Abstract. We prove the generalized Margulis lemma with a uniform index bound on an Alexandrov $n$-space $X$ with curvature bounded below, i.e., small loops at $p \in X$ generate a subgroup of the fundamental group of unit ball $B_1(p)$ that contains a nilpotent subgroup of index $\leq w(n)$, where $w(n)$ is a constant depending only on the dimension $n$. The proof is based on the main ideas of V. Kapovitch, A. Petrunin, and W. Tuschmann, and the following results:

(1) We prove that any regular almost Lipschitz submersion constructed by Yamaguchi on a collapsed Alexandrov space with curvature bounded below is a Hurewicz fibration. We also prove that such fibration is uniquely determined up to a homotopy equivalence.

(2) We give a detailed proof on the gradient push, improving the universal pushing time bound given by V. Kapovitch, A. Petrunin, and W. Tuschmann, and justifying in a specific way that the gradient push between regular points can always keep away from extremal subsets.

1. Introduction

In this paper we prove the Margulis lemma on Alexandrov spaces with curvature bounded below. A group $\Gamma$ is called $w$-nilpotent if there is a nilpotent subgroup $N < \Gamma$ whose index $[\Gamma : N] \leq w$. Let $B_r(p)$ denote a metric ball centered at $p$ of radius $r$.

**Theorem 1.1** (Generalized Margulis Lemma). There are $\epsilon(n), w(n) > 0$ such that for any Alexandrov space $X$ with curvature $\geq -1$ and any point $p \in X$, the subgroup $\Gamma_p(p; \epsilon)$ of fundamental group $\pi_1(B_1(p), p)$ generated by loops at $p$ lying in $B_\epsilon(p)$ with $0 < \epsilon \leq \epsilon(n)$ is $w(n)$-nilpotent.

The original Margulis lemma is also called Margulis-Heintze’s theorem, which was proved by Margulis (cf. [11]), and also independently discovered by Heintze [13] on manifolds of $-1 \leq K \leq 0$. Since then, it has been one of the fundamental facts in Riemannian geometry which has many applications, e.g., Gromov’s almost flat theorem [11], finiteness of closed negatively pinched manifolds [10] of bounded volume, and more recently the almost rigidity of maximal volume entropy [19] for manifolds of lower bounded Ricci curvature to be hyperbolic.
For manifolds with sectional curvature $K \geq -1$, it was proved by Fukaya-Yamaguchi [9] that $\Gamma(\epsilon)$ is almost nilpotent without a uniform bound on the index. For Alexandrov spaces, the earlier version of Theorem 1.1 was proved by Yamaguchi in [32], also without a uniform bound on the index of the nilpotent subgroup, where the proof was based on the Lipschitz submersion Theorem 1.3 and arguments in [9]. Later a global version of Theorem 1.1 for manifolds of almost nonnegative curvature was proved by Kapovitch-Petrunin-Tuschmann [15], where $\Gamma(\epsilon)$ admits a nilpotent subgroup with uniformly bounded index. Theorem 1.1 also follows from the main ideas of Kapovitch-Petrunin-Tuschmann [15].

Gromov conjectured that the Margulis lemma with a universal bounded index holds for manifolds of lower bounded Ricci curvature. A breakthrough on this conjecture was made by Cheeger-Colding [6], and it has been finally confirmed recently by Kapovitch-Wilking [16].

We point it out that the uniform index bound is very important to some geometric applications, for example, in Gromov’s almost flat theorem [11], the uniform index bound corresponds to the holonomy gap which is crucial in Gromov’s and Ruh’s proof (see [11], [26], [5]). The uniformly index bound is also crucial for the almost rigidity of maximal volume entropy [19] in deriving that the connectedness component of a Gromov-Hausdorff limit group of deck-transformations is a nilpotent Lie group.

Remark 1.2. More generally, one may further consider a metric space $X$ of $K$-bounded packing, i.e., there is $K > 0$ such that every ball of radius 4 in $X$ can be covered by at most $K$ balls of radius 1. In [12, §5.F] Gromov proposed a question whether a discrete isometric subgroup $\Gamma$ acting on a metric space with $K$-bounded packing is virtually nilpotent, if $\Gamma$ is generated by finite elements whose displacement at one point $< \epsilon(K)$? It has been answered affirmatively by [3] recently. However, the uniform index bound as in Theorem 1.1 is beyond their approach (see [3, Section 11]).

Our proof relies on Theorem 1.3 and Theorem 1.6 below.

For small $0 < \delta < \delta(n, \kappa)$, the $\delta$-strained radius [32] at a point $p$ in an $n$-dimensional Alexandrov space $Y$ of curv $\geq \kappa$ is defined to be

$$r_{\delta,\text{str}}(p) = \sup \{r \mid \text{there exists an } (n, \delta)\text{-strainer at } p \text{ of length } r \}.$$ 

Let $r_{\delta,\text{str}}(Y) = \inf \{r_{\delta,\text{str}}(p) : p \in Y \}$. Let $\varphi(\delta, \epsilon|n)$ denote a positive function depending on $n$, $\delta$ and $\epsilon$ satisfying $\varphi(\delta, \epsilon) \to 0$ as $\delta, \epsilon \to 0$. A map $f : X \to Y$ between Alexandrov spaces is an $\epsilon$-almost Lipschitz submersion [32] if

(i) $f$ is an $\epsilon$-Gromov-Hausdorff approximation (GHA for simplicity), i.e., for any $p, q \in X$, $|f(p)f(q)| - |pq| \leq \epsilon$ and $f(X)$ is $\epsilon$-dense in $Y$, where $|pq| = d(p, q)$ denote the distance between two points $p, q$; and
(ii) for any \( p, q \in X \),
\[
\left| \frac{|f(p)f(q)|}{|pq|} - \sin \theta \right| < \epsilon,
\]
where \( \theta(p, q) \) is the infimum of \( \angle qpx \) when \( x \) runs over \( f^{-1}(f(p)) \).

We call an \( \epsilon \)-almost Lipschitz submersion is regular, if in addition,

(iii) for any \( y, z \in Y \), there are points \( p \in f^{-1}(y), q \in f^{-1}(z) \) such that
\[
|\theta(p, q) - \frac{\pi}{2}| \leq \epsilon.
\]

**Theorem 1.3** (Lipschitz submersion & fibration). *For any dimension \( n \) and positive number \( \mu_0 \), there exist positive numbers \( \delta(n) \) and \( \epsilon(n, \mu_0) \) such that for any \( m \)-dimensional Alexandrov space \( X \) with \( \text{curv} \geq -1 \) and any \( n \)-dimensional Alexandrov space \( Y \) with \( \text{curv} \geq -1 \), if

(1.3.1) the \( \delta \)-strained radius of \( Y \), \( r_{\delta, \text{str}}(Y) \geq \mu_0 \) with \( 0 < \delta < \delta(n) \), and

(1.3.2) the Gromov-Hausdorff distance \( d_{GH}(X, Y) \leq \epsilon < \epsilon(n, \mu_0) \),

then there exists a regular \( \varepsilon(\delta, \epsilon(n)) \)-almost Lipschitz submersion \( f : X \to Y \) that is a Hurewicz fibration.*

**Remark 1.4.** If in addition, every \( f \)-fiber is a topological manifold without boundary of co-dimension \( n \), then \( f \) is a locally trivial fibration; see [25].

We also prove that the fibration in Theorem 1.3 is uniquely determined in the homotopic sense; see Theorem 2.8.

Theorem 1.3 can be traced back to the fibration theorem [8], [31], [21] for manifolds, which has played a fundamental role in the study of collapsed manifolds. The existence of regular almost Lipschitz in Theorem 1.3 is due to Yamaguchi [32], where he conjectured that it should be a locally trivial fibration. Here we partly verify his conjecture.

**Remark 1.5.** A direct corollary of Theorem 1.3 is a long exact sequence arising from the fibration:

\[
\cdots \to \pi_l(F, x) \to \pi_l(X, x) \xrightarrow{f_*} \pi_l(Y, f(x)) \to \pi_{l-1}(F, x) \to \cdots \to \pi_1(Y, f(x)) \to 0.
\]

In [22] Perelman concluded the same long exact sequence under a much weaker situation, that is, when a sequence of Alexandrov spaces \( X_i \) with \( \text{curv} \geq \kappa \) collapses to a limit space \( Y \), if \( Y \) contains no proper extremal subsets, then (1.5.1) holds for \( i \) large and a regular fiber \( F \) (i.e., the fiber of a lifting map to \( X_i \) of regular admissible maps locally defined on \( Y \) to \( \mathbb{R}^n \), see [22]).

By the proof of Theorem 1.1, both the homotopy fiber in Theorem 1.3 and Perelman’s regular fiber admit a \( w(m-n) \)-nilpotent fundamental group, where \( w \) depends on the codimension.

The gradient push developed by Kapovitch-Petrunin-Tuschmann [15] is important for us to deduce the uniform index bound in Theorem 1.1, as what happened for almost nonnegatively curved manifolds in [15].
Theorem 1.6 (Gradient push, [15, Lemma 2.5.1]). There are $\delta(n), T(n) > 0$ such that if the metric ball $B_1(p_0)$ centered at $p_0$ of radius 1 is relative compact in an Alexandrov $n$-space $X$ with curvature $\geq -1$, then there are regular points $\{a_j, b_j\}_{j=1}^n$ and $q_0$ in $B_1(p_0)$ such that $\{a_j, b_j\}_{j=1}^n$ is a $(n, \delta)$-strainer at $q_0$ and any point $q$ in $B_{\delta(a_n b_n)}(q_0)$ can be pushed successively by the gradient flows of $\frac{1}{2}\text{dist}^2_{p_0}, \frac{1}{2}\text{dist}^2_{a_j}, \frac{1}{2}\text{dist}^2_{b_j}$ ($j = 1, \ldots, n$) to any point $p \in B_{\frac{1}{2}}(q)$ in total time $\leq T(n)$.

Compared to the case of manifolds, a crucial difference on an Alexandrov $n$-space $X$ is that, there may be proper extremal subsets and no gradient curves can get out of them. When pushing a loop at a regular point to another regular point, it is a subtle point whether the successive gradient curve at base point do not pass any proper extremal subset in $X$.

Since it is hard for us by following [15] to check this directly, in the appendix we give a detailed proof of Theorem 1.6, by constructing a specific gradient pushing broken line, which consists of $k$-regular (i.e., the tangent cone $T_pX$ at least splits off $\mathbb{R}^k$) or $(n, \delta)$-strained points when $a_j, b_j$ and the ending point $p$ are $k$-regular. In particular, the gradient push between regular points can always keep away from extremal subsets. We also sharpen the universal time bound $T(n)$ to $n^2\delta^{-1}$, improving the universal time bound $\delta^{-n^2}$ in [15]. This provides a detailed justification for the gradient push in proving the Margulis lemma on an Alexandrov space.

Remark 1.7. Kapovitch-Wilking [16] developed a replacement (see the zooming in property and rescaling theorem in [16]) of Yamaguchi’s fibration theorem [31] and gradient push [15] in proving the Margulis lemma for manifolds with lower bounded Ricci curvature.

Note that it is necessary to change base points many times when the rescaling theorem is applied. Since a fixed base point is chosen to be valid for our case at every scale, the proof of Theorem 1.1 is more direct than [16].

Now let us briefly explain ideas of the proofs. According to [15], a finite generated group $G$ is $w$-nilpotent, if it admits a filtration $G_1 = G \triangleright G_2 \triangleright \cdots \triangleright G_l = \{e\}$, where $l \leq n$, each $G_i < G_1$, $G_i/G_{i+1}$ is $c$-abelian, and the conjugate action of $G_1$ on $G_i/G_{i+1}$, namely $\rho_i : G_1 \to \text{Out}(G_i/G_{i+1})$, has a finite image, whose order is bounded by $C$. By a contradicting argument and an iterated blowing-up process, we will prove that around any $p \in X$, there is a nearby point $q$ at which the local fundamental group corresponding to different collapsing scales (see Definition 3.4) has a filtration as above. Then Theorem 1.1 follows from a compact packing argument as in [16]. Theorem 1.3 is used in proving $G_{i+1} < G_i$ (for an alternative proof, see [9] or [32]). The normal property $G_i < G_1$ and a uniform bound on $\#\rho_i(G_1)$ follow from the universal time bound in Theorem 1.6.
According to Ferry’s result ([7], see also Theorem 2.2), the homotopy lifting property holds for the map in Theorem 1.3 if there are controlled homotopy equivalences between nearby fibers (called strong regular, see Section 2.2) and all fibers are abstract neighborhood retracts. As a generalization of the tubular neighborhood of fibers and horizontal curves of an $\epsilon$-Riemannian submersion, a neighborhood retraction $\varphi_p$ to a fiber $f^{-1}(p)$ of a LcL was constructed in [25] (see also Proposition 2.7, Section 2.4), which is defined via iterated gradient deformations of distance functions. By this neighborhood retraction associated to every fiber, we are able to define controlled homotopy equivalences between nearby fibers and prove the fiber is locally contractible.

The remaining of the paper is divided into three parts. In Section 2, we will review some topological results and prove Theorem 1.3. In Sections 3, 4 and 5 we prove Theorem 1.1. In the Appendix we give an elementary construction of the gradient push in Theorem 1.6 with a sharpened time estimate improving that in [15].

**Acknowledgements.** The first author would like to thank Xiaochun Rong and Hao Fang for helpful discussions, and thank the University of Iowa for hospitality and support during a visit in which a part of the work was completed. The second author would like to thank Yin Jiang and Liman Chen for helpful suggestions. We are grateful to Fuquan Fang for pointing out the nilpotency result in [3] to us. This work is supported partially by National Natural Science Foundation of China [11871349], [11821101], by research funds of Beijing Municipal Education Commission and Youth Innovative Research Team of Capital Normal University.

2. Homotopy lifting properties

2.1. **Proof of Theorem 1.3.** A map $f : X \rightarrow Y$ between two metric spaces is called an $\epsilon$-Lipschitz and co-Lipschitz [14], [25] (briefly, $\epsilon$-LcL), if for any $p \in X$, and any $r > 0$, the metric balls satisfy

$$B_{\epsilon^{-\epsilon}r}(f(p)) \subseteq f(B_r(p)) \subseteq B_{\epsilon^{\epsilon}r}(f(p)).$$

A 1-LcL preserves metric balls exactly and is called a *submetry* [1]. Clearly, a regular $\epsilon$-almost Lipschitz submersion is an $\epsilon^{C\epsilon}$-LcL for some universal constant $C$.

Since by definition, a regular almost Lipschitz submersion satisfies the LcL property, it suffices to show Theorem 2.1 below.

In order to simplify constant dependence, we introduce another terminology other than $\delta$-strained radius.

An $n$-dimensional Alexandrov space $Y$ is called $\epsilon$-*almost Euclidean* if for any point $p \in Y$, there is a neighborhood $U$ containing $p$ and a bi-Lipschitz map $\varphi : U \rightarrow \varphi(U) \subset \mathbb{R}^n$ onto an open neighborhood in $\mathbb{R}^n$ such that for any $x, y \in U$,

$$e^{-\epsilon}\|xy\| \leq |\varphi(x)\varphi(y)| \leq e^{\epsilon}\|xy\|,$$
If (2.1) holds on every r-ball in Y, then Y is called $(r, \epsilon)$-almost Euclidean. By [4, Theorem 5.4], an Alexandrov space with $\text{curv} \geq -1$ and $\delta$-strained radius $\geq \mu_0$ is $(\mu_0, \kappa(\delta))$-almost Euclidean.

