the project ANR Blanc ANR-10-BLAN 0116, acronyme GGAA.

2010 *Mathematics Subject Classification.* Primary 20F65, 53C23, 57M07. Secondary 20F10, 20F67, 20F69, 57N10.

groups. For general 3-manifold groups, sharp estimates are provided. We also consider mapping class groups, where we provide a new streamlined proof of the length of shortest conjugators which contains the corresponding results of Masur–Minsky in the pseudo-Anosov case (*Geom. Funct. Anal.* **10** (2000) 902–974) and Tao in the reducible case (*Geom. Funct. Anal.* **23** (2013) 415–466).

## 1. Introduction

One of the main purposes of this paper is to provide a connection between two invariants: the divergence and the order of thickness. The *divergence* arose in the study of non-positively curved manifolds and metric spaces and roughly speaking it measures the spread of geodesics. More precisely, given two geodesic rays $r, r'$ with $r(0) = r'(0)$ their divergence is defined as a map $\mathrm{div}_{r,r'}\colon \mathbb{R}_+ \to \mathbb{R}_+$, where $\mathrm{div}_{r,r'}(t)$ is the infimum of the lengths of paths joining $r(t)$ to $r'(t)$ outside the open ball centered at $r(0)$ and of radius $\lambda t$. Here, $\lambda$ is a fixed parameter in $(0, 1)$ whose choice turns out to be irrelevant for the order of the divergence.

For the divergence and distance functions, we will often want to compare how two such functions behave asymptotically. In particular, given two non-decreasing functions $\mathbb{R}_+ \to \mathbb{R}_+$, $f$ and $g$, we write $f \preceq g$ if there exists a constant $C \geq 1$ for which $f(x) \leq Cg(Cx + C) + Cx + C$ for all $x \in \mathbb{R}_+$; if we want to emphasize the particular constant, we write $f \preceq_C g$. One obtains an equivalence relation on the set of functions $\mathbb{R}_+ \to \mathbb{R}_+$ by setting $f \asymp g$ if and only if there exists constants so that $f \preceq g$ and $g \preceq f$. We note that the *order of growth* of any non-constant polynomial, or an *upper bound* for such an order, is an invariant of the equivalence class.

In symmetric spaces of non-compact type, the order of the divergence of geodesic rays is either exponential (when the rank is one) or linear (when the rank is at least two). This inspired an initial thought that in the presence of non-positive curvature the divergence must be either linear or exponential. See [Gro93] for a discussion—an explicit statement of this conjecture appears in $6.B_2$, subsection "Geometry of $\partial_T$ and Morse landscape at infinity," Example (h). In particular, Gromov stated an expectation that all pairs of geodesic rays in the universal cover of a closed Riemannian manifold of non-positive curvature diverge either linearly or exponentially [Gro93].

As an aside, we note that without the hypothesis of non-positive curvature the situation is more complicated. For instance, in nilpotent groups with left invariant metrics, while the maximal rate of divergence of geodesics is linear [DMS10], there exist geodesic rays that diverge sublinearly [Pau01, Lemma 7.1].

Gersten provided the first examples of CAT(0) spaces whose divergence did not satisfy the linear/exponential dichotomy and showed that such examples are closely tied to other areas in mathematics. The first such example was a CAT(0) space admitting a cocompact action of the group $F_2 \rtimes_{\varphi} \mathbb{Z}$ with $\varphi(a) = ab$, $\varphi(b) = b$ [Ger94b]; the space constructed by Gersten contains rays with quadratic divergence and he proves that no two rays in this space diverge faster than quadratically. Extending that work, Gersten then used divergence to distinguish classes of closed 3-manifolds [Ger94a]. Modulo the geometrization conjecture, he proved that the divergence of a 3-manifold is either linear, quadratic, or exponential; where quadratic divergence occurs precisely for graph manifolds and exponential divergence occurs precisely when at least one geometric component is hyperbolic. Gersten asked explicitly in [Ger94a] which orders of polynomial divergence were possible in a CAT(0) group.

In a different direction, the authors of the present paper together with L. Mosher [BDM09] introduced a geometric property called *thickness* which was proved to hold for many interesting spaces. The definition is an inductive one and, roughly speaking, characterizes a space as *thick of order* $n$ if *it is a network of subsets* which are each thick of order $n-1$, i.e. any two points in the space can be connected by a chain of subsets thick of order $n-1$ with each intersecting the next in an infinite diameter set. The base level of the induction, that is, spaces defined to be thick of order 0, are metric spaces with linear divergence. The precise definition of *thick* is given in Section 4.

The very structure of a network turns out to be well adapted to estimates on divergence. Indeed let $X$ be a geodesic metric space which is a $(\tau, \eta)$-tight network with respect to a collection of subsets $\mathcal{L}$, in the sense of Definition 4.1. Let $\delta$ be a number in $(0,1)$ and let $\gamma \geq 0$. For every subset $L \in \mathcal{L}$ let $\mathrm{Div}_{\gamma}^{L}(n;\delta)$ be the divergence function for the tubular neighborhood of $L$ of radius $\tau$, with the induced metric (see Definition 3.1 for the notion of divergence function of a metric space); define the *network divergence* of $X$ as

$$\mathrm{Div}_{\gamma}^{\mathcal{L}}(n;\delta) = \sup_{L \in \mathcal{L}} \mathrm{Div}_{\gamma}^{L}(n;\delta).$$

The following holds.

Theorem 4.9. *The divergence function in $X$ satisfies*

$$\mathrm{Div}_{\gamma}^{X}(n;\delta) \preceq_C n\,\mathrm{Div}_{\gamma}^{\mathcal{L}}(n;\delta),$$

*where the constant $C$ only depends on the constants $\tau, \eta, \delta$ and $\gamma$.*

Groups and spaces which are thick of order 0 or 1 both yield very rich classes of examples, see e.g. [BDM09], [BC], [BM08], [DS05]. In the present paper, we give the first constructions of groups which are thick of order greater than 1; indeed, for every positive integer $n$ we produce infinitely many quasi-isometry classes of groups which are thick of order $n$, as explained in the following theorem. Moreover, using a close connection between order of thickness and order of divergence we establish that the very same classes of examples have polynomial divergence of degree $n+1$. We note that the case $n=0$ of this theorem is trivial and the case $n=1$ follows from the above mentioned results in [Ger94a] combined with results from [BDM09] and [BN08]. The following is established in Section 5.

THEOREM 1.1. *For every positive integer $n$ there exists an infinite family of pairwise non-quasi-isometric finitely generated groups which are each*:

(1) CAT(0) *groups*;
(2) *thick of order $n$*;
(3) *with divergence of order $n+1$.*

Natasa Macura has given an independent construction of examples of CAT(0) groups with divergence of order $n$ for all positive integers [Mac].

The upper bound on divergence in Theorem 1.1 will follow from Theorem 4.9. The lower bound both for divergence and for the order of thickness is proved by exhibiting a bi-infinite geodesic with divergence precisely $x^{n+1}$.

It would be interesting to know if either in general, or under some reasonable hypotheses, the order of thickness and the divergence are directly correlated, i.e. can the order of thickness be shown to provide a lower bound in addition to the upper bound which in this paper we show holds in general. A homogeneous version of this question is:

QUESTION 1.2. If a group is thick of order $n$ must its divergence be polynomial of degree exactly $n+1$?

More specific questions on the possible orders of divergence include:

QUESTION 1.3. Are there examples of CAT(0)-groups whose divergence is strictly between $x^n$ and $x^{n+1}$ for some $n$?

QUESTION 1.4. What are the $\asymp$-equivalence classes of divergence functions of CAT(0)-groups?

Our inductive construction in Theorem 1.1 can be made to yield infinitely many quasi-isometry classes because the quasi-isometry type of the base is an invariant of the space, and this base is a CAT(0) 3-dimensional graph manifold. According to the main result of [BN08], we have infinitely many quasi-isometry classes of 3-dimensional graph manifolds to choose from.

Another geometric feature relevant for divergence is the presence of *Morse quasi-geodesics.* These are quasi-geodesics which represent in some sense "hyperbolic directions" in that they satisfy the Morse lemma from hyperbolic geometry. More precisely, we call a quasi-geodesics $\mathfrak{q}$ a *Morse quasi-geodesics* if any $(K,C)$-quasi-geodesic $\gamma$ with endpoints on $\mathfrak{q}$ is contained in a $M$-tubular neighborhood of $\mathfrak{q}$, where $M$ is a uniform bound depending only on $K$ and $C$, but otherwise not dependent on $\gamma$.

We begin by observing in Proposition 3.9(b), that in CAT(0) spaces one can associate *Morse parameters* to a Morse quasi-geodesic, measuring "how hyperbolic" that quasi-geodesic is (informally speaking: a Morse quasi-geodesic may be contained in a flat strip: the larger such a strip, the larger the Morse parameter; for a precise definition of these parameters see Definition 3.10).

Morse quasi-geodesics and their relationship to divergence are studied in Section 6. One topic discussed there is the following natural refinement of Question 1.3.

QUESTION 1.5. If $X$ is a CAT(0) space, can the divergence of a Morse geodesic be greater than $x^n$ and less than $x^{n+1}$?

The following theorem provides a negative answer for the case $n = 1$; its statement is the most general version of previous known results which required extra assumptions such as periodicity of the geodesic or properness of the space $X$, see [KL98] or Proposition 3.12, and also [BF].

THEOREM 6.6. *Let $\mathfrak{q}$ be a Morse quasi-geodesic in a* CAT(0) *metric space* $(X, \mathrm{dist})$. *Then the divergence* $\mathrm{Div}^{\mathfrak{q}} \geq (\kappa x - \kappa)^2$, *where $\kappa$ is a constant depending only on the constants chosen in the definition of the divergence and on the Morse parameters* (*see Definition* 3.10).

Further results and questions on the relation between Morse (quasi-) geodesics and divergence may be found in Section 6.

In Section 7, we study the question of finding shortest conjugators for Morse elements, in CAT(0) groups and in groups with "(non-positive curvature)-like behavior." We generalize results from the CAT(0) setting to Morse geodesics in other groups. In that section, we prove the following, which we then apply to graph manifolds in Corollary 7.5. Recall that an action of a group $G$ on a graph $X$ is called $l$-*acylindrical* for some $l > 0$ (or simply *acylindrical*) if the stabilizers in $G$ of pairs of points in $X$ at distance $\geq l$ are finite of uniformly bounded sizes. Recall also that in a finitely generated group $G$, for a finite generating set $S$ which we often do not explicitly mention, we denote by $|g|_S$ or simply by $|g|$ the distance from 1 to $g \in G$ in the word metric corresponding to $S$.

THEOREM 7.4. *Let $G$ be a group acting cocompactly and $l$-acylindrically on a simplicial tree $T$. For every $R > 0$ and for a fixed word metric on $G$ let $f(R)$ denote the supremum of all diameters of intersections* $\mathrm{stab}(a) \cap \mathcal{N}_R(g\,\mathrm{stab}(b))$, *where $a$ and $b$ are vertices in $T$ at distance at least $l$, and $g \in G$ is at distance $\leq$ $R$ from* 1.

*There exists a constant $K$ such that if two loxodromic elements $u, v$ are conjugate in $G$ then there exists $g$ conjugating $u, v$ such that*

$$|g| \leq f\big(|u| + |v| + K\big) + |u| + |v| + 2K.$$

Note that in this theorem we cannot simply replace "loxodromic" by "Morse," since there might exist Morse elements of $G$ in the stabilizer of a vertex, for example, this is the case if $G$ is free and $T$ is the quotient by a free factor.

Two natural questions related to the above result can be asked.

QUESTION 1.6. Can Theorem 7.4 be extended to actions that are not cocompact?

QUESTION 1.7. What are the possible values of the function $f(R)$ in Theorem 7.4?

As a consequence of Theorem 7.4, we obtain the following.

COROLLARY 1.8. *Let $M$ be a 3-dimensional prime manifold, and let $G$ be its fundamental group. For every word metric on $G$ there exists a constant $K$ such that the following holds*:

(1) *if $u,v$ are two Morse elements conjugate in $G$ then there exists $g$ conjugating $u,v$ such that*
$$|g| \leq K\big(|u| + |v|\big).$$
(2) *If $u,v$ are two arbitrary elements conjugate in $G$, then there exists $g$ conjugating $u,v$ such that*
$$|g| \leq K\big(|u| + |v|\big)^2.$$

The fact that 3-manifolds have solvable conjugacy problem was established in [Pre06], but without a bound on the complexity.

We also give a new unified proof of the following theorem, first proved in the pseudo-Anosov case by Masur–Minsky [MM00, Theorem 7.2] and later extended to the reducible case by Tao [Tao11, Theorem B].

THEOREM 7.8. *Given a surface $S$ and a finite generating set $F$ of its mapping class group $\mathcal{MCG}(S)$ there exists a constant $C$ depending only on $S$ and on $F$ such that for every two conjugate pure elements of infinite order $u$ and $v$ in $\mathcal{MCG}(S)$ there exists $g \in \mathcal{MCG}(S)$ satisfying $v = gug^{-1}$ and*
$$|g| \leq C\big(|u| + |v|\big).$$

## 2. General preliminaries

We recall some standard definitions and establish our notation.

We use the notation $\mathcal{N}_R(A)$ for the *(open) $R$-neighborhood* of a subset $A$ in a metric space $(X, \mathrm{dist})$, i.e. $\mathcal{N}_R(A) = \{x \in X : \mathrm{dist}(x,A) < R\}$. If $A = \{a\}$ then $\mathcal{N}_R(A) = B(a,R)$ is the open $R$-ball centered at $a$.

We use the notation $\overline{\mathcal{N}}_R(A)$ and $\bar{B}(a,R)$ to designate the corresponding *closed neighborhoods* and *closed balls* defined by non-strict inequalities.

We make the convention that $B(a,R)$ and $\bar{B}(a,R)$ are the empty set for $R<0$ and any $a\in X$. The terms "neighborhood" and "ball" will always mean an open neighborhood, respectively, ball.

NOTATION 2.1. Let $a>1$, $b,x,y$ be positive real numbers. We write $x\leq_{a,b} y$ if

$$x\leq ay+b.$$

We write $x\approx_{a,b} y$ if $x\leq_{a,b} y$ and $y\leq_{a,b} x$.

Consider two constants $L\geq 1$ and $C\geq 0$.

An $(L,C)$-*coarse Lipschitz map* is a map $f:X\to Y$ of a metric space $X$ to a metric space $Y$ such that

$$\operatorname{dist}\big(f(x),f\big(x'\big)\big)\leq_{L,C}\operatorname{dist}\big(x,x'\big),\quad \text{for all } x,x'\in X.$$

An $(L,C)$-*quasi-isometric embedding* is a map $f:X\to Y$ that satisfies

$$\operatorname{dist}\big(f(x),f\big(x'\big)\big)\approx_{L,C}\operatorname{dist}\big(x,x'\big),\quad \text{for all } x,x'\in X.$$

If moreover $Y\subseteq\mathcal{N}_C(f(X))$ the map $f$ is called a *quasi-isometry*.

An $(L,C)$-*quasi-geodesic* is an $(L,C)$-quasi-isometric embedding $\mathfrak{p}\colon I\to X$, where $I$ is a connected subset of the real line. A *sub-quasi-geodesic* of $\mathfrak{p}$ is a restriction $\mathfrak{p}|_J$, where $J$ is a connected subset of $I$.

When $I=\mathbb{R}$ we call both $\mathfrak{p}$ and its image *bi-infinite* $(L,C)$*-quasi-geodesic*.

We call $(L,0)$-quasi-isometries (quasi-geodesics) *$L$-bi-Lipschitz maps (paths)*.

When the constants $L,C$ are irrelevant they are often not mentioned.

When considering divergence, it is often useful to consider a particular type of metric spaces: tree-graded spaces, as these spaces, especially in their appearance as ultralimits, are particularly relevant. Recall the following definition from [DS05]: a complete geodesic metric space $\mathbb{F}$ is *tree-graded with respect to a collection* $\mathcal{P}$ of closed geodesic subsets (called *pieces*) when the following two properties are satisfied:

($T_1$) Every two different pieces have at most one common point.
($T_2$) Every simple geodesic triangle in $\mathbb{F}$ is contained in one piece.

LEMMA 2.2 (Druţu–Sapir [DS05]). *Let $\mathbb{F}$ be a space which is tree-graded with respect to a collection of pieces $\mathcal{P}$.*

(1) *For every point $x\in\mathbb{F}$, the set $T_x$ of topological arcs originating at $x$ and intersecting any piece in at most one point is a complete real tree (possibly reduced to a point). Moreover if $y\in T_x$ then $T_y=T_x$.*
(2) *Any topological arc joining two points in a piece is contained in the same piece. Any topological arc joining two points in a tree $T_x$ is contained in the same tree $T_x$.*

Lemma 2.3 (Druţu–Sapir [DS05], Lemma 2.31). *Let $X$ be a complete geodesic metric space containing at least two points and let $\mathcal{C}$ be a non-empty set of cut-points in $X$. There exists a uniquely defined* (*maximal in an appropriate sense*) *collection $\mathcal{P}$ of subsets of $X$ such that*

- *$X$ is tree-graded with respect to $\mathcal{P}$;*
- *any piece in $\mathcal{P}$ is either a singleton or a set with no cut-point in $\mathcal{C}$.*

*Moreover, the intersection of any two distinct pieces from $\mathcal{P}$ is either empty or a point from $\mathcal{C}$.*

## 3. Divergence

Throughout this section $(X, \mathrm{dist})$, or just $X$, will denote a geodesic metric space.

**3.1. Equivalent definitions for divergence.** We recall the various definitions of the divergence and the fact that under some mild conditions all these functions are equivalent. The main reference for the first part of this section is [DMS10, §3.1].

Consider two constants $0 < \delta < 1$ and $\gamma \geq 0$.

For an arbitrary triple of points $a, b, c \in X$ with $\mathrm{dist}(c, \{a, b\}) = r > 0$, define $\mathrm{div}_\gamma(a, b, c; \delta)$ as the infimum of the lengths of paths connecting $a, b$ and avoiding the ball $\mathrm{B}(c, \delta r - \gamma)$.

If no such path exists, define $\mathrm{div}_\gamma(a, b, c; \delta) = \infty$.

Definition 3.1. The *divergence function* $\mathrm{Div}^X_\gamma(n, \delta)$ of the space $X$ is defined as the supremum of all numbers $\mathrm{div}_\gamma(a, b, c; \delta)$ with $\mathrm{dist}(a, b) \leq n$. When there is no danger of confusion, we drop the superscript $X$.

A particular type of divergence will be useful to obtain lower bounds for the function Div defined as above. More precisely, let $\mathfrak{q}$ be a bi-infinite quasi-geodesic in the space $X$, seen as a map $\mathfrak{q} \colon \mathbb{R} \to X$ satisfying the required two inequalities. We define *the divergence of this quasi-geodesic* as the function

$$\mathrm{Div}^{\mathfrak{q}}_\gamma \colon (0, +\infty) \to (0, +\infty), \qquad \mathrm{Div}^{\mathfrak{q}}_\gamma(r) = \mathrm{div}_\gamma\big(\mathfrak{q}(r), \mathfrak{q}(-r), \mathfrak{q}(0); \delta\big).$$

Clearly for every bi-infinite quasi-geodesic $\mathfrak{q}$ in a space $X$, $\mathrm{Div}^{\mathfrak{q}}_\gamma \preceq \mathrm{Div}^X_\gamma$.

