QUANTITATIVE ESTIMATES ON THE SINGULAR SETS OF ALEXANDROV SPACES

NAN LI AND AARON NABER

ABSTRACT. Let $X \in \mathrm{Alex}^n(-1)$ be an n-dimensional Alexandrov space with curvature ≥ -1. Let the r-scale (k,ϵ)-singular set $\mathcal{S}^k_{\epsilon,r}(X)$ be the collection of $x \in X$ so that $B_r(x)$ is not ϵr-close to a ball in any splitting space $\mathbb{R}^{k+1} \times Z$. We show that there exists $C(n,\epsilon) > 0$ and $\beta(n,\epsilon) > 0$, independent of the volume, so that for any disjoint collection $\{B_{r_i}(x_i) : x_i \in \mathcal{S}^k_{\epsilon,\beta r_i}(X) \cap B_1,\ r_i \leq 1\}$, the packing estimate $\sum r_i^k \leq C$ holds. Consequently, we obtain the Hausdorff measure estimates $\mathcal{H}^k(\mathcal{S}^k_\epsilon(X) \cap B_1) \leq C$ and $\mathcal{H}^n(B_r(\mathcal{S}^k_{\epsilon,r}(X)) \cap B_1(p)) \leq C\, r^{n-k}$. This answers an open question in [8]. We also show that the k-singular set $\mathcal{S}^k(X) = \bigcup_{\epsilon>0} \left(\bigcap_{r>0} \mathcal{S}^k_{\epsilon,r} \right)$ is k-rectifiable and construct examples to show that such a structure is sharp. For instance, in the $k = 1$ case we can build for any closed set $T \subseteq \mathbb{S}^1$ and $\epsilon > 0$ a space $Y \in \mathrm{Alex}^3(0)$ with $\mathcal{S}^1_\epsilon(Y) = \phi(T)$, where $\phi\colon \mathbb{S}^1 \to Y$ is a bi-Lipschitz embedding. Taking T to be a Cantor set it gives rise to an example where the singular set is a 1-rectifiable, 1-Cantor set with positive 1-Hausdorff measure.

CONTENTS

The first author was supported by PSC-CUNY Grant 61533-0049. The second author was supported by NSF grant DMS-1809011.

1. Introduction

Let $\mathrm{Alex}^n(\kappa)$ be the collection of n-dimensional Alexandrov spaces with (sectional) curvature $\geq \kappa$. The aim of this paper is to study the quantitative stratifications of $X \in \mathrm{Alex}^n(\kappa)$. Given $X \in \mathrm{Alex}^n(\kappa)$, it is known that the tangent cone $T_p(X)$ at every point $p \in X$ is a metric cone $C(\Sigma)$, where $\Sigma \in \mathrm{Alex}^{n-1}(1)$ and it is unique. The singular set $\mathcal{S}(X)$ is the collection of points whose tangent cones are not isometric to $\mathbb{R}^n$. It has a natural stratification

$$\mathcal{S}(X) = \mathcal{S}^{n-1}(X) \supseteq \mathcal{S}^{n-2}(X) \supseteq \cdots \supseteq \mathcal{S}^1(X) \supseteq \mathcal{S}^0(X)\,,$$

where

$$\mathcal{S}^k(X) \equiv \{p \in X : \; T_p(X) \text{ is not isometric to } \mathbb{R}^{k+1} \times C(\Sigma) \text{ for any metric space } \Sigma\}\,. \tag{1.1}$$

We may omit the X and write for example $\mathcal{S}^k = \mathcal{S}^k(X)$ if it doesn't cause any ambiguity. Let us first state a notion of strong quantitative singular sets. We will then compare it with those used for the Ricci cases.

Definition 1.1 (Quantitative splitting)**.**

(1) Given a metric space Y and $k \in \mathbb{N}$, we say that Y is k-splitting if Y is isometric to $\mathbb{R}^k \times Z$ for some metric space Z.
(2) Given a metric space X we say that a metric ball $B_r(x) \subseteq X$ is (k, ϵ)-splitting if there exists a k-splitting space Y and $y \in Y$ such that $d_{GH}(B_r(x), B_r(y)) \leq \epsilon r$.

Definition 1.2 (Strong quantitative singular sets)**.** Given $k, \epsilon,\ r > 0$ and metric space X.

(1) The r-scale (k, ϵ)-singular set

$$\mathcal{S}^k_{\epsilon,r}(X) \equiv \{x \in X : B_r(x) \text{ is not } (k+1, \epsilon)\text{-splitting}\}\,. \tag{1.2}$$

(2) The (k, ϵ)-singular set

$$\mathcal{S}^k_\epsilon \equiv \bigcap_{r>0} \mathcal{S}^k_{\epsilon,r} = \{x \in X : B_r(x) \text{ is not } (k+1, \epsilon)\text{-splitting for every } 0 < r \leq 1\,\}. \tag{1.3}$$

It's easy to see that $\mathcal{S}^k = \bigcup_{\epsilon>0} \mathcal{S}^k_\epsilon = \bigcup_{\epsilon>0} \left(\bigcap_{r>0} \mathcal{S}^k_{\epsilon,r}\right)$. A weaker notion of quantitative singular sets, which we will denote by $\mathcal{WS}^k_{\epsilon,r}$, was introduced in [5] for manifolds with lower Ricci curvature bounds, see (5.10) for a definition. A significance for (1.2) is that it requires $B_s(x)$ to be (k, ϵ)-non-splitting only at the scale $s = r$, but not for all $r \leq s \leq 1$ as required in (5.10). It is worth pointing out that notion (1.2) is strictly stronger than (5.10) on manifolds with Ricci curvature bounds, while they are equivalent in some sense on Alexandrov spaces (see Section 5.2). The singular sets defined as in (1.2) are not known to satisfy the estimates established in [4], [5] or [6] for the Ricci cases.

It was proved in [3] that if X is a Gromov-Hausdorff limit of n-dimensional, v-noncollapsed Riemannian manifolds with $\mathrm{Ric} \geq -(n-1)$, then the Hausdorff dimension $\dim_{\mathcal{H}}(\mathcal{WS}^k) \leq k$. Under the same assumptions, it was proved in [4] that for any $0 < r,\ \epsilon \leq 1$, there exists a constant $C(n, v, \epsilon) > 0$ such that for any $p \in X$, it holds that

$$\mathrm{vol}(B_r(\mathcal{WS}^k_{\epsilon,r}(X)) \cap B_{1/2}(p)) \leq C(n, v, \epsilon) r^{n-k}. \tag{1.4}$$

It was also proved in [4] that $\mathcal{WS}^k_\epsilon(X)$ is k-rectifiable. For $X \in \text{Alex}^n(\kappa)$, it is proved in [2] that the Hausdorff dimension $\dim_{\mathcal{H}}(\mathcal{S}^k(X)) \le k$, and it was asked in [8] wether the $(n-2)$-dimensional packing estimate holds for $\mathcal{S}^{n-2}_\epsilon(X)$. In this paper, we prove the k-packing estimates and the k-rectifiability of $\mathcal{S}^k_\epsilon(X)$ for every $0 \le k \le n$. Moreover, all of our estimate are independent on the volume of unit balls in X. Note that it is crucial to have a positive lower volume bound in [4], [5] and [6], to obtain estimates such as (1.4) for manifolds with lower Ricci curvature bounds. It is not known whether the volume dependence can be removed for the Ricci cases.

Theorem 1.3 (Packing estimate)**.** *For any $n \in \mathbb{N}$ and $\epsilon > 0$ there exists $C = C(n,\epsilon) > 0$ and $\beta = \beta(n,\epsilon) > 0$ such that the following hold for any $(X,p) \in Alex^n(-1)$. If $x_i \in \mathcal{S}^k_{\epsilon,\beta r_i}(X) \cap B_1(p)$ and $\{B_{r_i}(x_i)\}$ are disjoint with $r_i \le 1$ for all $i \in \mathbb{I}$, then*

$$\sum_{i\in\mathbb{I}} r_i^k < C. \tag{1.5}$$

In particular, if $x_i \in \mathcal{S}^k_{\epsilon,r}(X) \cap B_1(p)$ and $\{B_r(x_i)\}$ are disjoint with $r \le 1$, then $|\mathbb{I}| < Cr^{-k}$.

Example 1.1. There exists Alexandrov spaces (in fact non-collapsed Gromov-Hausdorff limits of manifolds with sec ≥ 0) whose singular set is dense. Such a space was constructed in [10]. Begin with a regular tetrahedron X_1 in $\mathbb{R}^3$. Suppose convex polyhedra X_k with triangular faces Δ_i, $i = 1, 2 \dots, 4\cdot 3^{k-1}$ has been constructed. Let x_i be the centroid of face Δ_i. Let $y_i \in \mathbb{R}^3$ so that $d(y_i, X_k) = d(y_i, x_i) = d^i_k > 0$. Let Y_i be the tetrahedron formed by y^i and Δ_i. Define $X_{k+1} = X_k \cup (\cup_i Y_i)$. The constants $d^i_k = d^i_k(X_k)$ can be chosen small enough so that X_{k+1} is convex. We have that $\partial X_k \in \text{Alex}^2(0)$ for all k. Thus $Y = \lim_{i\to\infty} \partial X_k \in \text{Alex}^2(0)$. It's easy to see that if all X_k are convex, then $\max_i\{d^i_k\} \to 0$ as $k \to \infty$.

The set of singular points $\mathcal{S}^0(Y) \supseteq \bigcup_{i,k}\{x^i_k\}$ is dense in Y. However, $|\mathcal{S}^0_\epsilon| < N(\epsilon)$, asserted by Theorem 1.3. For this example, we can get an explicit estimate using Gauss-Bonnet formula. For each $p \in Y$, we have that the tangent cone $T_p(Y) = C(\mathbb{S}^1_\beta)$ with $0 < \beta \le 1$. Let $\theta_p = 2\pi\beta$ be the cone angle. Then we have $\mathcal{S}^0_\epsilon = \left\{p \in Y\colon \theta_p \le 2\pi - \epsilon\right\}$. Note that for any $p \in Y$ the Gaussian curvature $K_p \ge 0$ and $K_p = (2\pi - \theta_p)\delta_p$ if $p \in \mathcal{S}^0_\epsilon$, where δ_p is the Dirac delta function at p. By Gauss-Bonnet formula, we have for $\epsilon_i = 2^{-i}$

$$4\pi = \int_Y K \ge \sum_{i=0}^{\infty} \sum_{p\in\mathcal{S}^0_{\epsilon_{i+1}}\setminus\mathcal{S}^0_{\epsilon_i}} (2\pi - \theta_p) \ge \sum_{i=0}^{\infty} \epsilon_{i+1}\, |\mathcal{S}^0_{\epsilon_{i+1}} \setminus \mathcal{S}^0_{\epsilon_i}|.$$

In particular, we have the estimate $|\mathcal{S}^0_\epsilon| \le \frac{4\pi}{\epsilon}$. □

The statement (1.5) doesn't hold without a quantitative control of $\beta = \beta(n,\epsilon)$, if $\inf\{r_i\} = 0$. See the following example.

Example 1.2. Let $X = C(S^1_\rho)$ be a metric cone over a circle with radius $\rho = \frac{1}{20}$. Let p be the cone point and choose points $x_i \in X$, so that $d(p, x_i) = 3^{-i}$, $i = 0, 1, 2, \dots$. Consider disjoint collection $\mathcal{C} = \{B_{r_i}(x_i) : r_i = \frac{1}{2}\cdot 3^{-i}\}$. By the cone structure, we have $d_{GH}(B_{r_i}(x_i), Z\times[-r_i, r_i]) \ge \frac{1}{10} r_i \sin(\pi\rho) > \frac{1}{100} r_i$ for any metric space Z. Thus $x_i \in \mathcal{S}^0_{\epsilon,r_i}(X)$ for any $0 < \epsilon < \frac{1}{200}$. However, $|\mathcal{C}| = \infty$.

By a standard covering technique, Theorem 1.3 implies the following Hausdorff measure estimate.

Corollary 1.4 (Hausdorff measure estimate)**.** *For any $n \in \mathbb{N}$ and $\epsilon > 0$ there exists $C = C(n, \epsilon) > 0$ such that for any $X \in Alex^n(-1)$ and $p \in X$, we have the Hausdorff measure estimate*

$$\mathcal{H}^k(\mathcal{S}^k_\epsilon(X) \cap B_1(p)) < C(n, \epsilon). \tag{1.6}$$

We also have the following conjectural form of the constant in the above theorem:

Conjecture 1.1. For any $(X, p) \in \mathrm{Alex}^n(-1)$, we have

$$\mathcal{H}^k(\mathcal{S}^k_\epsilon \cap B_1(p)) < C(n)\epsilon^{1-(n-k)}.$$

Indeed, we may even have the following stronger summable form, see Example 1.1:

Conjecture 1.2. For any $(X, p) \in \mathrm{Alex}^n(-1)$ and let $\epsilon_i = 2^{-i}$, then we have

$$\sum_{i=0}^{\infty} \epsilon_{i+1}^{(n-k)-1} \mathcal{H}^k\Big((\mathcal{S}^k_{\epsilon_{i+1}} \setminus \mathcal{S}^k_{\epsilon_i}) \cap B_1(p)\Big) < C(n).$$

Now let $\{B_r(x_i)\}_{i=1}^N$ be a Vitali covering of $B_r(\mathcal{S}^k_{\epsilon,r}(X) \cap B_1(p))$ with $x_i \in \mathcal{S}^k_{\epsilon,r}(X)$. By Theorem 1.3, we have that $N \le C(n, \epsilon)\, r^{-k}$. Combining it with $\mathcal{H}^n(B_r(x)) \le C(n)\, r^n$ for every $x \in X$ and $r \le 1$, we have the following estimate, which only matters in the noncollapsing setting:

Corollary 1.5 (Volume estimate)**.** *For any $n \in \mathbb{N}$ and $\epsilon > 0$ there exists $C = C(n, \epsilon) > 0$ such that the following estimate holds for any $X \in Alex^n(-1)$ and $p \in X$.*

$$\mathcal{H}^n(B_r(\mathcal{S}^k_{\epsilon,r}(X)) \cap B_1(p)) \le C\, r^{n-k}\,. \tag{1.7}$$

We also show that $\mathcal{S}^k_\epsilon$ is k-rectifiable.

Theorem 1.6 (k-rectifiability)**.** *For any $X \in Alex^n(-1)$ and $0 \le k \le n$ we have that $\mathcal{S}^k(X)$ is k-rectifiable.*

In [9], similar results as Corollary 1.4 and Theorem 1.6 are also proved for geodesically complete spaces with upper curvature bounds.

It was asked for both Ricci and Alexandrov cases wether $\mathcal{S}^k_\epsilon$ carries with a k-manifold structure, away from a zero $\mathcal{H}^k$-measure subset. It was proved in [2] that for any $X \in \mathrm{Alex}^n(\kappa)$, if $p \in X \setminus \mathcal{S}^{n-1}_\epsilon$, then there exists $r > 0$ so that $B_r(p)$ is bi-Lipschitz to $B_r(0) \subset \mathbb{R}^n$. If $p \in \mathcal{S}^{n-1}_\epsilon \setminus \mathcal{S}^{n-2}_\epsilon$, then there exists $r > 0$ so that $B_r(p)$ is bi-Lipschitz to a ball centered at the origin in the half space $\mathbb{R}^{n-1} \times \mathbb{R}_{\ge 0}$. For $\mathcal{S}^{n-2}_\epsilon$, we construct examples $X \in \mathrm{Alex}^n(\kappa)$ to show that it may contain no manifold point.