**Theorem 2.1.** Let $f : X \to Y$ be a $\sqrt{1.023}$-LcL between finite-dimensional Alexandrov spaces with $\text{curv} \geq \kappa$. If $f$ is proper and the base space $Y$ is $\ln \sqrt{1.023}$-almost Euclidean, then $f$ is a Hurewicz fibration, i.e., satisfying the homotopy lifting property with respect to any space.

Theorem 2.1 has appeared in an earlier preprint [29].

**Proof of Theorem 1.3.**

The existence of a regular almost Lipschitz submersion is proven by Yamaguchi [32]. By Theorem 2.1 and the discussion above, any regular almost Lipschitz submersion $f : X \to Y$ is a Hurewicz fibration. □

The remaining of this section is devoted to prove Theorem 2.1.

### 2.2. A sufficient condition for a fibration.

The following topological results are used in the proof of Theorem 2.1.

For any Hurewicz fibration $f : X \to Y$, if $Y$ is path-connected, then by definition the fibers are homotopy equivalent to each other. In [7] Ferry proved that the inverse is also true, if the homotopy equivalences between nearby fibers and the homotopies are under control in the following sense.

A map $f : X \to Y$ between metric spaces is said to be strongly regular [7] if $f$ is proper and if for each $p \in Y$ and any $\epsilon > 0$ there is a $\delta > 0$ such that if $|pp_1| < \delta$, then there are homotopy equivalences between fibers $\varphi_{pp_1} : f^{-1}(p) \to f^{-1}(p_1)$, $\varphi_{p_1p} : f^{-1}(p_1) \to f^{-1}(p)$ which together with the homotopies move points in distance $< \epsilon$.

A topological space $X$ is an absolute neighborhood retract (ANR) if there is an embedding of $X$ as a closed subspace of the Hilbert cube $I^\infty$ such that some neighborhood $N$ of $X$ retracts onto $X$. If $X$ is finite covering dimensional and locally contractible, then $X$ is an ANR ([2]).

**Theorem 2.2 ([7]).** If $f : E \to B$ is a strongly regular map onto a complete finite covering dimensional space $B$ and all fibers are ANRs, then $f$ is a Hurewicz fibration.

**Remark 2.3.** Note that the properties of being an ANR or a Hurewicz fibration are local properties (cf. [7]), Theorem 2.2 was proved locally in [7]. Moreover, the Lipschitz submersion in Theorem 1.3 can be constructed locally ([32]). Hence, both of them holds over $\epsilon$-almost Euclidean points in a complete Alexandrov space. And so are Theorem 2.1 and Theorem 1.3.

According to Theorem 2.2 and the discussion above, Theorem 2.1 holds if an $\epsilon^r$-LcL between Alexandrov spaces with almost Euclidean base space is strongly regular, and all its fibers are locally contractible.
2.3. Gradient estimate for an LcL. Let us first recall a basis property of an $\varepsilon'$-LcL $f : X \to Y$. For any compact subset $S \subset Y$, let $\text{dist}_S$ be the distance function to $S$ in $Y$,

$$
\text{dist}_S(y) = |yS| = \inf\{d(y, s) : s \in S\}.
$$

Then the two functions $\text{dist}_S \circ f$ and $\text{dist}_{f^{-1}(S)} : X \to \mathbb{R}_+$ satisfy (see Lemma 1.4 in [25])

(2.3) \quad $e^{-\varepsilon} \cdot \text{dist}_S \circ f \leq \text{dist}_{f^{-1}(S)} \leq e^{\varepsilon} \cdot \text{dist}_S \circ f$.

Since LcL property is rescaling invariant, from now on we assume that $X$ is an Alexandrov space with $\text{curv} \geq -1$, $Y$ is an $n$-dimensional Alexandrov space with $\text{curv} \geq -1$ that is $\varepsilon'$-almost Euclidean. Let $f : X \to Y$ be an $\varepsilon'$-LcL. Under the assumption that $Y$ is a Riemannian manifold, we constructed in [25] a neighborhood retraction $\varphi_p$ of $f$-fiber over $p \in Y$, which is continuously depending on $p$ and can be used as a weaker replacement of the horizontal lifting of minimal geodesics. In the proof of Theorem 2.1 we will apply it to define controlled homotopy equivalences between nearby fibers. Because now $Y$ is an Alexandrov space, for reader’s convenience we recall its construction and point out the differences to [25] in below.

For an $\varepsilon$-almost Euclidean point $p \in Y$, let $r_p$ denote the maximal number that there is a map $\varphi : B(p, r_p) \to \mathbb{R}^n$ satisfying (2.1). Let $S_{r}(p) = \partial B_r(p)$ be the metric sphere around $p$ and let $x$ be any point in $B_r(p) \setminus \{p\}$. We have the following estimate on the gradient of distance function $\text{dist}_{f^{-1}(S_{r}(p))}$.

**Lemma 2.4.** Let $f : X \to Y$ and $0 < r \leq \min\{r_p, 1\}$ be as above. Let $x$ be a point in $f^{-1}(B_r(p)) \setminus f^{-1}(p)$. The gradient vector of $\text{dist}_{f^{-1}(S_{r}(p))}$ satisfies

(2.4) \quad $1 \geq |\nabla_x \text{dist}_{f^{-1}(S_{r}(p))}| \geq 1 - (e^{2\varepsilon} - 1) \cdot \frac{2r^2}{|xf^{-1}(p)| - |xf^{-1}(S_{r}(p))|}$.

**Proof.** The proof is similar to Lemma 1.5 (1.5.1) in [25]. Let $z \in f^{-1}(S_{r}(p))$, $y \in f^{-1}(p)$ be such that $|xz| = |xf^{-1}(S_{r}(p))|$ and $|xy| = |xf^{-1}(p)|$. Let $v$ be the direction at $x$ of a minimal geodesic from $x$ to $y$. It suffices to bound $\cos \angle(v, w)$ from above for any direction $w$ from $x$ to $f^{-1}(S_{r}(p))$.

Since $f$ and $\varphi$ are $\varepsilon'$-LcLs, by (2.3) we directly see

- $|xy| \leq e^{2\varepsilon} \cdot |\varphi(f(x))\varphi(p)|$,
- $|xz| \leq e^{2\varepsilon} \cdot |\varphi(f(x))\varphi(S_{r}(p))|$,
- $|yz| \geq |zf^{-1}(p)| \geq e^{-\varepsilon} |f(z)p| = e^{-\varepsilon} \cdot r$.

Moreover,

\[ |\varphi(f(x))\varphi(p)| + |\varphi(f(x))\varphi(S_{r}(p))| \leq |\varphi(f(x))\varphi(p)| + |\varphi(f(x))S_{r}(\varphi(p))| \]

\[ = e^{\varepsilon} r. \]

Thus

\[ |yz| \leq |xy| + |xz| \leq e^{2\varepsilon} \cdot (|\varphi(f(x))\varphi(S_{r}(p))| + |\varphi(f(x))\varphi(S_{r}(p))|) \leq 3^{\varepsilon} r. \]
Since the proof below is similar for different curvature lower bound, for simplicity we only prove for $\kappa = 0$. By the Euclidean cosine law, we derive

$$
\cos \angle_0(zxy) = \frac{|xz|^2 + |xy|^2 - |yz|^2}{2|xz| \cdot |xy|} \\
= \frac{(|xz| + |xy|)^2 - |yz|^2}{2|xz| \cdot |xy|} - 1 \\
= \frac{(|xz| + |xy| - |yz|) \cdot (|xz| + |xy| + |yz|)}{2|xz| \cdot |xy|} - 1 \\
\leq (e^{2\epsilon} - 1) \cdot \frac{r^2}{|xz| \cdot |xy|} - 1.
$$

By Lemma 2.4 and a standard argument, for sufficient small $\epsilon$ ($e^{2\epsilon} \leq 1.02368$), points in $f^{-1}(B_{2\epsilon}(p))$ can be flowed into $f^{-1}(B_{\frac{2}{3}}(p))$ along gradient curves of $\text{dist}_{f^{-1}(S_r(p))}$ in a definite time.

Lemma 2.5 (Lemma 1.5 in [25]). For any $p \in Y$ and $r < \min\{r_p, \frac{1}{2\epsilon}\}$, there is a constant $C_0(\epsilon) > 0$ depending on $\epsilon$ such that for all $x \in f^{-1}(B_{\frac{2}{3}}(p))$, the gradient curve $\Phi(t, x)$ of the function $\text{dist}_{f^{-1}(S_r(p))}$ satisfies

$$
\Phi(x, t) \in f^{-1}(B_{\frac{2}{3}}(p)), \quad t \geq C_0^{-1} \cdot \left(\frac{2}{3} e^\epsilon r - \left|x f^{-1}(S_r(p))\right|\right).
$$

2.4. Neighborhood retraction of a fiber. In this part we construct a neighborhood retraction around a fiber $f^{-1}(p)$ which continuously depends on $p$.

We first define a gradient deformation of $\text{id}_{f^{-1}(B_{2\epsilon}(p))}$ which maps $f^{-1}(B_{\frac{2}{3}}(p))$ into $f^{-1}(B_{\frac{2}{3}}(p))$ and fixes $f^{-1}(B_{0.3\epsilon}(p))$. Let

$$
T_{p, r}(x) = \max \left\{0, C_0^{-1} \cdot \left(\frac{2}{3} e^\epsilon r - \left|x f^{-1}(S_r(p))\right|\right)\right\},
$$

and $\Phi_{p, r}^{T_p, r}(x) = \Phi(x, T_{p, r}(x))$ be the gradient deformation of $\text{id}_{f^{-1}(B_{2\epsilon}(p))}$ with respect to $\text{dist}_{f^{-1}(S_r(p))}$. Then by Lemma 2.5 and direct calculation, for $e^{2\epsilon} \leq 1.02368$ and $r < \min\{r_p, \frac{1}{2\epsilon}\}$, we have

$$
\begin{align*}
\Phi_{p, r}^{T_p, r}(x) &\in f^{-1}(B_{\frac{2}{3}}(p)), \quad \forall x \in f^{-1}(B_{2\epsilon}(p)), \\
T_{p, r}(x) &= 0, \quad \forall x \in f^{-1}(B_{0.3\epsilon}(p)).
\end{align*}
$$

In [25, Proposition 1.6] we proved that $\Phi_{p, r}^{T_p, r}(x)$ is continuous both in $p$ and $x$, provided that $Y$ is a Riemannian manifold and $r$ is smaller than the injectivity radius of $Y$. In the following we prove the same holds when $Y$ is an almost Euclidean Alexandrov space.
Lemma 2.6. Let $0 < \epsilon < \ln \sqrt{1.02368}$, and let $f : X \to Y$ be an $\epsilon'$-LeL between Alexandrov spaces such that $Y$ is $(\mu_0, \epsilon)$-almost Euclidean. Then for any $0 < r < \frac{1}{2} \cdot \min\{\mu_0, 1\}$,

$$\Psi : \bigcup_{p \in Y} \{p\} \times f^{-1}(B_{2r}(p)) \subset Y \times X \to X, \quad \Psi(p,x) = \Phi_{p}^{T_\epsilon(x)}(x)$$

is a continuous map.

Proof. Since the proof is similar to [25, Proposition 1.6], we give a sketch proof by pointing out the difference.

Because the gradient curves are stable as function converges ([23]), it suffices to show that the distance functions $d_{f^{-1}(S_r(p))}$, $d_{f^{-1}(S_r(q))}$ (to $f^{-1}(S_r(p))$ and $f^{-1}(S_r(q))$ respectively) are $C|pq|$-close for small $|pq|$ and a constant $C$.

By the definition of LeL, it is easy to verify (see [25, Lemma 1.4, Lemma 1.7]) that the Hausdorff distance and the difference between $d_{f^{-1}(S_r(p))}$ and $d_{f^{-1}(S_r(q))}$ satisfy

$$d_{f^{-1}(S_r(p))}, d_{f^{-1}(S_r(q))}) = d_H(f^{-1}(S_r(p)), f^{-1}(S_r(q))),$$

$$d_H(f^{-1}(S_r(p)), f^{-1}(S_r(q))) \leq \epsilon' \cdot d_H(S_r(p), S_r(q)).$$

Let $d(p,q) = \epsilon_1$. By (2.6.1) and (2.6.2), what remains is to show $S_r(p)$ and $S_r(q)$ are $C\epsilon_1$-close in Hausdorff distance.

Let $z$ be a middle point in a minimal geodesic $[pq]$. Since both $S_r(p)$ and $S_r(q)$ lie in the annulus $B_{r+\epsilon_1}(z) \setminus B_{r-\epsilon_1}(z)$, it is easy to see that one only needs to bound the Hausdorff distance between metric spheres $S_{r+\epsilon_1}(z)$ and $S_{r-\epsilon_1}(z)$, i.e., for some constant $C$, $d_H(S_{r-\epsilon_1}(z), S_{r+\epsilon_1}(z)) \leq C\epsilon_1$.

Indeed, for any point $x \in S_{r-\epsilon_1}(z)$, since the point $x_1$ in a minimal geodesic $[xz]$ with distance $|x_1x| = 2\epsilon_1$ lies in $S_{r-\epsilon_1}(z)$, $S_{r+\epsilon_1}(z)$ lies in $2\epsilon_1$-neighborhood of $S_{r-\epsilon_1}(z)$.

Conversely, let $x \in S_{r-\epsilon_1}$. By the proof of Lemma 2.4, there is a point $y$ in $S_{2\epsilon_1}(z)$ such that the comparison triangle $\tilde{\Delta}_{-1}(zxy)$ is larger than $\pi/2$ by a positive definite error $\theta > 0$. By the triangle version of Toponogov theorem, there exists $y_1$ in $[xy]$ with distance $|xy_1| \leq \frac{2\epsilon_1}{\cos \tilde{\Delta}_{-1}(zxy)}$ such that $|y_1z| = r + \epsilon_1$.

Next, let us repeat the construction above for the sequence $\{r_i = \frac{r}{2^i}\}_{i=0,1,2,...}$ and let $\Phi_{p,i}^{T_{\epsilon_i}(x)}(x, T_{p,r_i}(x))$ be the gradient curves of $d_{f^{-1}(S_\epsilon(p))}$ at $x$ with time $T_{p,r_i}(x)$. By (2.5), $\Phi_{p,i}^{T_{\epsilon_i}(x)} : f^{-1}(B_{r_i}(p)) \to X$ takes $f^{-1}(B_{2\epsilon_i}(p))$ into $f^{-1}(B_{2\epsilon_i}(p))$, and

$$\Phi_{p,i}^{T_{\epsilon_i}(x)}|_{f^{-1}(B_{\epsilon_i}(p))} = \text{id}.$$

Hence the iterated gradient deformations

$$\Phi_{p,i}^{T_{\epsilon_i}(x)} \circ \Phi_{p,i-1}^{T_{\epsilon_i}(x)} \circ \cdots \circ \Phi_{p,0}^{T_{\epsilon_i}(x)}$$
is well-defined on \( f^{-1}(B_{\frac{2r}{3}}(p)) \) and its restriction on \( f^{-1}(B_{0,\frac{2r}{3}}(p)) \) is identity. Because

\[
T_{p,r}(x) \leq \frac{r}{2^{i-1}} \cdot \frac{e^r}{3} C_0^{-1},
\]

it can be directly verified that the sequence of maps

\[
\Psi_i : \bigcup_{p \in Y} \{p\} \times f^{-1}(B_{\frac{2r}{3}}(p)) \to X,
\]

\[
\Psi_i(p,x) = \Phi_{p,i}^{T_i} \circ \Phi_{p,i-1}^{T_i} \circ \cdots \Phi_{p,0}(x)
\]

uniformly converges. The limit \( \varphi_p(x) = \lim_{i \to \infty} \Psi_i(p,x) \) gives a retraction from the neighborhood \( f^{-1}(B_{\frac{2r}{3}}(p)) \) to \( f^{-1}(p) \), which by Lemma 2.6 is continuous both in \( p \) and \( x \). We summarize it to the following proposition.