In what follows we call a metric space $X$ *proper* if all its closed balls are compact. We call it *periodic* if for fixed constants $L \geq 1$ and $C \geq 0$ the orbit of some ball under the set of $(L, C)$-quasi-isometries covers $X$.

A geodesic metric space is said *to satisfy the hypothesis* $(\mathrm{Hyp}_{\kappa, L})$ for some $\kappa \geq 0$ and $L \geq 1$ if it is one-ended, proper, periodic, and every point is at distance less than $\kappa$ from a bi-infinite $L$-bi-Lipschitz path.

Example 3.2. A Cayley graph of a finitely generated one-ended group satisfies the hypothesis $(\mathrm{Hyp}_{\frac{1}{2}, 1})$.

LEMMA 3.3 (Lemma 3.4 in [DMS10]). *Assume that $X$ satisfies* ($\mathrm{Hyp}_{\kappa,L}$) *for some $\kappa \geq 0$ and $L \geq 1$. Then for $\delta_0 = \frac{1}{1+L^2}$ and every $\gamma \geq 4\kappa$ the function $\mathrm{Div}_\gamma(n, \delta_0)$ takes only finite values.*

A new divergence function, more restrictive as to the set of triples considered, is the following.

DEFINITION 3.4. Let $\lambda \geq 2$. The *small divergence function* $\mathrm{div}_\gamma(n; \lambda, \delta)$ is the supremum of all numbers $\mathrm{div}_\gamma(a, b, c; \delta)$ with $0 \leq \mathrm{dist}(a, b) \leq n$ and

$$\text{(i)} \qquad \lambda \,\mathrm{dist}\big(c, \{a, b\}\big) \geq \mathrm{dist}(a, b).$$

We define two more versions of divergence functions, with a further restriction on the choice of $c$. For every pair of points $a, b \in X$, we choose and fix a geodesic $[a, b]$ joining them such that if $x, y$ are points on a geodesic $[a, b]$ chosen to join $a, b$ the subgeodesic $[x, y] \subseteq [a, b]$ is chosen for $x, y$.

We say that a point $c$ is *between* $a$ and $b$ if $c$ is on the fixed geodesic segment $[a, b]$.

We define $\mathrm{Div}'_\gamma(n; \delta)$ and $\mathrm{div}'_\gamma(n; \lambda, \delta)$ same as $\mathrm{Div}_\gamma$ and $\mathrm{div}_\gamma$ before, but restricting $c$ to the set of points between $a$ and $b$. Clearly, $\mathrm{Div}'_\gamma(n; \delta) \leq \mathrm{Div}_\gamma(n; \delta)$ and $\mathrm{div}'_\gamma(n; \lambda, \delta) \leq \mathrm{div}_\gamma(n; \lambda, \delta)$ for every $\lambda, \delta$.

All these versions of divergence are now shown to be equivalent under appropriate conditions.

PROPOSITION 3.5 (Corollary 3.12 in [DMS10]). *Let $X$ be a space satisfying the hypothesis* ($\mathrm{Hyp}_{\kappa,L}$) *for some constants $\kappa \geq 0$ and $L \geq 1$, and let $\delta_0 = \frac{1}{1+L^2}$ and $\gamma_0 = 4\kappa$.*

(i) *Up to the equivalence relation $\asymp$, the functions $\mathrm{div}'_\gamma(n; \lambda, \delta)$ and $\mathrm{Div}'_\gamma(n; \delta)$ with $\delta \leq \delta_0$ and $\gamma \geq \gamma_0$ are independent of the choice of geodesics $[a, b]$ for every pair of points $a, b$.*

(ii) *For every $\delta \leq \delta_0$, $\gamma \geq \gamma_0$, and $\lambda \geq 2$*

$$\mathrm{Div}_\gamma(n; \delta) \asymp \mathrm{Div}'_\gamma(n; \delta) \asymp \mathrm{div}_\gamma(n; \lambda, \delta) \asymp \mathrm{div}'_\gamma(n; \lambda, \delta).$$

*Moreover, all the functions in this equation are independent of $\delta \leq \delta_0$ and $\gamma \geq \gamma_0$ (up to the equivalence relation $\asymp$).*

(iii) *The function $\mathrm{Div}_\gamma(n; \delta)$ is equivalent to $\mathrm{div}'_\gamma(n; 2, \delta)$ as a function in $n$. Thus in order to estimate $\mathrm{Div}_\gamma(n, \delta)$ for $\delta \leq \delta_0$ it is enough to consider points $a, b, c$ where $c$ is the midpoint of a (fixed) geodesic segment connecting $a$ and $b$.*

Proposition 3.5 implies that the $\asymp$-equivalence class of the divergence function(s) is a quasi-isometry invariant in the class of metric spaces satisfying the hypothesis ($\mathrm{Hyp}_{\kappa,L}$) for some constants $\kappa \geq 0$ and $L \geq 1$.

The equivalent notions of divergence introduced previously are closely related to the divergence as defined by S. Gersten in [Ger94b] and [Ger94a]. We refer to [DMS10] for a detailed discussion.

There exists a close connection between the linearity of divergence and the existence of global cut-points in asymptotic cones; see [Beh06] for an early example and [DMS10] for a general theory.

Definition 3.6. A metric space $B$ is *unconstricted* if the following properties hold:

(1) for some constants $c, \lambda, \kappa$, every point in $B$ is at distance at most $c$ from a bi-infinite $(\lambda, \kappa)$-quasi-geodesic in $B$;
(2) there exists an ultrafilter $\omega$ and a sequence $d$ such that for every sequence of observation points $b$, $\mathrm{Cone}_\omega(B, b, d)$ does not have cut-points.

If (2) is replaced by the condition that every asymptotic cone is without cut-points then the space $B$ is called *wide*.

Proposition 3.7 (Proposition 1.1 in [DMS10]). *Let $X$ be a geodesic metric space.*

(i) *If there exists $\delta \in (0, 1)$ and $\gamma \geq 0$ such that the function $\mathrm{Div}_\gamma(n; \delta)$ is bounded by a linear function, then every asymptotic cone of $X$ is without cut-points.*
(ii) *If $X$ is wide, then for every $0 < \delta < \frac{1}{54}$ and every $\gamma \geq 0$, the function $\mathrm{Div}_\gamma(n; \delta)$ is bounded by a linear function.*
(iii) *Let $\mathfrak{g}\colon \mathbb{R} \to X$ be a periodic geodesic. If $\mathfrak{g}$ has superlinear divergence then in any asymptotic cone, $\mathrm{Cone}_\omega(X)$, for which the limit of $\mathfrak{g}$ is nonempty there exists a collection of proper subsets of $\mathrm{Cone}_\omega(X)$ with respect to which it is tree-graded and the limit of $\mathfrak{g}$ is a transversal geodesic.*

Remark 3.8. (1) In [DMS10, Proposition 1.1], "wide" means a geodesic metric space satisfying condition (2) only. For this reason in statement (ii) of Proposition 1.1 in [DMS10], it is assumed that $X$ is periodic. However that condition is only used to ensure that condition (1) in our definition of wideness is satisfied.
(2) In Proposition 3.7(ii), the hypothesis that $X$ is wide cannot be replaced by the hypothesis that $X$ is unconstricted. Indeed in [OOS05] can be found examples of unconstricted groups with super-linear divergence.

**3.2. Morse quasi-geodesics and divergence.** Examples of groups with linear divergence include groups satisfying a law, groups with a central element of infinite order, groups acting properly discontinuously and cocompactly on products of spaces, uniform lattices in higher rank symmetric spaces or Euclidean buildings and some non-uniform lattices too. Conjecturally, all non-uniform lattices in higher rank have linear divergence [DMS10].

As shown by Proposition 3.7, super-linear divergence is equivalent to the existence of cut-points in at least one asymptotic cone. Nothing more consistent can be said on divergence in this very general setting. On the other hand, as we will see in detail in Section 6 the existence of Morse quasi-geodesics—a stronger property than existence of cut-points—allows us to produce better estimates on divergence many situations. The most commonly used definition for such quasi-geodesics is, as recalled in the Introduction, that the Morse lemma is satisfied. We will use an equivalent definition, as given by the following proposition. For simplicity, we will henceforth assume that quasi-geodesics are Lipschitz, and hence continuous paths (up to bounded perturbation this can always be assumed in a geodesic metric space).

Given two points $a, b$ in the image of a quasi-geodesic, $\mathfrak{q}$, we denote by $\mathfrak{q}_{ab}$ the restriction of $\mathfrak{q}$ defined on a maximal interval such that its endpoints are $a$ and $b$.

PROPOSITION 3.9. (a) *A bi-infinite quasi-geodesic* $\mathfrak{q}$ *in* $X$ *is* Morse *if and only if for every* $C \geq 1$ *there exists* $D \geq 0$ *such that every path of length at most* $Cn$ *connecting two points* $a, b$ *on* $\mathfrak{q}$ *at distance* $\geq n$ *contains* $\mathfrak{q}_{ab}$ *in its* $D$*-neighborhood.*

(b) *If, moreover,* $X$ *is a* CAT(0) *space then it suffices to know that for some* $C > 1$ *there exists* $D \geq 0$ *such that every path of length at most* $Cn$ *connecting two points* $a, b$ *on* $\mathfrak{q}$ *at distance* $\geq n$ *contains* $\mathfrak{q}_{ab}$ *in its* $D$*-neighborhood.*

*Proof.* Proposition 3.24 of [DMS10] provides a number of equivalent conditions for Morse quasi-geodesics; statement (a) here is the equality of conditions (2) and (5) of that proposition. It remains to prove statement (b).

By conditions (1) and (2) of Proposition 3.24 of [DMS10], $\mathfrak{q}$ being Morse is equivalent to the property that in every asymptotic cone the limit set of $\mathfrak{q}$ is either empty or contained in the transversal tree. We will proceed for a contradiction, by assuming that $\mathfrak{q}$ satisfies (b) and is not Morse, or, equivalently, it satisfies (b) and has the property that for some cone the limit $\mathfrak{q}_\omega$ of $\mathfrak{q}$ is neither empty nor contained in the transversal tree. The latter assumption implies that for some sequence of pairs of points $a_n, b_n$ on $\mathfrak{q}$ with their respective distances $\delta_n$ diverging to infinity, the limit points $a_\omega$ and $b_\omega$, at distance $\delta > 0$, can be joined by a piecewise geodesic path $\mathfrak{p}_\omega$ of length $K \geq \delta$ intersecting $\mathfrak{q}_\omega$ only in its endpoints.

Note that the condition in (b) implies that every geodesic connecting two points $a, b$ on $\mathfrak{q}$ contains $\mathfrak{q}_{ab}$ in its $D$-neighborhood. Hence, in what follows we may assume without loss of generality that the arc $\mathfrak{q}'_\omega$ of $\mathfrak{q}_\omega$ of endpoints $a_\omega, b_\omega$ is a geodesic.

The asymptotic cone of a CAT(0) space is a CAT(0) space: hence the nearest point projection onto $\mathfrak{q}'_\omega$ is well defined and is a contraction. Pushing the piecewise geodesic path $\mathfrak{p}_\omega$ along the geodesics joining each of its points with its projection on $\mathfrak{q}'_\omega$ we obtain a continuous path of arcs from $\mathfrak{p}$ to $\mathfrak{q}'_\omega$ with they property that each arc in this family only intersects $\mathfrak{q}'_\omega$ at its endpoints $a_\omega$, $b_\omega$ and also the lengths of the arcs vary continuously between $K$ and $\delta$. By replacing $\mathfrak{p}$ with another arc in the path, one may then assume that the length of $\mathfrak{p}$ is at most $C/2$ for the constant $C$ given in (b). By again, replacing $\mathfrak{p}$ with a path that is piecewise geodesic, we may assume that $\mathfrak{p}$ is a limit of paths $\mathfrak{p}_n$ of length at most $C\delta_n$, such that the minimal tubular neighbourhood of $\mathfrak{p}_n$ containing $\mathfrak{q}_{a_n b_n}$ is $\varepsilon\delta_n$, for some $\varepsilon > 0$. This contradicts the hypothesis, therefore our assumption, that the limit of $\mathfrak{q}$ is neither empty nor transversal in some asymptotic cone can not hold and it follows that $\mathfrak{q}$ is Morse. □

DEFINITION 3.10. (1) In a finitely generated group $G$, an element is called *Morse* if it has infinite order and the cyclic subgroup generated by it is a Morse quasi-geodesic.

(2) For a Morse quasi-geodesic or a Morse element in a CAT(0) space or group, we consider the quasi-geodesic constants together with the constants $C$ and $D$ as in Proposition 3.9(b) the *Morse parameters*.

Several important classes of groups contain Morse elements. Behrstock proved in [Beh06] that every pseudo-Anosov element in a mapping class group is Morse. Yael Algom-Kfir proved the same thing for fully irreducible elements of the outer automorphism group $\mathrm{Out}(F_n)$ in [AK], see also [Ham09]. In a relatively hyperbolic group, every non-parabolic element is Morse ([DS05], [Osi06]).

In [DMS10], it is proved that in a finitely generated group acting acylindrically on a simplicial tree or on a uniformly locally finite hyperbolic graph, any loxodromic element is Morse. An action on a graph is called *$l$-acylindrical* for some $l > 0$ if stabilizers of pairs of points at distance $\geq l$ are finite of uniformly bounded sizes. At times the constant $l$ need not be mentioned. Note that a group acting by isometries on a simplicial tree with unbounded orbits always contains loxodromic elements [Bow08].

Existence of Morse quasi-geodesics implies existence of cut-points in all asymptotic cones. The converse is only known to be true for universal covers of non-positively curved compact non-flat de Rham irreducible manifolds due to the following two results combined with Proposition 3.7(iii).

THEOREM 3.11 ([Bal85], [Bal95], [BS87]). *Let $M$ be a non-positively curved de Rham irreducible manifold with a group of isometries acting co-compactly. Then either $M$ is a higher rank symmetric space or $M$ contains a periodic geodesic which does not bound a half-plane.*

PROPOSITION 3.12 ([KL98, Proposition 3.3]). *Let $X$ be a locally compact, complete, simply connected geodesic metric space which is locally* CAT(0). *A periodic geodesic $\mathfrak{g}$ in $X$ which does not bound a flat half-plane satisfies*

$$\mathrm{Div}^{\mathfrak{g}}(r) \succeq r^2.$$

The lower estimate on divergence that enters as a main ingredient in this converse is as important as the converse itself. In Section 6, we prove that the estimate in Proposition 3.12 holds in a considerably more general CAT(0) setting, as well as for many of the examples of groups with Morse elements quoted above.

## 4. Divergence and networks of spaces

**4.1. Tight networks.** We strengthen the definitions of networks of subspaces and subgroups from [BDM09], with a view toward divergence estimate problems.

A subset $A$ in a metric space is called $C$-*path connected* if any two points in $A$ can be connected by a path in $\mathcal{N}_C(A)$. We say that $A$ is $(C,L)$-*quasi-convex* if any two points in $A$ can be connected in $\mathcal{N}_C(A)$ by a $(L,L)$-quasi-geodesic. When $C=L$, we simply say that $A$ is $C$-quasi-convex.

Definition 4.1 (Tight network of subspaces). Given $\tau$ and $\eta$ two non-negative real numbers we say that a metric space $X$ is a $(\tau,\eta)$-*tight network with respect to a collection* $\mathcal{L}$ *of subsets* or that $\mathcal{L}$ *forms a* $(\tau,\eta)$-*tight network inside* $X$ if every subset $L$ in $\mathcal{L}$ with the induced metric is $(\tau,\eta)$-quasi-convex, $X$ is covered by $\tau$-neighborhoods of the sets $L\in\mathcal{L}$, and the following condition is satisfied: for any two elements $L,L'\in\mathcal{L}$ and any point $x$ such that $B(x,3\tau)$ intersects both $L$ and $L'$, there exists a sequence of length $n\leq\eta$

$$L_1=L,L_2,\ldots,L_{n-1},L_n=L',\quad \text{with } L_i\in\mathcal{L}$$

such that for all $1\leq i<n$, $\mathcal{N}_\tau(L_i)\cap\mathcal{N}_\tau(L_{i+1})$ is of infinite diameter, $\eta$-path connected and it intersects $B(x,\eta)$. We write (N) to refer to the above condition about arbitrary pairs of elements in $\mathcal{L}$.

When $G$ is a finitely generated group and $\mathcal{L}$ is the collection of left cosets of $\mathcal{H}$, a collection of undistorted subgroups, the following strengthening of the above definition is sometimes easier to verify.

Definition 4.2 (Tight algebraic network of subgroups). We say a finitely generated group $G$ is an $M$-*tight algebraic network with respect to* $\mathcal{H}$ or that $\mathcal{H}$ *forms an* $M$-*tight algebraic network inside* $G$, if $\mathcal{H}$ is a collection of $M$-quasi-convex subgroups whose union generates a finite-index subgroup of $G$ and for any two subgroups $H,H'\in\mathcal{H}$ there exists a finite sequence $H=H_1,\ldots,H_n=H'$ of subgroups in $\mathcal{H}$ such that for all $1\leq i<n$, the intersection $H_i\cap H_{i+1}$ is infinite and $M$-path connected. We write (AN) to refer to the above condition about arbitrary pairs $H,H'\in\mathcal{H}$.

A modification of the proof of [BDM09, Proposition 5.3] yields the following.

Proposition 4.3. *Let $\mathcal{H}$ be a collection of subgroups that forms a tight algebraic network inside a finitely generated group $G$ and let $G_1$ be the finite-index subgroup of $G$ generated by the subgroups in $\mathcal{H}$. Then $G$ is a tight network with respect to the collection of left cosets*

$$\mathcal{L}=\{gH\mid g\in G_1, H\in\mathcal{H}\}.$$

*Proof.* Since cosets cover the subgroup $G_1$ and $G \subset \mathcal{N}_\tau(G_1)$ for some $\tau > 0$, it remains to prove (N). By left translation, we may assume that $x = 1$; also, it clearly suffices to prove condition (N) for $L = H$, and $L' = gH'$, where $H, H' \in \mathcal{H}$ and $g \in G_1 \cap \mathrm{B}(1, 3\tau)$.

The argument in Proposition 5.3 in [BDM09] as written there yields that for every such pair there exists a sequence $L_1 = H, L_2, \ldots, L_n = gH'$ in $\mathcal{L}$ composed of concatenations of left translations of sequences as in (AN). That this sequence also satisfies the property that all $L_i$ intersect $B(1, 3\tau)$ is by construction, since each of element of this sequence is in $\mathcal{H}$, or else is a left translate of such a subgroup by an element of $G$ which is represented by a subword of the element $g$ and hence this coset intersects $B(1, |g|) = B(1, 3\tau)$, as desired. In particular, every pair $L_i, L_{i+1}$ is a left translation of a pair of subgroups as in (AN) of the form $g'H_1, g'H_2$ with $H_1 \cap H_2$ infinite and $M$-path connected and $g' \in G_1$ closer to 1 than $g$.

By [MSW05, Lemma 2.2] the intersection $\mathcal{N}_\tau(L_i) \cap \mathcal{N}_\tau(L_{i+1})$ is at finite Hausdorff distance from $g(H_1 \cap H_2)$, hence it is of infinite diameter and $\eta$-path connected for $\eta$ large enough. $\square$

The following yields a large family of examples.

