Geometry & Topology 22 (2018) 1961–2026

Quantitative bi-Lipschitz embeddings of bounded-curvature manifolds and orbifolds

SYLVESTER ERIKSSON-BIQUE

We construct bi-Lipschitz embeddings into Euclidean space for bounded-diameter subsets of manifolds and orbifolds of bounded curvature. The distortion and dimension of such embeddings is bounded by diameter, curvature and dimension alone. We also construct global bi-Lipschitz embeddings for spaces of the form $\mathbb{R}^n/\Gamma$, where Γ is a discrete group acting properly discontinuously and by isometries on $\mathbb{R}^n$. This generalizes results of Naor and Khot. Our approach is based on analyzing the structure of a bounded-curvature manifold at various scales by specializing methods from collapsing theory to a certain class of model spaces. In the process, we develop tools to prove collapsing theory results using algebraic techniques.

30L05, 51F99, 53C21; 20H15, 53B20

1 Introduction

1.1 Statement of results

It is an old problem to describe, in some insightful manner, the metric spaces which admit a bi-Lipschitz embedding into Euclidean space; see Heinonen [27].

Definition 1.1 Let $L > 0$. A map $f\colon X \to Y$ between metric spaces is called an (L–)*bi-Lipschitz* function (also embedding) if

$$\frac{1}{L} d(a,b) \le d(f(a), f(b)) \le L\, d(a,b) \tag{1.2}$$

for all $a, b \in X$.[1] The smallest constant L satisfying the previous equation is referred to as *the distortion* of f. By a bi-Lipschitz function of (or with) distortion L, we mean a function whose distortion can be bounded by L.

We are generally not concerned with the optimal distortions of a bi-Lipschitz map, but simply some upper bounds. When $Y = \mathbb{R}^N$, a necessary condition for embeddability

[1] Where ambiguity does not arise, we will use the letter d for the distance on any metric space that arises in the course of this paper.

Published: 11 April 2018 DOI: 10.2140/gt.2018.22.1961

is that (X,d) is metric doubling; see Definition 3.6. As a partial converse, Assouad proved that any doubling metric space admits an α–bi-Hölder embedding into Euclidean space for any $0<\alpha<1$; see Assouad [3] and also Naor and Neiman [45]. But due to counterexamples by Pansu [48] (as observed by Semmes [54]) and Laakso [35], additional assumptions are necessary for bi-Lipschitz embeddability.

Most of the spaces considered in this paper admit some bi-Lipschitz embedding, but our interest is in proving uniform, or quantitative, embedding theorems where the distortion is bounded in terms of the parameters defining the space. While a number of authors have addressed the question of quantitative bi-Lipschitz embeddability, natural classes of spaces remain for which such embeddings haven't been constructed.

Inspired by a question originally posed by Pete Storm, we focus on embedding certain classes of manifolds and orbifolds. In our discussion, we will always view Riemannian manifolds (and orbifolds) as metric spaces by equipping them with the Riemannian distance function. We emphasize that the problem of finding *quantitative* bi-Lipschitz embeddings of Riemannian manifolds into Euclidean spaces has little to do with classical results such as the Nash isometric embedding theorem for Riemannian manifolds. The notion of isometric embedding in that context only requires the map to preserve distances infinitesimally, whereas a bi-Lipschitz map controls the distances at any scale up to a factor. A Nash embedding for a *compact* manifold/subset will, however, result in a bi-Lipschitz embedding with a possibly very large distortion. To give a bound for the distortion of our embedding, we will need to resort to very different techniques. We also note that Nash embeddings don't directly give bi-Lipschitz embeddings for noncompact spaces, while some of our results apply for such examples as well.

Example To illustrate the difference between isometric and *uniform* bi-Lipschitz embeddings, consider the family of surfaces

$$S_N = \{x^2 + y^2 + N^2 z^2 = 1\}.$$

Denote by g_N the restricted metric on this two-dimensional surface. This defines boundaries of ellipsoids in $\mathbb{R}^3$ which are isometrically embedded in $\mathbb{R}^3$, but the natural embeddings are not uniformly bi-Lipschitz. The distance along the surface between the points $x_\pm = (0,0,\pm 1/N)$ is $d(x_+,x_-) \sim 2$, but in $\mathbb{R}^3$ the straight-line distance is $|x_+ - x_-| = 2/N$. Thus, the natural embedding of S_N does not have a uniformly bounded bi-Lipschitz constant. We might ask if there exist different embeddings where the bi-Lipschitz distortion does not deteriorate. Indeed, consider the maps $L_N\colon S_N \to \mathbb{R}^3$ which are defined by sending

$$(x,y,z) \mapsto \bigl(x,y,z+\bigl(1-\sqrt{x^2+y^2}\bigr)H(z)\bigr).$$

Here, set $H(z)=1$ for $z \geq 0$ and otherwise $H(z)=0$. Such a map can be shown to be C–bi-Lipschitz, where the constant C can be chosen independent of N. In other words, the map L_N controls distances at all scales up to a factor independent of N.

Manifolds comprise a large class of metric spaces, and we need to place some assumptions in order to ensure uniform embeddability. In order to ensure doubling, it is natural to assume a diameter bound as well as a lower Ricci-curvature bound. Our results require the somewhat stronger assumption of bounding the sectional curvature in absolute value. Thus, we are led to the following theorem.

Theorem 1.3 *Let $D > 0$. Every bounded subset A with $\operatorname{diam}(A) \leq D$ in an n–dimensional complete Riemannian manifold M^n with sectional curvature $|K| \leq 1$ admits a bi-Lipschitz embedding $f\colon A \to \mathbb{R}^N$ with distortion less than $C(D,n)$ and dimension of the image N at most $N(D,n)$.*

We emphasize that there are no assumptions on the injectivity radius or lower volume bound. If such an assumption were placed, the result would follow directly from Cheeger–Gromov compactness or a straightforward doubling argument; see Lemma 3.7 and Greene and Wu [22]. While somewhat unexpected, we are also able to prove a version of the previous theorem for orbifolds.