**Proposition 2.7.** For any \( 0 < r < \frac{1}{2} \min\{\mu_0, 1\} \), there is a deformation retraction \( \varphi_p(x) \) from a neighborhood \( f^{-1}(B_{\frac{2r}{3}}(p)) \) to the fiber \( f^{-1}(p) \) such that

\[
\varphi : \bigcup_{p \in Y} \{p\} \times f^{-1}(B_{\frac{2r}{3}}(p)) \to X, \quad \varphi(p,x) = \varphi_p(x)
\]

is continuous both in \( p \) and \( x \), and satisfies

(2.7.1) \( \varphi_p(x) = x \) for any \( x \in f^{-1}(p) \), and

(2.7.2) \( |x - \varphi_p(x)| \leq 2C_1 \rho \), for some constant \( C_1(\epsilon) \) depending only on \( \epsilon \).

**Proof of Theorem 2.1.**

Up to a rescaling we assume that the lower curvature bounds of both \( X \) and \( Y \) are \(-1\). By Theorem 2.2, it suffices to show that \( f \) is strong regular and any fiber is an ANRs. For any \( p, q \in B \) with small distance \( 0 < |pq| < \frac{1}{2} \min\{\mu_0, 1\} \), let \( \rho = 2 |pq| \). By the definition of \( \text{LcL} \), it is easy to see that

\[
e^{-\epsilon} \cdot |pq| \leq d_H(f^{-1}(p), f^{-1}(q)) \leq e^\epsilon \cdot |pq|.
\]

Thus \( f^{-1}(q) \) lies in \( e^\epsilon \frac{2r}{3} \)-neighborhood of \( f^{-1}(p) \) and vice versa. By Proposition 2.7, there are neighborhood retractions \( \varphi_p : f^{-1}(B_{\frac{2r}{3}}(p)) \to f^{-1}(p) \) and \( \varphi_q : f^{-1}(B_{\frac{2r}{3}}(q)) \to f^{-1}(q) \) around \( f^{-1}(p) \) and \( f^{-1}(q) \) respectively. Then the homotopy equivalences between fibers can be chosen to be \( \varphi_p|_{f^{-1}(q)} : f^{-1}(q) \to f^{-1}(p) \) and \( \varphi_q|_{f^{-1}(p)} : f^{-1}(p) \to f^{-1}(q) \), and the homotopies are

\[
H_t = \varphi_p \circ \varphi_{\gamma(t)} : f^{-1}(p) \to f^{-1}(p) \\
K_t = \varphi_q \circ \varphi_{\gamma(1-t)} : f^{-1}(q) \to f^{-1}(q),
\]

where \( \gamma : [0,1] \to B \) is a minimal geodesic from \( p \) to \( q \). By (2.7.2), \( |H_t(x)| \leq 4C_1 \rho \) and \( |K_t(x)| \leq 4C_1 \rho \). Therefore \( f \) is strongly regular.

According to [20] (cf. [14], [23]), an Alexandrov space with curvature bounded below is locally contractible. For \( x \in f^{-1}(p) \), let \( U_x \ni x \) be a contractible neighborhood around \( x \) and \( H_t : U_x \to U_x \) be the homotopy from \( \text{id}_{U_x} \) to the retraction \( r : U_x \to \{x\} \) such that \( H_t(x) = x \). Then \( \varphi_p \circ H_t \) is a homotopy from \( \text{id}_{U_x \cap f^{-1}(p)} \) to the retraction \( r : U_x \cap f^{-1}(p) \to \{x\} \). Therefore \( f^{-1}(p) \) is locally contractible and thus an absolute neighborhood retract. \( \square \)
2.5. Homotopic uniqueness of fibration. Recently it is proved in \cite{30} that two collapsed metrics \( g_i \) \((i = 0, 1)\) on \( M \) induces the same nilpotent Killing structure up to a diffeomorphism, provided \( g_i \) are \( L_0 \)-Lipschitz equivalent and sufficiently collapsed.

In the following we prove that in the homotopic sense, the collapsing fibration in Theorem 1.3 is unique.

We say that two Hurewicz fibrations \( f_i : X_i \to Y \) \((i = 0, 1)\) are fibrewise homotopy equivalent if there are fiber-preserving maps \( h : X_0 \to X_1 \) and \( g : X_1 \to X_0 \) and fiber-preserving homotopies between \( g \circ h \) and identity \( 1_{X_0} \), and between \( h \circ g \) and \( 1_{X_1} \). We say that Hurewicz fibrations \( f_i : X_i \to Y_i \) \((i = 0, 1)\) are equivalent if there is a homeomorphism \( \psi : Y_0 \to Y_1 \) such that \( \psi \circ f_0 : X_0 \to Y_1 \) is fiber-homotopy equivalent to \( f_1 : X_1 \to Y_1 \).

**Theorem 2.8.** Let \( X, Y_i \) \((i = 0, 1)\) be Alexandrov spaces with \( \text{curv} \geq -1 \) such that \( Y_i \) satisfies (1.1.1), the dimension \( \dim Y_1 = \dim Y_2 \), and (1.1.2) holds for \( d_{GH}(X, Y_i) \). Then any two Hurewicz fibrations \( f_i \) from \( X \) to \( Y_i \) \((i = 0, 1)\) provided by Theorem 1.3 are equivalent.

It follows either from \cite[Theorem 9.8]{4} (a key lemma of its proof has a flaw, for a correct proof see \cite{28}), or from Theorem 1.3, that there is \( e^{\kappa(\kappa(n))} \)-bi-Lipschitz map \( \varphi : Y_0 \to Y_1 \) such that \( \varphi \circ f_0 \) is \( 100e \)-close to \( f_1 \). Thus, the uniqueness in Theorem 2.8 is reduced to a stability result below.

**Proposition 2.9.** Let \( X \) and \( Y \) be two Alexandrov spaces with \( \text{curv} \geq \kappa \), where \( Y \) is \((\mu_0, \ln \sqrt{1.023})\)-almost Euclidean. If two \( \sqrt{1.023} \)-LcLs \( f_0, f_1 : X \to Y \) are \( \frac{\mu_0}{3} \)-close, i.e.,

\[
(2.2) \quad d(f_0, f_1) = \sup_{x \in X} |f_0(x)f_1(x)| < \frac{1}{3} \mu_0,
\]

then they are equivalent as Hurewicz fibrations.

Theorem 2.9 is an improvement of a stability result in \cite{29}.

In the proof of Proposition 2.9, we need a “canonical” pointed contraction on the base space \( Y \), which are constructed similarly as in Proposition 2.7.

**Lemma 2.10.** Let \( Y \) be a \((\mu_0, \ln 1.02368)\)-almost Euclidean Alexandrov space with \( \text{curv} \geq -1 \). There is a continuous pointwise contraction on \( Y \),

\[
\tau : \bigcup_{p \in Y} \{p\} \times B_{\mu_0/2}(p) \times [0, 1] \to Y, \quad \tau(p, x, 0) = x, \quad \tau(p, x, 1) = p.
\]

**Proof.** Note that the estimates in Lemma 2.4 and 2.5 also holds for the distance function \( \text{dist}_{S_r(p)} \) for \( 0 < r \leq \min\{\mu_0, 1\} \). Let \( \psi(p, x, t) \) be the limit of iterated gradient flows of \( \text{dist}_{S_{r_i}(p)} \) for \( r_i = 2^{-i}\mu_0 \) with time \( t \in [0, T_{p,r_i}(x)] \), where \( C_0 \) is the constant in Lemma 2.5 and \( T_{p,r_i}(x) = \max\{0, C_0^{-1}(\frac{r_i}{4}e^{-r_i} - |x|_S(p))\} \). Let \( T(p, x) = \sum_{i=0}^{\infty} T_{p,r_i}(x) \). It follows from the proof of Proposition 2.7 directly that the map \( \tau(p, x, t) = \psi(p, x, tT(p, x)) \) satisfies the requirement of Lemma 2.10. \qed
Proof of Proposition 2.9.

Let \( f_0, f_1 : X \rightarrow Y \) be the \( 1.023 \)-LcLs between Alexandrov spaces with a \((\mu_0, \ln 1.023)\)-almost Euclidean base \( Y \). We now construct fiber-preserving maps \( h, g : X \rightarrow X \) and fiber-preserving homotopies \( g \circ h \) to the identity \( 1_X \) and from \( h \circ g \) to \( 1_X \) as follows.

For any point \( x \in X \), let \( p = f_0(x) \in B \), let \( F_0(p) \) be the fiber \( f_0^{-1}(p) \) and \( F_1(p) = f_1^{-1}(p) \). Suppose that \( d(f_0, f_1) < \frac{1}{\mu_0} \). Then by (2.3), \( F_0(p) \) lies in the \( 1.023 \)-neighborhood of \( F_1(p) \). Let \( \varphi_p \) be the neighborhood retraction of \( F_1(p) \) in Proposition 2.7 with respect to \( f_1 \), we define \( h : X \rightarrow X \) by \( h(x) = \varphi_p(x) = \varphi_{f_0(x)}(x) \). Then the continuous map \( h : X \rightarrow X \) is globally defined and maps all fibers of \( f_0 \) into that of \( f_1 \). Similarly we define \( g : X \rightarrow X \) through the neighborhood retraction of \( f_0 \)-fibers such that \( f_0 \circ g = f_1 \), where \( g(x) = \psi_{f_1(x)}(x) \) and \( \psi_g \) is the neighborhood retraction of \( f_0^{-1}(q) \) with respect to \( f_0 \).

Note that \( f_1(\varphi_{f_0(x)}(x)) = f_0(x) \), thus
\[
g \circ h(x) = \psi_{f_1(\varphi_{f_0(x)}(x))}(\varphi_{f_0(x)}(x)) = \psi_{f_0(x)} \circ \varphi_{f_0(x)}(x).
\]
Moreover, since \( \varphi_{f_1(x)} \) is a neighborhood retract to \( F_1(f_1(x)) \), \( \varphi_{f_1(x)}(x) = x \).
Similarly, \( \psi_{f_0(x)}(x) = x \), and thus
\[
\psi_{f_0(x)} \circ \varphi_{f_1(x)} = 1_X : X \rightarrow X.
\]
For \( p_0 = f_0(x) \), let \( p_1 = f_1(x) \) and let \( p_t = \tau(p_1, p_0, t) \) be the map provided by Lemma 2.10. Then \( p_t \) is a curve from \( p_0 \) to \( p_1 \) continuously depending on \( x \) and \( t \). We define the fiber-preserving homotopy \( H_t : X \rightarrow X \) by \( H_t(x) = \psi_{f_0(x)} \circ \varphi_{p_t}(x) \). Then \( H : [0,1] \times X \rightarrow X \) is a \( f_0 \)-fiber-preserving continuous map such that \( H_0 = g \circ h \) and \( H_1 = 1_X \). A fiber-preserving homotopy from \( h \circ g \) to the identity \( 1_X \) can be defined similarly. \( \square \)

3. Margulis lemma on Alexandrov spaces

3.1. Proof of Theorem 1.1. Let \( X \) be a locally complete Alexandrov \( n \)-space of curv \( \geq -1 \). Let \( B_1(p) \) be the 1-ball centered at some point \( p \in X \).

It is well-known that sufficient away from where \( X \) is non-complete, the global Toponogov comparison on Alexandrov space holds ([4], [24]). To be precisely, there is a constant \( C_T \), such that Toponogov comparison holds for any triangle in \( B_{\frac{1}{C_T}}(p) \), provided that \( B_1(p) \) is relative compact. According to the proof of global Toponogov theorem in [17] or [27], it is enough to choose \( C_T = 100 \).

Note that in a locally complete local Alexandrov \( n \)-space of curv \( \geq -1 \), the convex hull of a triangle may not be bounded. However, by the proof of Toponogov comparison (cf. [17], [27]), any contradicting triangle can be reduced successively to other ones, whose perimeters decay in a definite ratio to form a converging geometric progression, such that a contradiction can be derived in a neighborhood of the initial triangle whose radius is not more than 12-times of the initial perimeter.
Due to the above discussion, the fundamental facts on a complete Alexandrov space will be freely applied locally in this section without further mention.

We first reduce Theorem 1.1 to the following special case. For any \( p \in X \) and \( q \in B_1(p) \), \( 0 < r \leq 1 - |pq| \), let \( \Gamma_p(q; r) \) be the subgroup of the fundamental group \( \pi_1(B_1(p), q) \) generated by loops at \( q \) lying in \( B_r(q) \). As before, we will use \( d(p, q) \) or \( |pq| \) to denote the distance between two points.

**Theorem 3.1.** Suppose that \( B_1(p) \) is relative compact in \( X \). Then there are positive constants \( \epsilon(n), w(n) > 0 \), both depending only on the dimension \( n \), such that there is a “good” point \( q \in B_{\frac{1}{2\epsilon T}}(p) \) satisfying \( \Gamma_p(q; \epsilon(n)) \) is \( w(n) \)-nilpotent.

For general points in \( X \), we need the following result in [16].

**Lemma 3.2 ([16, Step 2 in §7]).** For any positive integer \( n \) and \( 0 < \epsilon \leq \epsilon(n) \) with \( \epsilon(n) \) in Proposition 3.1, there is \( L(\epsilon, n) > 0 \) such that the following holds.

Let \( X \) be an Alexandrov \( n \)-space of \( \text{curv} \geq -1 \), and \( p \in X \) be a point such that \( B_1(p) \) is relative compact in \( X \). Let \( \Gamma = \langle \beta_1, \cdots , \beta_k : d(\beta_ip, p) \leq \frac{1}{100C_TL(\epsilon , n)} \rangle \) be a discrete subgroup of isometries of \( X \) that acts freely. Then the subgroup

\[
H = \left\{ g \in \Gamma : d(gx, x) \leq \epsilon, \forall x \in B_{\frac{1}{2\epsilon T}}(p) \right\}
\]

has finite index \( [\Gamma : H] \leq (2k + 1)^{L(\epsilon, n)} \).

Note that for any isometry \( \gamma \) of \( X \) which moves \( p \) not farther than \( \frac{1}{100C_T} \) but a point in \( B_{\frac{1}{2\epsilon T}}(p) \) farther than \( \epsilon \), \( \gamma \) should move a point in any maximal \( \frac{1}{8} \)-net of \( B_{\frac{1}{2\epsilon T}}(p) \) farther than \( \frac{1}{8} \). Thus the total possibility of such isometries can be reduced to permutations of lattice, whose total number is under control by the relative volume comparison (see [18]). By considering the naturally extended action of \( \Gamma \) on the \( m \) times direct product space by \( X \) itself, the total number of cosets of \( H \) can be counting via a wordlength-cutting-off argument with

\[
L(\epsilon, n) = \frac{(\text{vol } B_{\frac{1}{2T}}(x))^m(\epsilon, n)}{\text{vol } B_{-1}(\frac{2}{3\epsilon T})} \quad \text{where} \quad m(\epsilon, n) = \frac{\text{vol } B_{\frac{1}{2T}}(x)^n}{\text{vol } B_{-1}(\frac{2}{3})},
\]

and \( B_{-1}(r) \) denotes a ball in the Hyperbolic space \( \mathbb{H}^n \). For details, see [16, §7].

Assuming Theorem 3.1, we now prove Theorem 1.1.