PROPOSITION 4.4. *Let $G$ be a fundamental group of a graph of groups where all the vertex groups are quasi-convex and the edge groups are infinite. Then $G$ is a tight network with respect to the family of all left cosets of vertex groups. Moreover, if the graph of groups is simply connected, then $G$ is a tight algebraic network with respect to the family of vertex groups.*

*Proof.* Since left cosets of a subgroup cover and the vertex sets are quasi-convex by hypothesis, to show that $G$ is a tight network it remains to verify property (N). Fix two left cosets $L, L'$ of vertex subgroups of $G$ and a point $x$ in $\mathcal{N}_{3\tau}(L) \cap \mathcal{N}_{3\tau}(L')$. By left translation we may assume that $x = 1$. In the Bass–Serre tree, $L$ and $L'$ correspond to vertices at distance at most $n$ apart with $n \leq 6\tau$ and letting $L = L_0, L_1, \ldots, L_n = L'$ be the sequence of left cosets corresponding to the vertices in the shortest path in the tree from $L$ to $L'$, it satisfies all the properties of (N). In particular, all $L_i$ intersect $B(1, \eta)$ for $\eta$ large enough, because there are finitely many possibilities for $L, L'$ (left cosets of vertex groups intersecting $B(1, 3\tau)$) and therefore for $L_i$.

If the graph of groups is simply connected, then the vertex groups generate $G$. The remaining properties of tight algebraic network follow immediately. $\square$

Tight networks define natural decompositions of geodesics, as described below.

LEMMA 4.5. *Let $X$ be a geodesic metric space and $\mathcal{L}$ a collection of subsets of $X$ such that $X = \bigcup_{L \in \mathcal{L}} \mathcal{N}_\tau(L)$. Every geodesic $[x, y]$ contains a finite sequence of consecutive points $x_0 = x, x_1, x_2, \ldots, x_{n-1}, x_n = y$ such that:*

(1) *for every* $i \in \{0,1,\dots,n-1\}$ *there exists* $L_i \in \mathcal{L}$ *such that* $x_i, x_{i+1} \in \mathcal{N}_{3\tau}(L_i)$;
(2) *for every* $i \in \{0,1,\dots,n-2\}$, $\mathrm{dist}(x_i,x_{i+1}) \geq \tau$.

*Proof.* We inductively construct a sequence of consecutive points $x_0 = x, x_1, x_2, \dots, x_n$ on $[x,y]$ such that for every $i \in \{0,1,\dots,n-1\}$ there exists $L_i \in \mathcal{L}$ with the property that $x_i \in \mathcal{N}_\tau(L_i)$ and $x_{i+1}$ is the farthest point from $x$ on $[x,y]$ contained in $\overline{\mathcal{N}}_{2\tau}(L_i)$. Assume that we found $x_0 = x, x_1, \dots, x_k$. Since $x_k \in X = \bigcup_{L\in\mathcal{L}} \mathcal{N}_\tau(L)$ there exists $L_{k+1} \neq L_k$ such that $x_k \in \mathcal{N}_\tau(L_{k+1})$. Pick $x_{k+1}$ to be the farthest point from $x$ on $[x,y]$ contained in $\overline{\mathcal{N}}_{2\tau}(L_{k+1})$. By our choice of $x_k$, it follows that this process will terminate with $n \leq \lceil \frac{\mathrm{dist}(x,y)}{\tau} \rceil$. $\square$

LEMMA 4.6. *Let $X$ be a geodesic metric space which is a $(\tau,\eta)$-tight network with respect to a collection of subsets $\mathcal{L}$, let $[x,y]$ be a geodesic in $X$ and let $L, L' \in \mathcal{L}$ be such that $x \in \mathcal{N}_\tau(L)$ and $y \in \mathcal{N}_\tau(L')$. There exists a finite sequence of consecutive points $x_0 = x, x_1, x_2, \dots, x_{n-1}, x_n = y$ on $[x,y]$ and a finite sequence of subsets $L_j \in \mathcal{L}$, satisfying $L_0 = L, L_1, \dots, L_q = L'$ with $q \leq n\eta$ such that*

(1) *for every* $i \in \{0,1,\dots,n-2\}$, $\mathrm{dist}(x_i,x_{i+1}) \geq \tau$;
(2) *for every* $j \in \{0,1,\dots,q-1\}$ *the intersection* $\mathcal{N}_\tau(L_j) \cap \mathcal{N}_\tau(L_{j+1})$ *is of infinite diameter and $\eta$-path connected*;
(3) *there exist* $j_0 = 0 < j_1 < \cdots < j_{n-1} < j_n = q$ *such that if* $j_{i-1} \leq j \leq j_i$ *then* $\mathcal{N}_\tau(L_j) \cap \mathcal{N}_\tau(L_{j+1})$ *intersects* $\mathrm{B}(x_i,\eta)$.

*Proof.* For the geodesic $[x,y]$ consider a sequence $x_0 = x, x_1, x_2, \dots, x_{n-1}, x_n = y$ as in Lemma 4.5, and the sequence $L_0, L_1, \dots, L_{n-1}$ determined by the condition Lemma 4.5(1), which can be taken with $L_0 = L$.

Property (N) applied to each of the pairs $L_i, L_{i+1}$ with $i \in \{0,1,\dots,n-2\}$, and to $L_{n-1}, L'$ provides a sequence $L = J_0, J_1, \dots, J_q = L'$ with $q \leq n\eta$, such that for all $j \in \{0,1,\dots,q-1\}$ the intersection $\mathcal{N}_\tau(J_j) \cap \mathcal{N}_\tau(J_{j+1})$ is of infinite diameter, $\eta$-path connected and it intersects $\mathrm{B}(x_i,\eta)$ for some $0 \leq i \leq n$. $\square$

The following shows that tight networks are a uniform version of the *networks* in [BDM09, Definition 5.1].

COROLLARY 4.7. *Let $X$ be a geodesic metric space which is a $(\tau,\eta)$-tight network with respect to a collection of subsets $\mathcal{L}$.*

*For every $M \geq 0$ there exists $R = R(M)$ such that for every $L, L' \in \mathcal{L}$ with $\mathcal{N}_M(L) \cap \mathcal{N}_M(L') \neq \emptyset$ and any point $a \in \mathcal{N}_M(L) \cap \mathcal{N}_M(L')$ there exists a sequence, $L_1 = L, L_2, \dots, L_{n-1}, L_n = L'$, with $L_i \in \mathcal{L}$ and $n \leq R$ such that for all $1 \leq i < n$, $\mathcal{N}_\tau(L_i) \cap \mathcal{N}_\tau(L_{i+1})$ is of infinite diameter, $\eta$-path connected, and intersects $B(a,R)$.*

*Proof.* Let $L, L' \in \mathcal{L}$ be such that $\mathcal{N}_M(L) \cap \mathcal{N}_M(L') \neq \emptyset$ and let $a$ be a point in the intersection. Take $x \in L$ and $y \in L'$ such that $\mathrm{dist}(x,a) < M$ and

$\mathrm{dist}(y,a) < M$. Lemma 4.6 applied to the geodesic $[x,y]$ yields a sequence of $L_i$ for which $\mathcal{N}_\tau(L_i) \cap \mathcal{N}_\tau(L_{i+1}) \cap \mathcal{N}_\eta([x,y]) \neq \emptyset$ and hence $\mathcal{N}_\tau(L_i) \cap \mathcal{N}_\tau(L_{i+1}) \cap B(a, \eta + 2M) \neq \emptyset$, yielding the desired conclusion with $R(M) = \eta + 2M$. □

**4.2. Network divergence.** We defined divergence functions in Definition 3.1; for a network of spaces, we now define an auxiliary function in order to bound the divergence of $X$.

DEFINITION 4.8. Let $X$ be a $(\tau, \eta)$-tight network with respect to a collection $\mathcal{L}$ of subsets, let $\delta$ be a number in $(0,1)$ and let $\gamma \geq 0$. For every subset $L \in \mathcal{L}$, we denote by $\mathrm{Div}_\gamma^L(n;\delta)$ the divergence function for $\mathcal{N}_\tau(L)$ with the induced metric.

The *network divergence of* $X$ is defined as

$$\mathrm{Div}_\gamma^{\mathcal{L}}(n;\delta) = \sup_{L \in \mathcal{L}} \mathrm{Div}_\gamma^L(n;\delta).$$

THEOREM 4.9. *Let $X$ be a geodesic metric space, let $\mathcal{L}$ be a collection of subsets which forms a $(\tau,\eta)$-tight network inside $X$, let $\delta$ be a number in* $(0,1)$ *and let $\gamma \geq 0$. The divergence in $X$ satisfies*

$$\mathrm{Div}_\gamma^X(n;\delta) \preceq_C n \,\mathrm{Div}_\gamma^{\mathcal{L}}(n;\delta), \tag{ii}$$

*where the constant $C$ only depends on the constants $\tau, \eta, \delta$ and $\gamma$.*

*Proof.* Let $a, b, c$ be three points such that $\mathrm{dist}(a,b) = n$ and $\mathrm{dist}(c, \{a,b\}) = r > \frac{\gamma}{\delta}$. If the ball $\mathrm{B}(c, \delta r - \gamma)$ does not intersect a geodesic $[a,b]$, then $\mathrm{div}_\gamma(a,b,c;\delta) = n$. Assume therefore that $\mathrm{B}(c, \delta r - \gamma)$ intersects $[a,b]$. This in particular implies that $r \leq \delta r - \gamma + \frac{n}{2}$, whence $r \leq \frac{n}{2(1-\delta)}$.

Lemma 4.6 applied to the geodesic $[a,b]$ implies the existence of a finite sequence of points $x_0 = a, x_1, \ldots, x_k = b$ with $k \leq \frac{n}{\tau} + 2$, consecutive on the geodesic, and of a finite sequence of subsets in $\mathcal{L}$, $L_0 = L, L_1, \ldots, L_q = L'$ with $q \leq k\eta \leq \eta(\frac{n}{\tau} + 2)$ such that:

(1) for every $j \in \{0, 1, \ldots, q-1\}$ the intersection $\mathcal{N}_\tau(L_j) \cap \mathcal{N}_\tau(L_{j+1})$ is of infinite diameter and $\eta$-path connected;
(2) there exist $j_0 = 0 < j_1 < \cdots < j_{n-1} < j_n = q$ such that if $j_{i-1} \leq j \leq j_i$ then $\mathcal{N}_\tau(L_j) \cap \mathcal{N}_\tau(L_{j+1})$ intersects $\mathrm{B}(x_i, \eta)$.

Let $\sigma > 0$ be large enough. Conditions on it will be added later on. Consider an arbitrary $j \in \{0, 1, \ldots, q-1\}$. There exists $i \in \{1, 2, \ldots, k-1\}$ such that $j_{i-1} \leq j \leq j_i$. This implies that there exists a point $z_{j+1}$ in $\mathcal{N}_\tau(L_j) \cap \mathcal{N}_\tau(L_{j+1})$ at distance at most $\eta$ from $x_i$.

Since we are in a network, $\mathcal{N}_\tau(L_j) \cap \mathcal{N}_\tau(L_{j+1})$ is of infinite diameter and thus contains an element $z'_{j+1}$ at distance $\geq \sigma r$ from $c$. Since the intersection is $\eta$-path connected there exists a path joining $z_{j+1}$ and $z'_{j+1}$ in the $\eta$ neighborhood of the intersection, and one can find on it a point at distance $\sigma r$ from $c$. Thus, there exists $y_{j+1}$ in $\mathcal{N}_\tau(L_j) \cap \mathcal{N}_\tau(L_{j+1})$ at distance $\sigma r + O(1)$ from $c$.

If $x_0 = a$ is in $B(c, \sigma r)$, then we may find a point $y_0$ in $\mathcal{N}_\tau(L_0)$ at distance $\sigma r + O(1)$ from $c$. Likewise we may have to replace $b$ by another point $y_{q+1}$.

We thus obtain a new sequence of points $y_0, y_1, \ldots, y_{q+1}$ all at distance $\sigma r + O(1)$ from $c$. For every $j \in \{0, 1, \ldots, q\}$ the pair $y_j, y_{j+1}$ is inside $\mathcal{N}_\tau(L_j)$. If $B(c, \delta r - \gamma)$ does not intersect $\mathcal{N}_\tau(L_j)$ then simply join $y_j, y_{j+1}$ by short geodesics to points in $L_j$ and join those points by a $(\eta, \eta)$-quasi-geodesic in $\mathcal{N}_\tau(L_j)$. Otherwise, the intersection of $B(c, \delta r - \gamma)$ with $\mathcal{N}_\tau(L_j)$ contains a point, which we denote $c_j$. The ball $B(c_j, 2\delta r - \gamma)$ contains $B(c, \delta r - \gamma)$. We choose $\sigma$ large enough so that $2r \leq \mathrm{dist}(c_j, \{y_j, y_{j+1}\})$. Thus, $\mathrm{dist}(c_j, \{y_j, y_{j+1}\}) \geq \sigma r + O(1)$ allows us to join $y_j$ and $y_{j+1}$ outside the ball $B(c_j, 2\delta r - \gamma)$ by a path of length at most $\mathrm{Div}_\gamma^{L_j}(2\sigma r + O(1); \delta) \leq \mathrm{Div}_\gamma^{L_j}(\frac{\sigma}{1-\delta} n + O(1); \delta)$.

The concatenation of all these curves gives a curve joining $a$ and $b$ outside $B(c, \delta r - \gamma)$ and of length at most $2\alpha n + (q+1)\,\mathrm{Div}_\gamma^{\mathcal{L}}(\alpha n + O(1); \delta)$, where $\alpha = \frac{\sigma}{1-\delta}$ and $q \leq \frac{\eta}{\tau} n + 2\eta$. Note that the first term stands for the lengths of the geodesics joining $a$ and $y_0$, respectively, $y_q$ and $b$. $\square$

Corollary 4.10. *Let $G$ be a tight algebraic network with respect to the collection of subgroups $\mathcal{H}$. For every $\delta \in (0,1)$ and $\gamma \geq 0$,*

$$\mathrm{Div}_\gamma^G(n; \delta) \preceq n \sup_{H \in \mathcal{H}} \mathrm{Div}_\gamma^H(n; \delta).$$

**4.3. Thick spaces and groups.** In some sense, the subsets forming a network are building blocks and the ambient space is constructed out of them. By iterating this construction, we obtain thick spaces, a notion introduced in [BDM09]. The initial step in [BDM09] (thick spaces of order zero) were taken to be unconstricted spaces. In this paper, we adapt the notion with a view to relate the order of thickness to the order of the divergence function. To this purpose, we introduce below *strongly thick spaces* by taking as initial step a subclass of unconstricted spaces, namely, wide spaces.

Definition 4.11. A collection of metric spaces, $\mathcal{B}$, is *uniformly wide* if:

(1) for some positive real constants $\lambda, \kappa$, every point in every space $B \in \mathcal{B}$ is at distance at most $\kappa$ from a bi-infinite $(\lambda, \lambda)$-quasi-geodesic in $B$;
(2) for every sequence of spaces $(B_i, \mathrm{dist}_i)$ in $\mathcal{B}$, every ultrafilter $\omega$, sequence of scaling constants $d = (d_i)$ and sequence of basepoints $b = (b_i)$ with $b_i \in B_i$, the ultralimit $\lim_\omega(B_i, b_i, \frac{1}{d_i}\mathrm{dist}_i)$ does not have cut-points.

All the examples of unconstricted spaces listed in [BDM09, p. 555] are in fact examples of uniformly wide collections of metric spaces (with "wide" replacing "unconstricted" in Example 5).

The following uniform version of Proposition 3.7(ii) can be easily obtained by adapting the proof of [DMS10, Lemma 3.17(ii)] and considering ultralimits of rescaled spaces in $\mathcal{B}$ instead of asymptotic cones.

PROPOSITION 4.12. *Let $\mathcal{B}$ be a collection of uniformly wide metric spaces. For every $0 < \delta < \frac{1}{54}$ and every $\gamma \geq 0$, the function $\sup_{B \in \mathcal{B}} \mathrm{Div}_\gamma^B(n;\delta)$ is bounded by a linear function.*

DEFINITION 4.13 (Metric thickness and uniform thickness).

($\mathrm{M}_0$) A metric space is called *strongly thick of order zero* if it is wide. A family of metric spaces is *uniformly strongly thick of order zero* if it is uniformly wide.

($\mathrm{M}_{n+1}$) Given $\tau \geq 0$ and $n \in \mathbb{N}$ we say that a metric space $X$ is $(\tau, \eta)$-*strongly thick of order at most $n+1$ with respect to a collection of subsets $\mathcal{L}$* if $X$ is a $(\tau, \eta)$-tight network with respect to $\mathcal{L}$, and moreover:

- ($\theta$) the subsets in $\mathcal{L}$ endowed with the restriction of the metric on $X$ compose a family uniformly strongly thick of order at most $n$.

Further, $X$ is said to be $(\tau, \eta)$-*strongly thick of order $n$ (with respect to $\mathcal{L}$)* if it is $(\tau, \eta)$-*strongly thick of order at most $n$ (with respect to $\mathcal{L}$)* and for no choices of $\tau, \eta$ and $\mathcal{L}$ is it strongly thick of order at most $n-1$.

When $\mathcal{L}$, $\tau$, $\eta$ are irrelevant, we say that $X$ is *strongly thick of order (at most) $n$* or simply that $X$ is *strongly thick.*

($\mathrm{M}_{\text{uniform}}$) A family $\{X_i \mid i \in I\}$ of metric spaces is *uniformly strongly thick of order at most $n+1$* if the following hold.

- ($\upsilon\theta_1$) There exist $\tau > 0$ and $\eta > 0$ such that every $X_i$ is a $(\tau, \eta)$-tight network with respect to a collection $\mathcal{L}_i$ of subsets;
- ($\upsilon\theta_2$) $\bigcup_{i \in I} \mathcal{L}_i$ is uniformly strongly thick of order at most $n$, where each $L \in \mathcal{L}_i$ is endowed with the induced metric.

The order of strong thickness is a quasi-isometry invariant, cf. [BDM09, Remark 7.2].

For finitely generated groups and subgroups with word metrics a stronger version of thickness can be defined.

DEFINITION 4.14 (Strong algebraic thickness). Consider a finitely generated group $G$.

($\mathrm{A}_0$) $G$ is called *strongly algebraically thick of order zero* if it is wide.

($\mathrm{A}_n$) Given $M > 0$, a group $G$ is called *$M$-strongly algebraically thick of order at most $n+1$ with respect to a finite collection of subgroups $\mathcal{H}$*, if:

- $G$ is an $M$-tight algebraic network with respect to $\mathcal{H}$;
- all subgroups in $\mathcal{H}$ are strongly algebraically thick of order at most $n$.

$G$ is said to be *strongly algebraically thick of order $n+1$ with respect to $\mathcal{H}$*, when $n$ is the smallest value for which this statement holds.

REMARK 4.15. The property of strong algebraic thickness does not depend on the choice of the word metric on the group $G$. This raised the question, asked in [BDM09, Question 7.5], whether strong algebraic thickness is invariant by quasi-isometry. The following example, due to Alessandro Sisto, answers this question by showing that algebraic thickness is not a quasi-isometry invariant. Let $G$ be the fundamental group of a closed graph manifold whose associated graph of groups consists of one vertex and one edge. Since any element acting hyperbolically on the Bass-Serre tree is a Morse element [DMS10], it follows that any subgroup that contains such an element has cut-points in all its asymptotic cones. Hence, any subgroup of $G$ which is both quasi-convex and wide (or even unconstricted) is contained in a conjugate of the vertex group. As any finite set of subgroups which are each contained in a particular conjugate of a vertex group generate an infinite index subgroup, it follows that no collection of unconstricted subgroups can constitute a tight algebraic network in $G$. Hence, $G$ is not algebraically thick. On the other hand, all fundamental groups of closed graph manifold are quasi-isometric [BN08] and some of them are algebraically thick, for example, the graph manifold built by gluing together two Seifert fibered spaces each with one boundary component.