Theorem 1.4 *Let $D > 0$. Every bounded subset A with $\operatorname{diam}(A) \leq D$ in an n–dimensional complete Riemannian orbifold O^n with sectional curvature $|K| \leq 1$ admits a bi-Lipschitz embedding $f\colon A \to \mathbb{R}^N$ with distortion less than $C(D,n)$ and dimension of the image N at most $N(D,n)$.*

In particular, we have the following nontrivial result.

Theorem 1.5 *Every connected complete flat and elliptic orbifold O of dimension n admits a bi-Lipschitz mapping $f\colon O \to \mathbb{R}^N$ with distortion less than $D(n)$ and dimension of the image N at most $N(n)$. The constants $D(n)$ and $N(n)$ depend only on the dimension.*

Remark We can give explicit bounds for the case of flat and elliptic orbifolds, but not for arbitrary bounded-curvature orbifolds. This is because the proof of Fukaya's fibration theorem in Cheeger, Fukaya and Gromov [12] uses compactness in a few steps and thus does not give an explicit bound for the parameters $\epsilon(n)$, $\rho(n,\eta,\epsilon)$. Some related work in Ghanaat, Min-Oo and Ruh [21], when combined with techniques from [12], is likely to give an explicit bound, but to our knowledge, this has not been published. In any case, the bounds extracted by known means would grow extremely rapidly; see the bounds in Buser and Karcher [10] and Ruh [52]. For flat orbifolds and manifolds, we attain a bound on the distortion $D(n)$ which is of the order $O(e^{Cn^4 \ln(n)})$.

For mostly technical reasons, we prove an embedding result for certain classes of "model" spaces which occur in our proofs. These involve the notion of a quasiflat space which is a slight generalization of a complete flat manifold to the context of quotients of nilpotent Lie groups. They correspond to model spaces for collapsing phenomena. The precise definition is in Definition 3.21.

Theorem 1.6 *For sufficiently small $\epsilon(n)$, every $\epsilon(n)$–quasiflat n–dimensional orbifold M admits a bi-Lipschitz embedding into $\mathbb{R}^N$ with distortion and dimension N depending only on n. Further, every locally flat Riemannian orbivector bundle over such a base with its natural metric admits such an embedding.*

It may be of interest that some of these spaces are not compact and do not have nonnegative sectional curvature. Yet, they admit global bi-Lipschitz embeddings. This theorem could also be stated for orbifolds, and the proof below will directly apply to them. For terminology related to orbifolds, we recommend consulting, for example, Ding [15], Kleiner and Lott [34], Ratcliffe [49] and Thurston [58]. Some of this terminology is covered in the appendix.

We remark that the results above can trivially be generalized to Gromov–Hausdorff limits of bounded-curvature orbifolds, which themselves may not be orbifolds. The Gromov–Hausdorff limits of bounded-curvature Riemannian manifolds have been intrinsically described using weak Alexandrov-type curvature bounds by Nikolaev [47], and as such, our results also lead to embeddings of such spaces.

1.2 Outline of method

Bi-Lipschitz embedding problems can be divided to two subproblems: embedding locally at a fixed scale and embedding globally. Similar schemes for constructing embeddings have been employed elsewhere, such as in Lang and Plaut [36], Naor, Peres, Schramm and Sheffield [46] and Seo [55]. Constructing global embeddings may be difficult, but by assuming a diameter bound and a lower curvature bound, we can reduce it to embedding a certain fine ϵ–net.

We are thus reduced to a local embedding problem at a definite scale. In particular, since we don't assume any lower volume bound, we need to apply collapsing theory in order to establish a description of manifolds at small but fixed scales. For general manifolds, we first use a result from Fukaya [19]; see also [12]. Recall that if g is a metric tensor of a Riemannian manifold, and T is any tensor, then its norm with respect to the metric is denoted by $\|T\|_g$. There are different tensor norms, but all of them are quantitatively equivalent. To fix a convention, $\|T\|_g$ will refer to the operator norm.

Theorem 1.7 (Cheeger, Fukaya and Gromov [12], Fukaya [20] and Rong [51]) *Let (M, g) be a complete Riemannian manifold of dimension n and sectional curvature $|K| \leq 1$, and let $\epsilon > 0$ be an arbitrary constant. For every $\eta > 0$, there exists a universal $\rho(n, \eta, \epsilon) > 0$ such that for any point $p \in M$ there exists a metric g' on the ball $B_p(10\rho(n, \eta, \epsilon))$ and a complete Riemannian manifold M' with the following properties:*

- $\|g - g'\|_g < \eta$.
- *$(B_p(5\rho(n, \eta, \epsilon)), g')$ is isometric[2] to a subset of M'.*
- *M' is either an ϵ–quasiflat manifold or a locally flat vector bundle over an ϵ–quasiflat manifold.*

The constant ϵ will always be very small; we will set it equal to some dimension-dependent $\epsilon(n)$, which is fixed below in Section 3. This choice guarantees that the model spaces are sufficiently nice. In other words, their collapsing can be studied algebraically. Throughout the paper, the constant $\epsilon(n)$ will have the same value, which depends on n. The constant η will be fixed as $\frac{1}{4}$, and it is used to guarantee that $(B_p(\rho(n, \eta, \epsilon)), g')$ is bi-Lipschitz to $(B_p(\rho(n, \eta, \epsilon)), g)$. Since ϵ and η are chosen fixed, dependent on the dimension, the scale $\delta = \rho(n, \eta, \epsilon)$ at which we have a nice model space depends only on the dimension.

This theorem isn't stated as such in the references. Our terminology is different and is introduced in Definitions 3.21 and 3.30. In [20], a topological version is stated, and in [12], the main results concern a fibration structure.