**Proof of Theorem 1.1.**

Let \( \tilde{X}, \tilde{p} \rightarrow (B_1(p), p) \) be the universal cover of \( B_1(p) \). Now we take \( \epsilon(n) \) and \( q \) to be the constant and a corresponding “good” point given by Theorem 3.1. Let

\[
H = \left\{ g \in \pi_1(B_1(p), p) : d(gx, x) \leq \epsilon(n) \right\}
\]

for any \( x \in B_{\frac{1}{2\epsilon T}}(\tilde{p}) \).
Since \( q \in B_{\frac{1}{2c_{T}}} (p) \), \( H \) can be viewed as a subgroup of \( \Gamma_p(q; \epsilon(n)) \), hence \( H \) is \( w(n) \)-nilpotent.

Since for some \( 0 < \delta \leq \frac{1}{2c_{T}} \), \( B_{2\delta}(p) \) is locally contractible, any loop lying in \( B_{\delta}(p) \) at \( p \) is homotopic to a joining of loops not longer than \( 3\delta \) at \( p \). By a standard argument of Gromov’s short basis, the generating set of \( \Gamma_p(p; \delta) \) can be chosen to have at most \( k(n) \) elements. By Lemma 3.2, let \( \delta(n) = \frac{1}{300C_{T}L(\epsilon(n), n)} \), then for any \( 0 < \delta \leq \delta(n) \), \( [\Gamma_p(p; \delta) : \Gamma_p(p; \delta) \cap H] \leq c(n) \).

Therefore, \( \Gamma_p(p; \delta) \cap H \) is a subgroup of \( H \), which is \( w(n) \)-nilpotent and has finite index \( \leq c(n) \) in \( \Gamma_p(p; \delta) \), we derive that \( \Gamma_p(p; \delta) \) itself is \( w'(n) \)-nilpotent.

What remains in this paper is devoted to prove Theorem 3.1. We will argue by contradiction. Assuming the contrary, then there is a sequence \( (X_\alpha, p_\alpha) \) of Alexandrov \( n \)-spaces with \( \text{curv} \geq -1 \), such that for any \( q_\alpha \in B_{\frac{1}{2c_{T}}} (p_\alpha) \), \( \Gamma_{p_\alpha}(q_\alpha; \alpha^{-1}) \) fails to be \( w(n) \)-nilpotent.

By passing to a subsequence, we may assume that \( (X_\alpha, p_\alpha) \overset{GH}{\to} (X_\infty, p_\infty) \), i.e., \( (X_\alpha, p_\alpha) \) Gromov-Hausdorff converges to a limit space \( (X_\infty, p_\infty) \). Following [15], we will show in the remaining sections that:

**Claim 3.3.** By passing to a subsequence, there is \( 0 < R_1 \leq \frac{1}{100C_{T}} \) such that for each sufficient large \( \alpha \), there is a point \( q_\alpha \in B_{\frac{1}{2c_{T}}} (p_\alpha) \) and a chain of subgroups

\[
G_{1,\alpha} = \Gamma_{p_\alpha}(q_\alpha; R_1) \triangleright G_{2,\alpha} \triangleright \cdots \triangleright G_{l,\alpha} = \{e\} \quad (l \leq n)
\]

satisfying

(A) \( G_{i,\alpha}/G_{i+1,\alpha} \) is \( C \)-abelian;

(B) \( G_{i,\alpha} \triangleleft G_{1,\alpha} \);

(C) By (A) and (B), \( G_{1,\alpha} \) acts on \( G_{i,\alpha}/G_{i+1,\alpha} \) by conjugation, which induces a homomorphism \( \rho_{i,\alpha} : G_{1,\alpha} \to \text{Out}(G_{i,\alpha}/G_{i+1,\alpha}) \). The image \( \rho_{i,\alpha}(G_{1,\alpha}) \) has finite elements, \( \#\rho_{i,\alpha}(G_{1,\alpha}) < N_0 \).

Then by [15, Lemma 4.2.1], we derive that \( G_{1,\alpha} \) is \( w(C, N_0) \)-nilpotent, a contradiction.

In order to construct each \( G_{i,\alpha} \), we define the local fundamental groups (Definition 3.5 below). Then (B) and (C) would follow from the leveled gap property (Definition 3.2) and a universal estimate of gradient push associated to a \( \delta^2 \)-maximal frame (Definition 3.8); see Section 4. (A) will be guaranteed by the construction and the generalized Bieberbach theorem ([9], cf. [32]); see Proposition 5.1.

### 3.2. Local fundamental group and numerical maximal frame

We first introduce the local fundamental group that will realize \( G_{i,\alpha} \).

**Definition 3.4.** Let \( X \) be a locally complete Alexandrov \( n \)-space with \( \text{curv} \geq -1 \). Let \( p \) be a point in \( X \) such that the metric ball \( B_1(p) \) is relative
compact in $X$. For $0 < r \leq \frac{1}{2}$, the $r$-local fundamental group $\pi_1^r(p; r)$ at $p$ is defined to be

$$\pi_1^r(p; r) = \langle \text{loop } \gamma \text{ at } p : \text{im } \gamma \in B_r(p) \rangle \sim \gamma_1$$

where $\gamma_1 \sim \gamma_2$ if they are homotopic in $B_{2r}(p)$.

For $r_1 > r_2 > 0$, let $\iota : \pi_1^r(p; r_2) \to \pi_1^r(p; r_1)$ be the inclusion homomorphism. A key property used in proving (B) and (C) is certain “leveled” gap between local fundamental groups at different scales as follow.

**Definition 3.5.** We say that $\pi_1^r(p; R_1)$ satisfies $(\epsilon, \sigma, l)$-leveled gap property, if there is a sequence of intervals $[\tau_1 = 0, \tau_2, \ldots, \tau_1, \tau_1]$ such that

1. $r_i \leq \epsilon R_i$, and $r_i/R_{i+1} \leq \sigma$,
2. $\iota : \pi_1^r(p; r_i) \to \pi_1^r(p; R_i)$ is an isomorphism,
3. $\iota(\pi_1^r(p; R_{i+1})) \leq \pi_1^r(p; r_i)$.

In practice, $r_i = 3 \text{diam} Y_i+1$, where $Y_i+1$ is a “regular fiber” at $i$-level, which by definition, is a level set of $F_{k_1+\cdots+k_i} = (\text{dist}_{a_1}, \ldots, \text{dist}_{a_{k_1+\cdots+k_i}})$, where $a_j$ are from a maximal $(k_1 + \cdots + k_i)$-frame (for definition see below), dist$_{a_j}$ is the distance function to $a_j$, and $2R_i$ is the radius of a Perelman’s fibration $F_{k_1+\cdots+k_i}$’s base disk around a regular point in a limit space.

Secondly, we introduce a $\delta^2$-maximal frame. Let $X$ be an Alexandrov $n$-space and let $k$ be a positive integer $\leq n$. Let $0 < \delta \leq \frac{1}{102}$. By [4], given a pair of points $(a_1, b_1)$ and a minimal geodesic segment $[a_1b_1]$ between them, a $k$-frame $\{[a_jb_j]\}_{j=1}^k$, which consists of $k$ minimal geodesic segment $[a_jb_j]$, can be built up successively (and non-uniquely) on $X$: Assuming $[a_{i-1}b_{i-1}]$ is well-defined, then take $[a_ib_i]$ on $X$ that satisfies the following

1. $b_i$ is the middle point of the geodesic segment $[a_{i-1}b_{i-1}]$,
2. $|a_i a_j| = |b_i a_j|$, for all $1 \leq j \leq i - 1$,
3. the edge $[a_i b_i]$ is $\delta^2$-collapsed, i.e., $|a_i b_i| \leq \delta^2 |a_{i-1} b_{i-1}|$.

A little more generally, we will consider $k$-frames where $b_i$ is not far away from the middle point $m_{i-1}$ of $[a_{i-1} b_{i-1}]$. Let $\{[a_jb_j]\}_{j=1}^k$ be a $k$-frame. Let $F_k = (\text{dist}_{a_1}, \ldots, \text{dist}_{a_k}) : X \to \mathbb{R}^k$. A new pair $(a_{k+1}, b_{k+1})$ is called $\delta^2$-maximal relative to a $k$-frame $\{[a_jb_j]\}_{j=1}^k$ if

1. $b_{k+1}$ is $\delta\min\{|a_kb_k|\}-close to the middle point $m_k$ of $[a_kb_k]$,
2. $F_k(a_{k+1}) = F_k(b_{k+1})$,
3. $|a_{k+1}b_{k+1}| = d_{k+1}$, where
   $$d_{k+1} = \max\{|b_{k+1}x| : x \in F_k^{-1}(F_k(b_{k+1})) \text{ and } |xb_{k+1}| \leq \delta^2 \min_{i=1,\ldots,k} |a_i b_i|\}.$$

Note that by (3.7.3), one always has $|a_{k+1}b_{k+1}| \leq \delta^2 \min_{i=1,\ldots,k} |a_i b_i|$.

**Definition 3.8.** A $k$-frame is called $\delta^2$-maximal if for each $2 \leq i \leq k$, $(a_i, b_i)$ is $\delta^2$-maximal relative to the $(i-1)$-frame $\{[a_jb_j]\}_{j=1}^{i-1}$. For an $\delta^2$-maximal $n$-frame $\{[a_jb_j]\}_{j=1}^n$, we say that $\{[a_jb_j]\}_{j=1}^n$ is centered at $x \in X$, if the point $x$ is $\delta\min\{|a_kb_k|\}$-close to $m_n$. 
By the construction above, Theorem 1.6 is reduced to a universal estimate of gradient push associated to a $\delta^2$-maximal frame; see Theorem 6.1 in the appendix.

The following fact on the gradient flow of $\lambda$-concave functions on Alexandrov space is applied in proving Theorem 3.1.

**Theorem 3.9** ([23]). Let $\Phi_t$ be the gradient flow of a $\lambda$-concave function on a complete Alexandrov space. Then $\Phi_t : X \to X$ is $e^{\lambda t}$-Lipschitz.

**Remark 3.10.** We remark that all results on gradient push with respect to a $\delta^2$-maximal frame also hold for a $(n, \delta)$-strainer with suitable maximum property. We only use maximal frames in this paper for simplicity.

4. **Proofs of Claims (B) and (C)**

We now prove that the existence of $(\epsilon, \sigma, l)$-leveled gap property and a $\delta^2$-maximal frame centered at $p$ would implies (B) and (C) hold for $G_i = \pi_1^l(q; R_i)$ with $q \in B_{\frac{1}{2C_T}}(p)$.

Throughout this subsection, we always assume that $X$ is a locally complete Alexandrov $n$-space with $\text{curv} \geq -1$ such that the metric ball $B_1(p)$ is relative compact in $X$.

Let $q \in B_{\frac{1}{2C_T}}(p)$, $0 < R_1 \leq \frac{1}{16C_T}$. Let $\pi_1^l(q; R_i)$ be a local fundamental group satisfying the $(\epsilon, \sigma, l)$-leveled gap property. Let $G_i = \pi_1^l(q; R_i)$ for each $1 \leq i \leq l$. Then by the proofs in [15], (B) and (C) hold for $G_i$. We give a proof for completeness.

**Proposition 4.1** ([15]). For $0 < R_1 \leq \frac{1}{16C_T}$ and $\sigma > 0$, there is $\epsilon(n) > 0$ such that for $0 < \epsilon < \epsilon(n)$, any local fundamental group $\pi_1^l(q; R_i)$ with $(\epsilon, \sigma, l)$-leveled gap property for intervals $[r_1 = 0, R_i], \cdots, [r_1, R_1]$, if there is a $\delta^2$-maximal frame $\{(a_j b_j)\}_{j=1}^n$ centered at $q$ such that

$$|a_1 b_1| = \min \left\{ \frac{1}{2C_T} \frac{1}{2} \text{diam } X \right\},$$

then the chain of groups $G_i = \pi_1^l(q; R_i)$, namely

$$G_1 \triangleright G_2 \triangleright \cdots \triangleright G_l = \{ e \},$$

satisfies (B) and (C).

Let $S_0$ be a short basis of $\pi_1(B_1(p), q)$ and $S_i = (S_0 \cap G_i) \cup (S_0 \cap G_i)^{-1}$. For any $\gamma \in G_1$, the norm $|\gamma|$ is defined to be the minimal length of its representative loops. The following elementary fact will be used in proving Lemma 4.3 below and (B), (C).

**Lemma 4.2.** Any element $\gamma \in S_i \setminus S_{i+1}$ has norm

$$2R_{i+1} \leq |\gamma| \leq \frac{2\epsilon}{3} \min_{\beta \in S_{i+1} \setminus S_i} |\beta|$$

and $G_i = \langle S_i \rangle$. 

Proof. Since $i : \pi_1^i(q; r_i) \to \pi_1^i(q; R_i)$ is an isomorphism, any loop lying in $B_{R_i}(q)$ at $q$ is homotopic to a loop lying in $B_{r_i}(q)$ at $q$. Furthermore, since $B_{2r_i}(q)$ is locally contractible, any loop lying in $B_{r_i}(q)$ at $q$ is homotopic to a joining of loops no longer than $3r_i$ at $q$.

Because $S_0$ is a short basis of $\pi_1(B_1(p), q)$, it can be seen that for any $\gamma \in S_i \setminus S_{i+1}$,
\[ 2R_{i+1} \leq |\gamma| \leq 3r_i \]
and $G_i = \langle S_i \rangle$.

Via gradient push by a $\delta^2$-maximal $n$-frame on certain cover $\tilde{X}$ of $B_1(p)$ and a $\delta^2$-maximal $n$-frame centered at $q$, up to a conjugation any loop in $G_1$, whose action on $\tilde{X}$ has a definite displacement, admits the following control in Lemma 4.3, which is essential in proving (B) and (C).

**Lemma 4.3** ([15]). Assume that there is a $\delta^2$-maximal frame $\{[a_j b_j] \}_{j=1}^n$ centered at $q$ such that $|a_1 b_1| = \min \{ \frac{1}{2\pi}, \frac{1}{2} \text{diam } X \}$. Suppose that $G_i \vartriangleleft G_1$.

Then for any element $\gamma = \gamma_1 \cdots \gamma_m$ with $|\gamma_j| \leq \frac{B_1}{100C(n)}$, there is $\beta \in G_i$ such that for any loop $\alpha \in G_i$ with $|\alpha| \leq 3r_i$,
\[ |(\gamma * \beta)^{-1} * \alpha * (\gamma * \beta)| \leq e^{\frac{2\text{tanh}^2(2 T(n) + C(n))}{\pi(n)}}|\alpha|, \]
where $C(n)$ is the constant in Remark 6.2, $T(n)$ is the constant in Theorem 6.1, and $(\tilde{X}_i, \tilde{q}_i) \xrightarrow{\pi_i} (B_1(p), q)$ is a suitable defined cover with $\pi_i \pi_1(\tilde{X}_i, \tilde{q}_i) = G_i$.

**Proof.** Up to a lifting to a cover $(\tilde{X}_1, \tilde{q}_1) \xrightarrow{\pi'_1} (B_1(p), q)$ with $\pi'_1 \pi_1(\tilde{X}_1, \tilde{q}_1) = G_1$,
we assume that $\pi_1(B_1(p), q) = G_1$. Indeed, by the definition of $G_1$, $\pi'_1$ maps $B_{R_1}(\tilde{q}_1)$ homeomorphically onto $B_{R_1}(q)$. If we want to construct a homotopy lying in $B_{R_1}(\tilde{q}_1)$ of a short loop, we can actually do the construction in $X$ with the resulting homotopy lies in $B_{R_1}(q)$, then composite this homotopy by $(\pi'_1|B_{R_1}(\tilde{q}_1))^{-1}$.