EXAMPLES 4.16. The following are some known examples of thick and algebraically thick spaces and groups:

(1) Mapping class groups of surfaces $S$ with complexity $\xi(S) = 3 \times \text{genus} + \#(\text{boundary components}) - 3 > 1$ are strongly algebraically thick of order 1 [BDM09], [Beh06];
(2) $\mathrm{Aut}(F_n)$ and $\mathrm{Out}(F_n)$, for $n \geq 3$, are strongly algebraically thick of order at most 1 with respect to a family of quasi-flats of dimension 2 [BDM09]; $\mathrm{Out}(F_n)$ is strongly algebraically thick of order exactly 1 due to the existence of Morse elements proved in [AK];
(3) various Artin groups are strongly algebraically thick of order at most one [BDM09, §10], right-angled Artin groups which are thick of order 1 are classified in [BC];
(4) graphs of groups with infinite edge groups and whose vertex groups are thick of order $n$, are thick of order at most $n+1$, by Proposition 4.4. In particular, the fundamental group $G = \pi_1(M)$ of a non-geometric graph manifold is strongly thick of order 1;
(5) for every surface $S$ of finite type with complexity $\xi(S) \geq 6$, the Teichmüller space with the Weil–Petersson metric is strongly thick of order one with respect to a family of quasi-flats of dimension two [BDM09, §12], [Beh06].

A connection between order of thickness and order of the divergence function can be easily established using Theorem 4.9.

COROLLARY 4.17. *If a family $\mathcal{B}$ of metric spaces is uniformly strongly thick of order at most $n$, then for every $0 < \delta < \frac{1}{54}$ and every $\gamma \geq 0$,*

$$\sup_{B \in \mathcal{B}} \mathrm{Div}_\gamma^B(x;\delta) \preceq x^{n+1}.$$

*In particular, if a metric space $X$ is strongly thick of order at most $n$, then for every $\delta$ and $\gamma$ as above:*

$$\mathrm{Div}_\gamma(x;\delta) \preceq x^{n+1}.$$

*Proof.* The statement follows by induction on $n$. For $n = 0$, since wide spaces have linear divergence the result holds, see Proposition 4.12. If the result holds for order $n$, then it follows immediately from Theorem 4.9 that the result holds for order $n+1$. □

Corollary 4.17 yields upper estimates for divergence functions of several spaces and groups. In the two corollaries below, we record these estimates in two cases that are not in the literature. The estimates are sharp (for the upper bounds see Section 6), with one exception when the exact order of divergence is unknown.

COROLLARY 4.18. *If $S$ is a compact oriented surface of genus $g$ and with $p$ boundary components such that $3g+p-3 \geq 4$ and $(g,p) \neq (2,1)$ then the Weil–Petersson metric on the Teichmüller space has at most quadratic divergence. When $(g,p) = (2,1)$ the divergence is at most cubic.*

*Proof.* It is an immediate consequence of Corollary 4.17 combined with [BDM09, Theorem 12.3] and [BM08, Theorem 18]. □

Quadratic lower bounds on the divergence of the Weil–Petersson metric is implicit in the results of [Beh06], see also Section 6. The following question remains open, which if answered negatively would provide an interesting quasi-isometry invariant differentiating the Weil–Petersson metric on the Teichmüller space of a surface of genus two with one boundary component from the other two Teichmüller spaces of surfaces of the same complexity (i.e., the four-punctured torus and the seven-punctured sphere).

QUESTION 4.19. Does the Weil–Petersson metric on the Teichmüller space of a surface of genus two with one boundary component have quadratic divergence?

COROLLARY 4.20. *For $n \geq 3$ both $\mathrm{Aut}(F_n)$ and $\mathrm{Out}(F_n)$ have divergence at most quadratic.*

This bound is believed to be sharp, see Question 6.8.

A natural question raised by Theorem 4.9 and supported by all known examples, including the two above, is the following:

QUESTION 4.21. Is a group $G$ strongly algebraically thick of order $n$ if and only if it has polynomial divergence of degree $n+1$?

## 5. Higher order thickness and polynomial divergence

In this section, we construct CAT(0) groups that are strongly algebraically thick of order $n$ and with polynomial divergence of degree $n+1$. This illustrates the richness of the thickness hierarchy introduced in [BDM09] and provides another construction answering Gersten's question [Ger94a] of whether a CAT(0) group can have polynomial divergence of degree 3 or greater, for an alternate construction illustrating polynomial divergence see Macura [Mac].

We construct, by induction on $n$, a compact locally CAT(0) space, $M_n$, whose fundamental group $G_n = \pi_1(M_n)$ is torsion-free. In Proposition 5.1, we show that $G_n$ is strongly algebraically thick of order at most $n$. In Proposition 5.2, we show that $M_n$ contains a closed geodesic $\mathfrak{g}_n$ such that in the universal cover $\widetilde{M_n}$ the lift $\widetilde{\mathfrak{g}}_n$ is Morse and has divergence $\asymp x^{n+1}$. Corollary 4.17 then implies that $G_n$ is strongly thick of order exactly $n$. Also, by [BN08, Theorem 3.2], our construction of $M_1$ can be chosen with fundamental group of any one of an infinite family of pairwise non-quasi-isometric classes. In our construction, the quasi-isometry type of $M_n$ is an invariant of the quasi-isometry type of $M_{n+1}$: hence our construction yields infinitely many quasi-isometry types of groups. Thus, this family of groups will yield Theorem 1.1.

For $n=1$ take $M_1$ to be a CAT(0), non-geometric graph manifold; these are easily constructed by taking a pair of hyperbolic surfaces each with at least one boundary component, crossing each with a circle, and then gluing these two 3-manifolds together along a boundary torus by flipping the base and fiber directions. It was proven by Gersten that these manifolds have quadratic divergence [Ger94a]. These groups are all thick of order 1 [BDM09], algebraically thick examples are easy to produce by using Example 4.16(4) and asking that the corresponding graph of groups is simply connected (as in the explicit example above). The remaining properties are easily verified.

Construction. Assume now that for a fixed integer $n \geq 1$ we have constructed a compact locally CAT(0) space $M_n$ with a closed geodesic $\mathfrak{g}_n$, such that the lifts $\widetilde{\mathfrak{g}}_n$ in the universal cover have divergence $\asymp x^{n+1}$ (and thus, in particular, are Morse); moreover such that the fundamental group, $G_n = \pi_1(M_n)$, is thick of order at most $n$. We obtain $M_{n+1}$ by gluing two isometric copies of $M_n$ (denoted $M_n$ and $M'_n$) by identifying the two copies of the closed geodesic $\mathfrak{g}_n$.

To check that $M_{n+1}$ is locally CAT(0), we note that this clearly holds in the neighborhood of each point $y$ not on $\mathfrak{g}_n$. If $y \in \mathfrak{g}_n$, then any geodesic triangle with endpoints in $B(y, \varepsilon)$ is either contained in one of the two copies of $M_n$ or two of its edges cross $\mathfrak{g}_n$. In either of the cases, it is easily checked that the triangle satisfies the CAT(0) condition.

It follows that $\widetilde{M}_{n+1}$ is a CAT(0) space on which the fundamental group $G_{n+1}$ acts cocompactly. The group $G_{n+1}$ is an amalgamated product of two copies of $G_n$ along the cyclic group $C_n$ generated by the element corresponding to $\mathfrak{g}_n$. We write this as $G_{n+1} = G_n *_{C_n} G'_n$ (where $G_n$ and $G'_n$ are isomorphic). The inductive hypothesis that $G_n$ is torsion-free implies that $G_{n+1}$ is torsion-free. Let $T_n$ be the simplicial tree corresponding to this splitting.

Proposition 5.1. *$G_{n+1}$ is strongly algebraically thick of order at most $n+1$.*

*Proof.* Since each of $G_n$ and $G'_n$ are thick of order at most $n$ and intersect in an infinite cyclic group, it only remains to prove that $G_n$ and $G'_n$ are both quasi-convex in $G_{n+1}$. This is equivalent to proving that in $\widetilde{M}_{n+1}$ the lifts $\widetilde{M}_n$ and $\widetilde{M}'_n$ of $M_n$ and $M'_n$, respectively are quasi-convex. Note that $\widetilde{M}_{n+1}$ is obtained by gluing all the translates $G_n\widetilde{M}_n$ and $G_n\widetilde{M}'_n$ along geodesics $G_n\widetilde{\mathfrak{g}}_n$. In particular, the geodesics in $G_n\widetilde{\mathfrak{g}}_n$ separate $\widetilde{M}_{n+1}$.

We first prove that $\widetilde{\mathfrak{g}}_n$ is totally geodesic. We prove by induction on $k$ that an arbitrary geodesic $[x,y]$ joining two points $x,y \in \widetilde{\mathfrak{g}}_n$ and crossing at most $k$ geodesics in $G_n\widetilde{\mathfrak{g}}_n$ must be contained in $\widetilde{\mathfrak{g}}_n$. Here, and in what follows, when we say that a subset $A$ of $\widetilde{M}_{n+1}$ *crosses* a geodesic $\overline{\mathfrak{g}}_n$ in $G_n\widetilde{\mathfrak{g}}_n$ we mean that $A$ intersects at least two connected components of $\widetilde{M}_{n+1} \setminus \overline{\mathfrak{g}}_n$.

For $k=0$, the statement follows from the fact that $\widetilde{\mathfrak{g}}_n$ is totally geodesic both in $\widetilde{M}_n$ and in $\widetilde{M}'_n$ and that, since the metric on $\widetilde{M}_{n+1}$ locally coincides with the metric on $\widetilde{M}_n$ (respectively $\widetilde{M}'_n$) the length of a path contained in $\widetilde{M}_n$ (respectively $\widetilde{M}'_n$) is the same in that space as in $\widetilde{M}_{n+1}$. Now assume that the statement is true for all integers less than $k$ and consider an arbitrary geodesic $[x,y]$ with endpoints $x,y \in \widetilde{\mathfrak{g}}_n$ and crossing at most $k$ geodesics in $G_n\widetilde{\mathfrak{g}}_n$. Let $\mathfrak{g}'_n$ be a geodesic crossed by $[x,y]$ such that the corresponding edge in $T_n$ is at maximal distance from the edge corresponding to $\widetilde{\mathfrak{g}}_n$. It follows that there exists $[a,b]$ subgeodesic of $[x,y]$ with endpoints on the geodesic $\mathfrak{g}'_n$ and not crossing any other geodesic in $G_n\widetilde{\mathfrak{g}}_n$. Then it must be entirely contained in $\mathfrak{g}'_n$ according to the initial step for $k=0$. Hence, the geodesic $\mathfrak{g}'_n$ is not crossed and we can use the inductive hypothesis.

We have thus proved that all geodesics in $G_n\widetilde{\mathfrak{g}}_n$ are totally geodesic in $\widetilde{M}_{n+1}$, in particular they are geodesics in $\widetilde{M}_{n+1}$. From this, it immediately follows that each of the subspaces in the orbits $G_n\widetilde{M}_n$ and $G_n\widetilde{M}'_n$ is totally geodesic. □

Proposition 5.2. *There exists a closed geodesic $\mathfrak{g}_{n+1}$ in $M_{n+1}$ such that in the universal cover $\widetilde{M}_{n+1}$ the lift $\widetilde{\mathfrak{g}}_{n+1}$ is Morse and has divergence $\asymp x^{n+2}$.*

*Proof.* The group $G_{n+1}$ acts on the tree $T_n$ with quotient an edge; therefore there exists a loxodromic element $\gamma \in G_{n+1}$. The action is acylindrical, moreover the stabilizers of two distinct edges have trivial intersection. Indeed consider two edges $e$ and $he$, with $h \in G_{n+1}$. Assume that their stabilizers $C_n$ and $hC_nh^{-1}$ intersect non-trivially. Then they intersect in some finite index cyclic subgroup $C_n'$ of $C_n$. In particular, there exist two integers $r, s$ such that if $\gamma_n$ is the generator of $C_n$ then $\gamma_n^r = h\gamma_n^s h^{-1}$. If $r \neq \pm s$ then it can be easily proved that $C_n$ must be distorted in $G_{n+1}$, contradicting the previous lemma. It follows that $r = \pm s$, and up to replacing $h$ by $h^2$ we may assume that $r = s$. It follows that $h$ is an element of infinite order in the center of $C_n'$, and this contradicts the fact that $C_n$ (and hence $C_n'$) is a Morse quasi-geodesic.

By Theorem 4.1 in [DMS10], since the cyclic subgroup $C = \langle\gamma\rangle$ acts acylindrically by isometries on a simplicial tree it is a Morse quasi-geodesic. Consider a point $x \in \widetilde{M}_{n+1}$. Since the map $G_{n+1} \to \widetilde{M}_{n+1}$ defined by $g \mapsto gx$ is a quasi-isometry, it follows that $Cx$ is a Morse quasi-geodesic. The sequence of geodesics $[\gamma^{-n}x, \gamma^n x]$ is contained in $\mathcal{N}_d(Cx)$ for a fixed $d$, hence it has a subsequence converging to a bi-infinite geodesic $\mathfrak{p}$ entirely contained in $\mathcal{N}_d(Cx)$. For every $k \in \mathbb{Z}$, $\gamma^k\mathfrak{p}$ is also inside $\mathcal{N}_d(Cx)$. It follows that the two bi-infinite geodesics $\mathfrak{p}$ and $\gamma^k\mathfrak{p}$ are at finite Hausdorff distance. Since the function $t \mapsto \mathrm{dist}(\mathfrak{p}(t), \gamma^k\mathfrak{p}(t))$ is convex positive and bounded it follows that it is constant, hence the two geodesics are parallel. According to [BH99], the set of bi-infinite geodesics parallel to $\mathfrak{p}$ compose a set isometric to $\mathfrak{p} \times K$, where $K$ is a convex subset. Since $C$ is a Morse quasi-geodesic, hence $\mathfrak{p}$ is a Morse geodesic, it follows that $K$ must be bounded. By possibly replacing it with a smaller set, we may assume that $\mathfrak{p} \times K$ is invariant with respect to $C$. If $b$ denotes the barycenter of $K$, then $\mathfrak{p} \times \{b\}$ is invariant with respect to $C$. Take $\widetilde{\mathfrak{g}}_{n+1} = \mathfrak{p} \times \{b\}$ and $\mathfrak{g}_{n+1} = \widetilde{\mathfrak{g}}_{n+1}/C$.

The only thing which remains to be proven is that the divergence of $\widetilde{\mathfrak{g}}_{n+1}$ is equivalent to $x^{n+2}$. Consider a shortest path, which we call $\mathfrak{c}$, joining $\widetilde{\mathfrak{g}}_{n+1}(-x)$ and $\widetilde{\mathfrak{g}}_{n+1}(x)$ outside the ball $\mathrm{B}(\widetilde{\mathfrak{g}}_{n+1}(0), \delta x - \kappa)$. In particular, the path $\mathfrak{c}$ is at distance at least $\frac{\delta}{3}x$ from the geodesic $\widetilde{\mathfrak{g}}_{n+1}$ restricted to $[-\frac{\delta}{3}x, \frac{\delta}{3}x]$; also, $\mathfrak{c}$ has to cross the same separating geodesics in $G_{n+1}\widetilde{\mathfrak{g}}_n$.

There exists a constant $m$, depending only on the space $M_{n+1}$, such that two separating geodesics crossed consecutively by $\widetilde{\mathfrak{g}}_{n+1}$ are at distance at most $m$ and such that the pair of points realizing the distance between two such consecutive geodesics is inside $\mathcal{N}_m(\widetilde{\mathfrak{g}}_{n+1})$. It follows that the number of separating geodesics in $G_{n+1}\widetilde{\mathfrak{g}}_n$ crossed by $\widetilde{\mathfrak{g}}_{n+1}$ restricted to $[-\frac{\delta}{3}x, \frac{\delta}{3}x]$ is $\asymp \frac{2\delta}{3}x$. Let $\alpha$ and $\alpha'$ be the two intersection points of $\mathfrak{c}$ with two geodesics in $G_{n+1}\widetilde{\mathfrak{g}}_n$ crossed consecutively, $\mathfrak{g}$ and $\mathfrak{g}'$, and let $\mathfrak{c}'$ be the subpath of $\mathfrak{c}$ of endpoints $\alpha$ and $\alpha'$. Let $[a, b]$ be a geodesic which is the shortest path joining $\mathfrak{g}$ and $\mathfrak{g}'$. For $\delta'$ small enough, we may assume that $\mathfrak{c}'$ is outside $\mathcal{N}_{\delta' x}([a, b])$. Let $\beta$ be the nearest point projection of $\alpha'$ onto $\mathfrak{g}$. Lemma 6.1 implies that $\mathrm{dist}(\alpha', \beta) \geq \lambda x$ for some constant $\lambda > 0$ independent of the point $\alpha'$.

Both the geodesics $[a, b]$ and $[\beta, \alpha']$ make an angle of at least $\frac{\pi}{2}$ with $\mathfrak{g}$, since one of the endpoints of each is the nearest point projection on $\mathfrak{g}$ of the other endpoint. This and the CAT(0)-property implies that $[\alpha',\beta]\cup[\beta,a]\cup[a,b]$ is an $(L,0)$-quasi-geodesic, for large enough $L$ (depending also on $\lambda$). It follows that $[\alpha',\beta]\cup[\beta,a]\cup[a,b]$ is contained in an $m$-neighborhood of $\mathfrak{g}'$. In particular, there exists $\beta'$ on $\mathfrak{g}'$ with $\operatorname{dist}(\beta',\beta)\asymp\lambda x$ and such that $[\alpha',\beta']\subseteq\mathfrak{g}'$ has nearest point projection on $\mathfrak{g}$ at distance $O(1)$ from $\beta$. Also $\beta$ is at distance $O(1)$ from $a$, otherwise we would obtain again that some finite index subgroup of $C$ has a non-trivial element $h\notin C$ in its center. We choose a point $\mu$ on $\mathfrak{g}$ between $\alpha$ and $\beta$, at distance $\varepsilon x$ from $\beta$ and a ball $\mathrm{B}(\mu,\frac{\varepsilon}{10}x)\subset\mathcal{N}_{\delta' x}([a,b])$ with $\varepsilon$ small enough. The ball $\mathrm{B}(\mu,\frac{\varepsilon}{10}x)$ does not intersect the path $\mathfrak{c}'\cup[\alpha',\beta']\cup[\beta',\beta]$. This and the fact that $\mathfrak{g}$ has divergence $\asymp x^{n+1}$ implies that the length of $\mathfrak{c}'\succeq x^{n+1}$. Note, to see that the divergence of $\mathfrak{g}$ (and $\tilde{\mathfrak{g}}_n$) is still $\asymp x^{n+1}$ in $\tilde{\mathfrak{M}}_{n+1}$, one applies the same argument as above to the sequence of geodesics from $G_{n+1}\widetilde{g}_n$ which are crossed by $\mathfrak{c}$ and which intersect $\mathrm{B}(\mathfrak{g}(0),\frac{\delta}{2}x)$.

We conclude that the length of the path $\mathfrak{c}\succeq x^{n+2}$. $\square$

## 6. Morse quasi-geodesics and divergence

In this section, we improve the result in Proposition 3.12 and generalize it to the utmost in the CAT(0) setting in Theorem 6.6. We also show that the quadratic lower bound on divergence occurs for many concrete examples of Morse elements in groups. This together with the estimate on divergence coming from the structure of thick metric space yields a divergence precisely quadratic for several groups and spaces.