Theorem 1.7. *For any closed subset $T \subseteq \mathbb{S}^1$ and $\epsilon > 0$, there exists a sequence of 3-dimensional manifolds M_i with $\sec_{M_i} \ge 0$ and $M_i \to Y \in Alex^3(0)$, for which $\mathcal{S}^1_\epsilon(Y) = \phi(T)$, where $\phi\colon \mathbb{S}^1 \to Y$ is a bi-Lipschitz embedding.*

In particular, $\mathcal{S}^1_\epsilon(Y)$ can be a Cantor set with $\mathcal{H}^1(\mathcal{S}^1_\epsilon(Y)) > 0$, which contains no manifold points. Let $Y_n = Y \times \mathbb{R}^{n-3} \in \mathrm{Alex}^n(0)$. We have that $\mathcal{S}^{n-2}_\epsilon(Y_n)$ contains no $(n-2)$-dimensional manifold point. This shows that the rectifiable structure in Theorem 1.6 is sharp. Examples for which $\mathcal{S}^k_\epsilon$ contains no k-manifold point, where $n \geq 4$ and $1 \leq k \leq n-3$, can be similarly constructed, with a good amount of extra technical work.

2. Outline of the Proof

We begin with the notion of bad scales $Bad^{\epsilon}(p)$. Fix a point $p \in X$ and $\epsilon > 0$, then we define a $\mathbb{Z}_2$-valued function $T^{\epsilon}_p(r,R)$ to describe the symmetry of metric balls $B_s(p)$ over scales $0 \leq r \leq s \leq R$. Define $T^{\epsilon}_p(r,R) = 0$ if there exists a cone space $C(\Sigma)$, depending on p, r, R, ϵ but not on $s \in [r,R]$, so that

$$d_{GH}\Big(B_s(p),\, B_s(p^*)\Big) \leq \epsilon s\,, \tag{2.1}$$

for every $s \in [r,R]$, where $p^* \in C(\Sigma)$ is the cone point. Otherwise we define $T^{\epsilon}_p(r,R) = 1$. In the case that $T^{\epsilon}_p(r,R) = 0$, we say that the metric ball $B_s(p)$ is uniformly $(0,\epsilon)$-symmetric for $r \leq s \leq R$. It is clear that if $[a_1,a_2] \subseteq [r,R]$ and $T^{\epsilon}_p(r,R) = 0$, then $T^{\epsilon}_p(a_1,a_2) = 0$. Contrapositively, if we have $[r,R] \subseteq [b_1,b_2]$ with $T^{\epsilon}_p(r,R) = 1$, then we also have $T^{\epsilon}_p(b_1,b_2) = 1$.

Definition 2.1 (Bad scales)**.** Let $r_\alpha = 2^{-\alpha}$, where $\alpha \in \mathbb{N}$. The *ϵ-bad scales* $\{r_{\beta_{(j)}}\} \subseteq \{r_\alpha,\, \alpha \in \mathbb{N}\}$ of p, denoted by $Bad^{\epsilon}(p)$, are defined recursively as follows. Let $r_{\beta_{(0)}} = r_0 = 1$ and

$$r_{\beta_{(k+1)}} = \begin{cases} r_{\beta_{(k)}+1}, & \text{if} \quad T^{\epsilon}_p(r_{\beta_{(k)}+1}, r_{\beta_{(k)}}) = 1; \\ r_\alpha, & \text{if there exists } \alpha \geq \beta_{(k)}+1 \text{ such that } T^{\epsilon}_p(r_\alpha, r_{\beta_{(k)}}) = 0 \text{ but } T^{\epsilon}_p(r_{\alpha+1}, r_{\beta_{(k)}}) = 1. \end{cases}$$

Note that if $R \geq 2r$ and $[r,R]$ contains no ϵ-bad scale of p, then $B_s(p)$ is uniformly $(0,\epsilon)$-symmetric for $r \leq s \leq R$. This definition is strictly stronger than the corresponding definitions in the Ricci curvature context.

The following is a key lemma to build up our covering techniques.

Lemma 2.2 (Finiteness of the number of bad scales)**.** *For any $n \in \mathbb{N}$ and $\epsilon > 0$, there exists $N(n,\epsilon) > 0$ such that for any $(X,p) \in Alex^n(-1)$, the number of ϵ-bad scales satisfies $|Bad^{\epsilon}(p)| < N(n,\epsilon)$.*

The proof of this lemma is based on various point-wise monotonic properties of Alexandrov spaces. In particular, we prove Lemma 4.3, which we call "almost packing cone implies almost metric cone". It is an analogy of "almost volume cone implies almost metric cone", which is the monotonic formula used for manifolds with lower Ricci curvature bound. Note that both our monotonicity and the corresponding rigidity are strictly stronger than in the Ricci curvature context.

In order to state and prove our rigidity results we will need a splitting theory for Alexandrov spaces.

Definition 2.3 (Strong splitting maps)**.** Let $u_1, u_2, \ldots, u_k \colon B_R(p) \to \mathbb{R}$ be ϵ-concave functions. The map $u = (u_1, \ldots, u_k) \colon B_R(p) \to \mathbb{R}^k$ is called a (k,ϵ)-splitting map if the following are satisfied.

(i) $|\langle \nabla u_i, \nabla u_j \rangle - \delta_{ij}| \leq \epsilon$.

(ii) For any $x, y \in B_R(p)$ and any minimizing geodesic γ connecting x and y, it holds that

$$\langle \uparrow_x^y, \nabla_x u_i \rangle + \langle \uparrow_y^x, \nabla_y u_i \rangle \le \epsilon.$$

Here $\uparrow_x^y$ and $\uparrow_y^x$ denote the unit tangent directions of γ at x and y respectively.

Remark 2.1. If X is a smooth Riemannian manifold, the condition (ii) in the above definition says that on each geodesic, u has a lower integral hessian bound.

By the definition, we have that if $u\colon B_R(p) \to \mathbb{R}^k$ is a (k, ϵ)-splitting map, then $u|_{B_r}$ is also a (k, ϵ)-splitting map for any $B_r \subset B_R(p)$. This restriction property of splitting maps is false in the context of manifolds with lower Ricci curvature bounds. The existence and the properties of the strong splitting maps are discussed in Section 5.1.

As in the standard dimension reduction, let us observe that for a metric cone $C(Z)$, the tangent cone of any point away from the cone tip splits off an extra $\mathbb{R}$-factor comparing to $C(Z)$. We prove an effective version of this property in Lemma 5.7.

The monotonic property and the splitting theory lead to Theorem 6.2. It says that there exist $\delta(n, \epsilon)$ and $\beta(n, \epsilon) > 0$ so that if $u\colon B_{50}(p) \to \mathbb{R}^k$ is a (k, δ)-splitting function, and $\{B_{r_i}(x_i)\}$ is a disjoint collection with $x_i \in \mathcal{S}^k_{\epsilon, \beta r_i}$, then for any $z \in \mathbb{R}^k$, we have

$$\left|\left\{ i \in \mathbb{I}\colon B_{\beta r_i}(x_i) \cap u^{-1}(z) \neq \varnothing \right\}\right| < N(n, \epsilon). \tag{2.2}$$

In particular, this Theorem implies that if we look at the associated collection of balls $\{B_{\beta r_i/4}(u(x_i))\} \subseteq \mathbb{R}^k$ then its intersection number is at most $N(n, \epsilon)$. That is, given any ball $B_{\beta r_j/4}(u(x_j)) \in \{B_{\beta r_i/4}(u(x_i))\}$ it intersects at most $N - 1$ other balls from the collection. This shows that Theorem 1.3 holds if $B_1(p)$ is (k, ϵ)-splitting. We will then complete the proof by an induction on k.

In Section 7 we construct examples to prove Theorem 1.7. Let us explain the moral of the construction below. The technical details will be added to make it rigorous in Section 7.

Let $Z = \bar{B}_1 \subset \mathbb{R}^2$ be a closed unit disk and $X_0 = Z \times [0, 1] \in \text{Alex}^3(0)$ be a solid cylinder. For $\epsilon > 0$ small, we have $\mathcal{S}^0(X_0) = \varnothing$ and $\mathcal{S}^1_\epsilon(X_0) = \partial Z \times \{0, 1\}$ is a union of two unit circles.

Now let $T \subseteq \partial Z$ be a closed subset, and thus $\partial Z \setminus T = \cup_\ell U_\ell$ is a collection of disjoint open intervals. Let p be the center of Z and define $C_\ell = \cup_{x \in U_\ell} \gamma_{px}$, where $\gamma_{x,y}$ denotes a line connecting x and y, be the collection of sectors associated to the open sets U_ℓ. Let us observe for any $x \in \partial Z$ that the curvature at $(x, 1) \in X_0$ is $+\infty$ along the normal direction of $\partial Z \times \{1\}$ and strictly positive along its tangential direction. This will allow us to smoothly "sand off" each of $U_\ell \times \{1\}$ inside its convex hull $C_\ell \times [0, 1]$, so that both the convexity of X_0 and the tangent cones at points in $X_0 \setminus (\cup C_\ell \times [0, 1])$ are preserved. Let $X_1 \in \text{Alex}^3(0)$ be the resulted space. In particular, the tangent cones at the points of $T \times \{1\}$ are preserved, and thus we have that $\mathcal{S}^1_\epsilon(X_1) = (T \times \{1\}) \cup (\partial Z \times \{0\})$. Similarly, we can smooth near $\partial Z \times \{0\}$ in order to construct X_2 with $\mathcal{S}^1_\epsilon(X_2) = T \times \{1\}$. Now let Y_2 be the doubling of X_2, which is now a boundary free Alexandrov space $Y_2 \in \text{Alex}^3(0)$ for which $\mathcal{S}(Y_2) = \mathcal{S}^1_\epsilon(Y_2) = T$ and $\mathcal{S}^0_\epsilon(Y_2) = \varnothing$.

3. Monotonicity and Packing Numbers

In this section we describe a monotone formula which plays an important role in the constructions of subsequent sections.

Definition 3.1 (Packing)**.** Let X be a metric space and $S \subseteq X$ with $\mathrm{diam}(S) < \infty$. For $\epsilon > 0$, we say that a subset $\mathbf{x} \equiv \{x_i\} \subseteq S$ is an *ϵ-subpacking* if

$$d(x_i, x_j) \geq \epsilon \, \mathrm{diam}(S) \text{ for every } i \neq j. \tag{3.1}$$

An ϵ-subpacking $\mathbf{x}$ is said to be a packing if it is also $\epsilon \, \mathrm{diam}(S)$-dense in S.

We write $|\mathbf{x}| = N$ as the number of elements in $\mathbf{x}$ if it is finite. If we want to signify the set in question we may write $\mathbf{x} = \mathbf{x}(S)$. We define the ϵ-packing number $P_\epsilon(S)$ by

$$P_\epsilon(S) \equiv \sup\{|\mathbf{x}| : \mathbf{x} \text{ is an } \epsilon\text{-subpacking for } S\}. \tag{3.2}$$

A packing $\mathbf{x}$ is called a maximal ϵ-packing of S if $|\mathbf{x}| = P_\epsilon(S) < \infty$.

In the case that $S = B_r(p)$ is a metric ball we may write $\mathbf{x}(p, r) = \mathbf{x}(B_r(p))$ and the ϵ-packing number $P_\epsilon(p, r) \equiv P_\epsilon(B_r(p))$. Let us record some easy but useful properties which hold for general metric spaces.

Lemma 3.2. *Let X be a metric space with $\epsilon > 0$ fixed. Then the following hold:*

(i) (Enlargement) If $\mathbf{x}$ is an ϵ-subpacking of $B_r(x)$, then either $\mathbf{x}$ is an ϵ-packing or there exists $x' \in B_r(x)$ such that $\mathbf{x}' \equiv \mathbf{x} \cup \{x'\}$ is also an ϵ-subpacking.
(ii) (Maximal subpacking $\implies$ packing) If $\mathbf{x}$ is an ϵ-subpacking of $B_r(x)$ with $|\mathbf{x}| = P_\epsilon(x, r) < \infty$, then $\mathbf{x}$ is an ϵ-packing.
(iii) (ϵ-monotonicity) If $\mathbf{x}$ is an ϵ-subpacking and $\epsilon' < \epsilon$, then $\mathbf{x}$ is an ϵ'-subpacking. In particular, for each $r > 0$ we have that $P_{\epsilon'}(x, r) \geq P_\epsilon(x, r)$.

We wish to now discuss some more refined properties of ϵ-packings and packing numbers for Alexandrov spaces. To do this let us introduce the induced subpacking. Indeed, this notation makes sense for any locally compact length metric space, but it is not so useful in general.

Definition 3.3 (Induced subpacking)**.** Let $p \in X$, $R > 0$ and for each $x \in \bar{B}_R(p) \setminus \{p\}$ we fix a geodesic $\gamma_{px} = \gamma^R_{px}$ connecting p and x. Given $0 < r < R$, we define the inducting function $\varphi^R_r : \bar{B}_R(p) \to \bar{B}_r(p)$, $x \mapsto \bar{x}$, where $\bar{x} \in \gamma^R_{px}$ is the point with $d(p, \bar{x}) = \frac{r}{R} \cdot d(p, x)$. Now let $\{x_i\}_{i=1}^N$ be an ϵ-subpacking of $\bar{B}_R(p)$ and $0 < r < R$, then we call the collection of points $\{\varphi^R_r(x_i)\}_{i=1}^N$ the induced subpacking in $\bar{B}_r(p)$ of $\{x_i\}_{i=1}^N$.

Note that the choice of geodesic γ_{px} in the definition of φ^R_r is certainly not unique. However in the above definition of φ^R_r, such a choice is fixed for a given $R > 0$ while independent of $0 < r \leq R$. If no confusion arises one may write $\gamma^R_{px} = \gamma_{px}$.

The proof of the following propositions are easy exercises based on the Toponogov comparisons.

Proposition 3.4. *Let $0 < \epsilon, R < 1$. The following hold for any $(X, p) \in Alex^n(-\epsilon)$ with $1 \geq c \geq 1 - R^2$. If $(X, p) \in Alex^n(0)$, then $c \equiv 1$ can be chosen as a constant.*

(i) *(Induced Packing) Let $\{x_i\}$ be an ϵ-subpacking of $B_R(p)$. For any $0 < r < R$, the induced subpacking in $B_r(p)$ is a $c\epsilon$-subpacking.*
(ii) *(Monotonicity) If $r \le R$, then the packing number $P_{c\epsilon}(x, r) \ge P_\epsilon(x, R)$.*
(iii) *(Bounds) If $\mathbf{x} = \{x_i\}$ is an ϵ-subpacking for $B_r(p)$ with $0 < r \le 1$, then $1 \le |\mathbf{x}| \le C(n)\epsilon^{-n}$. In particular, we have $1 \le P_\epsilon(p, r) \le C(n)\epsilon^{-n}$.*
(iv) *(Density) There exists a limit $\lim_{r\to 0} P_\epsilon(p, r) \equiv P_\epsilon(p) \le C(n)\epsilon^{-n}$, which we call the ϵ-density at x. In fact, $P_\epsilon(p) = P_\epsilon(p^*, 1)$, where p^* is the cone point in the tangent cone at p.*

4. Bad Scales

This section is dedicated to proving Lemma 2.2. It says that there are at most a finite number of bad scales at each point, and our space has a fixed cone structure which persists over all good scales. Let us begin with an easy proposition.