Let $(\tilde{X}_i, \tilde{q}_i) \xrightarrow{\pi_i} (B_1(p), q)$ be a cover with $\pi_i \pi_1(\tilde{X}_i, \tilde{q}_i) = G_i$. Then by our assumption, $\pi_i$ is a normal cover. (This assumption will also be used in proving (B) and (C).)

Let us construct a $\delta^2$-maximal frame $\{[\tilde{c}_j \tilde{d}_j] \}_{j=1}^n$ on $\tilde{X}_i$ such that $\pi_i(\tilde{c}_1) = q$ and $|\tilde{c}_1 \tilde{d}_1| = \min \{ \frac{B_1}{100C(n)}, \frac{1}{2} \text{diam } \tilde{X}_i \}$. Let $\tilde{q}'_i$ be a regular centered point of $\{[\tilde{c}_j \tilde{d}_j] \}_{j=1}^n$, i.e., $\tilde{q}'_i$ is close to the middle point $\tilde{m}_n$ of $[\tilde{c}_n \tilde{d}_n]$.

Since $|q \pi_i(\tilde{q}')| \leq \min \{ \frac{B_1}{100C(n)}, \frac{1}{2} \text{diam } X \}$, there is a gradient push $\varphi$ of $\{[a_j b_j] \}_{j=1}^n$ in time $\leq T(n)$ such that $\varphi(q) = \pi_i(\tilde{q}'_i)$, which gives rise to a homotopy $H$ from $\alpha$ to a loop $\varphi \circ \alpha$ at $\pi_i(\tilde{q}'_i)$. Moreover, the whole pushing line of broken geodesics has total length $\leq C(n) \cdot |q \pi_i(\tilde{q}')|$ (see Remark 6.2).

Since $G_i \vartriangleleft G_1$, there exists a lifting $\tilde{\alpha}$ of $\alpha$ at $\gamma \tilde{q}_i$, and a lifting homotopy $\tilde{H}$ of $H$ on $\tilde{X}_i$ from $\tilde{\alpha}$ to $\tilde{\alpha}' = \tilde{\varphi} \circ \tilde{\alpha}$, whose base points are $\gamma \tilde{q}_i$ and $\tilde{q}_i''$. Then $\tilde{H}$ and $\tilde{\alpha}'$ lie in $B_{R_1}(\gamma \tilde{q}_i)$. 


Moreover, there exists a deck transformation $\psi$ that maps $q''_i$ to $q'_i$. Let $\{\psi^{-1}[\hat{c}_j \hat{o}_j]\}$ be the pullback $\hat{\delta}^2$-frame at $q''_i$. Then there is a gradient push $\hat{\phi}$ of $\{\psi^{-1}[\hat{c}_j \hat{o}_j]\}$ in time $\leq T(n) + C(n)$, which gives rise to a homotopy from $\hat{\alpha}'$ to $\hat{\alpha}''$, whose base point is $q_i$.

Joining two homotopies above together, we get a homotopy from $\hat{\alpha}$ to $\hat{\alpha}''$, whose base points are $\gamma \hat{q}_i$ and $q_i$ respectively.

Note that any single step in these two homotopies are defined by a gradient flow of $\frac{1}{2} \text{dist}_x^2$ with $\text{dist}_x < 2$ for some $x$, hence the concavity of $\frac{1}{2} \text{dist}_x^2$ is bounded by $2 \cosh \frac{2}{\sinh 2}$. By Theorem 6.1 and Theorem 3.9, the length of $\pi(\hat{\alpha}'')$ satisfies

$$\text{length } \pi(\hat{\alpha}'') \leq e^{\frac{2 \cosh \frac{2}{\sinh 2}(2T(n)+C(n))}{\sinh 2}} \cdot \text{length } \alpha.$$ 

Let $\gamma'$ be the successive joining of push curves of $\varphi$ and $\pi \hat{\phi}$. Then it is clear that $\alpha''$ is homotopic to $\gamma' \ast \alpha \ast \gamma'$, and there is $\beta \in G_i$ such that $\gamma' = \gamma \ast \beta$.

Proof of (B) in Proposition 4.1.

By definition of leveled gap property (Definition 3.5), $G_2 \lhd G_1$. We now prove $G_3 \lhd G_1$.

Let $(\hat{X}_2, \hat{q}_2) \xrightarrow{\pi_2} (B_1(p), q)$ be the normal cover defined in the proof of Lemma 4.3. For any $\gamma \in S_1$, $\gamma$ satisfies that $|\gamma| \leq \frac{B_1}{100C(n)}$, as $\epsilon$ in Definition 3.5 is sufficient small. There is $\beta \in G_2$ such that $\gamma' = \gamma \ast \beta$, for any $\alpha \in S_3$,

$$|\gamma' \ast \beta| \leq e^{\frac{2 \cosh \frac{2}{\sinh 2}(2T(n)+C(n))}{\sinh 2}}|\alpha|.$$ 

Let us take $\epsilon^{-1} > 300C(n)e^{\frac{2 \cosh \frac{2}{\sinh 2}(2T(n)+C(n))}{\sinh 2}}$, then by Lemma 4.2, $\gamma^{-1} \ast \alpha \ast \gamma' \in G_3$. Since $G_3 \lhd G_2$, $\gamma^{-1} \ast \alpha \ast \gamma \in G_3$. This implies $G_3 \lhd G_1$.

Repeating the argument above for loops in each $G_i$ for $i \geq 4$ successively, we complete the proof.

Proof of (C) in Proposition 4.1.

For any fixed integer $m$, let $S_1^m = \{\gamma \in G_1 : \text{wordlength}(\gamma) \leq m\}$.

Firstly, similar to the proof of (B), let $(\hat{X}_1, \hat{q}_1) \xrightarrow{\pi_1} (B_1(p), q)$ be the normal cover defined in the proof of Lemma 4.3. For any $\gamma \in S_1^m$, $\gamma$ satisfies that $|\gamma| \leq \frac{B_1}{100C(n)}$, as $\epsilon$ in Definition 3.5 is sufficient small. There is $\beta \in G_i$ such that $\gamma' = \gamma \ast \beta$, for any $\alpha \in S_i$,

$$|\gamma' \ast \beta| \leq e^{\frac{2 \cosh \frac{2}{\sinh 2}(2T(n)+C(n))}{\sinh 2}}|\alpha|.$$ 

Secondly, let us consider the normal cover $(\hat{X}_{i+1}, \hat{q}_{i+1}) \xrightarrow{\pi_{i+1}} (B_1(p), q)$ defined in the proof of Lemma 4.3. Then the relative volume comparison holds in $B_{R_1}(\hat{q}_{i+1})$ (see [18]). By counting the lattice points $G_1(\hat{q}_{i+1})$ in balls of $(\hat{X}_{i+1}, \hat{q}_{i+1})$, up to an inner automorphism of $G_i/G_{i+1}$, the possibility of transformations $\rho_i(S_1^m)$ on $G_i/G_{i+1}$ is bounded by the following number

$$\left(\frac{\text{vol } B_{R_1}^{-1}(\epsilon 2 \cosh \frac{2}{\sinh 2}(2T(n)+C(n)) \cdot 3r_{i+1} + R_{i+2})}{\text{vol } B_{R_{i+2}}^{-1}}\right)^{\#(S_i \setminus S_{i+1})}.$$
Let \( N_0 = \sup_{0 < R_{i+2} \leq 1} \left[ \left( \frac{\text{vol} B^n_{\alpha}(c(n))}{\text{vol} B^n_{\alpha}(R_{i+2})} \right)^{c(n)} \right] + 2, \)
where \( c(n) \) is an upper bound of the total number of short basis \( \#S_1. \)
Let us take \( \epsilon^{-1} > 300C(n)N_0, \) then by Lemma 4.2, \( \#\rho_i(S_1^{N_0}) < N_0. \) Thus by [15, Trivial Lemma 4.2.2], \( \#\rho_i(G_1) < N_0. \) \( \square \)

5. Proof of Claim (A)

To finish the proof of Theorem 3.1, it suffices to construct the local fundamental groups and a maximal frames associated to a contradicting sequence \((X_\alpha, p_\alpha) \rightarrow (X_\infty, p_\infty), \) and then verify (A).

Proposition 5.1. Let \((X_\alpha, p_\alpha) \rightarrow (X_\infty, p_\infty)\) be a convergence sequence of Alexandrov \( n \)-spaces with \( \text{curv} \geq -1 \) such that \( \text{diam} X_\infty \geq 1. \) Then by passing to a subsequence of \((X_\alpha, p_\alpha), \) there are \( 0 < R_1 \leq \frac{1}{4C_T}, \) \( \sigma > 0, \) \( 1 \leq l \leq n, \) and for all large \( \alpha \in \mathbb{N} \) there exist a point \( q_\alpha \in B_{\frac{1}{2C_T}}(p_\alpha) \) such that

\begin{align*}
(5.1.1) & \text{there is an associated } \delta^2 \text{-maximal } n \text{-frame centered at } q_\alpha \text{ with } a_{1,\alpha} = p_\alpha, \ |a_{1,\alpha}b_{1,\alpha}| = \frac{1}{2C_T}; \\
(5.1.2) & \text{the } R_l \text{-local fundamental group } \pi^l_1(q_\alpha; R_l) \text{ satisfies } (\epsilon_\alpha, \sigma, l) \text{-leveled gap property with respect to } |r_{1,\alpha} = 0, R_{l,\alpha}|, \ldots, |r_{1,\alpha}, R_{l,\alpha} = R_l| \text{ and } \epsilon_\alpha \rightarrow 0; \\
(5.1.3) & \pi^l_1(q_\alpha; r_{1,\alpha})/\pi^l_1(q_\alpha; r_{l+1,\alpha}) \text{ is } C \text{-abelian for some constant } C. 
\end{align*}

Now Theorem 3.1 follows from earlier arguments in Section 4 and Proposition 5.1.

Proof of Theorem 3.1.

Continue from earlier discussion, we have assumed a contradicting sequence \((X_\alpha, p_\alpha)\) of Alexandrov \( n \)-spaces with \( \text{curv} \geq -1, \) such that for any \( q_\alpha \in B_{\frac{1}{2C_T}}(p_\alpha), \) \( \Gamma_{p_\alpha}(q_\alpha; \alpha^{-1}) \) fails to be \( w(n) \)-nilpotent, and \((X_\alpha, p_\alpha)\) Gromov-Hausdorff converges to a limit space \((X_\infty, p_\infty).\) Up to changing \( X_\alpha \) to \( X_\alpha \times \mathbb{R}^1, \) we further assume that \( \text{diam} X_\infty \geq 1. \)

By Section 4, it suffices to construct a sequence of local groups at some \( q_\alpha \in B_{\frac{1}{2C_T}}(p_\alpha) \) with leveled gap property such that (A) holds for \( G_{i,\alpha} = \pi^l_1(q_\alpha; R_i,\alpha), \) and there is \( \delta^2 \)-maximal frames at \( q_\alpha. \)

Since the construction follows from Proposition 5.1, the proof of Theorem 3.1 is complete. \( \square \)

What remains of the paper is proving Proposition 5.1. The following non-collapsing property of maximal frame will be used in its proof.

Lemma 5.2. Let \((X^n_\alpha, p_\alpha) \rightarrow (A^k, p_\infty)\) be a sequence of Alexandrov \( n \)-spaces that converges to an Alexandrov \( k \)-space \( A^k. \) Let \( \{[a_{i,\alpha}b_{i,\alpha}]\}_{i=1}^k \) be a \( \delta^2 \)-maximal frame in \( X^n_\alpha \) with \( a_{1,\alpha} = p_\alpha, \) \( 0 < d \leq d_{1,\alpha} = |a_{1,\alpha}b_{1,\alpha}| \leq 1. \) By
passing to a sequence, \( \{[a_{i,\alpha}b_{i,\alpha}]\}_{i=1}^{k} \) converges to a frame \( \{[a_{i,\infty}b_{i,\infty}]\}_{i=1}^{k} \) in \( A_{k} \). Then \( |a_{k,\infty}b_{k,\infty}| > 0 \).

**Proof.** Argue by induction on \( i \). Assume \( d_{i,\infty} = |a_{i,\infty}b_{i,\infty}| > 0 \) for \( i < k \), it suffices to show \( d_{i+1,\infty} > 0 \).

Since \( \dim A_{k} > i \), there is a point \( a'_{i+1,\infty} \in A_{k} \) such that \( 0 < |a'_{i+1,\infty}b_{i+1,\infty}| \leq \frac{1}{\delta} d_{i,\infty} \) and \( F_{i,\infty}(a'_{i+1,\infty}) = F_{i,\infty}(b_{i+1,\infty}) \), where \( F_{i,\infty} = (\text{dist}_{a_{1,\infty}}, \cdots, \text{dist}_{a_{i,\infty}}) \).

Take points \( a''_{i+1,\alpha} \in X_{\alpha}^{\infty} \) which converges to \( a'_{i+1,\infty} \). By [4, Theorem 5.4] (or see Theorem 6.6 below), without loss of generality we assume that \( F_{i,\alpha} \) is \( \frac{1-2\delta}{\sqrt{i}} \)-open, i.e., \( \frac{\sqrt{i}}{1-2\delta} \)-co-Lipschitz, on \( B_{dd_{i,\alpha}}(b_{i+1,\alpha}) \). Hence there is \( a''_{i+1,\alpha} \) in \( F_{i,\alpha}^{-1}(b_{i+1,\alpha}) \) such that \( |a''_{i+1,\alpha}a'_{i+1,\alpha}| \to 0 \) as \( \alpha \to \infty \). Since \( d_{i+1,\alpha} \) is maximal, we derive

\[
d_{i+1,\alpha} \geq |a''_{i+1,\alpha}b_{i+1,\alpha}| > \frac{1}{2}|a'_{i+1,\infty}b_{i+1,\infty}|.
\]

□

**Proof of Proposition 5.1.**

(5.1.1) The construction will be done inductively as follows.

The Starting Step. Assume \( \dim X_{\infty} = k_{1} \). Since \( \dim(X_{\infty}) \geq 1 \), we are able to directly construct a \( \delta^{2} \)-maximal \( k_{1} \)-frame \( \{[a_{j,\alpha}b_{j,\alpha}]\}_{j=1}^{k_{1}} \) in \( B_{1}(p_{\alpha}) \) such that \( a_{1,\alpha} = p_{\alpha} \) and \( b_{1,\alpha} \) is a point such that \( d_{1,\alpha} = |a_{1,\alpha}b_{1,\alpha}|_{X_{\alpha}} = \frac{1}{2\delta} \).

By Lemma 5.2, it converges to a \( k_{1} \)-frame \( \{[a_{j}b_{j}]\}_{j=1}^{k_{1}} \) in \( X_{\infty} \). Let \( m_{k_{1}} \) be the middle point of \( [a_{k_{1}}, b_{k_{1}}] \).

Let \( p_{1} \in X_{\infty} \) be a regular point such that \( |p_{1}m_{k_{1}}| \leq \frac{1}{200} \delta |a_{k_{1}}b_{k_{1}}| \). By [4, Theorem 5.4], there is \( 0 < R_{1} \leq \frac{1}{160\delta} \) such that \( B_{2R_{1}}(p_{1}) \) is bi-Lipschitz to an open ball \( D_{k_{1}}^{1} \) in the Euclidean space \( \mathbb{R}^{k_{1}} \) with bi-Lipschitz constant almost 1. By Perelman’s fibration theorem [21], the map \( F_{k_{1}}^{-1} \circ F_{k_{1},\alpha} \) is a locally trivial fibration, where \( F_{k} = (\text{dist}_{a_{1}}, \cdots, \text{dist}_{a_{k}}) \) is the map associate to the maximal \( k \)-frame.