We shall apply the following lemma.

LEMMA 6.1. *Let $X$ be a compact metric space which is locally* CAT(0). *Consider two periodic geodesics $\mathfrak{g}$ and $\mathfrak{g}'$ in $X$ that do not bound a flat strip, and lifts $\widetilde{\mathfrak{g}}$ and $\widetilde{\mathfrak{g}}'$ of these two geodesics such that* $\operatorname{dist}(\widetilde{\mathfrak{g}}(0),\widetilde{\mathfrak{g}}'(0))\le\kappa$. *Then there exists $\varepsilon>0$ and $x_0$ depending on $\kappa$ such that*

$$\operatorname{dist}\big(\widetilde{\mathfrak{g}}(x),\widetilde{\mathfrak{g}}'(x)\big)\ge\varepsilon x\quad\text{for every } x\ge x_0.$$

This result can be proved using an argument that is now standard; see, for instance, the proof of [KL98, Proposition 3.3].

LEMMA 6.2. *Let $\mathfrak{q}$ be a Morse quasi-geodesic in a metric space* $(X,\operatorname{dist})$. *Then for every $\lambda\in(0,1)$ and for every $M>0$ there exist $D>0$ such that the following holds. If $\mathfrak{c}$ is a sub-quasi-geodesic of $\mathfrak{q}$, $x$ and $y$ are points in $X$, $x'$ and $y'$ are points on $\mathfrak{c}$ minimizing the distance to $x$, respectively $y$, and* $\operatorname{dist}(x',y')\ge D$ *while* $\operatorname{dist}(x,x')+\operatorname{dist}(y',y)\le M\operatorname{dist}(x',y')$ *then* $\operatorname{dist}(x,y)\ge\lambda[\operatorname{dist}(x,x')+\operatorname{dist}(x',y')+\operatorname{dist}(y',y)]$.

*Proof.* We produce a constant $D$ which given $X$, $\lambda$, and $M$, will work for any other Morse quasi-geodesic with the same associated Morse constants as $\mathfrak{q}$.

Assume for a contradiction that there exists $\lambda\in(0,1)$ and $M>0$ such that for every $D_n>0$ there exist $\mathfrak{c}_n$ sub-quasi-geodesic of $\mathfrak{q}$, $x_n,y_n\in X$

and $x_n', y_n'$ points on $\mathfrak{c}_n$ minimizing the distance to $x_n$, respectively $y_n$ such that $\mathrm{dist}(x_n', y_n') \geq D_n$, $\mathrm{dist}(x_n, x_n') + \mathrm{dist}(y_n', y_n) \leq M\,\mathrm{dist}(x_n', y_n')$ while $\mathrm{dist}(x_n, y_n) \leq \lambda[\mathrm{dist}(x_n, x_n') + \mathrm{dist}(x_n', y_n') + \mathrm{dist}(y_n', y_n)]$.

We denote $\mathrm{dist}(x_n, x_n') + \mathrm{dist}(x_n', y_n') + \mathrm{dist}(y_n', y_n)$ by $\delta_n$. Note that by hypothesis

$$\mathrm{dist}\big(x_n', y_n'\big) \leq \delta_n \leq (M+1)\mathrm{dist}\big(x_n', y_n'\big).$$

Consider $\mathrm{Cone}_\omega(X, (x_n); (\delta_n))$. Since $\mathfrak{q}$ is a Morse quasi-geodesic $\mathfrak{q}_\omega = \lim_\omega(\mathfrak{q})$ is a bi-Lipschitz path in a transversal tree, cf. [DS05, Proposition 3.24]. Moreover, $\mathfrak{c}_\omega = \lim_\omega(\mathfrak{c}_n)$ is a subpath of $\mathfrak{q}_\omega$. Denote $x_\omega' = \lim_\omega(x_n')$, $y_\omega' = \lim_\omega(y_n')$, $x_\omega = \lim_\omega(x_n)$, and $y_\omega = \lim_\omega(y_n)$ the ultralimits of the points considered above. Any geodesics $[x_\omega, x_\omega']$ and $[y_\omega, y_\omega']$ intersects $\mathfrak{c}_\omega$ only in an endpoint. Let $[x_\omega'', x_\omega']$ be the intersection of the geodesic $[x_\omega, x_\omega']$ with the transversal tree containing $\mathfrak{q}_\omega$; let $[y_\omega'', y_\omega']$ be defined likewise. The union $[x_\omega'', x_\omega'] \cup [x_\omega', y_\omega'] \cup [y_\omega', y_\omega'']$ is a geodesic in a transversal tree, since it is a concatenation of three arcs such that $[x_\omega', y_\omega']$ does not reduce to a point, and it intersects its predecessor in $x_\omega'$ and its successor in $y_\omega'$.

This, the fact that transversal trees can always be added to the list of pieces in a tree-graded space [DS05, Remark 2.27], and Lemma 2.28 in [DS05] imply that $[x_\omega, x_\omega'] \cup [x_\omega', y_\omega'] \cup [y_\omega', y_\omega]$ is a geodesic. In particular, $\mathrm{dist}(x_n, y_n) = \delta_n + o(\delta_n)$. This contradicts the fact that $\mathrm{dist}(x_n, y_n) \leq \lambda\delta_n$. □

LEMMA 6.3. *Let $\mathfrak{q}$ be a Morse quasi-geodesic in a* $\mathrm{CAT}(0)$ *metric space* $(X, \mathrm{dist})$. *Then for every $\lambda \in (0,1)$ there exists $D > 0$ such that the following holds. If $\mathfrak{c}$ is a sub-quasi-geodesic of $\mathfrak{q}$, $x$ and $y$ are points in $X$, $x'$ and $y'$ are points on $\mathfrak{c}$ minimizing the distance to $x$, respectively $y$, and* $\mathrm{dist}(x', y') \geq D$ *then* $\mathrm{dist}(x, y) \geq \lambda[\mathrm{dist}(x, x') + \mathrm{dist}(x', y') + \mathrm{dist}(y', y)]$.

*Proof.* Again, we produce a constant $D$ which given $X$, $\lambda$, will work for any other Morse quasi-geodesic with the same associated Morse constants as $\mathfrak{q}$.

We argue by contradiction and assume that for some $\lambda$ there exist sequences of Morse quasi-geodesics $\mathfrak{q}_n$, with the same Morse constants as $\mathfrak{q}$, as well as sub-quasi-geodesics $\mathfrak{c}_n \subset \mathfrak{q}_n$ and pairs of points $x_n, y_n$ such that for some points $x_n'$ and $y_n'$ on $\mathfrak{c}_n$ minimizing the distance to $x_n$, respectively $y_n$ we have that $\mathrm{dist}(x_n', y_n') \geq n$ while $\mathrm{dist}(x_n, y_n) \leq \lambda[\mathrm{dist}(x_n, x_n') + \mathrm{dist}(x_n', y_n') + \mathrm{dist}(y_n', y_n)]$. In what follows, we fix some geodesics $[x_n, x_n']$ and $[y_n, y_n']$, and for every $u_n \in [x_n, x_n']$ and $v_n \in [y_n, y_n']$ we introduce the notation

$$\delta(u_n, v_n) := \mathrm{dist}\big(u_n, x_n'\big) + \mathrm{dist}\big(x_n', y_n'\big) + \mathrm{dist}\big(y_n', v_n\big).$$

We will break the argument into two cases.

*Case* (i): Assume $\lim_\omega \frac{\delta(x_n, y_n)}{\mathrm{dist}(x_n', y_n')} < \infty$.

Then in $\mathrm{Cone}_\omega(X; (x_n'), (\mathrm{dist}(x_n', y_n')))$, the limit $\mathfrak{c}_\omega = \lim_\omega \mathfrak{c}_n$ is a subarc in the transversal line $\mathfrak{q}_\omega = \lim_\omega \mathfrak{q}_n$ containing the two points $x_\omega'$ and $y_\omega'$ distance 1 apart, which points are the nearest points in $\mathfrak{c}_\omega$ to the points $x_\omega$ and $y_\omega$.

With an argument as in the proof of the previous lemma, we obtain that $[x_\omega, x'_\omega] \cup [x'_\omega, y'_\omega] \cup [y'_\omega, y_\omega]$ is a geodesic, in particular $\mathrm{dist}(x_n, y_n) = \delta(x_n, y_n) + o(\mathrm{dist}(x'_n, y'_n))$. This contradicts the fact that $\mathrm{dist}(x_n, y_n) \leq \lambda\delta(x_n, y_n)$.

*Case* (ii): Assume $\lim_\omega \frac{\delta(x_n,y_n)}{\mathrm{dist}(x'_n,y'_n)} = \infty$, or equivalently that

$$\lim_\omega \frac{\mathrm{dist}(x_n, x'_n) + \mathrm{dist}(y_n, y'_n)}{\mathrm{dist}(x'_n, y'_n)} = \infty.$$

We consider the parametrization proportional to the arc-length $\mathfrak{g}_x\colon [0,1] \to [x_n, x'_n]$ sending 0 to $x'_n$ and 1 to $x_n$. Similarly, define $\mathfrak{g}_y\colon [0,1] \to [y_n, y'_n]$.

Fix $\lambda' \in (\lambda, 1)$ and for every $n$ consider the maximal $t_n \in [0,1]$ such that

$$\mathrm{dist}\big(\mathfrak{g}_x(t_n), \mathfrak{g}_y(t_n)\big) \geq \lambda'\delta\big(\mathfrak{g}_x(t_n), \mathfrak{g}_y(t_n)\big).$$

Clearly, $t_n < 1$ and from the continuity of the two sides of the inequality above and the maximality of $t_n$ we deduce that

$$\text{(iii)} \qquad \mathrm{dist}\big(\mathfrak{g}_x(t_n), \mathfrak{g}_y(t_n)\big) = \lambda'\delta\big(\mathfrak{g}_x(t_n), \mathfrak{g}_y(t_n)\big).$$

Using the convexity of the distance, we have:

$$\begin{aligned}
\lambda'\delta\big(\mathfrak{g}_x(t_n), \mathfrak{g}_y(t_n)\big) &= \mathrm{dist}\big(\mathfrak{g}_x(t_n), \mathfrak{g}_y(t_n)\big) \\
&\leq (1 - t_n)\,\mathrm{dist}\big(x'_n, y'_n\big) + t_n\,\mathrm{dist}(x_n, y_n) \\
&\leq (1 - t_n)\,\mathrm{dist}\big(x'_n, y'_n\big) + t_n\lambda\delta(x_n, y_n) \\
&\leq \mathrm{dist}\big(x'_n, y'_n\big) + \lambda\delta\big(\mathfrak{g}_x(t_n), \mathfrak{g}_y(t_n)\big).
\end{aligned}$$

Whence it follows that

$$\big(\lambda' - \lambda\big)\delta\big(\mathfrak{g}_x(t_n), \mathfrak{g}_y(t_n)\big) \leq \mathrm{dist}\big(x'_n, y'_n\big).$$

In particular $\lim_\omega \frac{\delta(\mathfrak{g}_x(t_n), \mathfrak{g}_y(t_n))}{\mathrm{dist}(x'_n, y'_n)} \leq \frac{1}{\lambda' - \lambda} < \infty$, whence

$$\lim_\omega \frac{\mathrm{dist}(\mathfrak{g}_x(t_n), x'_n) + \mathrm{dist}(y'_n, \mathfrak{g}_y(t_n))}{\mathrm{dist}(x'_n, y'_n)} < \infty.$$

If the above limit is zero, then since

$$\begin{aligned}
\mathrm{dist}\big(x'_n, y'_n\big) - \big[\mathrm{dist}\big(\mathfrak{g}_x(t_n), x'_n\big) + \mathrm{dist}\big(y'_n, \mathfrak{g}_y(t_n)\big)\big] &\leq \mathrm{dist}\big(\mathfrak{g}_x(t_n), \mathfrak{g}_y(t_n)\big) \\
&\leq \delta\big(\mathfrak{g}_x(t_n), \mathfrak{g}_y(t_n)\big)
\end{aligned}$$

it follows that $\lim_\omega \frac{\delta(\mathfrak{g}_x(t_n), \mathfrak{g}_y(t_n))}{\mathrm{dist}(x'_n, y'_n)} = \lim_\omega \frac{\mathrm{dist}(\mathfrak{g}_x(t_n), \mathfrak{g}_y(t_n))}{\mathrm{dist}(x'_n, y'_n)} = 1$.

Thus if in equation (iii), we divide by $\mathrm{dist}(x'_n, y'_n)$ and take the $\omega$-limits we obtain $1 = \lambda'$, a contradiction.

We conclude that

$$0 < \lim_\omega \frac{\mathrm{dist}(\mathfrak{g}_x(t_n), x'_n) + \mathrm{dist}(y'_n, \mathfrak{g}_y(t_n))}{\mathrm{dist}(x'_n, y'_n)} < \infty.$$

In $\mathrm{Cone}_\omega(X; (x'_n), (\mathrm{dist}(x'_n, y'_n)))$, we again have that $\mathfrak{c}_\omega = \lim_\omega \mathfrak{c}_n$ is a subarc in the transversal line $\mathfrak{q}_\omega = \lim_\omega \mathfrak{q}_n$ containing the two points $x'_\omega$ and $y'_\omega$

which are distance 1 apart. The limits $\lim_\omega[x_n, x'_n]$ and $\lim_\omega[y_n, y'_n]$ are either two rays intersecting $\mathfrak{c}_\omega$ only in their origin or only one ray like this and one geodesic segment (possibly trivial) intersecting $\mathfrak{c}_\omega$ only in one point. The limits of the sequences of points $\mathfrak{g}_x(t_n)$ and $\mathfrak{g}_y(t_n)$ are respectively on each of the two rays (or on the ray and the segment), in particular it follows that $\mathrm{dist}(\mathfrak{g}_x(t_n), \mathfrak{g}_y(t_n)) = \delta(\mathfrak{g}_x(t_n), \mathfrak{g}_y(t_n)) + o(\mathrm{dist}(x'_n, y'_n))$.

This and equality (iii) yield a contradiction. ☐

LEMMA 6.4. *Let $\mathfrak{q}$ be a Morse quasi-geodesic in a* CAT(0) *metric space* $(X, \mathrm{dist})$. *There exists a constant $D_0$ such that if $\mathfrak{c}$ is a sub-quasi-geodesic of $\mathfrak{q}$ and two points $x, y \in X$ are such that both* $\mathrm{dist}(x, \mathfrak{c})$ *and* $\mathrm{dist}(y, \mathfrak{c})$ *are strictly larger than* $\mathrm{dist}(x, y)$ *and $x'$ and $y'$ are points on $\mathfrak{c}$ minimizing the distance to $x$, respectively $y$, then* $\mathrm{dist}(x', y') \le D_0$.

*Proof.* According to Lemma 6.2 there exists $D_0$ such that if $\mathfrak{c}$ is a sub-quasi-geodesic of $\mathfrak{q}$, $x, y \in X$ and $x'$ and $y'$ are points on $\mathfrak{c}$ minimizing the distance to $x$, respectively, $y$, $\mathrm{dist}(x, y) < \frac{1}{2}[\mathrm{dist}(x, x') + \mathrm{dist}(x', y') + \mathrm{dist}(y', y)]$ implies $\mathrm{dist}(x', y') \le D_0$. ☐

REMARK 6.5. When the space $(X, \mathrm{dist})$ is a CAT(0) space, in Lemmas 6.2, 6.3 and 6.4, the output constants only depend on the input constants and on the Morse parameters, as introduced in Definition 3.10.

Indeed, in the proofs of the three lemmas the argument is always by contradiction and the use of the geometry of asymptotic cones and their transversal trees. If the given Morse quasi-geodesic is replaced by a sequence of Morse quasi-geodesics with the same Morse parameters, then in the asymptotic cone the limit is still a transversal arc, and the argument works.

THEOREM 6.6. *Let $\mathfrak{q}$ be a Morse quasi-geodesic in a* CAT(0) *metric space* $(X, \mathrm{dist})$. *Then the divergence* $\mathrm{Div}^{\mathfrak{q}} \ge (\kappa x - \kappa)^2$, *where $\kappa$ is a constant depending only on the constants $\delta$ and $\gamma$ used to define the divergence and on the Morse parameters.*

*Proof.* Let $a = \mathfrak{q}(-x)$ and $b = \mathfrak{q}(x)$ and let $\mathfrak{p}$ be a path joining $a$ and $b$ outside $\mathrm{B}(\mathfrak{q}(0), \delta x - \gamma)$. Let $\mathfrak{c}$ be the maximal subpath of $\mathfrak{q}$ with endpoints contained in $\mathrm{B}(\mathfrak{q}(0), \frac{\delta}{2}x - 3\gamma)$. Then for $\gamma$ large enough we may assume that $\mathfrak{c}$ is entirely contained in $\mathrm{B}(\mathfrak{q}(0), \frac{\delta}{2}x - 2\gamma)$, as $\mathfrak{c}$ is at finite Hausdorff distance from the geodesic joining its endpoints. All the points in $\mathfrak{p}$ are at distance at least $\frac{\delta}{2}x + \gamma$ from $\mathfrak{c}$. Let $y_0 = a, y_1, \ldots, y_n = b$ be consecutive points on $\mathfrak{p}$ dividing it into subarcs of length $\frac{\delta}{2}x$ (except the last who might be shorter). For each of the points $y_i$, let $y'_i \in \mathfrak{c}$ be a point minimizing the distance to $y_i$. Lemma 6.4 implies that $\mathrm{dist}(y'_i, y'_{i+1}) \le D_0$ for every $i$. For some $\varepsilon > 0$ and for all $x$ sufficiently large, we have $\mathrm{dist}(y'_0, y'_n) \ge \varepsilon x$. Indeed if we would assume the contrary then there would exist a sequence of positive numbers $x_n$ diverging to infinity such that $\mathfrak{q}(-x_n)$ and $q(x_n)$ have their respective nearest points

$u_n$ and $v_n$ on $\mathfrak{c}_n$, maximal subpath of $\mathfrak{q}$ with endpoints in $\mathrm{B}(\mathfrak{q}(0), \frac{\delta}{2}x_n - 3\gamma)$, at distance $\mathrm{dist}(u_n, v_n)$ at most $\frac{1}{n}x_n$. Then in $\mathrm{Cone}_\omega(X, \mathfrak{q}(0), (x_n))$ we obtain two points on the transversal line $\mathfrak{q}_\omega$ separated by $B(\mathfrak{q}_\omega(0), \frac{\delta}{2})$ but whose nearest point projections onto $\mathfrak{q}_\omega \cap B(\mathfrak{q}_\omega(0), \frac{\delta}{2})$ coincide. This is impossible.

It follows that $\varepsilon x \leq n D_0$, whence $n \geq \frac{\varepsilon}{D_0}x$. It follows that the length of $\mathfrak{p}$ is larger than $(n-1)\frac{\delta}{2}x \geq (\frac{\varepsilon}{D_0}x - 1)\frac{\delta}{2}x$. In other words, the length of $\mathfrak{p} \geq (\kappa x - \kappa)^2$.

The fact that the constant $\kappa$ depends, besides $\delta$ and $\gamma$, only on the Morse parameters follows as in Remark 6.5. $\square$

COROLLARY 6.7. *Assume that a finitely generated group $G$ acts on a* CAT(0)*-space $X$ such that one* (*every*) *orbit map $G \to X$, $g \mapsto gx$, is a quasi-isometric embedding and its image contains a Morse quasi-geodesic. Then the divergence* $\mathrm{Div}^G \succeq x^2$.