Proposition 4.1. *For any $0 \le r < R/2 < R \le 1$, if $(r, R) \cap Bad^\epsilon(x) = \emptyset$, then $T^\epsilon_x(r, R) = 0$.*

Proof. The proof is almost taulogical. Let $r_{\beta_{(k)}} = \inf\{r_\beta \in Bad^\epsilon(x) : r_\beta \ge R\}$ and $r_{\beta_{(k+1)}}$ be the next ϵ-bad scale. Because $(r, R) \cap Bad^\epsilon(x) = \emptyset$ we have that $r_{\beta_{(k+1)}} \le r < R/2 < R \le r_{\beta_{(k)}}$. Therefore, $r_{\beta_{(k)}}/r_{\beta_{(k+1)}} > 2$ and $\beta_{(k+1)} - \beta_{(k)} \ge 2$. By the definition $T^\epsilon_x(r_{\beta_{(k+1)}}, r_{\beta_{(k)}}) = 0$. Note that $[r, R] \subseteq [r_{\beta_{(k+1)}}, r_{\beta_{(k)}}]$ and so we have that $T^\epsilon_x(r, R) = 0$. □

To prove Lemma 2.2, we need a result of the form "almost packing cone implies almost metric cone". We begin with the following proposition. It follows directly from the definitions of ϵ-packing and Hausdorff distance.

Proposition 4.2. *Let X and Y be metric spaces whose diameters are both no more than 1. Let $\{x_i\}_{i=1}^{N_1}$ be an ϵ-packing of X and $\{y_i\}_{i=1}^{N_2}$ be an ϵ-packing of Y. If $N_1 = N_2 = N$ and*

$$|d(x_i, x_j) - d(y_i, y_j)| \le \epsilon$$

for every $1 \le i, j \le N$, then $d_{GH}(X, Y) \le 4\epsilon$.

The first main result of this section is the following:

Lemma 4.3 (Almost packing cone implies almost metric cone)**.** *There is a universal constant $c > 0$ such that the following holds for any $n \in \mathbb{N}$ and $\epsilon \in (0, c)$. Let $(X, p) \in Alex^n(-\epsilon)$ and $0 \le r \le \frac{1}{2}R \le c$. Let $\mathbf{x}(p, R) = \{x_i\}_{i=1}^N$ be an ϵ-packing of $B_R(p)$. We have*

$$T_p^{\epsilon^{0.1}}(r, R) = 0 \tag{4.1}$$

if both of the following are satisfied.

(i) *$P_\epsilon(p, r) = N = P_\epsilon(p, R)$.*

(ii) $r^{-1}d(\varphi_r^R(x_i), \varphi_r^R(x_j)) \leq R^{-1}d(x_i, x_j) + \epsilon$, *for every* $1 \leq i, j \leq N$.

Here $\varphi_r^R \colon \bar{B}_R(p) \to \bar{B}_r(p)$ *is the inducing function defined as in Definition 3.3.*

Proof of Lemma 4.3. Let us introduce the notation $(\lambda)\bar{B}_s \equiv (\bar{B}_s, \lambda d)$ to denote the rescaled space. The proof consists of two points. First, we will see that it is almost immediate from the assumed conditions that the mapping $\varphi_s^R : \bar{B}_R(p) \to (R/s)\bar{B}_s(p)$ is a GH map. Second, we will show that $\bar{B}_R(p)$ is GH-close to a ball in a cone space $C(\Sigma)$, centered at the cone point. The combination of these two points prove the Lemma.

Let us discuss these points more carefully. For simplicity, we will only prove the result for $X \in \mathrm{Alex}^n(0)$. The general case is similar with a modification on c, which are just used to estimate the law of cosine formula in (4.17). Now by the assumptions and the monotonic property, the induced subpacking $\{\varphi_s^R(x_i) : x_i \in \mathbf{x}(p, R)\}$ is a packing of $\bar{B}_s(p)$ and

$$R^{-1}d(x_i, x_j) \leq s^{-1}d(\varphi_s^R(x_i), \varphi_s^R(x_j)) \leq R^{-1}d(x_i, x_j) + \epsilon, \tag{4.2}$$

for every $s \in [r, R)$ and every $1 \leq i \neq j \leq N$.

By Proposition 4.2, for all $s \in [r, R)$, we therefore have that

$$d_{GH}\Big(\bar{B}_R(p), (R/s)\bar{B}_s(p)\Big) \leq 4\epsilon R\,, \tag{4.3}$$

where $\varphi_s^R \colon \bar{B}_R(p) \to (R/s)\bar{B}_s(p)$ is a $8\epsilon R$-isometry. To prove (4.1) it therefore suffices to construct a metric cone $C(\Sigma)$ and show that

$$d_{GH}(\bar{B}_R(p), \bar{B}_R(p^*)) \leq \frac{1}{2}\epsilon^{0.1}R, \tag{4.4}$$

where $p^* \in C(\Sigma)$ is the cone point. Let us prove this by first assuming the following lemma, which we will prove later. Let $\iota_\lambda \colon \bar{B}_s \to (\lambda)\bar{B}_s$ be the identity map.

Lemma 4.4. *Let* $S_\rho = \{x \in \bar{B}_R(p) : d(p, x) = \rho\}$ *be the* ρ*-cross section in* X*.*

(i) For any $t \in [\epsilon^{0.5}, 1)$*, the restricted map* $\iota_{t^{-1}} \circ \varphi_{tR}^R|_{S_R} \colon S_R \to (t^{-1})S_{tR}$ *is* $\epsilon^{0.4}R$*-onto.*

(ii) For every $x, y \in S_R$ *and* $t_1, t_2 \in [\epsilon^{0.5}, 1)$*, we have*

$$\left|\cos\tilde{\angle}\left(p\,{}^{x}_{y}\right) - \cos\tilde{\angle}\left(p\,{}^{\varphi_{t_1R}^R(x)}_{\varphi_{t_2R}^R(y)}\right)\right| \leq \epsilon^{0.4}. \tag{4.5}$$

(iii) For any $x, y \in S_R$*, geodesic triangle* $\triangle pxy$ *is* $\epsilon^{0.3}R$*-close to a geodesic triangle in* $\mathbb{R}^2$*, equipped with the extrinsic metrics.*

Using the above we now construct a metric cone $C(\Sigma)$ and define a GH-map $f \colon \bar{B}_R(p^*) \to \bar{B}_R(p)$, where $p^* \in C(\Sigma)$ is the cone point. Define a distance function d_{S_R} on the R-cross section $S_R = \{x \in X : d(p, x) = R\}$ by

$$d_{S_R}(x, y) = \inf_{x_0, \ldots, x_L \in S_R}\left\{\sum_{\alpha=1}^{L} d_X(w_{\alpha-1}, w_\alpha) \colon w_0 = x,\ w_L = y,\ d_X(w_{\alpha-1}, w_\alpha) \leq \epsilon^{0.1}R\right\}. \tag{4.6}$$

Note that this is an approximation of the induced length space distance function on a subset. It's clear that $d_{S_R}(x, y) \geq d_X(x, y)$, and thus if $d_{S_R}(x, y) = 0$ then $x = y$. To verify the triangle inequality, we let $x, y, z \in S_R$.

By the definition we have for any $\eta > 0$ that there exists $w_\alpha \in S_R$ with $w_0 = x$, $w_{N_1} = y$, $w_{N_2} = z$, $d_X(w_{\alpha-1}, w_\alpha) \le \epsilon^{0.1} R$, so that $d_{S_R}(x,y) \ge \sum_{\alpha=1}^{N_1} d_X(w_{\alpha-1}, w_\alpha) - \eta$ and $d_{S_R}(y,z) \ge \sum_{\alpha=N_1+1}^{N_2} d_X(w_{\alpha-1}, w_\alpha) - \eta$. Then we have

$$d_{S_R}(x,y) + d_{S_R}(y,z) \ge \sum_{i=1}^{N_2} d_X(w_{\alpha-1}, w_\alpha) - 2\eta \ge d_{S_R}(x,z) - 2\eta. \tag{4.7}$$

Letting $\eta \to 0$ we then obtain the triangle inequality.

Now let $(\Sigma, d_\Sigma) = (S_R, \frac{1}{R} d_{S_R})$ and $C(\Sigma)$ be the metric cone over Σ and p^* is the cone point. Let $(\tilde S_R, d_{\tilde S_R}) = (\Sigma, R\, d_\Sigma)$ be the R-cross section in $C(\Sigma)$. Let $\Pi\colon C(\Sigma) \to \tilde S_R$ by $a = (\bar a, d(p^*, a)) \mapsto \bar a$ be the projection mapping. Identify $\tilde S_R$ with S_R and let us define

$$f\colon \bar B_R(p^*) \to \bar B_R(p) \text{ by } a \mapsto \varphi^R_{d(p^*, a)} \circ \Pi(a)\,. \tag{4.8}$$

We first show that f is $\epsilon^{0.4}R$-onto. Let $x \in \bar B_R(p)$. Note that for any $y \in \tilde S_R = S_R$, we have $(y, d_X(p,x)) \in C(\Sigma)$ and $f((y, d_X(p,x))) = \varphi^R_{d(p,x)}(y)$. Thus the $\epsilon^{0.4}R$-onto property of f follows from Lemma 4.4 (i).

Now we show that f is $\frac{1}{2}\epsilon^{0.1}R$-distance preserving. Let $a, b \in C(\Sigma)$, $x = f(a)$, $y = f(b)$ and $\gamma_{x,y}$ be a geodesic connecting x and y. By Lemma 4.4 (i), for any partition $\{u_i\}$ of $\gamma_{x,y}$, there exist $w_i \in S_R$ and $s_i > 0$ such that $d_X(\varphi^R_{s_iR}(w_i), u_i) \le \epsilon^{0.4}R$. Note that for any two points $x', y' \in S_R$ we have that $d_{S_R}(x', y') = d_X(x', y')$ if $d_X(x', y') \le \epsilon^{0.1}R$. By Lemma 4.4 (iii), the points u_i and w_i, $i = 1, \ldots, N$ can be chosen so that $\frac{1}{2}\epsilon^{0.1}R \ge d_X(w_{i-1}, w_i) = d_{S_R}(w_{i-1}, w_i) \ge \frac{1}{4}\epsilon^{0.1}R$. Thus for this partition we have that $N \le \frac{10R}{\epsilon^{0.1}R} = 10\epsilon^{-0.1}$.

Now let $\tilde\varphi$ be the inducting function on $C(\Sigma)$ defined in the same way as φ. By the cone metric, we have

$$d_{C(\Sigma)}\left(\tilde\varphi^R_{s_{i-1}R}(w_{i-1}), \tilde\varphi^R_{s_iR}(w_i)\right) = \sqrt{s_{i-1}s_i\, d^2_{S_R}(w_{i-1}, w_i) + (s_{i-1} - s_i)^2 R^2}. \tag{4.9}$$

By Lemma 4.4 (iii), we have

$$\left| d_X\left(\varphi^R_{s_{i-1}R}(w_{i-1}), \varphi^R_{s_iR}(w_i)\right) - \sqrt{s_{i-1}s_i\, d^2_X(w_{i-1}, w_i) + (s_{i-1} - s_i)^2 R^2} \right| < 10\epsilon^{0.3}R. \tag{4.10}$$

Therefore,

$$\begin{aligned} d_X(x,y) &= \sum_i d_X(u_{i-1},u_i) \\ &\geq \sum_{i=1}^{N} \left(d_X\left(\varphi^R_{s_{i-1}R}(w_{i-1}), \varphi^R_{s_iR}(w_i)\right) - 2\epsilon^{0.4}R \right) \\ &\geq -12N\cdot\epsilon^{0.3}R + \sum_{i=1}^{N} \sqrt{s_{i-1}s_i\, d_X^2(w_{i-1},w_i) + (s_{i-1}-s_i)^2R^2} \\ &\geq -120\epsilon^{0.2}R + \sum_{i=1}^{N} \sqrt{s_{i-1}s_i\, d_{S_R}^2(w_{i-1},w_i) + (s_{i-1}-s_i)^2R^2} \\ &= -120\epsilon^{0.2}R + \sum_{i=1}^{N} d_{C(\Sigma)}\left(\tilde\varphi^R_{s_{i-1}R}(w_{i-1}), \tilde\varphi^R_{s_iR}(w_i)\right) \\ &\geq -122\epsilon^{0.2}R + d_{C(\Sigma)}(a,b). \end{aligned} \tag{4.11}$$

The last inequality follows from the triangle inequality since w_1 and w_N can be chosen so that $d(a,\tilde\varphi^R_{s_1R}(w_1)) < \epsilon^{0.4}R$ and $d(b,\tilde\varphi^R_{s_NR}(w_N)) < \epsilon^{0.4}R$. Starting from a partition of $\gamma_{a,b}$ and apply the same arguments. We get

$$d_X(x,y) \leq d_{C(\Sigma)}(a,b) + 122\epsilon^{0.2}R. \tag{4.12}$$

Combining (4.11), (4.12) and the definition of f, we get the desired result. $\square$

Now let us finish the proof of Lemma 4.4:

Proof of Lemma 4.4. (i) By (4.3), for any $\lambda \in [r/R,1) \supseteq [1/2,1)$ we have that

$$\iota_{\lambda^{-1}} \circ \varphi^R_{\lambda R}|_{S_\rho} : S_\rho \to (\lambda^{-1})S_{\lambda\rho} \text{ is a } 24\epsilon R\text{-isometry for any } \rho \in (0,R]. \tag{4.13}$$

This in particular proves (i) for $t \in [r/R,1)$. For the case $t \ll r/R$, we need to inductively apply $\varphi^R_{R/2}$.

For any $t \in [\epsilon^{0.5},1)$, there is an integer $K = K(t) \leq \epsilon^{-0.1}$ such that $2^{-(K+1)} \leq t < 2^{-K}$. Since in X geodesics do not bifurcate and in the definition of induced packing, the choices of geodesics are a priori fixed in terms of p and R, we can write

$$\varphi^R_{tR} = \varphi^R_{2^KtR} \circ \underbrace{\varphi^R_{R/2} \circ \cdots \circ \varphi^R_{R/2}}_{K}. \tag{4.14}$$

Note that $2^Kt \in [1/2,1) \subseteq [r/R,1)$. Thus (4.13) applies to $\rho = R$ and $\lambda = 2^Kt$. Combining (4.13) and (4.14), we get that $\iota_{t^{-1}} \circ \varphi^R_{tR}$ is $24(K+1)\epsilon R$-onto. Then the result follows since $K \leq \epsilon^{-0.1}$.