The claims of our theorem are contained in a similar form in [12, Appendix 1], and the proof of our version can be derived from it. The main difference is that in [12], the role

[2] By isometry, we mean Riemannian isometry in the sense that the metric tensors are transformed to each other via the pushforward of the differential. What we really use is that the ball $(B_p(\rho(n, \eta, \epsilon)), g')$ will have a distance-preserving mapping to a subset of M'.

of M' is played by the normal bundle of a singular orbit (of the N–structure), $\nu(\mathcal{O}_q)$, with a natural metric. Here we have merely explicated that such a normal bundle can be described as a vector bundle over an $\epsilon(n)$–quasiflat manifold. In the appendix, we outline a proof of the analogous statement for orbifolds. We also use the results in [51] that guarantee us a dimension-independent bound on the sectional curvatures of the model spaces.

Fukaya's fibration theorem permits us to reduce the embedding problem to embedding certain vector bundles ("model spaces"). These spaces have the "twisted product" structure of a vector bundle over a quasiflat base. Further, these spaces also have the advantage of possessing an algebraic description. Thus, their metric geometry can be studied using the additional structure.

We next give a rough description of our approach to embedding the vector bundles M that arise. We first use an approximation argument, which consists of applying Lemmas 4.11, 4.9 and 3.10 to reduce the problem to embedding locally flat Riemannian vector bundles over flat manifolds. Next, the flat manifold M is decomposed into pieces, and embeddings for each piece are patched together using Lipschitz extension theorems. These decompositions are similar in spirit but were developed independently of arguments by Seo [55] and Romney [50].

To be more explicit, assume that the diameter of the base of the vector bundle is 1. Let us assume $L = V(S)$ is a locally flat Riemannian vector bundle over a flat base S. The space S is also the zero-section of the vector bundle L equipped with the restricted metric. For $i \geq 1$, define T_i to be the set of points at distance roughly 2^i from the zero-section. Additionally, define T_0 to be the set of points $x \in L$ with $d(x,S) \leq 4$. By using the radial function $r(x) = d(x,S)$, we are able to subdivide the embedding problem for the entire space to that of embedding each T_i individually.

To embed each T_i, we rescale the space by a definite amount to reveal a simpler vector bundle structure at a small, but definite scale. This is done using an algebraic collapsing theory argument. The difficult aspect of the proof is showing that these local descriptions exist at a scale comparable to the diameter of T_i, which is necessary for applying the doubling Lemma 3.7.[3] Further, the nature in which the scaling produces a simpler space is somewhat delicate. For $i = 0$, the resulting spaces are of the

[3] If T_i is scaled down to unit size, the curvature bound is scaled up by 2^{2i}, and thus applying collapsing theory is difficult. If one only wants to prove Theorem 1.3, then this scaling up could be avoided. In fact, the algebraic arguments below can be avoided by such an approach, but this would lead to other technicalities, as well as weaker results.

same dimension but resemble vector bundles over lower-dimensional spaces; ie their collapsing occurs more uniformly. For $i \geq 1$, the local descriptions split into a product of an interval and a lower-dimensional space. In either case, by induction, we can assume that the simpler spaces embed, and apply the doubling Lemma 3.7 to combine said embeddings to an embedding of T_i.

The algebraic collapsing theory argument involves quantitative versions of the local group arguments applied in collapsing theory [19]. Many of these arguments were inspired by proofs in [49] and by Thurston [59]. The core insight is that, up to a finite index, the fundamental groups are lattices. Lattices have "nice" bases, and the lengths of these basis elements give essentially canonically defined collapsing scales. These collapsing scales can be used to identify invariant submanifolds in the space.

We also construct embeddings for bounded-curvature orbifolds. For such spaces, we need a version of Fukaya's theorem for orbifolds. The proof of this result is essentially contained in [12, Appendix 1; 15], but we provide a rough outline in the appendix.

Theorem 1.8 *Let (O, g) be a complete Riemannian orbifold of dimension n and sectional curvature $|K| \leq 1$, and let $\epsilon > 0$ be an arbitrary constant. For every $\eta > 0$, there exists a universal $\rho(n, \eta, \epsilon) > 0$ such that for any point $p \in O$, there exists a metric g' on the ball $B_p(10\rho(n, \eta, \epsilon))$ and a complete Riemannian orbifold O' with the following properties:*

- $\|g - g'\|_g < \eta$.
- $(B_p(5\rho(n, \eta, \epsilon)), g')$ *is isometric to a subset of O'.*
- *O' is either an ϵ–quasiflat orbifold or a locally flat Riemannian orbivector bundle over an ϵ–quasiflat orbifold S.*

The same comments on the choice of these parameters apply as before. When we apply this theorem, we fix $\epsilon = \epsilon(n)$ and $\eta = \frac{1}{4}$, which allows us to use that the scale of the model spaces $\rho\big(n, \frac{1}{4}, \epsilon(n)\big)$ only depends on the dimension.

The methods to embed bounded-curvature orbifolds are essentially the same as for the manifold case, except for modifying terminology and adding some additional cases.

1.3 Previous work

Bi-Lipschitz embedding problems have been extensively studied for finite metric spaces; see Bourgain [6], Indyk and Matoušek [30] and Matoušek [41]. The problem has usually been to study the asymptotics of distortion for embedding certain finite metric spaces, either in general or restricting to certain classes of metrics (eg the earth-mover distance). The image is either a finite-dimensional l_p^n or an infinite-dimensional Banach space such as L^1. Motivation for such embeddings stems, for example, from approximation algorithms relating to approximate nearest-neighbor searches or querying distances. There has also been some study on bi-Lipschitz invariants, such as Markov type and Lipschitz extendability, which can be applied once a bi-Lipschitz map is constructed to a simple space; see Naor [44].