A phenomenon similar to that in Lemma 6.4 may occur in general in a metric space with a *$D$-contracting* (or simply *contracting*) quasi-geodesic, that is a quasi-geodesic $\mathfrak{q}$ such that for every ball $B$ disjoint from $\mathfrak{q}$ the points in $\mathfrak{q}$ nearest to points in $B$ compose a set of diameter $D$. It is easy to show that such a quasi-geodesic is Morse and has divergence $\mathrm{Div}^{\mathfrak{q}} \succeq x^2$.

Known examples of contracting quasi-geodesics include periodic geodesics in a graph manifold, which non-trivially intersect more than two Seifert-fibered components [KL98]; orbits of pseudo-Anosovs in the Teichmüller space endowed either with the Teichmüller metric [Min96] or with the Weil–Petersson metric [Beh06]; cyclic subgroups generated by pseudo-Anosov elements in mapping class groups [Beh06].

The divergence is quadratic in all of the above examples (except for the low complexity cases of the mapping class group and Teichmüller space where the divergence is larger). The above cited papers proved the upper bounds except in the case of the Weil–Petersson metric, which depends on [BDM09] and [BM08] and is still open in the case of genus two with one boundary component (see Question 4.19);[1] and for the Teichmüller metric, where the upper bound was established in [DR09].

We also note, that although the presence of Morse quasi-geodesics in $\mathrm{Aut}(F_n)$ and $\mathrm{Out}(F_n)$ has been established [AK], the following remains open:

QUESTION 6.8. In $\mathrm{Out}(F_n)$ for $n \geq 3$

- are the cyclic subgroups generated by fully irreducible elements contracting quasi-geodesics?
- is the divergence quadratic in the groups $\mathrm{Out}(F_n)$ and $\mathrm{Aut}(F_n)$?

[1] Since this paper was first circulated there has been progress on the WP metric on Teichmüller space of the genus two surface with one puncture in that it is now know to have super-quadratic divergence [Sul].

To sum up the relationship between divergence and existence of cut-points in asymptotic cones, the following are known:

- if all asymptotic cones are without cut-points then the divergence is linear;
- if we assume that at least one asymptotic cone is without cut-points then examples were constructed by Olshanskii–Osin–Sapir of groups $G$ with divergence satisfying $\mathrm{Div}^G(n) \leq Cn$ for a constant $C$ for all $n$ in an infinite subset of $\mathbb{N}$ and with $\mathrm{Div}^G \preceq f(n)$ for any $f$ such that $\frac{f(n)}{n}$ non-decreasing; in particular $\mathrm{Div}^G$ may be as close to linear as possible; but it is superlinear if one asymptotic cone has cut-points;
- if the space is CAT(0) and all asymptotic cones have cut-points coming from the limit set of a Morse quasi-geodesic then the divergence is at least quadratic.

This raises the following natural question.

QUESTION 6.9. If a CAT(0) (quasi-homogeneous) metric space has cut-points in every asymptotic cone, must the divergence of that metric space be at least quadratic?

An affirmative answer to this question would be an immediately corollary of an affirmative answer to the following (which we expect would be more difficult to establish):

QUESTION 6.10. Does the existence of cut-points in every asymptotic cone of a CAT(0) quasi-homogeneous metric space imply the existence of a Morse quasi-geodesic?

## 7. Morse elements and length of the shortest conjugator

It is known that given a group $G$ acting properly discontinuously and cocompactly on a CAT(0)-space, and two elements $u, v$ that are conjugate in $G$ there exists $K > 0$ depending on the choice of word metric in $G$ such that $v = gug^{-1}$ for some $g$ with $|g| \leq \exp(K(|u| + |v|))$. As shown below, a similar estimate on the length of the shortest conjugator holds in a more general context of groups with some non-positively curved or hyperbolic geometry associated to them.

**7.1. The CAT(0) set-up.** Through this section $X$ will be a locally compact CAT(0) space. A standard CAT(0) argument yields a bound on the length of the shortest conjugator of two axial Morse isometries $u$ and $v$ of a locally compact CAT(0) space in terms of two parameters of the geometry of the action of both $u$ and $v$. We define these parameters below.

(1) Recall that given an axial isometry $u$ of a CAT(0) space the set

$$\mathrm{Min}(u) = \left\{ x \in X \mid \mathrm{dist}(x, ux) = \inf_{y \in X} \mathrm{dist}(y, uy) \right\}$$

is isometric to a set $C \times \mathbb{R}$, where $u$ acts as a translation with translation length $t_u$ along each fiber $\{c\} \times \mathbb{R}$ and $C$ is a closed convex subset [BH99, Theorem 6.8, p. 231]. When the axial isometry is Morse, the set $C$ is bounded, and we denote by $D_u$ the diameter of the set $C \times [0, t_u]$.

(2) There exists $\theta_u > 0$ such that if $x$ is a point outside $\mathcal{N}_1(\mathrm{Min}(u))$ and $x'$ is its nearest point projection onto $\mathrm{Min}(u)$ then

$$\mathrm{dist}(x, ux) \geq \mathrm{dist}\big(x', ux'\big) + \theta_u \mathrm{dist}\big(x, x'\big). \tag{iv}$$

Indeed assume that on the contrary we have points $x_n$ outside $\mathcal{N}_1(\mathrm{Min}(u))$ with projections $x_n'$ such that $\mathrm{dist}(x_n, ux_n) \leq \mathrm{dist}(x_n', ux_n') + \frac{1}{n}\,\mathrm{dist}(x_n, x_n')$. By the convexity of the distance, we may assume that all $x_n$ are at distance 1 from $\mathrm{Min}(u)$ and by eventually applying powers of $u$ we may assume that all $x_n'$ are in the compact set $C \times [0, t_u]$. Since $X$ is locally compact the quadrangles of vertices $x_n, x_n', ux_n', ux_n$ converge on a subsequence in the Hausdorff distance to a flat quadrangle $a, a', ua', ua$ intersecting $\mathrm{Min}(u)$ in the edge $[a', ua']$ and such that for every $z \in [a, a']$, $\mathrm{dist}(z, uz) = t_u$. This contradicts the definition of $\mathrm{Min}(u)$.

Note that both constants $D_u$ and $\theta_u$ only depend on the conjugacy class of $u$, therefore we occasionally denote them as $D_{[u]}$ and $\theta_{[u]}$.

Proposition 7.1. *Let $X$ be a locally compact* CAT(0) *space, let $G$ be a group of isometries of $X$ and let $x_0$ be a point in $X$.*

*Let $u$ be a Morse axial isometry of $X$. For every element $v \in G$ conjugate to $u$, there exists an element $g \in G$ such that $v = gug^{-1}$ and such that*

$$\mathrm{dist}(x_0, gx_0) \leq \frac{1}{\theta_{[u]}}\big[\mathrm{dist}(x_0, ux_0) + \mathrm{dist}(x_0, vx_0)\big] + D_{[u]} + 2,$$

*where $\theta_{[u]}$ and $D_{[u]}$ are the constants dependent on the conjugacy class of $u$ previously defined.*

*In particular, if the map $g \mapsto gx_0$ is a quasi-isometric embedding of $G$ into $X$ then $\mathrm{dist}_G(1, g) \leq_{A,B} \mathrm{dist}(1, u) + \mathrm{dist}(1, v)$, where the constants $A, B$ depend on the Morse parameters of the above isometries, on $\theta_{[u]}$ and on $D_{[u]}$.*

*Proof.* Let $y_0$ be the nearest point projection of $x_0$ onto $\mathrm{Min}(u)$ and $z_0$ the nearest point projection of $x_0$ onto $\mathrm{Min}(v)$.

Let $g \in G$ be such that $v = gug^{-1}$. Then $\mathrm{Min}(v) = g\,\mathrm{Min}(u)$, in particular $gy_0 \in \mathrm{Min}(v)$. By eventually replacing $g$ with $v^k g$ for an appropriate $k \in \mathbb{Z}$ one may assume that both $z_0$ and $gy_0$ are in the same isometric copy of $C \times [0, t_u]$, which is a fundamental domain of the action of $v$ on $\mathrm{Min}(v)$, hence within distance at most $D_u$, with the notation introduced previously.

The distance $\mathrm{dist}(x_0, gx_0)$ is at most $\mathrm{dist}(x_0, z_0) + D_u + \mathrm{dist}(y_0, x_0)$. On the other hand either $\mathrm{dist}(x_0, y_0) \leq 1$ or $\mathrm{dist}(x_0, y_0) \leq \frac{1}{\theta_u}\,\mathrm{dist}(x_0, ux_0)$. Similarly, either $\mathrm{dist}(x_0, z_0) \leq 1$ or $\mathrm{dist}(x_0, z_0) \leq \frac{1}{\theta_u} \mathrm{dist}(x_0, vx_0)$. Hence,

$$\mathrm{dist}(x_0, gx_0) \leq \frac{1}{\theta_u}\left[\mathrm{dist}(x_0, ux_0) + \mathrm{dist}(x_0, vx_0)\right] + D_u + 2. \qquad \square$$

COROLLARY 7.2. *Let $X$ be the universal cover of a compact manifold of non-positive curvature $M$, and let $u, v$ be two conjugate rank-one elements in $G = \pi_1(M)$. Then there exists $g \in G$ such that* $\mathrm{dist}_G(1, g) \leq_{A,B} \mathrm{dist}_G(1, u) + \mathrm{dist}_G(1, v)$, *where the constants $A$ is the smallest constant such that both $u$ and $v$ are $A$-contractions and $B$ depends on $A$ and the translation length of $u$.*

*Proof.* Rank one elements such as $u, v$ are $A$-contractions for some $A$ (see [BF] and references therein). If in inequality (iv) the distance $\mathrm{dist}(x', ux')$ (and hence the translation length of $u$, which is equal to that of $v$) is larger than $A$, then by [BF, Lemma 3.5] we have that $\theta_u = 1$ in (iv). By replacing $u, v$ with $u, v$ raised to a sufficiently large power, which we call $B$, we can ensure their translation lengths are larger than $A$ and hence the required bound on $\mathrm{dist}_G(1, g)$ is obtained. $\square$

**7.2. Conjugators in groups acting acylindrically on trees.** In this subsection, we first prove a general result on the shortest length of conjugators of loxodromic elements in groups acting acylindrically on simplicial trees. We then apply the result to obtain sharp estimates of shortest lengths of conjugators in fundamental groups of 3-manifolds.

In what follows, we consider a finitely generated group $G$ acting cocompactly and $l$-acylindrically on a simplicial tree $T$.

Without loss of generality, by passing to a subtree if necessary, we may assume that $G$ acts without inversions of edges.

We fix an arbitrary left-invariant word metric dist on $G$ corresponding to a finite generating set $S$ with $S^{-1} = S$ and $1 \notin S$. We also fix a fundamental domain for the action of $G$ on $T$, in the shape of a finite sub-tree $D$ of $T$ (possibly without some endpoints) [Ser80, §3.1].

LEMMA 7.3. *For every $R \geq 0$, there exists a finite number $f(R)$ such that the following holds. For every vertex $o$ in the fundamental domain $D$ with the corresponding map $\pi_o \colon G \to T, \pi_o(g) = g \cdot o$, and every pair of vertices $a, b$ in the orbit $G \cdot o$ with* $\mathrm{dist}(a, b) \geq l$, *the set $V_{a,b}$ of elements $g \in \pi_o^{-1}(a)$ such that* $\mathrm{dist}(g, \pi_o^{-1}(b)) \leq R$ *is either empty or has diameter at most $f(R)$.*

*Proof.* Without loss of generality, we may assume that $a = o$, hence $\pi_o^{-1}(a) = \mathrm{stab}(o)$, and that there exists $g$ in $\mathrm{stab}(o)$ at distance at most $R$ from $\pi_o^{-1}(b)$. Without loss of generality we may also assume that $g = 1$. The set $\pi_o^{-1}(b)$ can then be written as $h\,\mathrm{stab}(o)$, where $h \in B(1, R)$, $h \cdot o = b$.

Every $g \in V_{o,b}$ is then in $\mathrm{stab}(o)$ and in $\mathcal{N}_R(h\,\mathrm{stab}(o))$. On the other hand, $\mathcal{N}_R(h\,\mathrm{stab}(o)) \subseteq \mathcal{N}_{2R}(h\,\mathrm{stab}(o)h^{-1}) = \mathcal{N}_{2R}(\mathrm{stab}(b))$. By [MSW05, Lemma 2.2], the intersection $\mathrm{stab}(o) \cap \mathcal{N}_{2R}(\mathrm{stab}(b))$ is at finite Hausdorff distance from $\mathrm{stab}(o) \cap \mathrm{stab}(b)$, hence it is of finite diameter, according to the hypothesis of $l$-acylindricity. The vertex $b$ is in the orbit of $B(1, R)$, hence in a finite set, therefore the diameter of $\mathrm{stab}(o) \cap \mathcal{N}_{2R}(\mathrm{stab}(b))$ has a maximum, $f(R)$. □

THEOREM 7.4. *Let $G$ be a group acting cocompactly and $l$-acylindrically on a simplicial tree $T$. There exists a constant $K$ such that if two loxodromic elements $u, v$ are conjugate in $G$ then there exists $g$ conjugating $u, v$ such that*

$$|g| \leq f\big(l|u| + l|v| + K\big) + l|u| + l|v| + 2K,$$

*where the function $f$ is the one defined in Lemma 7.3.*

*Proof.* Let $S$ be the fixed finite generating set defining the word metric on $G$. For every $s \in S$ and every vertex $o \in D$ consider a sequence $g_0 = 1, g_1, g_2, \ldots, g_{m-1}, g_m = s$ in $G$ such that $o, g_1 \cdot o, g_2 \cdot o, \ldots, g_{m-1} \cdot o, s \cdot o$ are the consecutive intersections of $[o, s \cdot o]$ with $G \cdot o$. We denote by $V(s, o)$ the set of elements $\{g_1, \ldots, g_m\}$.

Let $u, v$ be two loxodromic elements in $G$ such that $v = gug^{-1}$ for some $g \in G$. The element $u$ has a translation axis $A_u$ in $T$. Likewise $v$ has an axis $A_v \subset T$ and $gA_u = A_v$. Our goal is to control $|g|_S$ in terms of $|u|_S + |v|_S$.

If $D$ intersects $A_u$, then take a vertex $o$ in the intersection. If not, let $p$ be the nearest point to $D$ on $A_u$ and consider the unique vertex $o \in D$ and an element $h \in G$ such that $p = ho$.

For each $g \in G$ we write $\pi(g)$ to denote $g \cdot o$. Also for every geodesic $[a, b]$ in the Cayley graph of $G$, with consecutive vertices $g_1 = a, g_2, \ldots, g_m = b$ we denote by $\pi[a, b]$ the path in the tree $T$ composed by concatenation of the consecutive geodesics $[g_1 o, g_2 o], \ldots, [g_{m-1} o, g_m o]$.

Consider the geodesic $[1, u^m]$, for some fixed large enough power $m$. Its image by $\pi$ covers the geodesic $[o, u^m o]$. The latter geodesic contains the two points $ho$ and $u^m ho$. Then $[go, gu^m o]$ intersects $A_v$ in $[gho, gu^m ho]$, which can also be written as $[ko, v^m ko]$ for $k = gh$.

Likewise, the image under $\pi$ of the geodesic $[1, v^m]$ contains the geodesic $[o, v^m o]$ in $T$, and the latter geodesic contains a sub-geodesic of $A_v$ of the form $[ro, v^m ro]$, with $r = v^i ko$ for some $i \in \mathbb{Z}$. By possibly post-composing $g$ with $v^i$, we may assume that $ko$ and $ro$ coincide. If $m \geq l$, then $[ro, v^m ro]$ is of length at least $l$. This implies that the geodesic $[1, v^m]$ in the Cayley graph contains two pairs of consecutive vertices $v_1, v_1'$ and $v_2, v_2'$ such that $ro \in [v_1 o, v_1' o]$ and $v^m ro \in [v_2 o, v_2' o]$.

Likewise $[1, u^m]$ contains two pairs of consecutive vertices $u_1, u_1'$ and $u_2, u_2'$ such that $ho \in [u_1 o, u_1' o]$ and $u^m ho \in [u_2 o, u_2' o]$. We thus have that $ro = v_1 x_1 o = g u_1 x_1' o$, where $x_1 \in V(v_1^{-1} v_1', o)$ and $x_1' \in V(u_1^{-1} u_1', o)$; and that $v^m ro = v_2 x_2 o = g u_2 x_2' o$, where $x_2 \in V(v_2^{-1} v_2', o)$ and $x_2' \in V(u_2^{-1} u_2', o)$.

We denote by $M$ the maximum of all the $|x|_S$ for all $x$ in all sets of the form $V(o,s)$ for a vertex $o\in D$ and $s\in S$. The above implies that both $v_1x_1$ and $gu_1x_1'$ are in $\pi^{-1}(ro)$ and at distance at most $m|u|+m|v|+M$ from $\pi^{-1}(v^m ro)$. Lemma 7.3 implies that $v_1$ and $gu_1$ are within distance at most $f(m|u|+m|v|+M)+2M$. It follows that

$$|g|\leq m|v|+f\big(m|u|+m|v|+M\big)+m|u|+2M.$$

Since the only requirement on $m$ was that $m\geq l$, we may take $m=l$. □

It was established in [Pre06] that 3-manifolds have a solvable conjugacy problem, but no bounds on the complexity were provided. The general results we obtain above imply, in particular, a linear control of the shortest conjugator for Morse geodesics in (non-geometric) 3-manifolds. This allows us to obtain the following.

COROLLARY 7.5. *Let $M$ be a non-geometric prime 3-dimensional manifold and let $G$ be its fundamental group.*

*For every word metric on $G$ there exists a constant $K$ such that if two Morse elements $u,v$ are conjugate in $G$ then there exist $g$ conjugating $u,v$ such that*

$$|g|\leq K\big(|u|+|v|\big).$$

*Proof.* Since $M$ is non-geometric, it can be cut along tori and Klein bottles into finitely many geometric components that are either Seifert or hyperbolic. We will apply Theorem 7.4 by considering the Bass-Serre tree $T$ associated to the geometric splitting of $M$ described before and recalling that two elements in $\pi_1(M)$ are Morse if and only if they are either loxodromic elements for the action on $T$ or both contained in a hyperbolic component of $M/\pi_1(M)$.

The following proposition of Kapovich–Leeb, combined with the fact that for every non-geometric prime 3-dimensional manifold $M$ there exists a non-positively curved such manifold $N$, and a bi-Lipschitz homeomorphism between the universal covers $\widetilde{M}$ and $\widetilde{N}$ preserving the components [KL98, Theorem 1.1] implies that for $l\geq 3$ the function $f(R)$ given by Lemma 7.3 is at most $\lambda R+\kappa$.

PROPOSITION 7.6 ([KL97]). *Let $M$ be a non-geometric prime 3-dimensional manifold admitting a non-positively curved Riemannian metric. There exists a constant $\kappa>0$ dependent only on $M$ such that given two geometric components $C,C'$ of $\widetilde{M}$ separated by two flats, the nearest point projection of $C'$ onto $C$ has diameter at most $\kappa$.*

This and Theorem 7.4 settles the case when both $u$ and $v$ are loxodromic elements.