(ii) We first show that (4.5) is true for $t_1 = t_2 = t$. Fix $t \in [\epsilon^{0.5},1)$. Let $x_0 = x$, $x_i = \varphi^R_{R/2}(x_{i-1})$ for $1 \leq i \leq K$ and $x_{K+1} = \varphi^R_{2^KtR}(x_K) = \varphi^R_{tR}(x)$, where $K = K(t) \leq \epsilon^{-0.1}$ is defined as in (i). The sequence $\{y_i\}$ is defined similarly in terms of y. By (4.13) and because $2^Kt \in [1/2,1)$, for $1 \leq i \leq K$ we have

$$\left|\frac{1}{2}d(x_i,y_i) - d(x_{i-1},y_{i-1})\right| \leq 24\epsilon R \tag{4.15}$$

and

$$\left|2^K t\cdot d(x_{K+1},y_{K+1})-d(x_K,y_K)\right|\le 24\epsilon R. \tag{4.16}$$

Note that $d(p,x_{i-1})=d(p,y_{i-1})=\frac{1}{2}d(p,x_i)=\frac{1}{2}d(p,y_i)$ and $d(p,x_K)=d(p,y_K)=2^K t\cdot d(p,x_{K+1})=2^K t\cdot d(p,y_{K+1})$. By law of cosine, we have

$$\begin{aligned}\left|\cos\tilde\measuredangle\left(p\,{}^{x_i}_{y_i}\right)-\cos\tilde\measuredangle\left(p\,{}^{x_{i-1}}_{y_{i-1}}\right)\right| &= \frac{1}{2}\left|\left(\frac{d(x_i,\,y_i)}{d(p,x_i)}\right)^2-\left(\frac{d(x_{i-1},\,y_{i-1})}{d(p,x_{i-1})}\right)^2\right|\\ &\le \frac{50\epsilon R}{d(p,x_i)}\le\frac{50\epsilon R}{tR}\\ &\le 50\epsilon^{0.5}.\end{aligned} \tag{4.17}$$

Summing up (4.17) for $i=1,2,\ldots,K+1$, we get

$$\left|\cos\tilde\measuredangle\left(p\,{}^{x}_{y}\right)-\cos\tilde\measuredangle\left(p\,{}^{\varphi^R_{tR}(x)}_{\varphi^R_{tR}(y)}\right)\right|\le 50(K+1)\epsilon^{0.5}\le\epsilon^{0.4}. \tag{4.18}$$

Suppose $t_1\le t_2$. By Topnogov comparison, we have

$$\tilde\measuredangle\left(p\,{}^{x}_{y}\right)\le\tilde\measuredangle\left(p\,{}^{\varphi^R_{t_1R}(x)}_{\varphi^R_{t_2R}(y)}\right)\le\tilde\measuredangle\left(p\,{}^{\varphi^R_{t_1R}(x)}_{\varphi^R_{t_1R}(y)}\right). \tag{4.19}$$

Then (4.5) follows from (4.18) and (4.19).

The statement (iii) is a direct consequence of (ii). □

Lemma 4.3 implies that when passing an ϵ-bad scale, either the packing number, or the rescaled distance distortion is increased by at least a definite amount, depending on ϵ.

Corollary 4.5. *For any $\epsilon>0$, there is $\delta=\delta(\epsilon)>0$ such that the following holds. Let $(X,p)\in Alex^n(-\delta)$ be an Alexandrov space with $r_{\beta_{(k)}}$, $r_{\beta_{(k+1)}}$ ϵ-bad scales of p. Let $\{x_i\}$ be a maximal δ-packing of $B_{r_{\beta_{(k)}}}(p)$ and $\{y_i\}$ be the induced subpacking in $B_{r_{\beta_{(k+1)}+1}}(p)$. Then one of the following holds:*

(i) $P_\delta(p,r_{\beta_{(k+1)}+1})\ge P_\delta(p,r_{\beta_{(k)}})+1$;
(ii) there exist $i\ne j$ such that $r^{-1}_{\beta_{(k+1)}+1}d(y_i,y_j)>r^{-1}_{\beta_{(k)}}d(x_i,x_j)+\delta$.

Proof. By the definition of bad scales, we have that $T^\epsilon_p(r_{\beta_{(k+1)}+1},r_{\beta_{(k)}})=1$. Then the result follows from Lemma 4.3 with $\delta=\epsilon^{10}$. □

Now we give a proof of Lemma 2.2 using the above monotone property.

Proof of Lemma 2.2. We will only prove for $X\in\mathrm{Alex}^n(0)$ to keep notation simple, the general case is similar. Let $r_\alpha=2^{-\alpha}$, $\alpha\in\mathbb{N}$ and $K>J\ge 0$ be integers. Let $N_J=P_\delta(p,r_J)$ be the maximum δ-packing number of $B_{r_J}(p)$. Let $\mathbb{I}=\{\beta_{(k)}\in\mathbb{N}:r_{\beta_{(k)}}\in Bad^\epsilon(p)\cap[r_K,\,r_J]\}$. We claim that if $|\mathbb{I}|>10N_J^2\delta^{-1}$, then $P_\delta(p,r_K)\ge P_\delta(p,r_J)+1$. If this is not true, then $P_\delta(p,r_K)=P_\delta(p,r_\alpha)=P_\delta(p,r_J)$ for every $\alpha\in[J,K]\cap\mathbb{Z}$.

Let $\{x_i\}_{i=1}^{N_J}$ be a maximal δ-packing of $B_{r_J}(p)$ and $\{x_i^\alpha\}$ be the induced subpacking in $B_{r_\alpha}(p)$. By Corollary 4.5, for every $\beta_{(k)} \in \mathbb{I}$, there exist i and j, depending on $\beta_{(k)}$, such that

$$r^{-1}_{\beta_{(k+1)}+1} d(x_i^{\beta_{(k+1)}+1}, x_j^{\beta_{(k+1)}+1}) > r^{-1}_{\beta_{(k)}} d(x_i^{\beta_{(k)}}, x_j^{\beta_{(k)}}) + \delta. \tag{4.20}$$

Given a pair of indices (i, j), let $\mathbb{I}_{(i,j)}$ be the collection of $\beta_{(k)} \in \mathbb{I}$ such that (4.20) holds. Because $|\mathbb{I}| > 10N_J^2\,\delta^{-1}$, there exist $1 \le i_0, j_0 \le N_J$, such that $|\mathbb{I}_{(i_0,j_0)}| > 10\,\delta^{-1}$. Furthermore, there is a subset $\mathbb{J}_{(i_0,j_0)} \subseteq \mathbb{I}_{(i_0,j_0)}$ with $|\mathbb{J}_{(i_0,j_0)}| > 5\,\delta^{-1}$, so that the intervals $\{(\beta_{(k)}, \beta_{(k+1)} + 1) : \beta_{(k)} \in \mathbb{J}_{(i_0,j_0)}\}$ are disjoint. Note that by the monotonic property, we have

$$r^{-1}_{\alpha_1} d(x_i^{\alpha_1}, x_j^{\alpha_1}) \ge r^{-1}_{\alpha_2} d(x_i^{\alpha_2}, x_j^{\alpha_2}), \tag{4.21}$$

for every $\alpha_1 \ge \alpha_2$. Summing up (4.20) for $\beta_{(k)} \in \mathbb{J}_{(i_0,j_0)}$ and taking in account (4.21), we get

$$2 \ge r_K^{-1}\, d(x_i^K, x_j^K) > r_J^{-1} d(x_i^J, x_j^J) + \sum_{\beta_{(k)}\in\mathbb{J}_{(i_0,j_0)}} \delta \ge 5, \tag{4.22}$$

a contradiction.

Note now that for every $r > 0$ we have that $P_\delta(p, r) \le C(n, \delta)$. Thus it follows from the above claim that

$$|Bad^\epsilon(p)| \le (C(n, \delta) + 1)\,(10\,C(n, \delta)^2\delta^{-1} + 1). \tag{4.23}$$

□

For any $\lambda \in (0, 1/4)$, it follows from Lemma 2.2 that for any $x \in X$, there is at least one of the intervals

$$[\lambda^{2Bad^\epsilon(x)+1}, \lambda^{2Bad^\epsilon(x)}],\ \ldots,\ [\lambda^5, \lambda^4],\ [\lambda^3, \lambda^2],\ [\lambda, 1]$$

containing no ϵ-bad scale. Thus we have

Lemma 4.6. *For any $n \in \mathbb{N}$, $1/4 > \lambda > 0$ and $\delta > 0$, there exists $\eta = \eta(n, \delta, \lambda) > 0$ such that for any $x \in X \in Alex^n(-1)$ and any $0 < R \le 1$, there exists $r_x \ge \eta R$, such that $Bad^\delta(x) \cap [\lambda r_x, r_x] = \varnothing$ and thus $T_x^\delta(\lambda r_x, r_x) = 0$.*

5. Splitting Theory and Dimension Reduction

5.1. **Splitting theory.** In this subsection, we discuss the splitting theory in Alexandrov geometry. Proposition 5.1 is a key geometric property for spaces with lower sectional curvature bounds that distinguishes them from spaces with lesser geometric constraints, such as lower Ricci curvature bounds. In words, it says that if some ball almost-splits off a Euclidean factor, then all sub-balls continue to almost-split off this factor.

Proposition 5.1. *For any $n, \epsilon > 0$, there exist $\delta = \delta(n, \epsilon) > 0$ so that the following holds for any $X \in Alex^n(-\delta)$ and $R \in (0, 1]$.*

(i) *Let $u = (u_1, \ldots, u_k)\colon B_{5R}(p) \to \mathbb{R}^k$ be a (k,δ)-splitting map. For any $B_r \subseteq B_R(p)$ and any $\xi \in u(B_r)$, there exists a map $\phi\colon B_r \to u^{-1}(\xi)$ so that*

$$(u,\phi)\colon B_r \to \mathbb{R}^k \times u^{-1}(\xi)$$

is ϵr-isometry.

(ii) *If $f\colon B_{5R}(p) \to B_{5R}(z)$, where $z \in \mathbb{R}^k \times Z$, is a δR-isometry, then there exists a (k,ϵ)-splitting map $u\colon B_R(p) \to \mathbb{R}^k$.*

(iii) *If there is a (k,δ)-strainer $\{(a_i,b_i)\}$ with $d(p,a_i)$, $d(p,b_i) \geq 5R$ for every $1 \leq i \leq k$, then there exists a (k,ϵ)-splitting map $u\colon B_R(p) \to \mathbb{R}^k$.*

The above splitting theory in Alexandrov geometry is well understood. For completeness we outline the proof.

Proof. We argue (i) by contradiction. This argument can be made effective with some extra work. Note that if $X_i \in \mathrm{Alex}^n(\kappa)$ with $X_i \to X$ and $\delta_i \to \delta$, then the limit of (k,δ_i)-splitting functions on X_i is a (k,δ)-splitting function on X. Thus passing to a limit of contradictive rescalled sub-balls, it suffices to show that if $(X,p) \in \mathrm{Alex}^n(0)$ and there is a 0-splitting function $u = (u_1,\ldots,u_k)\colon B_5(p) \to \mathbb{R}^k$, then for any $\xi \in u(B_1(p))$, there exists a map $\phi\colon B_1(p) \to u^{-1}(\xi)$ so that

$$(u,\phi)\colon B_1(p) \to \mathbb{R}^k \times u^{-1}(\xi)$$

is an isometric embedding. For such a 0-splitting function u, the following hold for every i and j:

(1) u_i is 0-concave.
(2) $\langle \nabla u_i, \nabla u_j \rangle = \delta_{ij}$.
(3) For any $x, y \in B_R(p)$ and any minimizing geodesic γ connecting x and y, it holds that

$$\langle \Uparrow_x^y, \nabla_x u_i \rangle + \langle \Uparrow_y^x, \nabla_y u_i \rangle = 0.$$

We now prove the result by induction on k. Start with the base case $k = 1$. Let $\sigma_x(t)$ be a u-gradient flow with $\sigma_x(0) = x$. If no confusion arises one may write $x_t = \sigma_x(t)$. Because u is 0-concave and $|\nabla u| = 1$, we have $u(x_t) - u(x) = t$ and $d(u(x_t), u(x_s)) = |t - s|$. In particular, $\sigma_x(t)$ is a geodesic from x. It's clear that the directed tangent vectors $\sigma^+(t) = \nabla_{x_t} u$ and $\sigma^-(t) = -\nabla_{x_t} u$.

Let T_x be the time so that $\sigma_x(T_x) \in u^{-1}(\xi)$ and define $\phi(x) = \sigma_x(T_x) \in u^{-1}(\xi)$. We will show that

$$(u,\phi)\,|_{B_1}\colon B_1 \to \mathbb{R} \times u^{-1}(\xi)$$

is an isometric embedding. This follows from the following statements for arbitrary $\xi \in u(B_1(p))$ and $t, s \in [0,1]$.

(A) $|T_x| = d(x, u^{-1}(\xi))$.
(B) For any two u-gradient curves α and β, we have $d(\alpha(t),\beta(t)) = d(\alpha(s),\beta(s))$.
(C) The Pythagorean Theorem $d^2(x_t, y) = d^2(x,y) + t^2$.

We first prove (A). It's clear that $|T_x| \geq d(x, u^{-1}(\xi))$. Recall that if $X \in \text{Alex}^n(0)$ and $f\colon X \to \mathbb{R}$ is a λ-concave function, then

$$d(p,q) \cdot \left\langle \uparrow_p^q, \nabla_p f \right\rangle \geq f(q) - f(p) - \frac{\lambda}{2} \cdot d^2(p,q). \tag{5.1}$$

Thus if $u\colon X \to \mathbb{R}$ is a 0-concave function, then

$$d(p,q) \geq d(p,q) \cdot \left\langle \uparrow_p^q, \nabla_p u \right\rangle \geq u(q) - u(p). \tag{5.2}$$

Let $y \in u^{-1}(\xi)$ so that $d(x,y) = d(x, u^{-1}(\xi))$, then we have

$$d(x,y) \geq |u(x) - u(y)| = |u(\sigma_x(0)) - \xi| = |u(\sigma_x(0)) - u(\sigma_x(T_x))| = |T_x|.$$

To prove (B), we let $x_t = \alpha(t)$, $y_t = \beta(t)$ and $\ell(t) = d(x_t,\, y_t)$. Assume $t \geq s$. Let $\ell^+(t) = \lim\limits_{\eta \to 0^+} \dfrac{\ell(t+\eta) - \ell(t)}{\eta}$ and $\ell^-(t) = \lim\limits_{\eta \to 0^+} \dfrac{\ell(t-\eta) - \ell(t)}{\eta}$ be the one-sided derivatives. By the first variation formula and because u is 0-concave, we have

$$\ell^+(t) \leq -\langle \uparrow_{x_t}^{y_t}, \nabla_{x_t} u \rangle - \langle \uparrow_{y_t}^{x_t}, \nabla_{y_t} u \rangle \leq -\frac{u(y_t) - u(x_t)}{\ell(t)} - \frac{u(x_t) - u(y_t)}{\ell(t)} = 0. \tag{5.3}$$

Thus we get $\ell(t) \leq \ell(s)$. Since $\langle \uparrow_{x_t}^{y_t}, \alpha^+(t) \rangle + \langle \uparrow_{y_t}^{x_t}, \beta^+(t) \rangle = \langle \uparrow_{x_t}^{y_t}, \nabla_{x_t} u \rangle + \langle \uparrow_{y_t}^{x_t}, \nabla_{y_t} u \rangle \leq 0$, we have

$$\ell^-(t) \leq -\langle \uparrow_{x_t}^{y_t}, \alpha^-(t) \rangle - \langle \uparrow_{y_t}^{x_t}, \beta^-(t) \rangle = \langle \uparrow_{x_t}^{y_t}, \alpha^+(t) \rangle + \langle \uparrow_{y_t}^{x_t}, \beta^+(t) \rangle \leq 0. \tag{5.4}$$

Thus $\ell(t) \geq \ell(s)$.