For nonfinite metric spaces, the research is somewhat more limited. We mention only some of the contexts within which embedding problems have been studied: subsets of $\mathbb{R}^n$ with intrinsic metrics by Tatiana Toro [60; 61], ultrametric spaces by Luosto [38] and Luukkainen and Movahedi-Lankarani [39], certain weighted Euclidean spaces by David and Semmes [53; 54], and an embedding result for certain geodesic doubling spaces with bicombings by Lang and Plaut [36]. Some of these works concern bi-Lipschitz parametrizations, which is a stronger problem. Concurrently with this paper, Matthew Romney [50] has considered related embedding problems on Grushin-type spaces and applied methods from [55] as well as the current paper. The different abstract question of existence of any bi-Lipschitz parametrization for nonsmooth manifolds has been discussed by Heinonen and Keith [29]. In contrast to these previous results, our goal is to prove embeddability theorems for natural classes of smooth manifolds, with uniformly bounded distortion and target dimension.

The results closest to this work are from Naor and Khot [33], who construct embeddings for flat tori and estimate the worst possible distortion. Improvements to their results were obtained by Haviv and Regev [26]. In comparison, our work provides embeddings for any compact or noncompact complete flat orbifold. Our bounds are arguably weaker due to the higher generality. Also, noteworthy is the paper of Bonk and Lang [5], where they prove a bi-Lipschitz result for Alexandrov surfaces with bounded integral curvature. Related questions on bi-Lipschitz embeddability have been discussed by Andoni, Naor and Neiman [2], where they consider an Alexandrov space target.

A number of spaces can be shown not to admit any bi-Lipschitz embedding into Euclidean space. In addition to classical examples such as expanders, nontrivial examples were found by Pansu [48] and Laakso [35]. These examples are related to a large class of examples that are covered by Lipschitz differentiation theory, which was initially developed by Cheeger [11].

1.4 Open problems

A natural question is to further study the dependence of our results on the curvature bounds assumed. At first, one might ask whether the upper curvature bound is necessary. We conjecture that it may be dropped. In fact, in light of our results on orbifolds, it seems reasonable to presume that Alexandrov spaces admit such an embedding. For the definition of an Alexandrov space, we refer to Burago, Burago and Ivanov [7] and Burago, Gromov and Perelman [8]. The conjecture is also interesting within the context of Riemannian manifolds.

Conjecture 1.9 *Let $D > 0$. Every bounded subset A with $\operatorname{diam}(A) \leq D$ in an n–dimensional complete Alexandrov space X^n with curvature $K \geq -1$ admits a bi-Lipschitz embedding $f\colon A \to \mathbb{R}^N$ with distortion less than $C(D,n)$ and dimension of the image N at most $N(D,n)$.*

If we also assume that the space X^n is volume noncollapsed, the theorem follows from an argument by Alexander, Kapovitch and Petrunin [1]. One of the main obstacles in proving embedding theorems for these spaces is the lack of theorems in the Lipschitz category for Alexandrov spaces. Most notably, proving a version of Perelman's stability theorem would help in constructing embeddings; see Kapovitch [31]. Perelman claimed a proof without publishing it, and thus it is appropriately referred to as a conjecture. See Definition 3.9 for the definition of the Gromov–Hausdorff distance d_{GH}.

Conjecture 1.10 *Let X^n be a fixed n–dimensional compact Alexandrov space with curvature $K \geq -1$. Then there exists an $\epsilon_0 > 0$ (depending possibly on X^n) such that for any other Alexandrov space Y^n with*

$$d_{GH}(X^n, Y^n) < \epsilon_0,$$

we have a bi-Lipschitz map $f\colon X^n \to Y^n$ with distortion at most L (which may depend on X^n).

Both of these conjectures seem hard. But mostly for lack of counterexamples, we also suggest that the lower sectional-curvature bound could be weakened to a lower Ricci-curvature bound. Weakening the curvature assumption to a simple Ricci-curvature bound results in great difficulty in controlling collapsing phenomena. However, even without collapsing, the conjecture is interesting and remains open. Thus, we state the following conjecture.

Conjecture 1.11 *Let A be a subset, with $\operatorname{diam}(A) \leq D$, in an n–dimensional complete Riemannian manifold (M^n, g) with curvature $\operatorname{Ric}(g) \geq -(n-1)g$. Assume $\operatorname{Vol}(B_1(p)) > v$ for some $p \in A$. Then there exists a bi-Lipschitz embedding $f\colon A \to \mathbb{R}^N$ with distortion less than $C(D,n,v)$ and dimension of the image N at most $N(D,n,v)$.*

Finally, we remark on the problem of optimal bounds for our embeddings in the theorem for flat and elliptic orbifolds (Theorem 1.5). As remarked, we obtain a bound of $D(n)$ of the order $O(e^{Cn^4 \ln(n)})$. The main source of distortion is the repeated and inefficient use of doubling arguments at various scales, which result in multiplicative increases in distortion. In a related paper, Regev and Haviv [26] improve the superexponential upper bound from [33] and obtain $O(n\sqrt{\log(n)})$ distortion for n–dimensional flat tori. Thus, it seems reasonable to suspect that the true growth rate of $D(n)$ is polynomial. As pointed out to us by Assaf Naor, this problem is also related to [2] because certain finite approximations to Wasserstein spaces arise as quotients of permutation groups.

1.5 Outline

In Section 2, we give explicit embeddings for three types of bounded-curvature spaces. These examples illustrate the methods used to prove the embedding results. We will not explain all the details, as some of them are presented well in other references and will become more apparent in the course of the proof of the main theorem. In Section 3, we collect some general tools and lemmas that will be used frequently in the proofs of the main results. Some of these are very similar to [36] and [46]. Finally, in Section 4, we give full proofs of the main embedding theorems. This section proceeds by increasing generalities. First, flat manifolds are embedded, followed by flat orbifolds and quasiflat orbifolds. Ultimately, the results are applied to Riemannian manifolds and orbifolds. The appendix collects a few of the most technical results on quotients of nilpotent Lie groups and collapsed orbifolds.

Acknowledgements The author thanks his adviser Bruce Kleiner for suggesting the problem and for numerous discussions on the topic. Discussions and comments from Jeff Cheeger, Zahra Sinaei, Or Hershkovits, Matthew Romney and Tatiana Toro have also been tremendously useful. We are grateful for Anthony Cardillo for drawing Figure 1. We also thank the referee for many detailed comments that helped us improve the paper tremendously, as well as the reviewing committee for the patience in going through different versions of the paper. This research was supported by a NSF grant DGE 1342536.