Assume now that $u$ and $v$ both stabilize hyperbolic components. Assume that we have fixed a basepoint $x_0$ in the universal cover $\widetilde{M}$. The map $g\mapsto gx_0$

is a quasi-isometry with fixed constants depending only on the given word metric on $G$. Assume that $u$ stabilizes the hyperbolic component $H \subset \widetilde{M}$ and that it acts on this component as a loxodromic element. We see $H$ as a subset of $\mathbb{H}^3$. Let $A_u$ be the geodesic axis in $\mathbb{H}^3$ on which $u$ acts by translation, denote the translation length by $t$. Note that every segment of length $t$ on $A_u$ intersects $H$. We denote by $\mathrm{Sat}(A_u)$ the set obtained from $A_u$ by replacing its intersections with the open horoballs that compose $\mathbb{H}^3 \setminus H$, with the corresponding boundary horospheres.

The element $v = gug^{-1}$ stabilizes a hyperbolic component $H' = gH \subset \widetilde{M}$, and there exists a geodesic axis $A_v = gA_u$ in $\mathbb{H}^3 \supset H'$ such that $v$ acts on this axis by translation with translation length $t$. We define $\mathrm{Sat}(A_v)$ similarly.

Let $x_0'$ be the nearest point projection of $x_0$ on $H$, let $y_0'$ be the nearest point projection of $x_0'$ onto $A_u$ and let $y_0 \in \mathrm{Sat}(A_u)$ be either the intersection point of $[x_0', y_0']$ with a boundary horosphere if $y_0'$ is in $\mathbb{H}^3 \setminus H$, or equal to $y_0'$ if this latter point is in $H$. Note that $ux_0'$ will be on a different boundary horosphere than $x_0'$, and the same for $uy_0$ and $y_0$, if $y_0$ is on a boundary horosphere. According to [DS05, Lemma 4.26], the geodesic $\mathfrak{g}_{y_0, uy_0}$ joining $y_0$ to $uy_0$ is contained in a $\delta$-neighborhood of $\mathrm{Sat}(A_u)$, moreover due to the fact that every segment of length $t$ on $A_u$ intersects $H$, it follows that $\mathfrak{g}_{y_0, uy_0}$ intersects the $\delta$-neighborhood of $A_u$.

Due to the fact that the metric space $\widetilde{M}$ is hyperbolic relative to the connected components of $\widetilde{M} \setminus \mathrm{Interior}(H)$, it follows that the concatenation of the geodesics $[x_0, x_0'], [x_0', y_0], \mathfrak{g}_{y_0, uy_0}, [uy_0, ux_0'], [ux_0', ux_0]$ composes an $(L, C)$-quasi-geodesic, with $L \geq 1$ and $C \geq 0$ depending only on $M$ [DS05, Lemma 8.12]. We denote this quasi-geodesic $\mathfrak{q}_{x_0, ux_0}$. We construct in a similar manner an $(L, C)$-quasi-geodesic $\mathfrak{q}_{x_0, vx_0}$ joining $x_0$ and $vx_0$ and containing in its $\delta$-neighborhood a sub-segment of the axis $A_v$. Note that the $(L, C)$-quasi-geodesic $g\mathfrak{q}_{x_0, ux_0}$ joining $gx_0, gux_0$ contains in its $\delta$-neighborhood another sub-segment of $A_v$. By pre-composing $g$ with a power of $v$ and possibly replacing $u, v$ by large enough powers, we may assume that the two sub-segments of $A_v$ mentioned above are the same. In particular the $\delta$-neighborhoods of $\mathfrak{q}_{x_0, vx_0}$ and of $g\mathfrak{q}_{x_0, ux_0}$ intersect. It follows that

$$\mathrm{dist}(x_0, gx_0) \preceq \mathrm{dist}(x_0, vx_0) + \mathrm{dist}(x_0, ux_0). \qquad \square$$

Corollary 7.7. *Let $M$ be a 3-dimensional prime manifold, and let $G$ be its fundamental group.*

*For every word metric on $G$ there exists a constant $K$ such that if two elements $u, v$ are conjugate in $G$ then there exist $g$ conjugating $u, v$ such that*

$$|g| \leq K\big(|u| + |v|\big)^2.$$

*Proof.* Assume first that $M$ is non-geometric, hence decomposable by tori and Klein bottles into hyperbolic and Seifert components. The only case not covered by Corollary 7.5 is when both $u$ and $v$ stabilize a Seifert component, that is, are contained in two groups which are virtually $\mathbb{Z} \times F_n$. In this case one can easily find a conjugator of quadratic length.

When $M$ is a geometric nilmanifold, that is, when $\pi_1(M)$ is 2-step nilpotent the quadratic upper bound for a conjugator length is proved in [JOR10, Proposition 2.1.1].

When $M$ is a geometric solmanifold, the linear upper bound for a conjugator length is proved in [Sal].

The other geometric cases are easy. □

**7.3. Conjugators in mapping class groups.** In what follows $S$ denotes a compact oriented surface of genus $g$ and with $p$ boundary components and $\xi(S) = 3g + p - 3$ denotes the complexity of the surface.

We prove linear control of the shortest conjugator of infinite order pure elements in the mapping class group by providing a new proof of the following result which was established by Masur–Minsky [MM00, Theorem 7.2] in the pseudo-Anosov case and by J. Tao [Tao11, Theorem B] in the reducible case.

THEOREM 7.8. *There exists a constant $C$ depending only on the surface $S$ and the fixed generating set of $\mathcal{MCG}(S)$ such that for every two conjugate pure elements of infinite order $u$ and $v$ there exists $g$ such that $v = gug^{-1}$ and*

$$|g| \leq C\big[|u| + |v|\big].$$

It is worth noting that the mapping class group is not CAT(0), cf. [KL96] or [BH99]. Nevertheless, there exists a natural analogue of the inequality (iv) from the CAT(0) setting which holds here; this will be explained further in the proof below.

*Background.* We will use a quasi-isometric model of a mapping class group, the *marking complex*, $\mathcal{M}(S)$, defined as follows. Its vertices, called *markings*, consist of the following pair of data:

- *base curves*: a multicurve consisting of $\xi(S)$ components, that is, a maximal simplex in $\mathcal{C}(S)$. This collection is denoted base$(\mu)$.
- *transversal curves*: to each curve $\gamma \in \text{base}(\mu)$ is associated an essential curve. Letting $T$ denote the complexity 1 component of $S \setminus \bigcup_{\alpha \in \text{base}(\mu), \alpha \neq \gamma} \alpha$, the transversal curve to $\gamma$ is a curve $t(\gamma) \in \mathcal{C}(T)$ with $\text{dist}_{\mathcal{C}(T)}(\gamma, t(\gamma)) = 1$.

Two vertices $\mu, \nu$ in the marking complex $\mathcal{M}(S)$ are connected by an edge if either of the two conditions hold:

(1) *Twists*: $\mu$ and $\nu$ differ by a Dehn twist along one of the base curves: base$(\mu) =$ base$(\nu)$ and all their transversal curves agree except for $t_\mu(\gamma)$, obtained from $t_\nu(\gamma)$ by twisting once about the curve $\gamma$.
(2) *Flips*: The base curves and transversal curves of $\mu$ and $\nu$ agree except for one pair $(\gamma, t(\gamma)) \in \mu$ for which the corresponding pair in $\nu$ consists of the same pair but with the roles of base and transversal reversed.

Note that after performing one flip the new base curve may now intersect several other transversal curves. Nevertheless by [MM00, Lemma 2.4], there is a finite set of natural ways to resolve this issue which, in turn, yields a uniformly bounded on the diameter of possible markings which can be obtained by flipping the pair $(\gamma, t(\gamma)) \in \mu$; an edge connects each of these possible flips to $\mu$.

THEOREM 7.9 ([MM00]). *The graph $\mathcal{M}(S)$ is locally finite and the mapping class group acts cocompactly and properly discontinuously on it. In particular, the orbit map yields a quasi-isometry from $\mathcal{MCG}(S)$ to $\mathcal{M}(S)$.*

Given a simplex $\Delta$ in the curve complex $\mathcal{C}(S)$, we define $\mathcal{Q}(\Delta)$ to be the set of elements of $\mathcal{M}(S)$ whose bases contain $\Delta$. We recall that there is a coarsely defined closest point projection map from $\mathcal{M}(S)$ to $\mathcal{Q}(\Delta)$ which is coarsely Lipschitz.

*Proof of Theorem 7.8.* We assume that $S$ is a surface with $\xi(S) > 1$, otherwise the mapping class group is hyperbolic and the result is standard. We make use of two cocompact actions of $\mathcal{MCG}(S)$: the above mentioned properly discontinuous action on the marking complex $\mathcal{M}(S)$ and an action far from properly discontinuous on the curve complex $\mathcal{C}(S)$. Neither of the two complexes $\mathcal{C}(S)$ nor $\mathcal{M}(S)$ are CAT(0).

We begin with the case of two conjugate pseudo-Anosov elements. Our goal is to find a natural analogue of the inequality (iv) from the CAT(0) setting. The difficulty is that a pseudo-Anosov element is loxodromic with a translation axis in $\mathcal{C}(S)$, which makes it hard to find an appropriate definition of a projection of an element in $\mathcal{M}(S)$ to it.

Let $k$ be a pseudo-Anosov. According to [Bow08, Theorem 1.4], there exists $m = m(S)$ such that $k^m$ preserves a bi-infinite geodesic $\mathfrak{g}_k$ in $C(S)$. For every curve $\gamma$, denote by $\gamma'$ a closest point to it on $\mathfrak{g}_k$. Let $\widehat{k} = k^m$. A standard hyperbolic geometry argument implies that for every $i \geq 1$

$$\text{(v)} \qquad \operatorname{dist}_{\mathcal{C}(S)}\big(\gamma, \widehat{k}^i\gamma\big) \geq \operatorname{dist}_{\mathcal{C}(S)}\big(\gamma', \widehat{k}^i\gamma'\big) + O(1) \geq i + O(1).$$

Let $\mu$ be an arbitrary element in $\mathcal{M}(S)$ and let $\gamma$ be a closest point to $\pi_{\mathcal{C}(S)}(\mu)$ on $\mathfrak{g}_k$. A hierarchy path $\mathfrak{h}$ joining $\mu$ and $\widehat{k}\mu$ contains two points $\nu, \nu'$ such that:

- the subpath with endpoints $\mu, \nu$ is at $\mathcal{C}(S)$-distance $O(1)$ from any $\mathcal{C}(S)$-geodesic joining $\pi_{\mathcal{C}(S)}(\mu)$ and $\gamma$;
- the subpath with endpoints $\widehat{k}\mu, \nu'$ is at $\mathcal{C}(S)$-distance $O(1)$ from any $\mathcal{C}(S)$-geodesic joining $\pi_{\mathcal{C}(S)}(\widehat{k}\mu)$ and $\widehat{k}\gamma$;
- if the translation length of $\widehat{k}$ along $\mathfrak{g}_k$ is large enough then the subpath with endpoints $\nu, \nu'$ is at $\mathcal{C}(S)$-distance $O(1)$ from $\mathfrak{g}_k$;
- $\operatorname{dist}_{\mathcal{C}(S)}(\nu', \widehat{k}\nu)$ is $O(1)$.

Note that by equation (v) there exists an integer $N = N(S)$ such that for every pseudo-Anosov $k$ the power $k^N$ preserves a bi-infinite geodesic $\mathfrak{g}_k$ in $\mathcal{C}(S)$, and every subpath with endpoints $\nu, \nu'$ defined as above is at $\mathcal{C}(S)$-distance $O(1)$ from $\mathfrak{g}_k$.

The group $\mathcal{MCG}(S)$ acts co-compactly on $\mathcal{M}(S)$, therefore there exists a compact subset $K$ of $\mathcal{M}(S)$ such that $\mathcal{MCG}(S)K = \mathcal{M}(S)$. We pick a basepoint $\mu_0$ in $K$. The map $\mathcal{MCG}(S) \to \mathcal{M}(S), g \mapsto g\mu_0$ is a quasi-isometry, by Theorem 7.9.

Let $u$ and $v$ be an arbitrary pair of pseudo-Anosovs for which there exists $g \in \mathcal{MCG}(S)$ such that $v = gug^{-1}$. Our goal is to prove that for an appropriate choice of $g$, $\operatorname{dist}(\mu_0, g\mu_0)$ is controlled by a linear function of $\operatorname{dist}(\mu_0, u\mu_0) + \operatorname{dist}(\mu_0, v\mu_0)$.

If $\mathfrak{g}_u$ and $\mathfrak{g}_v$ are axes in $\mathcal{C}(S)$ defined as above, then $\mathfrak{g}_v = g\mathfrak{g}_u$. Up to replacing $u$ and $v$ by their $N$th powers, we may assume that both preserve their respective axes $\mathfrak{g}_u$ and $\mathfrak{g}_v$, and every subpath with endpoints $\nu, \nu'$ defined as above is at $\mathcal{C}(S)$-distance $O(1)$ from $\mathfrak{g}_u$, respectively $\mathfrak{g}_v$.

Let $\mathfrak{h}$ be a hierarchy path joining $\mu_0$ and $u\mu_0$ and let $\nu$ and $\nu'$ be two points on it defined as above. Then $g\mathfrak{h}$ is a hierarchy path joining $g\mu_0$ and $gu\mu_0 = vg\mu_0$ and $g\nu$, $g\nu'$ satisfy similar properties for the path $g\mathfrak{h}$, the pseudo-Anosov $v$ and its axis $\mathfrak{g}_v = g\mathfrak{g}_u$.

Now let $\mathfrak{k}$ be a hierarchy path joining $\mu_0$ and $v\mu_0$ and let $\xi$ and $\xi'$ be the two points on it defined as above.

By eventually replacing $g$ with $v^k g$, for an appropriate $k \in \mathbb{Z}$, we may assume that $g\nu$ and $\xi$ are at $\mathcal{C}(S)$ distance at most $t_v + O(1)$, where $t_v$ is the translation length of $v$ along $\mathfrak{g}_v$. There are two cases to discuss. In order to define the necessary parameters, recall the following result.

THEOREM 7.10 (Masur–Minsky; [MM00]). *If $\mu, \nu \in \mathcal{M}(S)$, then there exists a constant $K(S)$, depending only on $S$, such that for each $K > K(S)$ there exists $a \geq 1$ and $b \geq 0$ for which*:

$$\text{(vi)} \qquad \operatorname{dist}_{\mathcal{M}(S)}(\mu, \nu) \approx_{a,b} \sum_{Y \subseteq S} \left\{\!\!\left\{ \operatorname{dist}_{\mathcal{C}(Y)}\big(\pi_Y(\mu), \pi_Y(\nu)\big) \right\}\!\!\right\}_K .$$

In particular this implies that there exists $\kappa > 0$ and $A, B$ depending only on $S$ such that if $\operatorname{dist}_{\mathcal{M}(S)}(\mu, \nu) \geq \kappa \operatorname{dist}_{\mathcal{C}(S)}(\mu, \nu)$ then

$$\text{(vii)} \qquad \operatorname{dist}_{\mathcal{M}(S)}(\mu, \nu) \approx_{A,B} \sum_{Y \subsetneq S} \left\{\!\!\left\{ \operatorname{dist}_{\mathcal{C}(Y)}\big(\pi_Y(\mu), \pi_Y(\nu)\big) \right\}\!\!\right\}_K .$$

In the formulas above we use the following notation. For two numbers $d \geq 0$ and $K \geq 0$, $\{\!\{d\}\!\}_K$ is equal to $d$ if $d \geq K$, and it is zero otherwise.

The subsurfaces that appear in (vii) for a given pair $\mu, \nu$ and a given constant $K > K(S)$ are called $K$-*large domains* of that pair. The proper subsurfaces are called $K$-*large proper domains*. We omit $K$ when irrelevant.

*Case 1.* Assume that $\operatorname{dist}_{\mathcal{M}(S)}(\mu_0, g\mu_0) \leq \kappa \operatorname{dist}_{\mathcal{C}(S)}(\mu_0, g\mu_0)$.

Note that $\mathrm{dist}_{\mathcal{C}(S)}(\mu_0, g\mu_0) \leq \mathrm{dist}_{\mathcal{C}(S)}(\mu_0, \xi) + \mathrm{dist}_{\mathcal{C}(S)}(\mu_0, \nu) + t_v + O(1) \leq 2\,\mathrm{dist}_{\mathcal{C}(S)}(\mu_0, v\mu_0) + \mathrm{dist}_{\mathcal{C}(S)}(\mu_0, u\mu_0) + O(1)$, which implies that we have $\mathrm{dist}_{\mathcal{M}(S)}(\mu_0, g\mu_0) \leq_{A',B'} \mathrm{dist}(\mu_0, u\mu_0) + (\mu_0, v\mu_0)$, where $A', B'$ depend on $\kappa, a, b$ and $K$ from Theorem 7.10.

*Case 2.* Assume that $\mathrm{dist}_{\mathcal{M}(S)}(\mu_0, g\mu_0) \geq \kappa\,\mathrm{dist}_{\mathcal{C}(S)}(\mu_0, g\mu_0)$.

This together with equation (vii) then implies:

$$\mathrm{dist}_{\mathcal{M}(S)}(\mu_0, g\mu_0) \approx_{A,B} \sum_{Y \subsetneq S} \left\{\!\!\left\{ \mathrm{dist}_{\mathcal{C}(Y)}\big(\pi_Y(\mu_0), \pi_Y(g\mu_0)\big) \right\}\!\!\right\}_K . \tag{viii}$$

Recall that the point nearest to $\pi_{\mathcal{C}(S)}(\mu_0)$ on the axis $\mathfrak{g}_v$ is at $\mathcal{C}(S)$-distance $O(1)$ from $\xi$ (actually, this may not be a point, but since $\mathcal{C}(S)$ is hyperbolic the set of closest points form a bounded diameter; hence we abuse notation slightly, as this set is coarsely a point). Likewise the point nearest to $\pi_{\mathcal{C}(S)}(g\mu_0)$ on $\mathfrak{g}_v$ is at $\mathcal{C}(S)$-distance $O(1)$ from $g\nu$ and at distance $t_v + O(1)$ from $\xi$. Every proper subsurface $Y$ appearing in the sum (viii) has the property that every hierarchy path joining $\mu_0$ to $g\mu_0$ intersects $Q(\partial Y)$. Hence, $\partial Y$ is at $\mathcal{C}(S)$-distance $O(1)$ from the union of two geodesics in $\mathcal{C}(S)$ joining $\pi_{\mathcal{C}(S)}(\mu_0)$ respectively, $\pi_{\mathcal{C}(S)}(g\mu_0)$ to their nearest points on $\mathfrak{g}_v$ with the arc of $\mathfrak{g}_v$ with endpoints these two nearest points. It follows that a nearest point to $\pi_{\mathcal{C}(S)}(\partial Y)$ on the axis $\mathfrak{g}_v$ is at $\mathcal{C}(S)$-distance at most $t_v + O(1)$ from $\xi$.

The analogue of equation (viii) is also satisfied by $v\mu_0$ and $vg\mu_0 = gu\mu_0$. In particular for every proper subsurface $Y'$ appearing in that formula, a nearest point to $\pi_{\mathcal{C}(S)}(\partial Y')$ on the axis $\mathfrak{g}_v$ is at $\mathcal{C}(S)$-distance at most $t_v + O(1)$ from $v\xi$. Then by replacing both $u$ and $v$ with their $J$th power, for some $J = J(S)$ and arguing with the corresponding hierarchy paths $\mathfrak{h}$ joining $\mu_0$, $u^J(\mu_0)$, respectively $\mu_0$, $v^J(\mu_0)$, we may assume that the pairs $\mu_0$, $g\mu_0$ and respectively $v\mu_0$, $vg\mu_0$ have no large proper domain in common.