Now we prove (C). By Toponogov comparison and (5.2), we get that

$$\begin{aligned} d^2(x_t, y) &\leq d^2(x,y) + t^2 - 2t \cdot d(x,y) \cdot \left\langle \uparrow_x^y, \nabla_x u \right\rangle \\ &\leq d^2(x,y) + t^2 - 2t \cdot (u(y) - u(x)). \end{aligned} \tag{5.5}$$

Start with $x, y \in u^{-1}(\xi)$. Fix y and flow x by time t. By (5.5), we get

$$d^2(x_t, y) \leq d^2(x,y) + t^2. \tag{5.6}$$

Fix x_t and flow y by time t. That is, in (5.5), substitute y by x_t, x by y and x_t by y_t. We get

$$\begin{aligned} d^2(y_t, x_t) &\leq d^2(y, x_t) + t^2 - 2t \cdot (u(x_t) - u(y)) \\ &\leq d^2(x_t, y) + t^2 - 2t \cdot (u(x_t) - u(x)) \\ &= d^2(x_t, y) - t^2. \end{aligned} \tag{5.7}$$

Combine (5.6) and (5.7). We have

$$d^2(x_t, y_t) \leq d^2(x_t, y) - t^2 \leq d^2(x,y).$$

By (B), we have $d(x_t, y_t) = d(x,y)$. Thus the Pythagorean Theorem $d^2(x_t, y) = d^2(x,y) + t^2$ follows.

Suppose that the statement has been proved for k. Apply the previous argument on the 0-splitting function $u_{k+1}\colon B_5(p) \to \mathbb{R}$, we have that $B_1(p)$ is isometric to a ball in $Z \times \mathbb{R} \in \text{Alex}^n(0)$, and it splits off $\mathbb{R}^1$ along the direction ∇u_{k+1}. Note that $Z \times \mathbb{R} \in \text{Alex}^n(0)$ if and only if $Z \in \text{Alex}^{n-1}(0)$. Thus restricted on

$Z \times \{0\} \in \mathrm{Alex}^{n-1}(0)$, the map $(u_1, \ldots, u_k)$ is $(k, 0)$-splitting. Then the result follows from the inductive hypothesis on k.

Assertion (ii) is a consequence of (iii). The proof of (iii) is standard, for instance if $u_i(x) = d(a_i, x)$, then by the arguments used in Sections 5.6 – 5.7 in [2] we have that $u = (u_1, \ldots, u_k)$ is a $(k, 100\delta)$-splitting function on $B_{\delta R}(p)$, if $\delta = \delta(n)$ is chosen sufficiently small. □

The following statement is an easy consequence of Proposition 5.1.

Corollary 5.2. *For any $n, k \in \mathbb{N}$ and $\epsilon > 0$, there exists $\delta = \delta(n, \epsilon) > 0$ so that if $X \in Alex^n(-\delta)$ and $B_5(p)$ is (k, δ)-splitting, then $B_r(x)$ is (k, ϵ)-splitting for every $x \in B_1(p)$ and every $r \in (0, 1]$.*

5.2. **Strong and weak singularity.** In this Subsection we discuss the relations between the strong and weak quantitative singular sets. In fact, they are equivalent in some sense for Alexandrov spaces.

Define weak singular sets

$$\widetilde{\mathcal{S}}^k_{\epsilon,r}(X) = \{x \in X : B_s(x) \text{ is not } (k+1, \epsilon)\text{-splitting for every } s \in (r, 1]\}. \tag{5.8}$$

It's clear that $\widetilde{\mathcal{S}}^k_{\epsilon,r}(X) \subseteq \mathcal{S}^k_{\epsilon,r}(X)$.

By Corollary 5.2, we have

Proposition 5.3. *For any $n, \epsilon > 0$, there exists $\delta(n, \epsilon) > 0$ such that for any $X \in Alex^n(-\delta)$ and $0 < r \leq 1$, we have*

$$\widetilde{\mathcal{S}}^k_{\epsilon,r}(X) \subseteq \mathcal{S}^k_{\epsilon,r}(X) \subseteq \widetilde{\mathcal{S}}^k_{\delta,r}(X). \tag{5.9}$$

The quantitative singular sets defined for the Ricci cases in [5] is as follows. Note that we do not use it in this paper and it may be skipped. We are presenting this for comparison sake to the Ricci curvature context.

Definition 5.4 (Quantitative symmetric)**.**

(1) Given a metric space Y and $k \in \mathbb{N}$, we say that Y is k-symmetric if $Y \equiv \mathbb{R}^k \times C(\Sigma)$ for some metric space Σ.
(2) Given $x \in X$ we say that $B_r(x)$ is (k, ϵ)-symmetric if there exists a k-symmetric space Y such that $d_{GH}(B_r(x), B_r(y)) \leq \epsilon r$, where $y \in Y$ is a cone point.

Define

$$\mathcal{WS}^k_{\epsilon,r}(X) \equiv \{x \in X : B_s(x) \text{ is not } (k+1, \epsilon)\text{-symmetric, for every } s \in (r, 1]\}. \tag{5.10}$$

It's clear that $\widetilde{\mathcal{S}}^k_{\epsilon,r}(X) \subseteq \mathcal{WS}^k_{\epsilon,r}(X)$.

The following is an easy lemma, by a standard contradiction argument.

Lemma 5.5. *For each $n \in \mathbb{N}$ and $\epsilon > 0$ there exists $\delta(n, \epsilon) > 0$ such that the following holds for any metric space (X, p). If $B_r(p)$ is both $(0, \delta)$-symmetric and (k, δ)-splitting, then $B_r(p)$ is (k, ϵ)-symmetric.*

Proposition 5.6. *For any $\epsilon > 0$, there exist $\eta(n,\epsilon)$ and $\delta(n,\epsilon) > 0$ such that for any $X \in Alex^n(-\delta)$ and $0 < r \le 1$, we have*

$$\mathcal{WS}^k_{\epsilon,\eta r}(X) \subseteq \widetilde{\mathcal{S}}^k_{\delta,r}(X) \subseteq \mathcal{WS}^k_{\delta,r}(X). \tag{5.11}$$

Proof. If $x \notin \widetilde{\mathcal{S}}^k_{\delta,r}(X)$, then $B_s(x)$ is $(k+1,\delta)$-splitting for some $s \ge r$. By Corollary 5.2, we have that $B_t(x)$ is $(k+1,\delta_1)$-splitting for all $t \in (0, \frac{1}{5}r]$. On the other hand, by Lemma 4.6, there exists $\eta(n,\delta_1) > 0$ and $r_x \in [\eta r, \frac{1}{5}r]$ such that $B_{r_x}(x)$ is $(0,\delta_1)$-symmetric. Due to Lemma 5.5, with appropriately selected δ and δ_1, we have that $B_{r_x}(x)$ is $(k+1,\epsilon)$-symmetric. Therefore, $x \notin \mathcal{WS}^k_{\epsilon,\eta r}(X)$. □

Remark 5.1. Our notion of quantitative splitting for Alexandrov spaces is also equivalent to those defined using strainers. In particular, there exists $0 < \delta_1(n,\epsilon) < \delta_2(n,\epsilon)$ so that

$$\mathcal{S}^k_{\epsilon,r/5}(X) \subseteq \{x \in X : x \text{ does not admit any } (k+1,\delta_2)\text{-strainer with size } \ge r\} \subseteq \mathcal{S}^k_{\delta_1,r}(X) \tag{5.12}$$

for any $X \in \text{Alex}^n(-\delta_1)$ and $0 < r \le 1$.

Remark 5.2. By a similar argument, one can show that if X is a v-non-collapsed limit of n-dimensional manifolds with $\text{Ric} \ge -1$, then there exist $\eta_i(n,\epsilon,v) > 0$, $i = 1,2$, such that

$$\mathcal{WS}^k_{\epsilon,\eta_1 r}(X) \subseteq \widetilde{\mathcal{S}}^k_{\eta_2,r}(X) \subseteq \mathcal{S}^k_{\eta_2,r}(X).$$

However, the statement in the form $\mathcal{S}^k_{\epsilon,\eta_1 r}(X) \subseteq \mathcal{WS}^k_{\eta_2,r}(X)$ doesn't hold for the Ricci case.

5.3. **Dimension reduction.** Note that in a metric cone $C(\Sigma)$, the tangent cone at any point $p \in C(\Sigma)$ away from the cone point splits off an extra $\mathbb{R}$-factor in comparison to $C(\Sigma)$. This is the basis of Federer dimension reduction. The following lemma is a quantitative version of this on Alexandrov spaces.

Lemma 5.7. *For any $n,k \in \mathbb{N}$ and $\epsilon > 0$, there exists $\delta = \delta(n,\epsilon)$ and $\beta = \beta(n,\epsilon) > 0$ such that the following holds for any $(X,p) \in Alex^n(-\delta)$ and (k,δ)-splitting function $u = (u_1,\dots,u_k)\colon B_{50}(p) \to \mathbb{R}^k$. Let $x \in B_1(p)$ and $y \in X$ with $d(x,y) = r > 0$.*

(i) If $T^\delta_x(r,2r) = 0$ and $d(x,y) - d(u(x),u(y)) > \epsilon r$, then $B_s(y)$ is $(k+1,\epsilon)$-splitting for every $0 < s \le \beta r$.
(ii) If $T^\delta_x(r,2r) = 0$ and $B_s(y)$ is not $(k+1,\epsilon)$-splitting for some $0 < s \le \beta r$, then

$$\Big|d(u(x),u(y)) - d(x,y)\Big| \le \epsilon\, d(x,y). \tag{5.13}$$

Proof. We only need to prove (i) since it is equivalent to (ii), taking in account that $|\nabla u| < 1 + \delta$. Let $\delta = \delta(n,\epsilon)$, $\delta_i = \delta_i(n,\epsilon)$ be constants with $0 < \delta < \delta_1 < \delta_2 < \cdots < \epsilon$.

Let us take $z \equiv u(y)$. Choosing $\delta(n,\epsilon) > 0$ small we have by Proposition 5.1 that there exists $\phi\colon B_{10r}(x) \to u^{-1}(z)$ such that $(u,\phi)\colon B_{10r}(x) \to \mathbb{R}^k \times u^{-1}(z)$ is a $\delta_1 r$-isometry. Let $x_1 = \phi(x) \in u^{-1}(z)$ and $\rho = d(x_1,y)$. See Figure 1.

We first find a splitting function along the slice $u^{-1}(z)$, using that y is away from the cone point x_1. Because $T^\delta_x(r,2r) = 0$, there exists $y' \in X$ so that

$$|d(y,y') - r| \le \delta_1 r, \qquad |d(x,y') - 2r| \le \delta_1 r. \tag{5.14}$$

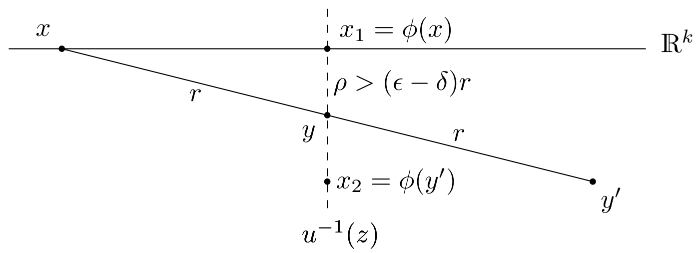

FIGURE 1.

Let $x_2 = \phi(y') \in u^{-1}(z)$. Combine (5.14), $d(x,y) = r$ and the δr-splitting structure on $B_{10r}(x)$. We have

$$|d(x_2, y) - \rho| \le 4\delta_1 r, \qquad |d(x_1, x_2) - 2\rho| \le 4\delta_1 r. \tag{5.15}$$

To see the maximal scales that $B_s(y)$ splits, we need a lower bound of ρ. Because (u, ϕ) is a $\delta_1 r$-isometry and by the assumptions, we have

$$\begin{aligned} \rho = d(x_1, y) &\ge d(x,y) - d(x, x_1) \\ &\ge d(x,y) - d\Big((u,\phi)(x),\ (u,\phi)(x_1)\Big) - \delta_1 r \\ &= d(x,y) - d\Big((u(x), \phi(x)),\ (u(y), \phi(x))\Big) - \delta_1 r \\ &= d(x,y) - d\Big(u(x),\ u(y)\Big) - \delta_1 r \\ &> (\epsilon - \delta_1) r. \end{aligned} \tag{5.16}$$

Choosing δ_1 small and by (5.15), we have that $\{x_1, x_2\}$ forms a $(1, \delta_2)$-strainer. Thus $u_{k+1}(q) \equiv d(q, x_1)$ is a $(1, \delta_3)$-splitting map on $B_{\frac{1}{5}\epsilon r}(y)$.

By the δ_1-almost splitting structure of $B_{10}(p)$ and the Toponogov comparison using (5.15), we have

$$|\langle \nabla_q u_i, \nabla_q u_{k+1} \rangle| < \delta_3. \tag{5.17}$$

for any $q \in B_{\frac{1}{5}\epsilon r}(y)$ and every $i = 1, 2, \dots k$. Thus the function $(u, u_{k+1})|_{B_{\frac{1}{5}\epsilon r}(y)}$ is a $(k+1, \delta_3)$-splitting map. By Proposition 5.1 (i), $B_s(y)$ is ϵ-splitting for every $0 < s \le \frac{1}{50}\epsilon r$. □

Using Lemma 5.7, we can prove the rectifiability.

Proof of Theorem 1.6. Note that $(\mathcal{S}^k_\epsilon \setminus \mathcal{S}^{k-1}_\delta) \cap B_1$ can be covered by countably many balls $\{B_{\delta r_i}(x_i)\}$ with $x_i \notin \mathcal{S}^k_{\delta, 50 r_i}$. That is, $B_{50 r_i}(x_i)$ is (k, δ)-splitting. By Proposition 5.1, for each of $B_{\delta r_i}(x_i)$, there exists a δ_1-splitting map $u_i \colon B_{50\delta r_i}(x_i) \to \mathbb{R}^k$. Note that $\dim_{\mathcal{H}}(\mathcal{S}^{k-1}_\delta) \le k - 1$. Thus it suffices to prove the following statement. There exists $\delta = \delta(n, \epsilon) > 0$, such that if there is a (k, δ)-splitting map $u \colon B_{50}(p) \to \mathbb{R}^k$, then $\mathcal{S}^k_\epsilon \cap B_1(p)$ is k-rectifiable.

Let $\delta(n, \epsilon) > 0$ be determined later. Recall that by Lemma 2.2, for every point $x \in X$, the number of δ-bad scales is at most $N(n, \delta)$. For each $x \in B_1(p)$, let $s_x \in (0, 1]$ be the minimum of 1 and the smallest δ-bad scale at x. Put $\Gamma^t_x = B_{t/2}(x) \cap \mathcal{S}^k_\epsilon \cap \{y \in B_1(p) : s_y > 2t\}$. We claim that for any $t > 0$, the map $u|_{\Gamma^t_x} \colon \Gamma^t_x \to \mathbb{R}^k$

is bi-Lipschitz onto its image. Once the claim is proved, we have that $\mathcal{S}^k_\epsilon \cap \{y \in B_1(p) : s_y > 2t\} = \cup_{x\in B_1(p)}\Gamma^t_x$ is k-rectifiable. Therefore,

$$\mathcal{S}^k_\epsilon \cap B_1(p) = \bigcup_{t>0}\Big(\mathcal{S}^k_\epsilon \cap \{y \in B_1(p) : s_y > 2t\}\Big)$$

is rectifiable.

Now we prove the claim. Let $x_1, y_1 \in \Gamma^s_x$. Then $d(x_1, y_1) \le t < s_{x_1}/2$. Because $Bad^\delta(x_1) \cap [0, s_{x_1}) = \varnothing$, by Proposition 4.1, we have $T^\delta_{x_1}(0, 2t) = T^\delta_{x_1}(0, s_{x_1}) = 0$. Note that $y_1 \in \mathcal{S}^k_\epsilon$ and thus $B_\rho(y_1)$ is not $(k+1, \epsilon)$-splitting for every $\rho \in (0, 1]$. By Lemma 5.7 (ii), we have

$$\Big|d(u(x_1), u(y_1)) - d(x_1, y_1)\Big| \le \epsilon\, d(x_1, y_1). \tag{5.18}$$

□

6. Packing Estimates

We prove Theorem 1.3 in this section. The following is the key lemma.