2 Embedding some key examples

In this section, we give nearly explicit embeddings for three types of bounded-curvature spaces. These examples illustrate the methods used to prove the embedding results.

2.1 Lens spaces

Take two distinct coprime numbers $p, q \in \mathbb{N}$. Consider the lens space $L(p,q)$, which is defined as the quotient of $S^3 = \{(z_1, z_2) \in \mathbb{C}^2 : |z_1|^2 + |z_2|^2 = 1\}$ by an action of $\mathbb{Z}_p$. This action is defined by $t(z_1, z_2) = (e^{2\pi t/pi} z_1, e^{2\pi tq/pi} z_2)$ for $t \in \mathbb{Z}_p$. Denote the equivalence class of an element (z_1, z_2) by $[(z_1, z_2)]$. The action is free and isometric, wherefore $L(p,q)$ inherits a constant-curvature metric. For an insightful discussion of these spaces, see [59]. We will construct a bi-Lipschitz embedding $f\colon L(p,q) \to \mathbb{R}^N$ for some N, and distortion, independent of p and q.

Collapsing theory states that the space can be covered by charts in which it resembles a normal bundle over some simpler space. For the space $L(p,q)$, such charts can be realized explicitly by the sets $U_1 = \{[(z_1, z_2)] \in L(p,q) : |z_2| < \sqrt{3}|z_1|\}$ and $U_2 = \{[(z_1, z_2)] \in L(p,q) : |z_1| < \sqrt{3}|z_2|\}$. These two sets cover $L(p,q)$. Next, we will describe their geometry.

The sets U_j, for $j = 1, 2$, contain the "polar" circles S_j defined by $|z_j| = 1$. Further, U_j can be expressed as normal bundles over S_j with a curved metric. On the lift of U_1 to S^3, we can locally introduce coordinates by $(\alpha, \theta, \phi) \mapsto (\cos(\alpha)e^{i\theta}, \sin(\alpha)e^{i\phi})$ for $\alpha \in [0, \pi/3)$ and $(\theta, \phi) \in S^1 \times S^1$. The induced metric becomes

$$\begin{aligned} g &= |\sin(\alpha)|^2|d\alpha|^2 + |\cos(\alpha)|^2|d\theta|^2 + |\cos(\alpha)|^2|d\alpha|^2 + |\sin(\alpha)|^2|d\phi|^2 \\ &= |\cos(\alpha)|^2|d\theta|^2 + |d\alpha|^2 + |\sin(\alpha)|^2|d\phi|^2. \end{aligned}$$

We can also define another metric on the lift of U_1 by

$$g_f = |d\theta|^2 + |d\alpha|^2 + \alpha^2|d\phi|^2.$$

Note that the norms induced by g and g_f differ by a factor at most 4 since $\sin(\alpha)^2 \geq \frac{1}{4}\alpha^2$ and $\cos(\alpha)^2 \geq \frac{1}{4}$ for $\alpha \in [0, \pi/3)$. The group $\mathbb{Z}_p$ acts on the lift of U_1 by isometries with respect to either of these metrics, and thus the latter descends to a metric g_f' on U_1, which is up to a factor 4 equal to the pushforward of the original metric g. In fact, (U_1, g_f') is isometric to a subset of the holonomy bundle, which is given by the $\mathbb{R}^2$–vector bundle over $(2\pi/p)S^1$ (the circle of radius $2\pi/p$), with holonomy generated by a $(2\pi q/p)$–rotation. Since the metric tensors differ only up to a factor and both (U_1, g) and (U_1, g'_f) are geodesically convex, we can see that the identity map $\iota_1\colon (U_1, g) \to (U_1, g'_f)$ is bi-Lipschitz with factor 4. Similar analysis can be performed for U_2, where the bundle has holonomy generated by a rotation of angle $2\pi s/p$, where $sq \equiv 1 \pmod p$. One is thus led to consider the problem of embedding flat vector bundles with nontrivial holonomy. These spaces will be discussed in the next subsection, where we indicate a construction for the desired bi-Lipschitz maps $F_j\colon (U_j, g'_f) \to \mathbb{R}^{n_j}$.

Finally, given the embeddings F_j, we define the map $f = (F_1, F_2, d(\cdot, U_1))$, where F_1 and F_2 are extended using the McShane lemma. To observe that this map is bi-Lipschitz, we refer to similar arguments in the proof of Lemma 3.7, or the proof of [36, Theorem 3.2], which we present below. We remark that there are many ways by which to patch up bi-Lipschitz embeddings from subsets into a bi-Lipschitz embedding of a whole. Our choice here is somewhat arbitrary. In fact, later in Lemma 3.7 a slightly different choice is employed.

For convenience of the reader, we restate and prove Theorem 3.2 from [36]. This result is not necessary for the proofs below as we will use slightly different arguments. However, understanding this argument is helpful.

Lemma 2.1 [36, Theorem 3.2] *Let X be a metric space, $A, B \subset X$ two sets such that $X \subset A \cup B$, and $f\colon A \to \mathbb{R}^n$ and $g\colon B \to \mathbb{R}^m$ two maps. If f and g are L–bi-Lipschitz and are extended arbitrarily to be L–Lipschitz on X, then $F(x) = (f(x), g(x), d(x, A))$ defines a CL–bi-Lipschitz map $F\colon X \to \mathbb{R}^{n+m+1}$ for some C, which depends only on L.*