Let us chose \( p_{1,\alpha} \in X_{\alpha} \to p_{1} \). Clearly, \( \{[a_{j,\alpha}b_{j,\alpha}]\}_{j=1}^{k_{1}} \) is also a \( \delta^{2} \)-maximal \( k_{1} \)-frame at \( p_{1,\alpha} \).

We define \( r_{1,\alpha} = 3 \diam X_{\alpha} F_{k_{1},\alpha}^{-1}(F_{k_{1},\alpha}(p_{1,\alpha})) \), i.e., the extrinsic diameter of a regular fiber of the Perelman’s fibration.

Step 1. Let \( Y_{2,\alpha} = F_{k_{1},\alpha}^{-1}(F_{k_{1},\alpha}(p_{1,\alpha})) \), and let \( \theta_{2,\alpha} = \frac{1}{7} r_{1,\alpha} = \diam X_{\alpha} Y_{2,\alpha} \).

Since \( \theta_{2,\alpha} \to 0 \), it is easy to see \( Y_{2,\alpha} \) is connected.

By passing to a subsequence, we assume the rescaled sequence

\[
\theta_{2,\alpha}^{-1}(X_{\alpha}, p_{1,\alpha}) \to (A_{2} \times \mathbb{R}^{k_{1}}, p_{2}').
\]

Moreover, as subsets, \( Y_{2,\alpha} \) converges to \( A_{2} \times \{0\} \).

Assume \( \dim A_{2} = k_{2} \). Let \( b_{k_{1}+1,\alpha} = p_{1,\alpha} \), and \( a_{k_{1}+1,\alpha} \in Y_{2,\alpha} \) be the farthest point away from \( p_{1,\alpha} \). Starting with \( \{[a_{k_{1}+1,\alpha}b_{k_{1}+1,\alpha}]\} \), we construct a \( \delta^{2} \)-maximal frame

\[
\{[a_{k_{1}+j,\alpha}b_{k_{1}+j,\alpha}]\}_{j=1}^{k_{2}}.
\]
which by the same argument as in the Starting Step, converges to a $k_2$-frame in $A_2 \times \{0\}$,
\[
\{|a_{k_1+j}, b_{k_1+j}|\}_{j=1}^{k_2}.
\]
Let $p_2$ in $A_2 \times \{0\}$ such that $|p_2 m_{k_1+k_2}| \leq \frac{1}{\delta_2} \delta |a_{k_1+k_2} b_{k_1+k_2}|$.
We similarly define $p_{2,\alpha} \in X_{\alpha} \rightarrow p_2, R_{2,\alpha} = \theta_{2,\alpha} \cdot R_2$, where there is a
Perelman’s fibration $F_{k_1+k_2}^{-1} \circ F_{k_1+k_2,\alpha} : \theta_{2,\alpha}^{-1} X_{\alpha} \rightarrow A_2 \times \mathbb{R}^{k_1}$ over $B_{2R_2}(p_2)$,
which is bi-Lipschitz to an open ball $D^{k_1+k_2}$ in the Euclidean space $\mathbb{R}^{k_1+k_2}$
with bi-Lipschitz constant almost 1. (Note that every component in $F_{k_1}$ is
a canonical Busemann function on $\mathbb{R}^{k_1}$.)
Let $r_{2,\alpha} = 3 \text{diam}_X F_{k_1+k_2,\alpha}^{-1}(F_{k_1+k_2,\alpha}(p_{2,\alpha}))$.

Step 2. Do the same process as in Step 1 for $Y_{3,\alpha} = F_{k_1+k_2,\alpha}(F_{k_1+k_2,\alpha}(p_{2,\alpha}))$, and
$\theta_{3,\alpha} = \frac{1}{3} r_{2,\alpha} = \text{diam}_X Y_{3,\alpha}$.
Let us repeat the process in Step 2 until $k_1 + \cdots + k_l = n$, then we have constructed a $\delta^2$-maximal $n$-frame
\[
\{|a_{j,\alpha} b_{j,\alpha}|\}_{j=1}^{n},
\]
centered at $p_{l,\alpha}$.
Let $q_{\alpha} = p_{l,\alpha}$, then (5.1.1) is complete.

(5.1.2) By definition, each $\pi_{i}^L(p_{l,\alpha}; R_{i+1,\alpha})$ satisfies $(\epsilon_\alpha, \sigma, l)$-leveled gap property, where $\epsilon_\alpha = \frac{\rho_{i,\alpha}}{\text{diam}_{i,\alpha}} \rightarrow 0, \sigma = \max \{|R_2^{-1}, \cdots, 3R_1^{-1}|\}.

Indeed, in order to verify (5.1.2), it suffices to show that
\[
\pi_{i}^L(p_{l,\alpha}; R_{i+1,\alpha}) < \pi_{i}^L(p_{l,\alpha}; r_{i,\alpha})
\]
(3.15) follows from the Hurewicz fibration Theorem 1.3. Indeed, let $D_{i,\alpha} = F_{k_1+\cdots+k_1,\alpha}(D^{k_1+\cdots+k_1})$. Then
\[
\pi_{i}^L(p_{l,\alpha}; R_{i,\alpha}) \equiv \pi_{i}^L(p_{l,\alpha}, r_{i,\alpha}) \equiv \pi_1(D_{i,\alpha}, p_{l,\alpha}).
\]
By the choice of $p_{l,\alpha}, \theta_{i+1,\alpha}^{-1}(D_{i,\alpha}, p_{l,\alpha}) \xrightarrow{GH} (A_{i+1} \times \mathbb{R}^{k_1+\cdots+k_1}, p_{i+1})$, where
$p_{i+1}$ is a regular point in $A_{i+1} \times \{0\}$. Let $(\hat{D}_{i,\alpha}, \hat{p}_{l,\alpha})$ $\pi_{\alpha_{\alpha}}^L (D_{i,\alpha}, p_{l,\alpha})$ be a
cover of $D_{i,\alpha}$ with
\[
\pi_1(\hat{D}_{i,\alpha}, \hat{p}_{l,\alpha}) = \pi_{i}^L(p_{l,\alpha}, r_{i,\alpha}).
\]
We are to show that $\pi_{i,\alpha}$ is normal.
Let $S_{i,\alpha}$ of be a short basis of $\pi_{i}^L(p_{l,\alpha}, r_{i,\alpha})$. By passing to a subsequence, we assume that for the same $t_i, \{\gamma_{i,\alpha,1}, \cdots, \gamma_{i,\alpha,t_i}\} = S_{i,\alpha} \setminus \pi_{i}^L(p_{l,\alpha}, r_{i+1,\alpha})$.
By the definition of a short basis, their lifting curve $\hat{\gamma}_{i,\alpha,1}, \cdots, \hat{\gamma}_{i,\alpha,t_i}$ are
minimal geodesics in $\hat{D}_{i,\alpha}$ from $\hat{p}_{l,\alpha}$ to some $\hat{q}_{i+1,\alpha,1}, \cdots, \hat{q}_{i+1,\alpha,t_i}$, respectively.

It suffices to show that for any loop $\gamma \in \pi_{i}^L(p_{l,\alpha}; r_{i,\alpha}),$ and any $\gamma_{i,\alpha,s},$
there is a homotopy with fixed endpoint from $\gamma_{i,\alpha,s} \ast \gamma_{i,\alpha,s}^{-1}$ to a loop in
\[
\pi_{i}^L(p_{l,\alpha}; r_{i+1,\alpha}).
\]
By passing to a subsequence, $\theta_{i+1,\alpha}^{-1}(\hat{D}_{i,\alpha}, \hat{p}_{l,\alpha}) \rightarrow (A_{i+1} \times \mathbb{R}^{k_1+\cdots+k_1}, \hat{p}_{i+1})$.
And each minimal geodesic $\hat{\gamma}_{i,\alpha,s}$ converges to $\hat{\gamma}_{i,\alpha}$ in $\hat{A}_{i+1} \times \{0\}$, which is a
minimal geodesic form $\hat{p}_{i+1}$ to $\hat{q}_{i+1,\alpha,s}$.
If \( \hat{\gamma}_{i,1}, \ldots, \hat{\gamma}_{i,t} \) pass only regular points, then by [4] there is a positive \( \eta > 0 \) such that the neighborhood \( B_{2\eta} = U_{2\eta}(\bigcup_{s=1}^{t} \hat{\gamma}_{i,s}) \subset \hat{A}_{i+1} \times \mathbb{R}^{k_1+\cdots+k_t} \) contains only \( (k_1 + \cdots + k_{i+1}, \delta) \)-strained points with a universal strainer radius. Thus by Theorem 1.3 and Remark 2.3, for \( \alpha \) large we have

\[
(5.3.1) \quad \text{any } F_{k_1+\cdots+k_i,\alpha} \text{-fiber in } D_{i,\alpha} \text{ has extrinsic diameter not larger than } \eta/4.
\]

(5.3.2) there is a Hurewicz fibration \( \varphi_{i,\alpha} \) which is close to the original GHA, over \( B_{2\eta} \), whose fiber’s diameter \( \leq \eta/4 \), such that \( \hat{\gamma}_{i,\alpha,s} \subset \varphi_{i,\alpha}^{-1}(U_{\eta}(\hat{\gamma}_{i,s})) \) and \( \varphi_{i,\alpha}(\hat{\gamma}_{i,\alpha,s}) \subset U_{\eta/4}(\hat{\gamma}_{i,s}) \).

Note that the lifting of \( \gamma_{i,\alpha,s} \ast \gamma \ast \gamma_{i,\alpha,s}^{-1} \) at \( \hat{q}_{i+1,\alpha,s} \) is \( \hat{\gamma}_{i,\alpha,s} \ast \hat{\gamma} \ast \hat{\gamma}_{i,\alpha,s}^{-1} \) with \( \hat{\gamma} \) a closed lifting of \( \gamma \) at \( \hat{p}_{l,\alpha} \). Then by the construction of \( \varphi_{i,\alpha} \) (see Proposition 2.7), there is a canonical contraction from tubular neighborhood of a \( \varphi_{i,\alpha} \)-fiber to itself. Thus, by (5.3.1) there is a homotopy \( \hat{H}_1 \) maps \( \hat{\gamma} \) to \( \hat{\gamma}' \subset \varphi_{i,\alpha}^{-1}(\varphi_{i,\alpha}(\hat{p}_{l,\alpha})) \) keeping \( \hat{p}_{l,\alpha} \) unmoved. Thus, we have a homotopy \( \hat{H}_1 \) maps \( \hat{\gamma}_{i,\alpha,s} \ast \hat{\gamma} \ast \hat{\gamma}_{i,\alpha,s}^{-1} \) to \( \hat{\gamma}_{i,\alpha,s} \ast \hat{\gamma} \ast \hat{\gamma}_{i,\alpha,s}^{-1} \), keeping \( \hat{\gamma}_{i,\alpha,s} \) and \( \hat{\gamma}_{i,\alpha,s}^{-1} \) unmoved.

Furthermore, by (5.3.2) there is a homotopy \( \hat{H}_2 \) maps \( \hat{\gamma}' \) to \( \hat{\gamma}'' \), which lies in \( \varphi_{i,\alpha}^{-1}(\varphi_{i,\alpha}(\hat{q}_{i+1,\alpha,s})) \), moving \( \hat{p}_{l,\alpha} \) to \( \hat{q}_{i+1,\alpha,s} \) along \( \hat{\gamma}_{i,\alpha,s} \) such that \( \hat{\gamma}'' \) is a loop at \( \hat{q}_{i+1,\alpha,s} \). Thus, we have a homotopy \( \hat{H}_2 \) maps \( \hat{\gamma}_{i,\alpha,s} \ast \hat{\gamma} \ast \hat{\gamma}_{i,\alpha,s}^{-1} \) to \( \hat{\gamma}'' \), keeping \( \hat{q}_{i+1,\alpha,s} \) unmoved.

Then \( \pi_{i,\alpha}(\hat{H}_2 \ast \hat{H}_1) \) is a homotopy maps \( \gamma_{i,\alpha,s} \ast \gamma \ast \gamma_{i,\alpha,s}^{-1} \) to \( \pi_{i,\alpha}(\hat{\gamma}'') \) keeping \( \hat{p}_{l,\alpha} \) unmoved, and \( \pi_{i,\alpha}(\hat{\gamma}'') \) lies in \( \pi_{i,\alpha}^{-1}(\pi_{i,\alpha}(\hat{q}_{i+1,\alpha})) \).

In order to complete the proof of (3.15), we now verify that all limit minimal geodesics \( \hat{\gamma}_{i,s} \) pass regular points. Firstly, it is clear that the limit projection \( \pi_{i} : \hat{A}_{i+1} \times \{0\} \to A_{i+1} \times \{0\} \) is a submetry (i.e., 1-LcL). Secondly, there is a neighborhood of \( \hat{p}_{i+1} \) restricted on which \( \pi_{i} \) is an isometry. This is because near \( \hat{p}_{l,\alpha} \), there is a homeomorphic lifting of \( D_{i,\alpha} \) in \( \hat{D}_{i,\alpha} \). Hence \( \dim \hat{A}_{i+1} = \dim A_{i+1} \), and all lift points \( \hat{q}_{i+1}, \hat{q}_{i+1,1}, \ldots, \hat{q}_{i+1,t} \) are regular. By [4], any minimal geodesic \( \hat{\gamma}_{i,s} \) between them contains only regular points.

The proof of (5.1.2) is now complete.

(5.1.3) By (3.15), \( (\hat{D}_{i,\alpha}, \hat{p}_{l,\alpha}) \xrightarrow{\pi_{i,\alpha}} (D_{i,\alpha}, p_{l,\alpha}) \) is a normal cover, whose deck-transformation group is \( \Lambda_{i,\alpha} = \pi_{i,\alpha}^{-1}(\pi_{i,\alpha}(r_{i,\alpha}))/\pi_{i,\alpha}^{-1}(\pi_{i,\alpha}(r_{i+1,\alpha})) \).

Since \( \Lambda_{i,\alpha} \) equivariantly converges, the limit group \( \Lambda_{i} \) acts on \( \hat{A}_{i+1} \times \{0\} \) isometrically. By the generalized Bieberbach theorem [9] (cf. [32]), \( \Lambda_{i} \) is \( C \)-abelian. Since \( \Lambda_{i} \) is a discrete group, the GHA \( p_{l,\alpha} \) between \( \Lambda_{i,\alpha} \) and \( \Lambda_{i} \) is a homomorphism.

We now prove that \( p_{l,\alpha} \) is an isomorphism. Firstly, since there is no non-trivial element of \( \Lambda_{i,\alpha} \) whose displacement is shorter than \( 2R_{i+1,\alpha} \), \( p_{l,\alpha}'s \) kernel should be a subgroup \( K_{i,\alpha} \), which moves \( \hat{p}_{l,\alpha} \) to infinity. Secondly, since \( \pi_{i,\alpha}^{-1}(p_{l,\alpha}(r_{i,\alpha})) \) is generated by all of its elements whose displacements are not longer than \( 3r_{i,\alpha} \) and any generating relation can be written as a word in these elements with wordlength \( \leq 3 \), the corresponding property holds for \( \Lambda_{i,\alpha} \). Because the relative volume comparison (see [18]) provides an uniform bound to the number of \( \Lambda_{i,\alpha} \)-orbit points in \( B_{g_{r_{i,\alpha}}} \), by
passing to a subsequence, the presentation of $\Lambda_{i,\alpha}$ is stable. Hence $\rho_{i,\alpha}$ is an isomorphism.