Lemma 6.1. *For any $n \in \mathbb{N}$ and $\epsilon > 0$, there exist $\delta(n, \epsilon) > 0$ and $\beta(n, \epsilon) > 0$ so that the following holds for any $(X, p) \in Alex^n(-\delta)$. Suppose $u\colon B_{50}(p) \to \mathbb{R}^k$ is a (k, δ)-splitting function, and let $\{B_{r_i}(x_i)\}$ with $i \in \mathbb{I}$ be a disjoint collection of balls living on a fixed level set $x_i \in u^{-1}(z) \cap B_1(p)$ for some $z \in \mathbb{R}^k$. Then if $x_i \in \mathcal{S}^k_{\epsilon,\beta r_i}$ we have the estimate $|\mathbb{I}| < N(n, \epsilon)$.*

Proof. We will construct a sequence Vitali coverings of $u^{-1}(z) \cap B_1(p)$, which "converges" to $\{B_{r_i}(x_i),\ i \in \mathbb{I}\}$. The constants $\delta(n, \epsilon)$, $\eta(n, \epsilon) > 0$ and $\lambda > 0$ will be determined later.

(Step 1.) Let $\bar{B}_\rho$ be an arbitrary closed ball with $W \subseteq \bar{B}_\rho$ be a closed subset and $\mathbb{I}(W) = \{i \in \mathbb{I} : x_i \in W\}$. For $x \in W$ and $0 < \epsilon, s \le 1$, define function

$$\sigma(x, \epsilon, s) = \begin{cases} \inf\Big\{\tau : T^\epsilon_x(\tau s, 2s) = 0\Big\}, & \text{if} \quad T^\epsilon_x(s, 2s) = 0; \\ 1, & \text{otherwise.} \end{cases} \tag{6.1}$$

By Lemma 4.6, for each $0 < \lambda < 4^{-1}$ there exists $\eta = \eta(n, \epsilon, \lambda) > 0$ such that for any $x \in W$, there exists $r_x \in [\eta\rho, \rho]$ such that $T^\epsilon_x(\lambda r_x, 2r_x) = 0$. Therefore, we have

$$\lambda_x \equiv \sigma(x, \epsilon,\ r_x) \le \lambda \tag{6.2}$$

Define $\mathcal{F}(W) = \{i \in \mathbb{I}(W) : r_i \ge \frac{1}{10}\eta\rho\}$ and $\mathcal{F}^c(W) = \mathbb{I}(W) \setminus \mathcal{F}(W)$. It is clear that $|\mathcal{F}(W)| \le N(n, \eta)$, since $r_i \ge \frac{1}{10}\eta\rho$ and $W \subseteq \bar{B}_\rho$. Now because $r_i < \frac{1}{10}\eta\rho \le \frac{1}{10}r_{x_i}$ for every $i \in \mathcal{F}^c(W)$, we have that $\{B_{\frac{1}{10}r_{x_i}}(x_i), i \in \mathcal{F}^c(W)\}$ is a covering of $\cup_{i\in\mathcal{F}^c(W)}B_{r_i}(x_i)$. Let $\mathcal{G}(W) \subseteq \mathcal{F}^c(W)$ be a collection of indices so that $\{B_{\frac{1}{10}r_{x_j}}(x_j),\ j \in \mathcal{G}(W)\}$ covers $\cup_{i\in\mathcal{F}^c(W)}B_{r_i}(x_i)$, while $\{B_{\frac{1}{50}r_{x_j}}(x_j),\ j \in \mathcal{G}(W)\}$ are disjoint. It's clear that $|\mathcal{G}(W)| \le N(n, \eta)$, since $r_{x_i} \ge \eta\rho$. Now we have

$$\mathbb{I}(W) \subseteq \mathcal{F}(W) \cup \left(\underset{j\in\mathcal{G}(W)}{\cup}\mathbb{I}\left(B_{\frac{1}{10}r_{x_j}}(x_j)\right)\right), \tag{6.3}$$

where $|\mathcal{F}(W)| + |\mathcal{G}(W)| \le N(n, \eta)$.

Note that function $\sigma(x, \epsilon, s)$ is semi-continuous in x. That is, $\liminf_{z\to y} \sigma(z, \epsilon, r_x) \ge \sigma(y, \epsilon, r_x)$. For each $j \in \mathcal{G}(W)$, there exists $y_j \in \bar{B}_{\lambda_{x_j} r_{x_j}}(x_j) \cap W$ so that

$$\sigma_{y_j} \equiv \sigma(y_j, \epsilon, r_{x_j}) = \inf\{\sigma(x, \epsilon, r_{x_j}) : x \in \bar{B}_{\lambda_{x_j} r_{x_j}}(x_j)\} \le \lambda_{x_j} \le \lambda. \tag{6.4}$$

We claim that $\mathcal{F}(W)$ and $\mathcal{G}(W)$ satisfy the following properties.

(1) $|\mathcal{F}(W)| + |\mathcal{G}(W)| \le N(n, \eta)$.
(2) $\mathbb{I}(W) = \mathcal{F}(W) \cup \left(\bigcup_{j\in\mathcal{G}(W)} \mathbb{I}\left(B_{\lambda_{x_j} r_{x_j}}(x_j) \cap B_{\sigma_{y_j} r_{x_j}}(y_j)\right)\right)$.
(3) If $\sigma_{y_j} > 0$, then for every $z \in \bar{B}_{\lambda_{x_j} r_{x_j}}(x_j) \cap \bar{B}_{\sigma_{y_j} r_{x_j}}(y_j)$

$$\left|Bad^\epsilon(z) \cap [\sigma_{y_j} r_{x_j}, 1]\right| \ge \left|Bad^\epsilon(z) \cap [\rho, 1]\right| + 1. \tag{6.5}$$

Statement (1) has been proved in the construction. To prove (2), we start with an obvious inclusion formula:

$$\begin{aligned} B_{\frac{1}{10} r_{x_j}}(x_j) \subseteq &\left(A^{\frac{1}{10} r_{x_j}}_{\lambda_{x_j} r_{x_j}}(x_j) \cup A^{\frac{1}{10} r_{x_j}}_{\sigma_{y_j} r_{x_j}}(y_j)\right) \cup \left(\bar{B}_{\lambda_{x_j} r_{x_j}}(x_j) \cap \bar{B}_{\sigma_{y_j} r_{x_j}}(y_j)\right) \\ &\cup \left(\bar{B}_{\lambda_{x_j} r_{x_j}}(x_j) \setminus \bar{B}_{r_{x_j}}(y_j)\right) \cup \left(\bar{B}_{\sigma_{x_j} r_{x_j}}(y_j) \setminus \bar{B}_{r_{x_j}}(x_j)\right). \end{aligned} \tag{6.6}$$

Let $\lambda < \frac{1}{10}$ be a constant. Note that $d(x_j, y_j) \le \sigma_{y_j} r_{x_j} \le \lambda_{x_j} r_{x_j} \le r_{x_j}/10$. Thus we have $\bar{B}_{\lambda_{x_j} r_{x_j}}(x_j) \subseteq \bar{B}_{r_{x_j}}(y_j)$ and $\bar{B}_{\sigma_{y_j} r_{x_j}}(y_j) \subseteq \bar{B}_{r_{x_j}}(x_j)$. In particular we then have the better inclusion

$$B_{\frac{1}{10} r_{x_j}}(x_j) \subseteq \left(A^{\frac{1}{10} r_{x_j}}_{\lambda_{x_j} r_{x_j}}(x_j) \cup A^{\frac{1}{10} r_{x_j}}_{\sigma_{y_j} r_{x_j}}(y_j)\right) \cup \left(\bar{B}_{\lambda_{x_j} r_{x_j}}(x_j) \cap \bar{B}_{\sigma_{y_j} r_{x_j}}(y_j)\right). \tag{6.7}$$

It remains to show that

$$\mathbb{I}\left(A^{\frac{1}{10} r_{x_j}}_{\lambda_{x_j} r_{x_j}}(x_j)\right) = \mathbb{I}\left(A^{\frac{1}{10} r_{x_j}}_{\sigma_{y_j} r_{x_j}}(y_j)\right) = \emptyset. \tag{6.8}$$

Suppose $\mathbb{I}\left(A^{\frac{1}{10} r_{x_j}}_{\lambda_{x_j} r_{x_j}}(x_j)\right) \neq \emptyset$. That is, there exists $i \in \mathcal{F}^c(W)$, so that $\lambda_{x_j} r_{x_j} \le d(x_i, x_j) \le \frac{1}{10} r_{x_j}$. Then from the definition of r_{x_j} we have

$$T^\epsilon_{x_j}\Big(d(x_i, x_j),\ 2d(x_i, x_j)\Big) = 0\,. \tag{6.9}$$

Now let $\delta = \delta(n, \epsilon) > 0$ and $\beta = \beta(n, \epsilon) > 0$ be the constants determined in Lemma 5.7. Because $B_{\beta r_i}(x_i)$ is not $(k+1, \epsilon)$-splitting and $r_i \le d(x_i, x_j)$, the restricted map $u|_{\{x_i, x_j\}}$ is $(1 \pm 4\epsilon)$-bi-Lipschitz. This contradicts to the assumption $u(x_i) = u(x_j) = z$. The proof for $\mathbb{I}\left(A^{\frac{1}{10} r_{x_j}}_{\sigma_{y_j} r_{x_j}}(y_j)\right) = \emptyset$ is similar.

To prove (3), let $z \in \bar{B}_{\lambda_{x_j} r_{x_j}}(x_j) \cap \bar{B}_{\sigma_{y_j} r_{x_j}}(y_j) \in \mathcal{D}(W_r)$. By the definition of σ_{y_j}, we have $\sigma(z, \epsilon, r_{x_j}) \ge \sigma_{y_j} > 0$. Thus $T^\epsilon_z\left(\frac{1}{2}\sigma_{y_j} r_{x_j}, r_{x_j}\right) = 1$. By the definition of bad scales, this implies $\left|Bad^\epsilon(z) \cap [\sigma_{y_j} r_{x_j}, r_{x_j}]\right| \ge 1$. Then (3) follows since $[\sigma_{y_j} r_{x_j}, r_{x_j}] \subseteq [\sigma_{y_j} r_{x_j}, \rho]$.

(Step 2.) In this step we construct a covering of $\mathbb{I}$ inductively. Let the decomposition functions $\mathcal{F}$ and $\mathcal{G}$ be defined in Step 1. Begin with $W = B_1(p)$. Let $\mathcal{C}_1 = \mathcal{F}(W)$ and $\mathcal{D}_1 = \mathcal{G}(W)$. Suppose $\mathcal{C}_k$ and $\mathcal{D}_k$ have been constructed and satisfy the following $(A_k) - (C_k)$:

(A_k) $|\mathcal{C}_k| \le kN(n,\eta)^k$, $|\mathcal{D}_k| \le N(n,\eta)^k$.

(B_k) $\mathbb{I} = \mathcal{C}_k \cup \left(\bigcup_{j\in\mathcal{D}_k} \mathbb{I}\left(B_{\lambda_{x_j} r_{x_j}}(x_j) \cap B_{\sigma_{y_j} r_{x_j}}(y_j) \right) \right)$.

(C_k) $\left| Bad^\epsilon(z) \cap [\sigma_{y_j} r_{x_j}, 1] \right| \ge k$ for any $j \in \mathcal{D}_k$ and $z \in B_{\lambda_{x_j} r_{x_j}}(x_j) \cap B_{\sigma_{y_j} r_{x_j}}(y_j)$, provided $\sigma_{y_j} > 0$.

For each $j \in \mathcal{D}_k$ and $W_j = \bar{B}_{\lambda_{x_j} r_{x_j}}(x_j) \cap \bar{B}_{\sigma_{y_j} r_{x_j}}(y_j)$, using the construction of Step 1 let

$$\mathcal{C}_{k+1} = \mathcal{C}_k \cup \left(\cup_{j\in\mathcal{D}_k} \mathcal{F}(W_j) \right)$$

and

$$\mathcal{D}_{k+1} = \cup_{j\in\mathcal{D}_k} \mathcal{G}(W_j).$$

Now we prove $(A_{k+1}) - (C_{k+1})$ for $\mathcal{C}_{k+1}$ and $\mathcal{D}_{k+1}$. By (1) in Step 1, we have $|\mathcal{F}(W_j)| + |\mathcal{G}(W_j)| \le N(n,\eta)$. Thus

$$|\mathcal{C}_{k+1}| \le |\mathcal{C}_k| + N|\mathcal{D}_k| \le kN^k + N^{k+1} \le (k+1)N^{k+1}$$

and

$$|\mathcal{D}_{k+1}| \le N|\mathcal{D}_k| \le N^{k+1}.$$

Statements (B_{k+1}) and (C_{k+1}) follow from (2) and (3) respectively.

(Step 3.) By Lemma 2.2, the number of ϵ-bad scales is at most $K = K(n,\epsilon)$. Thus due to (C_k), we have $\mathcal{D}_k = \varnothing$ if $k > K$. Therefore, $\mathbb{I} = \mathcal{C}_K$ and $|\mathbb{I}| = |\mathcal{C}_K| \le KN^K$. □

Furthermore, we have the following theorem.

Theorem 6.2. *For any $n \in \mathbb{N}$, $\epsilon > 0$ and $\Lambda \ge 1$, there exist $\delta(n,\epsilon) > 0$ and $\beta(n,\epsilon) > 0$ so that the following holds for any $(X,p) \in Alex^n(-1)$. Suppose that there is a (k,δ)-splitting function $u\colon B_{50}(p) \to \mathbb{R}^k$. If $\{B_{r_i}(x_i)\}$ are disjoint and $B_{\beta\Lambda r_i}(x_i) \cap \mathcal{S}^k_{\epsilon,\beta\Lambda r_i} \ne \varnothing$ for all $i \in \mathbb{I}$, then for any $z \in \mathbb{R}^k$, we have*

$$\left| \left\{ i \in \mathbb{I}\colon B_{\beta\Lambda r_i}(x_i) \cap u^{-1}(z) \ne \varnothing \right\} \right| < N(n,\epsilon,\Lambda). \tag{6.10}$$

Additionally, if $r_i = r$ with $B_{\Lambda r}(x_i) \cap \mathcal{S}^k_{\epsilon,\Lambda r} \ne \varnothing$ and $\{B_r(x_i)\}$ are disjoint for all $i \in \mathbb{I}$, then for any $z \in \mathbb{R}^k$, we have

$$\left| \left\{ i \in \mathbb{I}\colon B_{\Lambda r}(x_i) \cap u^{-1}(z) \ne \varnothing \right\} \right| < N(n,\epsilon,\Lambda). \tag{6.11}$$

Proof. Let $\bar{x}_i \in B_{\beta\Lambda r_i}(x_i) \cap u^{-1}(z)$ and $y_i \in B_{\beta\Lambda r_i}(x_i) \cap \mathcal{S}^k_{\epsilon,\beta\Lambda r_i}$. There exists $\eta(n,\epsilon) > 0$ such that $\bar{x}_i \in \mathcal{S}^k_{\eta,10\beta\Lambda r_i}$, since $B_{10\beta\Lambda r_i}(\bar{x}_i) \supseteq B_{\beta\Lambda r_i}(y_i)$ and $B_{\beta\Lambda r_i}(y_i)$ is not (k,ϵ)-splitting. Moreover, we have that $B_{r_i/2}(\bar{x}_i)$ are disjoint, because $B_{r_i/2}(\bar{x}_i) \subseteq B_{r_i}(x_i)$. Estimate (6.10) follows by applying Lemma 6.1 to the collection $\{B_{r_i/2}(\bar{x}_i)\}$.