Proof Clearly, F is $3L$–Lipschitz since its three components are L–Lipschitz. Thus, we only need to prove the lower bound in (1.2). Let $x, y \in X$ be arbitrary. If we have $x, y \in A$ or $x, y \in B$, then

$$|F(x) - F(y)| \geq |f(x) - f(y)| \geq \frac{1}{L} d(x, y)$$

or

$$|F(x) - F(y)| \geq |g(x) - g(y)| \geq \frac{1}{L} d(x, y),$$

respectively. Thus, assume that both x and y don't lie in the same subset A or B. By symmetry, assume $x \in A$ and $y \notin A$. There are two cases, $10L^2 d(y, A) \leq d(x, y)$, or $d(x, y) < 10L^2 d(y, A)$. In the latter case,

$$|F(x) - F(y)| \geq d(y, A) \geq \frac{1}{10L^2} d(x, y).$$

Consider next the first case. Define $y' \in A$ to be such that $d(y, y') \leq 2d(y, A)$. Then

$$|f(y) - f(y')| \leq L\,d(y, y') \leq 2L\,d(y, A) \leq \frac{1}{5L} d(x, y).$$

Hence

$$\begin{aligned} |F(x) - F(y)| &\geq |f(x) - f(y)| \\ &\geq |f(x) - f(y')| - |f(y') - f(y)| \\ &\geq \frac{1}{L} d(x, y) - \frac{1}{5L} d(x, y) \geq \frac{1}{2L} d(x, y). \end{aligned}$$

Combining the two lower bounds, and the upper bound, we get that F is $10L^2$–bi-Lipschitz. □

2.2 Flat vector bundles

In the previous example, we reduced the problem of embedding a lens space to that of embedding flat vector bundles. We are thus motivated to consider the embedding problem for them. Consider the space E^3_θ, which is a flat $\mathbb{R}^2$–bundle over S^1 with holonomy θ. One has $E^3_\theta = \mathbb{R} \times \mathbb{C}/\mathbb{Z}$, where the $\mathbb{Z}$–action is defined as $t(x, z) = (x + 2\pi t, e^{it\theta} z)$ for $t \in \mathbb{Z}$ and $z \in \mathbb{C}$. Denote the orbit of (x, z) in the quotient by $[x, z]$. Let r be the distance function to the zero-section, which is given by $r([x, z]) = |z|$. To embed this space, we use a decomposition argument. Let $T_0 = \{[x, z] : |z| \leq 4\}$ and $T_j = \{[x, z] : 2^{j-1} < |z| < 2^{j+1}\}$, for $j \geq 1$ an integer. Assume first that each T_j can be embedded with uniform bounds on distortion and dimension by mappings $f_j \colon T_j \to R^n$. Then we can collect the maps f_j with disjoint domains by defining four functions F_s for $s = 0, \dots, 3$ as follows. Consider the sets $D_s = \bigcup_{k=0}^{\infty} T_{s+4k}$, and define $F_s(x) = f_{s+4k}(x)$ for $x \in T_{s+4k}$. The resulting functions will be Lipschitz on their respective domains D_s. By an application of McShane extension Theorem 3.1, we can extend them to E^3_θ. The embedding we consider is $F = (r, F_0, F_1, F_2, F_3)$.

To see that $F(x)$ is bi-Lipschitz, we need to establish the lower bound in (1.2). Take $x, y \in E^3_\theta$ arbitrary. First, if $d(x, y) \leq 4|r(x) - r(y)|$, then the lower bound is trivial. Thus, assume $d(x, y) \geq 4|r(x) - r(y)|$. We will identify a j such that $x, y \in T_j$. This is obvious if $r(x) = r(y)$. We assume by symmetry that $r(x) > r(y)$ and $r(x) > 0$. Then the lower bound is obtained by using the lower-Lipschitz bounds for f_j on T_j. We have the estimate $4|r(x) - r(y)| \leq d(x, y) \leq r(x) + r(y) + 1$. Thus, $r(y) \geq r(x) - |r(x) - r(y)| \geq \frac{3}{4} r(x) - \frac{1}{4} r(y) - \frac{1}{4}$, and $r(y) \geq \frac{3}{5} r(x) - \frac{1}{5}$. In particular, if $x \in T_0$, then $r(y) \le r(x) \le 4$, and thus $y \in T_0$. But if $2^j \le r(x) \le 2^{j+1}$ for some $j \ge 2$, then $r(y) \ge 1/2r(x) \ge 2^{j-1}$, and thus $x, y \in T_j$.

Next, we discuss the construction of the embeddings f_j. On T_0 the injectivity radius is bounded from below. Thus, we can find f_0 by patching up a finite number of bi-Lipschitz charts. More specifically, we use doubling and the Lemma 3.7. For T_j when $j \ge 1$, we will first modify the metric. Use coordinates $(x, r, \theta) \mapsto (x, re^{i\theta}) \in \mathbb{R}^3$. The metric on T_j (or its lift in $\mathbb{R}^3$) can be expressed as

$$g = |dx|^2 + |dr|^2 + r^2|d\theta|^2.$$

We change the metric to $g_f = |dx|^2 + |dr|^2 + 2^{2j}|d\theta|^2 = g_1$. As metric spaces, the space (T_j, g) is 10–bi-Lipschitz to (T_j, g_f). The new metric space (T_j, g_f) is isometric to $(2^{j-1}, 2^{j+1}) \times T_\theta^j$, where T_θ^j is a torus defined by $\mathbb{R}^2/\mathbb{Z}^2$, where the $\mathbb{Z}^2$–action is given by $(n, m)(x, y) = (x + 2\pi n, y + n\theta + 2\pi m)$. The product manifold $(2^{j-1}, 2^{j+1}) \times T_\theta^j$ can be embedded by embedding each factor separately. Thus, the problem is reduced to embedding a flat torus. This can be done by choosing a short basis and is explained in detail in [33].

We comment briefly on the problem of embedding general flat vector bundles. Consider a flat $\mathbb{R}^n$–bundle over S^1. The argument remains unchanged for T_0, but for T_j when $j \ge 1$ one can no longer modify the metric to be a flat product metric. Instead, one needs a further decomposition argument of the $\mathbb{R}^d$ factor similar to that used for lens spaces. If the base is more complicated than S^1, we are led to an induction argument where these decompositions are used to reduce the embedding problem for spaces that are simpler in some well-defined sense.