It follows that for every large proper domain $Y$ of the pair $\mu_0, g\mu_0$,

$$\begin{aligned}\mathrm{dist}_{\mathcal{C}(Y)}\big(\pi_Y(\mu_0), \pi_Y(g\mu_0)\big) \leq {} & \mathrm{dist}_{\mathcal{C}(Y)}\big(\pi_Y(\mu_0), \pi_Y(v\mu_0)\big) \\ & + \mathrm{dist}_{\mathcal{C}(Y)}\big(\pi_Y(g\mu_0), \pi_Y(vg\mu_0)\big) + K.\end{aligned}$$

In particular for $K > K(S)$, where $K(S)$ is the constant from Theorem 7.10, if we consider $Y$ a $3K$-large proper domain for the pair $\mu_0, g\mu_0$, it must be a $K$-large proper domain either for $\mu_0, v\mu_0$ or for $g\mu_0, vg\mu_0$ or for both pairs. We may then write that

$$\begin{aligned}&\sum_{Y \subsetneq S} \left\{\!\!\left\{ \mathrm{dist}_{\mathcal{C}(Y)}\big(\pi_Y(\mu_0), \pi_Y(g\mu_0)\big) \right\}\!\!\right\}_{3K} \\ &\quad \leq 3 \sum_{Y \subsetneq S} \left\{\!\!\left\{ \mathrm{dist}_{\mathcal{C}(Y)}\big(\pi_Y(\mu_0), \pi_Y(v\mu_0)\big) \right\}\!\!\right\}_K \\ &\qquad + 3 \sum_{Y \subsetneq S} \left\{\!\!\left\{ \mathrm{dist}_{\mathcal{C}(Y)}\big(\pi_Y(g\mu_0), \pi_Y(gu\mu_0)\big) \right\}\!\!\right\}_K\end{aligned}$$

whence

$$\mathrm{dist}_{\mathcal{M}(S)}(\mu_0, g\mu_0) \preceq_{A',B'} \mathrm{dist}_{\mathcal{M}(S)}(\mu_0, v\mu_0) + \mathrm{dist}_{\mathcal{M}(S)}(\mu_0, u\mu_0).$$

This completes the proof of Theorem 7.8 for pairs of pseudo-Anosov elements. We now proceed to the full proof of Theorem 7.8.

Let $u, v \in \mathcal{M}$.

LEMMA 7.11. *Let $\nu$ and $\rho$ be two points in $\mathcal{M}(S)$, let $\Delta$ be a multicurve, and let $\nu', \rho'$ be respective nearest point projections of $\nu, \rho$ on $\mathcal{Q}(\Delta)$. Assume there exist $U_1, \ldots, U_k$ subsurfaces such that $\Delta = \partial U_1 \cup \cdots \cup \partial U_k$, and $\mathrm{dist}_{\mathcal{C}(U_i)}(\nu, \rho) > M$ for every $i = 1, \ldots, k$, where $M = M(S)$ is a large enough constant.*

*Then for every $\mathfrak{h}_1$, $\mathfrak{h}_2$ and $\mathfrak{h}_3$ hierarchy paths joining $\nu, \nu'$ respectively $\nu', \rho'$ and $\rho', \rho$, the path $\mathfrak{h}_1 \sqcup \mathfrak{h}_2 \sqcup \mathfrak{h}_3$ has length $\approx_{a,b} \mathrm{dist}_{\mathcal{M}(S)}(\nu, \rho)$, where $a, b$ depend only on the topological type of $\Delta$.*

*Proof.* This follows by a limiting argument from [BDS, Lemma 4.27] and [BDS, Theorem 4.16]. □

Let $\Delta_u$ be a multicurve such that if $U^1, \ldots, U^m$ are the connected components of $S \setminus \Delta_u$ and the annuli with core curve in $\Delta_u$ then $u$ is a pseudo-Anosov on $U^1, \ldots, U^k$ (Dehn twists are assumed to be pseudo-Anosovs on annuli) and the identity map on $U^{k+1}, \ldots, U^m$, and $\Delta_u = \partial U^1 \cup \cdots \cup \partial U^k$ (the latter condition may be achieved by deleting the boundary between two components on which $u$ acts as identity). Similarly, for $v$ we consider the multicurve $\Delta_v$ and $V^1, \ldots, V^m$.

Then $g\Delta_u = \Delta_v$ and $gU_i = V_i$, up to reordering $V^1, \ldots, V^m$.

Let $\nu$ and $\xi$ be nearest point projections of $\mu_0$ onto $\mathcal{Q}(\Delta_u)$ and respectively, $\mathcal{Q}(\Delta_v)$. By eventually replacing $u, v$ with large enough powers, we may assume that Lemma 7.11 applies to the pairs $\mu_0, u\mu_0$ and $\mu_0, v\mu_0$, respectively.

Like for pseudo-Anosovs, we have two cases.

*Case 1.* Assume that

$$\mathrm{dist}_{\mathcal{M}(S)}(\mu_0, g\mu_0) \approx_{A,B} \sum_{Y \subseteq S, Y \pitchfork \Delta_v} \left\{\!\!\left\{ \mathrm{dist}_{\mathcal{C}(Y)}\big(\pi_Y(\mu_0), \pi_Y(g\mu_0)\big) \right\}\!\!\right\}_K .$$

Then the same relation is true for $v\mu_0, vg\mu_0$. By eventually replacing $v$ by a power of itself, we may assume that the set of proper domains appearing in the sum above is disjoint from the corresponding set of proper domains for $v\mu_0, vg\mu_0$. It follows that all the large proper domains for the pair $\mu_0, g\mu_0$ are large proper domains either for $\mu_0, v\mu_0$ or for $g\mu_0, vg\mu_0 = gu\mu_0$.

We discuss the case of the whole surface $S$ separately. Consider a tight geodesic $\mathfrak{g}_u$ in $\mathcal{C}(S)$ joining $\pi_{\mathcal{C}(S)}(\mu_0)$ to $\Delta_u$. We state that by replacing $u$ with a large enough power we may ensure that points on $u\mathfrak{g}_u$ at $\delta$-distance from $\mathfrak{g}_u$ are at distance at most $D$ from $\Delta_u$, for some $D > 0$, where $\delta > 0$ is the hyperbolicity constant of $\mathcal{C}(S)$. Indeed assume that there exists a point

$a$ on $\mathfrak{g}_u$ at distance at least $D$ from $\Delta_u$ such that $B(a,\delta)$ intersects $u^i\mathfrak{g}_u$ for $i\in\{1,2,\dots,N\}$. This in particular implies that $\mathrm{dist}_{\mathcal{C}(S)}(a,u^ia)\le 2\delta$.

Let $b$ be the point on $\mathfrak{g}_u$ at distance $\frac{D}{2}$ from $\Delta_u$. By the above, every $u^ib$ is in $B(b,2\delta)$. On the other hand, by [Bow08, Theorems 1.1 and 1.2] if $D\ge D_0(S,\delta)$ then there exists $m=m(S,\delta)$ such that $B(b,2\delta)$ contains at most $m$ points from the union of tight geodesics $\mathfrak{g}_u\cup\bigcup_{i=1}^N u^i\mathfrak{g}_u$. It follows in particular that if $N=m+1$ then there exist $i<j$ with $i,j\in\{1,2,\dots,N\}$ such that $u^ib=u^jb$, hence $u^{j-i}b=b$. But, for $D$ large enough $b$ together with any curve from $\Delta_u$ fills the surface, hence it cannot be fixed by a power of $u$. We obtained a contradiction. Thus we conclude that there exists $k\le N=N(S,\delta)$ such that the intersection of the $\delta$-neighborhood of $\mathfrak{g}_u$ with the $\delta$-neighborhood of $u^k\mathfrak{g}_u$ is contained in the $D$-neighborhood of $\Delta_u$.

Likewise we argue that given a tight geodesic $\mathfrak{g}_v$ in $\mathcal{C}(S)$ joining $\pi_{\mathcal{C}(S)}(\mu_0)$ to $\Delta_v$ there exists $r\le N$ such that the intersection of the $\delta$-neighborhood of $\mathfrak{g}_v$ with the $\delta$-neighborhood of $v^r\mathfrak{g}_v$ is contained in the $D$-neighborhood of $\Delta_v$. We deduce that $\mathrm{dist}_{\mathcal{C}(S)}(\mu_0,\Delta_u)\le \mathrm{dist}_{\mathcal{C}(S)}(\mu_0,u^k\mu_0)+O(1)\le k\,\mathrm{dist}_{\mathcal{C}(S)}(\mu_0,u\mu_0)+O(1)$ and that $\mathrm{dist}_{\mathcal{C}(S)}(\mu_0,\Delta_v)\le r\,\mathrm{dist}_{\mathcal{C}(S)}(\mu_0,v\mu_0)+O(1)$. Then $\mathrm{dist}_{\mathcal{C}(S)}(\mu_0,g\mu_0)\le \mathrm{dist}_{\mathcal{C}(S)}(\mu_0,\Delta_v)+\mathrm{dist}_{\mathcal{C}(S)}(\Delta_v,g\mu_0)\le N[\mathrm{dist}_{\mathcal{C}(S)}(\mu_0,v\mu_0)+\mathrm{dist}_{\mathcal{C}(S)}(\mu_0,u\mu_0)]+O(1)$.

*Case 2.* Assume that

$$\mathrm{dist}_{\mathcal{M}(S)}(\mu_0,g\mu_0)\approx_{A,B}\sum_{Y\subseteq S,\,Y\not\pitchfork\Delta_v}\left\{\!\!\left\{\mathrm{dist}_{\mathcal{C}(Y)}\big(\pi_Y(\mu_0),\pi_Y(g\mu_0)\big)\right\}\!\!\right\}_K.$$

In other words $\mathrm{dist}_{\mathcal{M}(S)}(\mu_0,g\mu_0)\approx_{A',B'}\mathrm{dist}_{\mathcal{M}(S)}(\xi,g\nu)$. Note that it suffices to bound $\mathrm{dist}_{\mathcal{M}(S)}(\xi,g\nu)$ by a multiple of $\mathrm{dist}_{\mathcal{M}(S)}(\xi,v\xi)+\mathrm{dist}_{\mathcal{M}(S)}(g\nu,vg\nu)$. Since we are only considering pure elements, it follows that $\xi$ and $g\nu$ have projections on $\mathcal{M}(U^j), k+1\le j\le m$, at bounded distance. Thus, any large domain $Y$ for $\xi,g\nu$ must satisfy $Y\subseteq U^j$ for some $j$ in $\{1,2,\dots,k\}$.

For every $j\in\{1,2,\dots,k\}$ recall that $v$ restricted to $V^j$ coincides with a pseudo-Anosov $v_j$. We use the same notation $v_j$ to denote the mapping class that acts as $v_j$ on $V^j$ and as identity on $S\setminus V^j$. A hierarchy path joining $\xi$ to $g\nu$ projects onto a quasi-geodesic $\mathfrak{q}_j$ in $\mathcal{C}(V^j)$ containing in a tubular neighborhood of radius $O(1)$ all the multicurves $\partial Y$ where $Y\subsetneq V^j$ is a large domain for $\xi$ and $g\nu$. By eventually pre-composing $g$ with a power of $v_j$ (hence with an element in the centralizer of $v$), we may assume that the sub-arc of $\mathfrak{q}_j$ contained in a $O(1)$-tubular neighborhood of the translation axis of $v_j$ has length $\preceq t_{v_j}$, where $t_{v_j}$ is the translation length of $v_j$. Then, by eventually replacing $v$ with a large enough power, we may assume that the set of large domains for $\xi,g\nu$ that are proper sub-surfaces of $V^j$ has nothing in common with the set of large domains for $v\xi,vg\nu$ that are proper sub-surfaces of $V^j$. Hence they are all large domains either for $\xi,v\xi$ or for $g\nu,vg\nu$.

If $U^j$ is a large domain itself for $\xi,g\nu$, then by arguing as in the pseudo-Anosov case (and noting that the copy of $\mathbb{Z}^k$ generated by the pseudo-Anosov components $v_j$ is in the centralizer of $v$) we may prove that, by eventually post-composing $g$ with an element in the centralizer of $v$, we may ensure that $\mathrm{dist}_{C(U^j)}(\xi, g\nu) \preceq \mathrm{dist}_{C(U^j)}(\xi, v\xi) + \mathrm{dist}_{C(U^j)}(g\nu, vg\nu)$.

**Acknowledgments.** We thank L. Mosher, P. Papasoglu, A. Sale, A. Sisto, and J. Tao for useful conversations and corrections. The first author would also like to thank the law firms of Orrick, Herrington & Sutcliffe and Disability Rights Advocates, and in particular René Kathawala and Elina Druker for working to stop the New York City Department of Education from discriminating against children with disabilities. We thank the referee for a number of comments which helped improve the exposition.

## References

[AK] Y. Algom-Kfir, *Strongly contracting geodesics in outer space*, Geom. Topol. **15** (2011), 2181–2233, available at arXiv:0812.1555. MR 2862155

[Bal85] W. Ballmann, *Nonpositively curved manifolds of higher rank*, Ann. of Math. (2) **122** (1985), no. 3, 597–609. MR 0819559

[Bal95] W. Ballmann, *Lectures on spaces of nonpositive curvature*, DMV Seminar, vol. 25, Birkhäuser, Basel, 1995. With an appendix by Misha Brin. MR 1377265

[BC] J. Behrstock and R. Charney, *Divergence and quasi-morphisms of right-angled Artin groups*, Math. Ann. **352** (2012), 543–595. MR 2874959

[BDM09] J. Behrstock, C. Druţu and L. Mosher, *Thick metric spaces, relative hyperbolicity, and quasi-isometric rigidity*, Math. Ann. **344** (2009), 543–595. MR 2501302

[BDS] J. Behrstock, C. Druţu and M. Sapir, *Median structures on asymptotic cones and homomorphisms into mapping class groups*, Proc. Lond. Math. Soc. (3) **102** (2011), no. 3, 503–554. MR 2783135

[Beh06] J. A. Behrstock, *Asymptotic geometry of the mapping class group and Teichmüller space*, Geom. Topol. **10** (2006), 1523–1578. MR 2255505

[BF] M. Bestvina and K. Fujiwara, *A characterization of higher rank symmetric spaces via bounded cohomology*, Geom. Funct. Anal. **19** (2009), 11–40. MR 2507218

[BH99] M. Bridson and A. Haefliger, *Metric spaces of non-positive curvature*, Springer, Berlin, 1999. MR 1744486

[BM08] J. Brock and H. Masur, *Coarse and synthetic Weil–Petersson geometry: Quasi-flats, geodesics, and relative hyperbolicity*, Geom. Topol. **12** (2008), no. 3, 2453–2495. MR 2443970

[BN08] J. Behrstock and W. Neumann, *Quasi-isometric classification of graph manifold groups*, Duke Math. J. **141** (2008), no. 2, 217–240. MR 2376814

[Bow08] B. Bowditch, *Tight geodesics in the curve complex*, Invent. Math. **171** (2008), no. 2, 281–300. MR 2367021

[BS87] K. Burns and R. Spatzier, *Manifolds of nonpositive curvature and their buildings*, Publ. Math. Inst. Hautes Études Sci. **65** (1987), 35–59. MR 0908215

[DMS10] C. Druţu, S. Mozes and M. Sapir, *Divergence in lattices in semisimple Lie groups and graphs of groups*, Trans. Amer. Math. Soc. **362** (2010), no. 5, 2451–2505. MR 2584607

[DR09] M. Duchin and K. Rafi, *Divergence of geodesics in Teichmüller space and the mapping class group*, Geom. Funct. Anal. **19** (2009), 722–742. MR 2563768

[DS05] C. Druţu and M. Sapir, *Tree-graded spaces and asymptotic cones of groups*, Topology **44** (2005), 959–1058. With an appendix by D. Osin and M. Sapir. MR 2153979

[Ger94a] S. M. Gersten, *Divergence in 3-manifold groups*, Geom. Funct. Anal. **4** (1994), no. 6, 633–647. MR 1302334

[Ger94b] S. M. Gersten, *Quadratic divergence of geodesics in* CAT(0)*-spaces*, Geom. Funct. Anal. **4** (1994), no. 1, 37–51. MR 1254309

[Gro93] M. Gromov, *Asymptotic invariants of infinite groups*, Geometric group theory, vol. 2 (Sussex, 1991) (G. Niblo and M. Roller, eds.), LMS Lecture Notes, vol. 182, Cambridge Univ. Press, Cambridge, 1993, pp. 1–295. MR 1253544

[Ham09] U. Hamenstädt, *Lines of minima in outer space*, preprint, 2009, available at arXiv:math/0911.3620.

[JOR10] R. Ji, C. Ogle and B. Ramsey, *Relatively hyperbolic groups, rapid decay algebras and a generalization of the Bass conjecture*, J. Noncommut. Geom. **4** (2010), 83–124. MR 2575390

[KL96] M. Kapovich and B. Leeb, *Actions of discrete groups on nonpositively curved spaces*, Math. Ann. **306** (1996), no. 2, 341–352. MR 1411351

[KL97] M. Kapovich and B. Leeb, *Quasi-isometries preserve the geometric decomposition of Haken manifolds*, Invent. Math. **128** (1997), no. 2, 393–416. MR 1440310

[KL98] M. Kapovich and B. Leeb, 3*-manifold groups and nonpositive curvature*, Geom. Funct. Anal. **8** (1998), no. 5, 841–852. MR 1650098

[Mac] N. Macura, CAT(0) *spaces with polynomial divergence of geodesics*, preprint, 2011.

[MM00] H. Masur and Y. Minsky, *Geometry of the complex of curves II: Hierarchical structure*, Geom. Funct. Anal. **10** (2000), no. 4, 902–974. MR 1791145

[Min96] Y. Minsky, *Quasi-projections in Teichmüller space*, J. Reine Angew. Math. **473** (1996), 121–136. MR 1390685

[MSW05] L. Mosher, M. Sageev and K. Whyte, *Quasi-actions on trees II: Finite depth Bass–Serre trees*, Mem. Amer. Math. Soc. **214** (2011), vi+105. MR 2867450

[OOS05] A. Y. Olshanskii, D. V. Osin and M. V. Sapir, *Lacunary hyperbolic groups*, Geom. Topol. **13** (2009), no. 4, 2051–2140. With an appendix by M. Kapovich and B. Kleiner. MR 2507115

[Osi06] D. Osin, *Relatively hyperbolic groups: Intrinsic geometry, algebraic properties, and algorithmic problems*, Mem. Amer. Math. Soc. **179** (2006), no. 843, vi+100. MR 2182268

[Pau01] S. Pauls, *The large scale geometry in nilpotent Lie groups*, Comm. Anal. Geom. **9** (2001), no. 5, 951–982. MR 1883722

[Pre06] J.-P. Préaux, *Conjugacy problem in groups of oriented geometrizable 3-manifolds*, Topology **45** (2006), no. 1, 171–208. MR 2170497

[Sal] A. Sale, *Short conjugators in solvable groups*, preprint, 2011, available at arXiv:1112.2721.

[Ser80] J. P. Serre, *Trees*, Springer, New York, 1980. MR 0607504

[Sul] H. Sultan, *The asymptotic cones of Teichmüller space: Thickness and divergence*, to appear in Algebr. Geom. Topol.

[Tao11] J. Tao, *Linearly bounded conjugator property for mapping class groups*, Geom. Funct. Anal. **23** (2013), 415–466. MR 3037904

Jason Behrstock, Lehman College and The Graduate Center, CUNY, New York, New York, USA

*E-mail address*: jason.behrstock@lehman.cuny.edu

Cornelia Druţu, Mathematical Institute, 24-29 St Giles, Oxford OX1 3LB, UK

*E-mail address*: cornelia.drutu@maths.ox.ac.uk