To prove (6.11), one can go through the proof of Lemma 6.1 and (6.10) with small modifications, or use the following re-covering arguments. Let $r' = r/\beta$. Then we have $B_{\beta\Lambda r'}(x_i) = B_{\Lambda r}(x_i)$. The given conditions $B_{\Lambda r}(x_i) \cap \mathcal{S}^k_{\epsilon,\Lambda r} \ne \varnothing$ and $B_{\Lambda r}(x_i) \cap u^{-1}(z) \ne \varnothing$ are equivalent to $B_{\beta\Lambda r'}(x_i) \cap \mathcal{S}^k_{\epsilon,\beta\Lambda r'} \ne \varnothing$ and

$B_{\beta\Lambda r'}(x_i)\cap u^{-1}(z)\neq\emptyset$, respectively. The collection $\{B_{r'}(x_i)\}$ is not disjoint, so we can't use (6.10) directly. However, note that if $B_{r'}(x_i)\cap B_{r'}(x_j)\neq\emptyset$, then $B_r(x_j)\subseteq B_{2r'}(x_i)$. Because $\{B_r(x_i)\}$ are disjoint, for every i, there are at most $N(n, r'/r) = N(n,\beta)$ balls $B_{r'}(x_j)$ such that $B_{r'}(x_i)\cap B_{r'}(x_j)\neq\emptyset$. Therefore, the collection $\{B_{r'}(x_i)\}$ can be written as the union of $N(n,\beta)$ disjoint collections. Then the result follows from (6.10). □

Let us now remark on a standard covering argument. Let $\mathfrak{B}$ be a collection of sets. The intersection number $\mathcal{N}(\mathcal{B})$ of $\mathfrak{B}$ is the minimum number k so that $B_1\cap B_2\cap\cdots\cap B_{k+1}=\emptyset$ for any $B_1, B_2,\ldots,B_{k+1}\in\mathfrak{B}$. In particular, if $\mathcal{N}(\mathcal{B})=1$, then $\mathfrak{B}$ is a disjoint collection. We have the following easy lemma:

Lemma 6.3. *Let $B_R(0)\subset\mathbb{R}^k$ and $\mathfrak{B}=\{B_{r_i}(x_i)\subseteq B_R(0)\}$ be a collection of balls. If the intersection number $\mathcal{N}(\mathfrak{B})\leq N<\infty$, then $\sum r_i^k < N\cdot C(k)R^k$.*

Now let us prove a local version of Theorem 1.3.

Lemma 6.4 (Local packing estimate)**.** *For any $n\in\mathbb{N}$, $\epsilon>0$, $R\leq 1$ and $\Lambda\geq 1$, there exists $\delta(n,\epsilon)>0$ and $\beta(n,\epsilon)>0$ so that the following hold for any $(X,p)\in Alex^n(-1)$, provided that $B_{500R}(p)$ is (k,δ)-splitting.*

(i) If $x_i\in\mathcal{S}^k_{\epsilon,\beta r_i}\cap B_R(p)$ with $r_i\leq R$ and $\{B_{r_i}(x_i)\}$ are disjoint for all $i\in\mathbb{I}$, then $\sum_{i\in\mathbb{I}} r_i^k < C(n,\epsilon)R^k$.

(ii) If $x_i\in\mathcal{S}^k_{\epsilon,\Lambda r}\cap B_R(p)$ with $r\leq R$ and $\{B_r(x_i)\}$ are disjoint for all $i\in\mathbb{I}$, then $|\mathbb{I}|<C(n,\epsilon,\Lambda)(R/r)^k$.

Proof. We prove (i) only and the proof of (ii) is similar, modulo (6.11). By Proposition 5.1, there is a δ_1-splitting map $u\colon B_{50R}(p)\to\mathbb{R}^k$. Assume $u(p)=0^k\in\mathbb{R}^k$.

Consider the collection of balls $\mathfrak{B}=\{B_{\frac{1}{2}\beta r_i}(u(x_i)), i\in\mathbb{I}\}$ in $\mathbb{R}^k$. Because u is 1-Lipschitz, we have that $B_{\frac{1}{2}\beta r_i}(u(x_i))\subseteq B_{2R}(0^k)$. Given $z\in\mathbb{R}^k$, let $\mathbb{I}_z=\{i\in\mathbb{I}: z\in B_{\frac{1}{2}\beta r_i}(u(x_i))\}$. By Proposition 5.1 again, we have $u^{-1}(z)\cap B_{\beta r_i}(x_i)\neq\emptyset$. It follows from (6.10) that $|\mathbb{I}_z|\leq N(n,\epsilon)$. This shows that the intersection number $\mathcal{N}(\mathfrak{B})\leq N(n,\epsilon)$. Then the desired result follows from Lemma 6.3. □

Now we prove Theorem 1.3 by showing the following stronger statement.

Theorem 6.5 (Packing estimate)**.** *Lemma 6.4 still holds if the splitting assumption is dropped.*

Proof. We prove by induction on k. Let $0<\delta'(n,\epsilon)<\delta(n,\epsilon)<\delta_1(n,\epsilon)<\epsilon$ be determined latter. The constant C may vary line by line. Lemma 6.4 proves the case for $k=0$ as well as the case that $B_{500R}(p)$ is $(k+1,\delta_1)$-splitting. Assume that (i) and (ii) are true for $k<n$. We will prove them for $k+1$, assuming that $B_{500R}(p)$ is not $(k+1,\delta_1)$-splitting.

Not losing generality, assume $R=\frac{1}{500}$. That is, $B_1(p)$ is not $(k+1,\delta_1)$-splitting. We begin with a decomposition of $B_1(p)$. Let $R_\alpha=2^{-\alpha}$, $\alpha\in\mathbb{Z}$. Recall the definition of the weak (k,δ)-singular set $\widetilde{\mathcal{S}}^k_{\delta,r}$ in (5.8). By Proposition 5.3, we have $B_1(p)\subseteq\widetilde{\mathcal{S}}^k_{\delta,10}$. Thus

$$B_1(p)\setminus\widetilde{\mathcal{S}}^k_{\delta}\subseteq\widetilde{\mathcal{S}}^k_{\delta,10}\setminus\widetilde{\mathcal{S}}^k_{\delta}\subseteq\bigcup_{\alpha=-4}^{\infty}\left(\widetilde{\mathcal{S}}^k_{\delta,R_\alpha}\setminus\widetilde{\mathcal{S}}^k_{\delta,R_{\alpha+1}}\right). \tag{6.12}$$

For each α, let

$$\left\{B_{\rho_\alpha}(y_j^\alpha),\ j \in \mathbb{J}_\alpha\right\} \subseteq \left\{B_{\frac{1}{20}R_\alpha}(y),\ y \in \widetilde{\mathcal{S}}^k_{\delta, R_\alpha} \setminus \widetilde{\mathcal{S}}^k_{\delta, R_{\alpha+1}}\right\} \tag{6.13}$$

be a Vitali covering of $\left(\widetilde{\mathcal{S}}^k_{\delta, R_\alpha} \setminus \widetilde{\mathcal{S}}^k_{\delta, R_{\alpha+1}}\right)$, for which $\{B_{\frac{1}{5}\rho_\alpha}(y_j)\}$ are disjoint but $\{B_{\rho_\alpha}(y_j)\}$ is a covering. A useful property for this decomposition is that for each $y \in \{y_j^\alpha\}$, we have that $B_{20\rho_\alpha}(y) = B_{R_\alpha}(y)$ is not $(k+1, \delta)$-splitting, but $B_{10\rho_\alpha}(y) = B_{\frac{1}{2}R_\alpha}(y)$ is $(k+1, \delta)$-splitting.

We first prove (ii), which will be needed in the proof of (i). By the inductive hypothesis, we only need to consider the collection of balls $\left\{B_r(x_i) : x_i \in \mathcal{S}^{k+1}_{\epsilon, \Lambda r} \text{ but } x_i \notin \mathcal{S}^k_{\delta', r}\right\}$, where $\delta' = \delta'(n, \delta) > 0$ will be determined latter. For each $j \in \mathbb{J}_\alpha$, because $y_j^\alpha \in \widetilde{\mathcal{S}}^k_{\delta, R_\alpha} \subseteq \mathcal{S}^k_{\delta, R_\alpha}$, by the inductive hypothesis, we have an upper bound on the number of these balls:

$$|\mathbb{J}_\alpha| \leq C(n, \epsilon) R_\alpha^{-k}. \tag{6.14}$$

Recall that $x_i \in \mathcal{S}^k_{\epsilon, \Lambda r} \cap B_{1/500}(p)$ with $r \leq 1/500$ and $\{B_r(x_i)\}$ are disjoint. Given $j \in \mathbb{J}_\alpha$, let $\mathbb{I}_j^\alpha = \{i : x_i \in B_{\rho_\alpha}(y_j^\alpha)\}$. We claim that if $\rho_\alpha < r/1000$, then $\mathbb{I}_j^\alpha = \varnothing$ for every j. Suppose $\rho_\alpha < r/1000$ but there is $i \in \mathbb{I}_j^\alpha$ for some j. Note then that $R_\alpha = 20\rho_\alpha < \frac{1}{50} r$, we have $B_{r/5}(x_i) \supseteq B_{20\rho_\alpha}(y_j^\alpha)$. Because $B_{20\rho_\alpha}(y_j^\alpha)$ is not $(k+1, \delta)$-splitting, we have $x_i \in \mathcal{S}^k_{\delta', r}$, for some $\delta'(n, \delta) > 0$. This contradicts to the assumptions.

Now for each $i \in \mathbb{I}_j^\alpha$, we have $x_i \in \mathcal{S}^{k+1}_{\epsilon, \Lambda r}$, $r \leq 1000\rho_\alpha = 50R_\alpha$, and $B_{10\rho_\alpha}(y_j^\alpha) = B_{\frac{1}{2}R_\alpha}(y)$ is $(k+1, \delta)$-splitting. By Lemma 6.4 (ii) we have

$$|\mathbb{I}_j^\alpha| \leq C(n, \epsilon, \Lambda)(R_\alpha/r)^{k+1}. \tag{6.15}$$

Because $\cup_{\alpha \geq -4}\{B_{\rho_\alpha}(y_j^\alpha)\}$ is a covering of $B_1(p) \setminus \widetilde{\mathcal{S}}^k_\delta \supseteq B_1(p) \setminus \widetilde{\mathcal{S}}^k_{\delta'} \supseteq \{x_i : i \in \mathbb{I}\}$, by (6.14) and (6.15), we have

$$\begin{aligned} |\mathbb{I}| &\leq \sum_{\frac{1}{50}r \leq R_\alpha \leq 10} \sum_{j \in \mathbb{J}_\alpha} |\mathbb{I}_j^\alpha| \\ &\leq \sum_{\alpha=-4}^{\infty} C(n, \epsilon, \Lambda) R_\alpha^{-k} (R_\alpha/r)^{k+1} \leq C(n, \epsilon, \Lambda) r^{-(k+1)}. \end{aligned}$$

We prove (i) in a similar way. By the inductive hypothesis, we only need to consider the balls $\left\{B_{r_i}(x_i) : x_i \in \mathcal{S}^{k+1}_{\epsilon, \beta r_i} \text{ but } x_i \notin \mathcal{S}^k_{\delta', \beta r_i}\right\}$, for some $\delta'(n, \delta) > 0$. Given $j \in \mathbb{J}_\alpha$, let $\mathbb{I}_j^\alpha = \{i : x_i \in B_{\rho_\alpha}(y_j^\alpha)\}$. We claim that for every $i \in \mathbb{I}_j^\alpha$, we have $r_i \leq \frac{1000}{\beta}\rho_\alpha$. If this is not true, then $B_{\beta r_i/5}(x_i) \supseteq B_{20\rho_\alpha}(y_j^\alpha)$. Because $B_{20\rho_\alpha}(y_j^\alpha)$ is not $(k+1, \delta)$-splitting, we have that $B_{\beta r_i}(x_i)$ is not $(k+1, \delta')$-splitting for some $\delta' = \delta'(n, \delta) > 0$. Thus $x_i \in \mathcal{S}^k_{\delta', \beta r_i}$, which contradicts to the assumptions.

Note that $x_i \in \mathcal{S}^{k+1}_{\epsilon, \beta r_i}$ and $B_{10\rho_\alpha}(y_j^\alpha) = B_{\frac{1}{2}R_\alpha}(y)$ is $(k+1, \delta)$-splitting. We can apply Lemma 6.4 (i) and get

$$\sum_{i \in \mathbb{I}_j^\alpha} r_i^{k+1} \leq C(n, \epsilon)\rho_\alpha^{k+1} = C(n, \epsilon) R_\alpha^{k+1}. \tag{6.16}$$

Note that (6.14), which was proved in the course of proving (ii), still holds. Combine (6.14) and (6.16). We have

$$\begin{aligned}\sum_{i\in\mathbb{I}} r_i^{k+1} &\le \sum_{\alpha=-4}^{\infty}\sum_{j\in\mathbb{J}_\alpha}\sum_{i\in\mathbb{I}_j^\alpha} r_i^{k+1}\\ &\le \sum_{\alpha=-4}^{\infty}\sum_{j\in\mathbb{J}_\alpha} C(n,\epsilon)R_\alpha^{k+1}\\ &\le \sum_{\alpha=-4}^{\infty} C(n,\epsilon)R_\alpha^{-k}R_\alpha^{k+1}\\ &\le \sum_{\alpha=-4}^{\infty} C(n,\epsilon)R_\alpha \le C(n,\epsilon). \end{aligned} \tag{6.17}$$

□

7. SHARPNESS OF THE RECTIFIABILITY

In this section we prove Theorem 1.7. Let us begin with a smoothing lemma.

Lemma 7.1. *Let $\mathbb{U}\subset\mathbb{R}^n$ be a compact convex subset and $f\colon\mathbb{U}\to\mathbb{R}$ be a strictly convex function. Let $\Omega=\cup_{i=1}^{\infty}\Omega_i$, where Ω_i are disjoint open convex subsets in $\mathbb{U}$. For any $\delta>0$, there exists a strictly convex function $F\colon\mathbb{U}\to\mathbb{R}$ such that the following hold.*

(i) $F|_\Omega$ is C^∞.
(ii) $F|_{\mathbb{U}\setminus\Omega}=f|_{\mathbb{U}\setminus\Omega}$ and $|F-f|<\delta$ on Ω,
(iii) For any $x\notin\Omega$ and any vector v, it holds that

$$\lim_{t\to0^+}\frac{F(x+tv)-F(x)}{t}=\lim_{t\to0^+}\frac{f(x+tv)-f(x)}{t}. \tag{7.1}$$

In particular, if $Df(x)$ exists at $x\notin\Omega$, then $DF(x)=Df(x)$.