6. Appendix on gradient push

Let $\{(a_j, b_j)\}_{j=1}^n$ be a $\delta^2$-maximal $n$-frame with $|a_1 b_1| \leq 1$ on an Alexandrov $n$-space $X$ with curv $\geq -1$. Let $F_k(x) = (\text{dist}_{a_1}(x), \text{dist}_{a_2}(x), \cdots, \text{dist}_{a_k}(x))$. Recall that by the definition of maximal frame, $d_j = |a_j b_j|$ which satisfies

$$d_{k+1} = \max_{F_k^{-1}(F_k(b_{k+1}))} \{\min\{\text{dist}_{b_{k+1}}, \delta^2 \cdot \min_{i=1,\ldots,k} |a_i b_i|\}\}.$$  

(4.1)

In the following we always assume that $m_k$ is the middle point of $[a_k b_k]$, and $b_{n+1}$ is a point $\frac{\delta}{100}|a_n b_n|$-close to the middle point $m_n$ of $[a_n b_n]$.

We restate Theorem 1.6 and give a proof in the following form.

**Theorem 6.1** ([15]). There is $T(n) > 0$ such that for any $\delta^2$-maximal $n$-frame $\{(a_j, b_j)\}_{j=1}^n$ with $|a_1 b_1| \leq 1$ on an Alexandrov $n$-space $X$ with curvature $\geq -1$, any point $b_{n+1}$ that is $\frac{\delta}{100}|a_n b_n|$-close to the middle point $m_n$ of $[a_n b_n]$ can be pushed successively by the gradient flows of $\frac{\delta}{n} \text{dist}^2_{b_{n+1}}$, $\frac{1}{2} \text{dist}_{a_j}^2$, $\frac{1}{2} \text{dist}_{b_j}^2$ $(j = 1, \ldots, n)$ to any point $p \in B_{\frac{100}{|a_1 b_1|}(b_{n+1})}$ in total time $\leq T(n)$.

**Remark 6.2.** By the proof of Theorem 6.1 (or by replacing $\frac{\delta}{n} \text{dist}^2$ with dist in Theorem 6.1), the length of broken gradient curves between $b_{n+1}$ and $p$ is no more than $C(n) \cdot |b_{n+1} p|$, where $C(n) = 2n + \frac{4(n-1)}{\sin \sigma(n)} + 1$ with $\sigma(n)$ in Lemma 6.5. This fact is also used in the proof of Theorem 1.1; see Lemma 4.3.

Some partial motivations to write a proof other than just referring to [15, Lemma 2.5.1] are as follows.

Firstly, there is only a sketched proof for Theorem 6.1 in [15], where the ratio bound on the pushing time $\frac{b_{n+1}}{d_k} \leq \frac{1}{\delta^2}$ from level $d_k$ to $d_{k-1}$ is claimed without explanation in the proof of [15, Lemma 2.5.1].

Since it is hard for us to follow at that point, we write a detailed proof on the surjectivity and universal speed of gradient pushing-out (using the maximum property (4.1) and $\frac{1-2n \alpha}{\sqrt{n}}$-openness in [4, Theorem 5.4]). In particular, our proof leads to a sharpened universal time bound $n^2 \delta^{-1}$, improving the universal time bound $\delta^{-n^2}$ claimed in [15].

Secondly, a crucial difference between an Alexandrov space $X$ with curvature bounded below and a Riemannian manifold $M$ is that, there may be proper extremal subsets in $X$ such that no gradient curves can get out of them. Without further explanation, it is also hard for us to see from [15] that the gradient pushing-out process can be chosen to avoid extremal subsets.

We fill more details and construct a specific gradient pushing broken line, consisting of $k$-regular (i.e., the tangent cone $T_p X$ at least splits off $\mathbb{R}^k$) or $(n, \delta)$-strained points when $a_j, b_j$ and the ending point $p$ are $k$-regular.
Since all our estimates will hold for a new $n$-frame $\{[a'_j b'_j]\}_{j=1}^n$, where $a'_j$ and $b'_j$ are nearby regular points around $a_j$ and $b_j$. It follows that gradient push between regular points only passes through regular points.

This provides a detailed justification for the gradient push in proving the Margulis lemma on an Alexandrov space.

6.1. **Proof of Theorem 6.1.** The proof of Theorem 6.1 can be divided into two steps.

**Step 1.** Prove that in at most a definite time $T(\delta, n)$, $b_{n+1}$ can be pushed to any point in a ball $B_{\frac{\delta}{2}d_n}(m_n)$ whose radius is at a small but fixed relative scale, where $m_n$ is the middle point of $[a_nb_n]$.

**Lemma 6.3.** For $0 < \delta < \delta(n)$ and any $q \in B_{\frac{\delta}{2}d_n}(m_n)$, $b_{n+1}$ can be pushed by an at most countably succession of the gradient flows of $\frac{1}{2}\text{dist}_{a_j}^2, \frac{1}{2}\text{dist}_{b_j}^2$ to $q$ in time $\leq \frac{12n^2}{4n-1}\delta$.

Compared with the proof of [4, Theorem 5.4], Lemma 6.3 follows from certain reversing argument, which will be given at the end of the appendix.

**Step 2.** Prove that $B_{\frac{\delta}{2}d_n}(m_n)$ can be pushed outside further. If $\frac{d_m}{m}$ admits a definite lower bound $\tau$, then one may push $B_{\frac{\delta}{2}d_n}(m_n)$ onto $B_{100d}(m_n)$ by $\frac{1}{2}\text{dist}_{m_n}^2$ just one more time taking no more than $T(\tau, n)$. However, $d_n$ may be far less than $1$, or even $d_n-1$.

To overcome this difficulty, we divided an $n$-frame into several levels. We say that a $\delta^2$-maximal $n$-frame $\{[a_j b_j]\}_{j=1}^n$ is of $(\frac{\delta^2}{100}, l)$-leveling if there is $1 \leq k_1 < \cdots < k_l = n$ such that $[a_{k_i-1} b_{k_i-1+1}], \ldots, [a_{k_i} b_{k_i}]$ lies in the same level in the sense that $d_j > \frac{\delta^2}{100}d_{j-1}$ for any integer $k_{i-1} + 1 \leq j \leq k_i$, $1 \leq i \leq l$ ($k_0 = 0$), and $[a_{k_i} b_{k_i}], [a_{k_i+1} b_{k_i+1}]$ lie in different levels, i.e.,

$$d_{k_i+1} \leq \frac{\delta^2}{100}d_{k_i}$$

for any $1 \leq i \leq l$.

Inside each $i$-th level, it follows from elementary gradient estimate that $B_{\frac{\delta}{2}d_{k_i}}(b_{k_i+1})$ can be pushed by the center $b_{n+1}$, i.e., $\frac{1}{2}\text{dist}_{b_{n+1}}^2$, onto $B_{100d_{k_i+1}}(b_{n+1})$ in time $\leq \ln(\frac{100}{\delta^2})^{k_i-(k_{i-1}+1)}$.

In order to push $B_{100d_{k_i+1}}(b_{n+1})$ further outside onto a large leveled ball in a specific way, we need to prove the following lemma.

**Lemma 6.4.** If $d_{k+1} = |a_{k+1} b_{k+1}| \leq \frac{\delta^2}{100}d_k$, then for any $p \in B_{\frac{\delta}{2}d_k}(b_{k+1})$, there is some point $q \in B_{50d_{k+1}}(b_{k+1})$ which can be pushed successively along finitely-broken geodesics, each of which is pointing to one of $\{a_j, b_j\}_{j=1}^n$, by the gradient flows of $\frac{1}{2}\text{dist}_{a_j}^2, \frac{1}{2}\text{dist}_{b_j}^2$ to $p$ in time $\leq C(n)\delta$.

Note that in the case of Lemma 6.4 for different level, we are using endpoints of long edges in the frame, which lie outside the small ball $B_{50d_{k+1}}(b_{k+1})$.

In the proof of Lemma 6.4, the core is the following angle estimate, which follows from the numerical maximum property (4.1) of $\delta^2$-maximal frame.

**Lemma 6.5 (Angle Estimate).** There is $\delta(n) > 0$ such that the following holds for $0 < \delta < \delta(n)$. If $d_{k+1} = |a_{k+1} b_{k+1}| \leq \frac{\delta^2}{100}|a_k b_k|$, then for
any $p \in B_{\frac{4}{5}d_k}(b_{k+1}) \setminus B_{50d_k}(b_{k+1})$, there exists $e \in \{a_j, b_j\}_{j=1}^k$ such that

$$\angle (p; e, b_{k+1}) \leq \frac{\pi}{2} - \sigma(n).$$

**Proof.** Argue by contradiction. For any $\delta > 0$, there is a $\delta^2$-maximal k-frame such that the conclusion of Lemma 6.5 fails. Then by Toponogov comparison (cf. [4, Lemma 5.6]), there is $\sigma = \sigma(\delta) \to 0$ as $\delta \to 0$ such that for any $x \in [b_{k+1}]$ and every $1 \leq i \leq k$,

$$\angle(x; a_i, b_{k+1}) \geq \frac{\pi}{2} - \sigma,$$

and

$$\angle(x; b_i, b_{k+1}) \geq \frac{\pi}{2} - \sigma.$$

Then as $\sigma$ sufficient small,

(6.5.1) \[ |a_i x| - |a_j b_{k+1}| \leq |xb_{k+1}| \cdot \sin \sigma. \]

By [4, Theorem 5.4] (or see Theorem 6.6 below), for any $0 \leq \sigma \leq \frac{1}{2k}$ the partial distance coordinates map associated to $k$-subframe $\{[a_j b_j]\}_{j=1}^k$,

$$F_k : X \to \mathbb{R}^k, \quad F_k(x) = (|a_1 x|, |a_2 x|, \ldots, |a_k x|),$$

is $\frac{1-2\sigma}{\sqrt{k}}$-open, i.e., $\frac{1-2\sigma}{\sqrt{k}}$-co-Lipschitz, on $B_{50d_k}(b_{k+1})$. Hence there is $x' \in F_k^{-1}(F_k(b_{k+1})) \cap B_{3d_k}(b_{k+1})$ such that the distance

(6.5.2) \[ |x x'| \leq \frac{\sqrt{k}}{1 - 2k\delta} \cdot |F_k(x) - F_k(x')|, \]

which is, by (6.5.1), far less than $|xb_{k+1}|$. Let $|xb_{k+1}| = 50d_{k+1}$, then as $\delta = \delta(n)$ sufficient small,

$$|xx' b_{k+1}| >> d_{k+1},$$

a contradiction to the choice of $(a_{k+1}, b_{k+1})$ in (4.1). \qed

We now prove Lemma 6.4.

**Proof of Lemma 6.4.**

Let $e = e(p)$ be one of $\{a_j, b_j\}_{j=1}^k$ provided by Lemma 6.5, and let us connect $p$ and $e$ by a minimal geodesic $[pe]$. By Toponogov comparison and Lemma 6.5, there is a universal $\Delta r$ determined by the $(-1)$-law of cosine such that for any $p' \in [pe]$ with $|pp'| \leq \Delta r$, one has

$$0 < \frac{|b_{k+1}p| - |b_{k+1}p'|}{|pp'|} \leq \sin \sigma(n).$$

If $p'$ can be chosen that $|pp'| \cap B_{50d_{k+1}}(b_{k+1}) \neq \emptyset$, then $x$ is one of the intersection point and the geodesic $[xp]$ is the gradient flow of $\frac{1}{2}\text{dist}^2$. Otherwise, let $p' = p$ with $|pp'| = \Delta r$. By repeating the process above successively, we get a finitely-broken geodesic from $p$ to some point $q \in B_{50d_{k+1}}(b_{k+1})$, whose reverse realizes the geodesic flows from $q$ to $p$ by endpoints $\{a_j, b_j\}_{j=1}^k$.

Because for each $p'$ above, $|p' e(p')| \geq \frac{\delta}{2}d_k - \frac{\delta}{100}d_k$, and the total length of the broken geodesic is bounded by $\frac{1}{\sin \sigma}d_k$, this completes the proof of Lemma 6.4. \qed
Now we are ready to prove Theorem 6.1.

**Proof of Theorem 6.1.**

Let us assume that the $n$-frame $\{a_j b_j\}_{j=1}^n$ admit a $(\frac{\delta^2}{400}, l)$-leveling, $1 \leq k_1 < \cdots < k_l = n$. Let $k_0 = 0$.

By Lemma 6.3, $b_{n+1}$ can be pushed onto $B_\delta(b_{n+1})$ in time $\leq 16\delta \frac{4n^2}{4n-1}$.

In each $i$-th level, $B_\delta(d_{k_i}(b_{n+1})$ can be pushed by $\frac{1}{2}$ dist$_{b_{n+1}}$ onto $B_{100d_{k_{i-1}+1}}(b_{n+1})$ in time $\leq \ln(\frac{400}{4} \frac{100}{d_{k_i}} (k_i-1))$.

From $i$-th level to $(i-1)$-th level, note that for any $100d_{k_{i-1}+1} \leq r \leq \frac{\delta}{4}d_{k_{i-1}}$, $B_\delta(b_{k_{i-1}+1}) \subset B_r(b_{n+1}) \subset B_{2r}(b_{k_{i-1}+1})$. By Lemma 6.4, $B_{100d_{k_{i-1}+1}}(b_{n+1})$ can be pushed onto $B_{\frac{\delta}{4}d_{k_{i-1}}}(b_{n+1})$ in time $\leq C(n)\delta$.

Since it finishes after $2l$-steps, the proof completes. □

### 6.2. Tracing back process.

For completeness we give a proof for the co-Lipschitzness of $F_k : X \to \mathbb{R}^k$, which has been used in proving Lemma 6.5. Lemma 6.3 also follows similarly. The idea of proof is just the same as that of [4, Theorem 5.4].

**Theorem 6.6 ([4, Theorem 5.4]).** There is $\sigma(k) > 0$ such that the following holds for $0 \leq \sigma \leq \sigma(k)$.

Let $\{a_j, b_j\}_{j=1}^k$ be a $(k, \sigma)$-strainer at $x_0$ with radius

$$r_k = \min\{|a_j x_0|, |b_j x_0|\}_{j=1}^k \leq \max\{|a_j b_j|\}_{j=1}^k \leq 1.$$

Let $F_k : X \to \mathbb{R}^k$, $F_k(x) = (|a_1 x|, \ldots, |a_k x|)$, be the map associated to $\{a_j, b_j\}_{j=1}^k$ that forms a partial distance coordinates around $x_0$.

Let $p = p_0$ be a point in $B_{\frac{\delta}{4}d_{k}(x_0)}$ such that

$$(6.6.1) \quad ||a_j p| - |a_j x_0|| \leq \frac{1}{4k}|px_0| \quad (j = 1, \ldots, k).$$

Then there is a (infinitely-)broken geodesic $[p_0 p_1 \cdots p_k p_1 \cdots p_k p_2 \cdots]$, contained in $B_{\frac{\delta}{4}d_{k}(x_0)}$ such that the endpoint $p_l = p_k^l$ converges to a point $p'$ as $l \to \infty$, which satisfies

$$\text{(6.6.2)} \quad |pp'| \leq \frac{4k + 1}{3\sqrt{k}} |F_k(p)F_k(x_0)|, \quad F_k(p') = F_k(x_0).$$

Let $\delta > 0$ be a small number other than $\sigma$. Let $p$ be a point in $B_{5d_{k}(x_0)}$.