Finally, we highlight the main ideas used in the previous embedding constructions for lens spaces and holonomy-bundles. First, the space is decomposed into sets ("charts") on which we have in some sense a simpler geometry. The embeddings for these charts can be patched together to give an embedding of the space. In the case for lens spaces, the simpler geometry was that of a vector bundle. For vector bundles, the simpler geometry was that of a product manifold. The simpler geometries may need to be decomposed several times, but we can bound the number of iterated decompositions required by the dimension of the space. Ultimately, the problem is reduced to embedding something very simple such as a noncollapsed space with a lower bound on the injectivity radius. The main technical issues arise from the fact that high-dimensional spaces may require several decompositions and that for abstract spaces it is complicated to identify good charts for which a simpler structure exists.

2.3 Quotients of the Heisenberg group and nilpotent groups

The Heisenberg group may be described as the simply connected Lie group of upper-triangular 3×3–matrices with diagonal entries equal to one:

$$\mathbb{H} = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \,\middle|\, a, b, c \in \mathbb{R} \right\}. \tag{2.2}$$

Define a lattice

$$\Gamma = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \,\middle|\, a, b, c \in \mathbb{Z} \right\} \subset \mathbb{H}$$

which acts on $\mathbb{H}$ by left multiplication. We can define the quotient space $\mathbb{H}/\Gamma$. Next metrize this base as follows. Take a basis for the Lie algebra of left-invariant vector fields given by

$$X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad Z = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Consider the space $\mathbb{H}_\epsilon = (\mathbb{H}/\Gamma, g_\epsilon)$ with the metric defined by setting $X \perp Y \perp Z$, $|X| = |Y| = 1$ and $|Z| = \epsilon$. Initially, this metric is defined on $\mathbb{H}$, but since Γ acts by isometries the metric descends onto the quotient. A direct computation shows that the sectional curvatures of this space lie in

$$\left[-\tfrac{3}{4}\epsilon^2, \tfrac{1}{4}\epsilon^2\right].$$

As $\epsilon \to 0$, the spaces $\mathbb{H}_\epsilon$ Gromov–Hausdorff converge to a torus T^2. The required quasi-isometry is given by the S^1–fibration map $\pi\colon \mathbb{H}_\epsilon \to T^2$ defined as

$$\pi\colon \left[\begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \right] \mapsto ([a], [b]) \in \frac{1}{2\pi} S^1 \times \frac{1}{2\pi} S^1 = T^2,$$

where $[a]$ is the fractional part of $a \in \mathbb{R}$. By $(1/2\pi)S^1$, we mean the circle of unit circumference. This map is easily seen to be well defined on $\mathbb{H}_\epsilon$. We wish to find a bi-Lipschitz embedding for $\mathbb{H}_\epsilon$ with distortion independent of $\epsilon > 0$, as long as its sufficiently small.

An embedding for $\mathbb{H}_\epsilon$ can be visualized using Lemma 3.7. Consider a small ball $B_p(\delta) \subset T^2$, with some $\delta > 0$ independent of ϵ. Since T^2 is noncollapsed, there exists a $\delta > 0$ such that $B_p(\delta)$ is isometric to a ball in the plane. Then since this ball is contractible, we can conclude that $\pi^{-1}(B_p(\delta))$ is diffeomorphic to $B_p(\delta) \times S^1$. Because the curvature is almost flat, we can show that this diffeomorphism can be chosen to be a bi-Lipschitz map with small distortion. This claim can be proven by a similar argument to Lemma A.5 in the appendix. On the other hand, the space $B_p(\delta) \times S^1$ is easy to embed since it splits as a product of two simple spaces. The result can then be deduced by covering T^2, and thus $\mathbb{H}_\epsilon$, by a controlled number of these sets, and patching up the embeddings using Lemma 3.7 or a repeated application of Lemma 2.1.

3 Frequently used results

Below, there are a number of constants determined from specific theorems. Since we wish to prove embedding theorems with quantitative bounds on the distortions, we need to be careful about the dependence on these parameters. In order to keep track of these different constants and functions throughout the chapter, we fix their meaning. The special constants are as follows: $C'(n)$ is from Theorem 3.20 and Definition 3.21, $c(n)$ is the maximum of the constants necessary for Lemmas 3.32, 3.20 and 3.18, $\epsilon(n)$ is fixed via Lemmas 3.27, 4.9 and 4.11, $\delta(n)$ is from Lemma A.5, and $C(n)$ is from Lemma A.4. Also, we will choose $c(n)$ increasing and $\epsilon(n)$ decreasing in n. Their values will be considered fixed throughout the paper, although the precise value is left implicit. In general, we will use the convention that a constant $M(a_1, a_2, \dots, a_n)$ is a quantity depending on $a_1, \dots, a_n$ only.

3.1 Embedding lemmas

Consider any metric space X. Throughout this paper we will denote open balls by $B_x(r) = B(x,r) = \{y \in X \mid d(y,x) < r\}$. For a metric space X, we call $N \subset X$ a *ϵ–net* for X if for every $x \in X$, there is an $n \in N$ such that $d(x,n) \leq \epsilon$, and if $x, y \in N$ then $d(x,y) > \epsilon$. Such nets can always be constructed using Zorn's lemma. For a subset $A \subset X$ and $r > 0$, we denote by

$$N_r(A) = \{x \in X \mid d(x,A) < r\}$$

the r–tubular neighborhood of a subset A. Recall $d(x,A) = \inf_{y \in A} d(x,y)$. The metric spaces in this paper will be complete connected manifolds M and orbifolds O equipped with a Riemannian distance function d.

Our Lipschitz mappings are initially defined locally and for a global embedding we need some type of extension theorem.