Proof. Let

$$\epsilon(x)=e^{-\frac{\delta}{d(x,\mathbb{U}\setminus\Omega)}} \tag{7.2}$$

be an error function defined on Ω. By Theorem 1.1 in [7], for each i, there exists a strictly C^∞ convex function $g_i\colon\Omega_i\to\mathbb{R}$ such that for any $x\in\Omega_i$, we have

$$|f(x)-g_i(x)|\le\epsilon(x). \tag{7.3}$$

Let $F\colon\mathbb{U}\to\mathbb{R}$ be the gluing of all of g_i and $f|_{\mathbb{U}\setminus\Omega}$. That is,

$$F(x)=\begin{cases} g_i(x), & \text{if}\quad x\in\Omega_i;\\ f(x), & \text{if}\quad x\notin\Omega.\end{cases} \tag{7.4}$$

It is obvious that (i) and (ii) are satisfied. The following estimates (7.5) and (7.6) imply (iii). If $x, y \notin \Omega$, it is obvious that

$$\Big| |F(x) - F(y)| - |f(x) - f(y)| \Big| = 0. \tag{7.5}$$

For any $x \notin \Omega$ and $y \in \Omega$, we have $y \in \Omega_i$ for some i and thus

$$\begin{aligned} \Big| |F(x) - F(y)| - |f(x) - f(y)| \Big| &= \Big| |f(x) - g_i(y)| - |f(x) - f(y)| \Big| \\ &\le |g_i(y) - f(y)| \\ &\le e^{-\frac{\delta}{d(y, \mathbb{U}\setminus\Omega)}} \\ &\le e^{-\frac{\delta}{d(x,y)}}. \end{aligned} \tag{7.6}$$

It's clear that F is strictly convex on each of Ω_i. It remains to show that F is strictly convex, moving out from Ω_i. We need the following two lemmas which we will outline the proof latter. They are well known to the experts.

Lemma 7.2. *A Lipschitz function $h\colon [a,b] \to \mathbb{R}$ is convex if and only if for any non-negative smooth function $\phi\colon [a,b] \to \mathbb{R}$, it holds that*

$$\int_a^b h'(t)\,\phi'(t)\,dt \le h'_-(b)\,\phi(b) - h'_+(a)\,\phi(a). \tag{7.7}$$

Here $h'_\pm$ denote the one-sided derivatives.

Lemma 7.3. *Let $h\colon [a,c] \to \mathbb{R}$ be a Lipschitz function and $b \in [a,c]$. If $h|_{[a,b]}$ and $h|_{[b,c]}$ are both convex functions and $h'_-(b) \le h'_+(b)$, then h is a convex function over $[a,c]$.*

Now we show that F is a convex function. It is obvious that F is locally convex for any $x \notin \partial\Omega_i$. For $x \in \partial\Omega_i$, we show that F is convex along each line passing through x in $\mathbb{U}$. Let $\gamma(s) = x + sv$, $s \in (-\epsilon, \epsilon)$ be a unit speed geodesic in $\mathbb{U}$ and $h(s) = F(x + sv)$. By (7.1) and the fact that f is convex, we have

$$\begin{aligned} h'_-(0) &= \lim_{t\to 0^-} \frac{F(x+tv) - F(x)}{t} \\ &= \lim_{t\to 0^-} \frac{f(x+tv) - f(x)}{t} \le \lim_{t\to 0^+} \frac{f(x+tv) - f(x)}{t} \\ &= \lim_{t\to 0^+} \frac{F(x+tv) - F(x)}{t} = h'_+(0). \end{aligned} \tag{7.8}$$

Then the convexity of F follows from Lemma 7.3. By (iii), F is also strictly convex. □

Proof of Lemma 7.2. The necessity is obvious. To prove the sufficiency it is sufficient to verify

$$\frac{h(t_2) - h(t_1)}{t_2 - t_1} - \frac{h(t_3) - h(t_2)}{t_3 - t_2} \le 0$$

for every $a \le t_1 < t_2 < t_3 \le b$. This can be proved by a direct computation with $\phi(t)$ chosen as a smooth approximation to

$$\psi(t) = \begin{cases} 0, & \text{if} \quad t \le t_1, \\ \frac{t-t_1}{t_2-t_1}, & \text{if} \quad t_1 < t \le t_2, \\ \frac{t_3-t}{t_3-t_2}, & \text{if} \quad t_2 < t < t_3, \\ 0, & \text{if} \quad t \ge t_3. \end{cases}$$

□

Proof of Lemma 7.3. By Lemma 7.2, for any non-negative smooth function $\phi : [a,b] \to [0,\infty)$, we have

$$\int_a^b h'(t)\,\phi'(t)\,dt \le h'_-(b)\,\phi(b) - h'_+(a)\,\phi(a), \tag{7.9}$$

$$\int_b^c h'(t)\,\phi'(t)\,dt \le h'_-(c)\,\phi(c) - h'_+(b)\,\phi(b). \tag{7.10}$$

Sum up the two inequalities and apply Lemma 7.2 again, we get the desired result. □

Proof of Theorem 1.7. Let $Z = \bar{B}_1(O) \subset \mathbb{R}^2$ be a closed unit disk centered at p. Fix $0 < \delta < 1$ and define a strictly concave function on Z:

$$f_0(z) = \begin{cases} \sqrt{d(z,\partial Z)} & \text{if} \quad d(z,\partial Z) \le \frac{1}{4}, \\ \delta \cdot \sqrt{d(z,\partial Z)} + (1-\delta)\cdot\frac{1}{2} & \text{if} \quad d(z,\partial Z) > \frac{1}{4}. \end{cases} \tag{7.11}$$

Let $Z_t = \{z \in Z : f_0(z) \ge t\}$ be the sub-level set. We denote the subgraph of $f : Z \to \mathbb{R}^+$ by

$$G_{Z,f} = \Big\{(z,t) \in Z \times \mathbb{R} : 0 \le t \le f(z)\Big\}.$$

Because f_0 is strictly concave, we have $X_0 = G_{Z,f_0} \in \text{Alex}^3(0)$ with boundary. See Figure 2 below. For $\delta > 0$ small the following hold:

(1) $\mathcal{S}(X_0) = \partial X_0$,
(2) $\mathcal{S}^1_\epsilon(X_0) \setminus \mathcal{S}^0(X_0) = \Big(\partial Z_{1/2} \times \{\frac{1}{2}\}\Big) \cup (\partial Z \times \{0\})$
(3) $\mathcal{S}^0(X_0) = \{P\} = \Big\{\Big(O, \frac{1}{2}(1+\delta)\Big)\Big\}$ is the tip of the graph.

Not losing generality, let $T \subseteq \partial Z_{1/2}$ be any closed subset. Then $\partial Z_{1/2} \setminus T = \cup_{i=1}^\infty U_i$ is a union of disjoint open intervals. Let Ω_i be the open sectors in Z corresponding to the arc U_i. That is, $\Omega_i = \{x \in Z^\circ : \text{ray } \lambda \cdot \overrightarrow{Ox} \cap U_i \neq \emptyset\}$, as the shaded region in Figure 2. Clearly, $\{\Omega_i\}$ is a collection of disjoint open convex sets.

Now apply Lemma 7.1 to $f_0 : Z \to \mathbb{R}$ on $\cup_{i=1}^\infty \Omega_i$ to obtain a strictly convex function $f_1 : Z \to \mathbb{R}$ which is smooth on $\cup_{i=1}^\infty \Omega_i$ and $f_1 = f_0$ away from $\cup_{i=1}^\infty \Omega_i$. Now consider the new subgraph $X_1 = G_{Z,f_1} \in \text{Alex}^3(0)$. Note that if f_1 is smooth at a point $x \in Z^\circ$, then the tangent cone of X_1 at $(x, f_1(x)) \in \partial X_1$ is a three

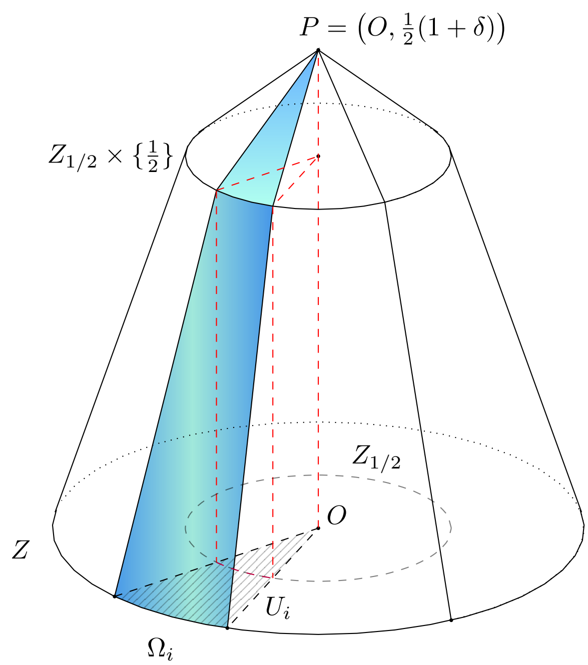

FIGURE 2. $X_0 = G_{Z,f_0} \in \text{Alex}^3(0)$

dimensional half space. If $x \in \partial Z_{1/2} \setminus \cup_{i=1}^{\infty} \Omega_i \equiv T$, then the tangent cone at $(x, f_1(x)) \in \partial X_1$, which is isometric to the tangent cone of X_0 at $(x, f_0(x))$, splits off only $\mathbb{R}^1$. Therefore we have

(4) $\mathcal{S}(X_1) = \partial X_1$,
(5) $\mathcal{S}^1_\epsilon(X_1) \setminus \mathcal{S}^0(X_1) = \big(T \times \{\frac{1}{2}\}\big) \cup (\partial Z \times \{0\})$,
(6) $\mathcal{S}^0(X_1) = \{P\}$ is the tip of the graph.

A similar, but less involved, smoothing procedure can be performed in a small neighborhood of $\partial Z \times \{0\}$ and P so that the resulting space $X_2 \in \text{Alex}^3(0)$ satisfies

(7) $\mathcal{S}(X_2) = \partial X_2$,
(8) $\mathcal{S}^1_\epsilon(X_2) = T \times \{\frac{1}{2}\}$,
(9) $\mathcal{S}^0(X_2) = \varnothing$.

Finally, we double X_2 and arrive at a boundary free space $Y \in \text{Alex}^3(0)$ which satisfies

(10) $\mathcal{S}(Y) = \mathcal{S}^1_\epsilon(Y) = T \times \{\frac{1}{2}\}$;
(11) $\mathcal{S}^0(Y) = \varnothing$.

Let us sketch a smoothing procedure to approximate Y by a sequence of non-collapsed manifolds with sec ≥ 0. By performing similar smoothing procedures, we can approximate X_2 by a sequence of smooth convex bodies in $\mathbb{R}^3$. Thus now it's sufficient to smooth the double, denoted by $\tilde{Y}$, of a compact smooth convex body $\tilde{X} \subset \mathbb{R}^n$.

Let $\mu_\epsilon \colon [0, \epsilon] \to \mathbb{R}$ be a strictly concave function that satisfies:

(1) μ_ϵ is smooth in the open interval $(0, \epsilon)$;
(2) $\mu_\epsilon(0) = 0$ and $\mu_\epsilon(\epsilon) = \epsilon$;
(3) the limits of derivatives $\lim_{t \to 0^+} \dfrac{d^k \mu^{-1}}{dt^k} = 0$ and $\lim_{t \to \epsilon^-} \dfrac{d^k \mu}{dt^k} = 0$ for every $k \geq 1$.

Define

$$h_\epsilon(x) = \begin{cases} \mu_\epsilon(d(x, \partial X)), & \text{if} \quad d(x, \partial X) < \epsilon; \\ \epsilon, & \text{if} \quad d(x, \partial X) \geq \epsilon. \end{cases}$$

Then $h_\epsilon(x) \colon \tilde{X} \to \mathbb{R}$ is a smooth concave function away from $\partial \tilde{X}$. Let us denote the graphs $\tilde{X}^{\pm}_\epsilon = \left\{(x, \pm h_\epsilon(x)) \in \tilde{X} \times \mathbb{R}\right\}$, equipped with the intrinsic metrics. It's clear that the subgraph of h_ϵ is a convex body in $\mathbb{R}^{n+1}$ by the concavity of h_ϵ, and has smooth boundary away from $\partial \tilde{X} \times \{0\}$. Because the boundary of a convex body in Euclidean space is an Alexandrov space with curvature ≥ 0, we have that the sectional curvature, with respect to the intrinsic metric of $\tilde{X}^{\pm}_\epsilon$, is non-negative on the interior of $\tilde{X}^{\pm}_\epsilon$.

Note that the doubling of $\tilde{X}^{+}_\epsilon$ is isometric to the union $\tilde{Y}_\epsilon = \tilde{X}^{+}_\epsilon \cup \tilde{X}^{-}_\epsilon$ in $\mathbb{R}^{n+1}$. By conditions (1), (2) and (3) for μ_ϵ, we have that $\tilde{Y}_\epsilon$ is a smooth manifold. Moreover, $\tilde{Y}_\epsilon$ is non-negatively curved because it is smooth and non-negatively curved on the interior of $\tilde{X}^{+}_\epsilon$ and $\tilde{X}^{-}_\epsilon$. Note that $h_\epsilon(x) \to 0$ as $\epsilon \to 0$. Thus $\tilde{X}^{\pm}_\epsilon \to \tilde{X}$ as $\epsilon \to 0$, and $\tilde{Y}_\epsilon$ Gromov-Hausdorff converges to $\tilde{Y}$.

□

References

[1] S. Alexander, V. Kapovitch, A. Petrunin, *Alexandrov geometry*, a draft avaliable at *www.math.psu.edu/petrunin*.

[2] Y. Burago, M. Gromov, G. Perel'man, *A.D. Alexandrov spaces with curvature bounded below*, Uspekhi Mat. Nauk, **47:2** (1992), 3-51; translation in Russian Math. Surveys, **47:2** (1992), 1-58.

[3] J. Cheeger, T. Colding, *On the structure of spaces with Ricci curvature bounded below I*, J. Differential Geom., **45** (1997), 406–480.

[4] J. Cheeger, W. Jiang, A. Naber, *Rectifiability of singular sets in noncollapsed spaces with Ricci curvature bounded below*, arXiv: 1805.07988.

[5] J. Cheeger, A. Naber, *Lower bounds on Ricci curvature and quantitative behavior of singular sets*, Invent. Math., **191** (2013) no. 2, 321 – 339.

[6] J. Cheeger, A. Naber, *Regularity of Einstein manifolds and the codimension 4 conjecture*, Ann. of Math. (2) **182** (2015), no. 3, 1093 – 1165.

[7] R. Greene, H. Wu, C^∞ *approximations of convex, subharmonic, and plurisubharmonic functions*, Ann. Sci. Éc. Norm. Supér. (12) **1** (1979), 47 – 84.

[8] V. Kapovitch, A. Lytchak, A. Petrunin *Metric-measure boundary and geodesic flow on Alexandrov spaces*, arXiv: 1705.04767.

[9] A. Lytchak, K. Nagano, *Geodesically complete spaces with an upper curvature bound*, Geom. Funct. Anal., **29** (2019), 295–342.

[10] Y. Otsu, T. Shioya, *The Riemannian structure of Alexandrov spaces*, J. Differential Geom., **39** (1994), 629–658.

N. Li, Department of Mathematics, The City University of New York - NYC College of Technology, 300 Jay St., Brooklyn, NY 11201

E-mail address: NLi@citytech.cuny.edu

A. Naber, Department Of Mathematics, Northwestern University, 2033 Sheridan Rd., Evanston, IL 60208-2370

E-mail address: anaber@math.northwestern.edu