Let us first define its $l$-th round $k$-tracing back point $p_l = p_k^l$ of $p$ towards $x_0$’s $F_k$-fiber inductively as follows. Here tracing back means moving along gradient curves of distance to $a_j$ or $b_j$ backwards.

Let $p_0 = p$ and let us assume that $p_{l-1}$ is well-defined. For the first coordinate function $f_1 = \text{dist}_{a_1}$, let $p_1^l$ be a point lies in the broken geodesic $[a_1 p_{l-1} b_1]$ such that

$$f_1(p_1^l) - f_1(p_{l-1}) = f_1(x_0) - f_1(p_{l-1}).$$

Let $p_2^l$ be a point lies in the broken geodesic $[a_2 p_{l-1} b_2]$ such that

$$f_2(p_2^l) - f_2(p_{l-1}^1) = f_2(x_0) - f_2(p_{l-1}).$$
Repeating $k$-times, we have $p_{l-1}^k$ in $[a_k p_{l-1}^{k-1} b_k]$ such that
\[ f_k(p_l) - f_k(p_{l-1}^{k-1}) = f_k(x_0) - f_k(p_{l-1}). \]

Then $p_l$ is defined to be $p_l^k$.

Since $p \in B_{\delta d_k}(x_0)$, by an elementary angle estimate [4, Lemma 5.6], the following holds for $0 < \delta < \frac{\pi}{20}$: for any $1 \leq i \leq k$, $1 \leq j \leq i - 1$,
\[ |\angle(p; a_j, a_i) - \frac{\pi}{2}| \leq 4\sigma, \quad |\angle(p; b_j, a_i) - \frac{\pi}{2}| \leq 4\sigma, \]
\[ |\angle(p; a_j, b_i) - \frac{\pi}{2}| \leq 4\sigma, \quad |\angle(p; b_j, b_i) - \frac{\pi}{2}| \leq 4\sigma. \]

Clearly, it follows that the relations below hold.

**Lemma 6.7.** For some positive function $\epsilon = \epsilon(\sigma) \rightarrow 0$ as $\sigma \rightarrow 0$,

\begin{align*}
(6.7.1) \quad |p_l^j p_l^{j-1}| & \leq (1 + \epsilon) \cdot |f_i(x_0) - f_i(p_l)|; \\
(6.7.2) \quad |f_j(p_l^j) - f_j(p_l^{j-1})| & \leq \epsilon \cdot |f_i(x_0) - f_i(p_{l-1})| \text{ for any } j \neq i.
\end{align*}

Now we are ready to prove Theorem 6.6.

**Proof of Theorem 6.6.**

Let $\delta = \frac{\pi}{20}$. Let $A_l = \sum_{j=1}^{k} |f_j(p_l) - f_j(x_0)|$ and $B_l = |p_{l+1} p_{l}|$. As long as the $l$-th tracing back point $p_l$ lies in $B_{2\delta d_k}(x_0)$, the estimates (6.7.1)-(6.7.2) hold. By triangle inequality, we derive $A_{l+1} = \epsilon(k-1) A_l$ and $B_l = (1 + \epsilon) A_l$. As $\delta$ sufficient small, $\epsilon \leq \frac{1}{10}$ so that $A_{l+1} \leq \frac{1}{4} A_l$ and $B_l \leq \frac{4k+1}{4k} A_l$, and thus $A_l \leq \frac{1}{4} A_0$ and $B_l \leq \frac{4k+1}{4k} \cdot \frac{1}{4} A_0$ are Cauchy sequences.

Now let us check that, by induction on $l$, each $p_l$ satisfies $|p_l x_0| \leq \frac{3}{2} \delta d_k$ so that $p_l \in B_{2\delta d_k}(x_0)$. By the assumption (6.6.1), $A_0 \leq \frac{1}{4} |p x_0|$, and thus $A_l \leq \frac{1}{4^l + 1} |p x_0|$, $B_l \leq \frac{4k+1}{4k} \cdot \frac{1}{4^l + 1} |p x_0|$. Then
\[ \sum_{l=0}^{\infty} B_l \leq \frac{4k+1}{4k} \cdot \frac{1}{3} |p x_0| \leq \frac{1}{2} |p x_0|, \]
which justifies $|p_l x_0| \leq \frac{3}{2} \delta d_k$.

Let $p'$ be the limit point of $p_l$, then
\[ |pp'| \leq \sum_{l=0}^{\infty} B_l \leq \frac{4k+1}{3k} A_0 \leq \frac{4k+1}{3\sqrt{k}} |F_k(x_0) F_k(p)|. \]

The conclusion of Theorem 6.6 now follows. \qed

### 6.3. Proof of Lemma 6.3

In this subsection we prove that a gradient push can be started from $b_{n+1}$ to any point in a very small ball in a definite short time.

Note that if we set $k = n$, $x_0 = q$ and $p = b_{n+1}$ in Theorem 6.6, then $b_{n+1}$ can be moved to $q$ along gradient curves of $\frac{1}{2} \text{dist}_{a_i}^2$, $\frac{1}{2} \text{dist}_{b_j}^2$ backwards. So we need to reverse the tracing back process defined in the proof of Theorem 6.6.
Let $m_n$ be the middle point of $[a_n, b_n]$ of a $\delta^2$-maximal $n$-frame $\{[a_j, b_j]\}_{j=1}^n$.
Let $b_{n+1}$ be a point $\frac{d}{10n}d_n$ close to $m_n$. For any $q \in B_{\frac{d}{2}d_n}(b_{n+1})$, the $l$-th round pushing forward point $O_l$ from $O_0 = b_{n+1}$ towards $q$ is defined inductively as follows.

Assume that $O_l$ is well-defined. By tracing back $q_{l+1}^0 = q$ to $O_l$ by a single round, we have the $n$-tracing points and tacking broken geodesic $[q_{l+1}^0, q_{l+1}^1 \cdots q_{l+1}^n]$, where $q_{l+1}^i \in [a_i, q_{l+1}^{i-1}]$ or $[b_i, q_{l+1}^{i-1}]$ ($i = 1, \ldots, n$). Let $\Phi_{l+1}$ be the successive gradient flow defined by

$$\Phi_{l+1} = \Phi_{l+1,1} \circ \Phi_{l+1,2} \circ \cdots \circ \Phi_{l+1,n} : X \to X,$$

where $\Phi_{l+1,i}$ is the gradient flow of $\frac{1}{2} \text{dist}_{a_i}^2$ or $\frac{1}{2} \text{dist}_{b_i}^2$ which maps $q_{l+1}^i$ to $q_{l+1}^{i-1}$. We define $O_{l+1} = \Phi_{l+1}(O_l)$.

By (6.7.1), it is easy to see that the total time satisfies

$$T_{l+1} = \sum_{i=1}^{n} t_{l+1,i} \leq \frac{4(1 + \epsilon)}{d_n} \sum_{i=1}^{n} |f_i(q) - f_i(O_l)|.$$

**Proof of Lemma 6.3.**

It suffices to show that the $l$-th round pushing forward point $O_l$ towards $q$ converges to $q$, and the total time admits the bound in Lemma 6.3.

Let $A_t = |qO_t|$ and $B_t = \sum_{i=1}^{n} |f_i(q) - f_i(O_t)|$. Then (4.8) can be rewritten as $T_{l+1} \leq \frac{4(1+\epsilon)}{d_n} B_t$.

We first assume that $O_t$ always lies in the cube

$$J_{\delta d_n}(m_n) = \{x \in X : |f_i(x) - f_i(m_n)| \leq \delta d_n, \forall 1 \leq i \leq n\}.$$

Since the Lipschitz constant of distant coordinate function $F_n$ on $J_{\delta d_n}(m_n)$ is almost 1,

$$|q_{l+1}^n O_l| \leq 2 \sum_{i=1}^{n} |f_i(q_{l+1}^n) - f_i(O_l)|,$$

where by (6.7.2)

$$\sum_{i=1}^{n} |f_i(q_{l+1}^n) - f_i(O_l)| \leq \epsilon(n-1) \sum_{i=1}^{n} |f_i(q) - f_i(O_l)| \leq \epsilon(n-1)n|qO_l|.$$

By $|f_i(q) - f_i(O_{l+1})| \leq |qO_{l+1}|$,

$$B_{l+1} = \sum_{i=1}^{n} |f_i(q) - f_i(O_{l+1})| \leq nA_{l+1}.$$

The concavity of $\frac{1}{2} \text{dist}_{x}^2$ with $\text{dist}_{x} < 2$ is bounded by $2\cosh \frac{2}{\sinh}$. By Theorem 3.9, (4.8) and (4.9)-(4.10),

$$A_{l+1} = d(q, O_{l+1}) \leq e^{\frac{-2}{\sinh} \frac{2}{\sinh} T_{l+1}} |q_{l+1}^n O_l|$$

$$\leq e^{\frac{-2}{\sinh} \frac{2}{\sinh} B_{l+1}/d_n} \cdot 2\epsilon(n-1)n|qO_l|$$

$$= e^{\frac{-2}{\sinh} \frac{2}{\sinh} (1+\epsilon)nA_l/d_n} \cdot 2\epsilon(n-1)nA_l.$$
Let us take $\delta(n) > 0$ such that for $0 < \delta \leq \delta(n)$, $A_0/d_n \leq \delta/2 \leq \frac{1}{(1+\epsilon)n}$, and $\epsilon \leq \frac{1}{8n^3 e^{\frac{\delta}{2\sinh^2\delta}}}$.

Then $(1 + \epsilon)n A_0 / d_n \leq 1$. Moreover, $A_1 \leq \frac{1}{4n} A_0 \leq A_0$. By induction, for any $l$, $A_l \leq \frac{1}{(4n)^l} A_0$, $B_l \leq \frac{1}{(4n)^l} A_0$, and $O_l$ lies in $I_{\delta d_n}(m_n)$.

Therefore, all estimates above are valid for $0 < \delta < \delta(n)$, and $A_l \to 0$ as $l \to \infty$, i.e., $O_l \to q$. Moreover, the total time

$$T = \sum_{i=1}^{\infty} T_i \leq 4(1 + \epsilon)n \sum_{i=1}^{\infty} \frac{A_i}{d_n} \leq 2(1 + \epsilon)\delta \frac{4n^2}{4n - 1}.$$

□

References

[1] V. N. Berestovskii and Luis Guijarro. A metric characterization of riemannian submersions. *Ann. Global Anal. Geom.*, 18(6):577–588, 2000.
[2] K. Borsuk. On some metrizations of the hyperspace of compact sets. *Fund. Math.*, 41:168–201, 1955.
[3] E. Breuillard, B. Green, and T. Tao. The structure of approximate groups. *Publications Mathématiques de l' IHÉS*, 116:115–221, 2012.
[4] Y. Burago, M. Gromov, and G. Perelman. A.d. alexandrov spaces with curvature bounded below. *Uspekhi Mat. Nauk*, 47(2(284)):3–51, 1992.
[5] P. Buser and H. Karcher. Gromov’s almost flat manifolds, 1981.
[6] J. Cheeger and T. H. Colding. Lower bounds on ricci curvature and the almost rigidity of warped products. *Ann. of Math.*, 144(1):189–237, 1996.
[7] S. Ferry. Strongly regular mappings with compact anr fibers are hurewicz fiberings. *Pacific J. Math.*, 75(2):373–382, 1978.
[8] K. Fukaya. Collapsing of riemannian manifolds to ones of lower dimensions. *J. Differential Geom.*, 25:139–156, 1987.
[9] K. Fukaya and T. Yamaguchi. The fundamental groups of almost non-negatively curved manifolds. *Ann. of Math.*, 136:253–333, 1992.
[10] M. Gromov. Almost flat manifolds. *J. Differential Geom.*, 13:231–241, 1978.
[11] M. Gromov. Manifolds of negative curvature. *J. Differential Geom.*, 13:223–230, 1978.
[12] M. Gromov. *Metric Structures for Riemannian and Non-Riemannian Spaces*. Modern Birkhäuser Classics, Birkhäuser, Boston, Based on the 1981 French original, With appendices by M. Katz, P. Pansu and S. Semmes, Translated from the French by Sean Michael Bates., 2007.
[13] E. Heintze. *Mannigfaltigkeiten negativer Krümmung*. PhD thesis, Universitit Bonn Habilitationsschrift, 1976.
[14] V. Kapovitch. Perelman’s stability theorem. In *Surveys of Differential Geometry, Metric and Comparison Geometry*, XI, pages 103–136. Int. Press, Somerville, 2007.
[15] V. Kapovitch, A. Petrunin, and W. Tuschmann. Nilpotency, almost nonnegative curvature, and the gradient flow on alexandrov spaces. *Ann. of Math.*, 171:343–373, 2010.
[16] V. Kapovitch and B. Wilking. Structure of fundamental groups of manifolds with ricci curvature bounded below. *Preprint*, arXiv:1105.5955, 2011.
[17] U. Lang and V. Schroeder. On toponogov’s comparison theorem for alexandrov spaces. *arXiv:1207.3668 [math.MG]*, 2012.
[18] N. Li and X.C. Rong. Relatively maximum volume rigidity in alexandrov geometry. *Pacific J. Math.*, 259:387–420, 2012.
[19] X.C. Rong; S.C. Xu L.N. Chen. Quantitative volume space form rigidity under lower ricci curvature bound i. arXiv:1604.06986 [math.DG], to appear in J. Diff. Geom., 2016.
[20] G. Perelman. Alexandrov spaces with curvatures bounded from below ii. Preprint, 1991.
[21] G. Perelman. Elements of Morse theory on Aleksandrov spaces. St. Petersburg Math. J., 5:205–213, 1993.
[22] G. Perelman. Collapsing with no proper extremal subsets. In K. Grove and P. Petersen, editors, Comparison Geometry, volume 30 of MSRI Books, pages 149–155. Cambridge Univ. Press, Cambridge, 1997.
[23] A. Petrunin. Semiconcave functions in alexandrov’s geometry. In J. Cheeger and K. Grove, editors, Metric and Comparison Geometry, volume XI of Surveys in Differential Geometry, pages 137–202. Int. Press, Somerville, 2007.
[24] C. Plaut. Spaces of wald-berestovskii curvature bounded below. J. Geom. Anal., 6:113–134, 1996.
[25] X. Rong and S. Xu. Stability of $\epsilon$-lipschitz and co-lipschitz maps in gromov-hausdorff topology. Advances in Mathematics, 231:774–797, 1 October 2012.
[26] E. A. Ruh. Almost flat manifolds. J. Differential Geom., 17:1–14, 1986.
[27] Y.S. Wang. A Schur-Toponogov theorem in riemannian geometry & a new proof of Toponogov’s theorem in alexandrov geometry. arXiv:1809.09818 [math.DG], 2018.
[28] H.W. Sun; Y.S. Wang X.L. Su. A new proof of almost isometry theorem in alexandrov geometry with curvature bounded below. Asian J. Math., 17(4):715–728, 2013.
[29] S. Xu. Homotopy lifting property of an $\epsilon$-lipschitz and co-lipschitz map. preprint, 2013.
[30] Z.H. Jiang; S.C. Xu. Stability of pure nilpotent structures on collapsed manifolds. arXiv:1805.06139 [math.DG], to appear in Int. Math. Res. Not., 2018.
[31] T. Yamaguchi. Collapsing and pinching under a lower curvature bound. Ann. of Math., 133:317–357, 1991.
[32] T. Yamaguchi. A convergence theorem in the geometry of Alexandrov spaces, 1996.
E-mail address: shichengxu@gmail.com

School of Mathematical Sciences, Capital Normal University, Beijing, China
E-mail address: qiujuli@vip.sina.com

School of Mathematical Sciences, Capital Normal University, Beijing, China