Theorem 3.1 (McShane–Whitney, [28, Theorem 2.3]) *Let X be a metric space with $A \subset X$. Then any L–Lipschitz function $f\colon A \to \mathbb{R}^n$ has a $\sqrt{n}L$–Lipschitz extension $\widetilde{f}\colon X \to \mathbb{R}^n$ such that $\widetilde{f}|_A = f$.*

Remark McShane–Whitney is a good extension result to use because of its generality. However, for bounded-curvature manifolds, better constants could be attained by the use of a generalized Kirszbraun's theorem [37]. For most of the paper, Kirszbraun could be used instead and slightly better constants would ensue.

In addition to extension, we will use decomposition arguments in two ways. On the one hand, we have spaces that admit certain splittings, such as cones and products.

Lemma 3.2 *Let (X, d_X), (Y, d_Y) be metric spaces and take $X \times Y$ with the product metric $d((x_1, y_1), (x_2, y_2)) = \sqrt{d_X(x_1, x_2)^2 + d_Y(y_1, y_2)^2}$. If X and Y admit bi-Lipschitz embeddings $f\colon X \to \mathbb{R}^N$ and $g\colon Y \to \mathbb{R}^M$ with distortion L, then $X \times Y$ admits a bi-Lipschitz embedding $(f, g)\colon X \times Y \to \mathbb{R}^{N+M}$ with distortion at most $\sqrt{2}L$.*

The proof of this result is trivial. The product space could also be equipped with different bi-Lipschitz equivalent metrics. In other cases, the space admits a conical splitting. We define a cone over a metric space.

Definition 3.3 Let X be a compact metric space with $\operatorname{diam}(X) \le \pi$. Then we define the cone over X as the space $C(X) = \{(t, x) \in \mathbb{R} \times X\}/\sim$, where $(0, x) \sim (0, y)$ for any $x, y \in X$, equipped with the metric

$$d([t, x], [s, y]) = \sqrt{t^2 + s^2 - 2ts\cos(d(x, y))}.$$

The metric cones which arise in our work will always be quotients of $\mathbb{R}^n$ by a group with a fixed point. Thus, we state the following lemma whose proof is immediate. Note that all the groups of the following lemma are forced to be finite.

Lemma 3.4 *Let Γ be a group acting properly discontinuously and by isometries on $\mathbb{R}^n$ which fixes $0 \in \mathbb{R}^n$. Then the group Γ acts properly discontinuously and by isometries on $S^{n-1} \subset \mathbb{R}^n$, and $\mathbb{R}^n/\Gamma$ is isometric to $C(S^{n-1}/\Gamma)$.*

Lemma 3.5 *Let X be a compact metric space with* $\operatorname{diam}(X) \le \pi$, *and let* $Y = C(X)$. *If X admits a bi-Lipschitz embedding to $\mathbb{R}^n$ with distortion L, then Y admits a bi-Lipschitz embedding to $\mathbb{R}^{n+1}$ with distortion $20L$.*

Proof Let r be the radial function on Y, where points on Y are given by equivalence classes $[(r,x)]$ for $r \ge 0$ and $x \in X$. We will simplify notation by suppressing the brackets. Assume that $f\colon X \to \mathbb{R}^n$ is bi-Lipschitz with distortion L and $0 \in \operatorname{Im}(f)$. Define $g(r,x) = (Lr, rf(x))$ on Y and show that it is the desired embedding. If (s,x) and (t,y) are points in Y and $s \le t$, then

$$|g(s,x) - g(t,y)| \le L|s-t| + s|f(x) - f(y)| + |s-t||f(y)|.$$

Note that $|f(y)| \le L\pi$ (because $\operatorname{diam}(X) \le \pi$). Further,

$$d((s,x),(t,y)) = \sqrt{t^2 + s^2 - 2ts\cos(d(x,y))} \ge 2s\,d(x,y)/\pi.$$

Using $s|f(x) - f(y)| \le Ls\,d(x,y)$ we get

$$|g(s,x) - g(t,y)| \le \bigl(L\pi + L + \tfrac{1}{2}\pi L\bigr) d((s,x),(t,y)) \le 10L\,d((s,x),(t,y)).$$

Next, we derive the necessary lower bound from (1.2). Take any two points (s,x) and (t,y) in Y with $s \le t$. Either $|s-t| \ge (1/(20L^2)) d((s,x),(t,y))$ or not. If so,

$$|g(s,x) - g(t,y)| \ge L|s-t| \ge \frac{1}{20L} d((s,x),(t,y)).$$

If not, then $|s-t| \le (1/(20L^2)) d((s,x),(t,y))$, and simple estimates give

$$\begin{aligned}
|g(s,x) - g(t,y)| &\ge |g(t,x) - g(t,y)| - |s-t||f(x)| - |s-t|L \\
&\ge \frac{1}{L} d((t,x),(t,y)) - |s-t||f(x)| - |s-t|L \\
&\ge \frac{1}{L} d((s,x),(t,y)) - |s-t|/L - |s-t||f(x)| - |s-t|L \\
&\ge \frac{1}{L} d((s,x),(t,y)) - (L + L\pi + 1/L)|s-t| \\
&\ge \frac{1}{L} d((s,x),(t,y)) - \frac{L + L\pi + 1/L}{20L^2} d((s,x),(t,y)) \\
&\ge \frac{1}{2L} d((s,x),(t,y)).
\end{aligned}$$

These estimates complete the proof. □

Remark Similar constructions would also apply for a spherical suspension of a metric space (see [8]), but we do not need that result.

When a simple splitting structure doesn't exist at a given scale, it may exist at a smaller scale. To enable us to accommodate for this we need the following "doubling argument". To state it, we need to define doubling metric spaces. We emphasize that a metric space must be doubling in order to admit a bi-Lipschitz embedding into $\mathbb{R}^N$.

Definition 3.6 A doubling metric space with doubling constant D is a metric space X such that for any ball $B_x(r) \subset X$ there exist points $p_1, \dots, p_D$ such that

$$B_x(r) \subset \bigcup_{i=1}^{D} B_{p_i}\left(\tfrac{1}{2}r\right).$$

The following lemma could be proven by repeatedly applying Lemma 2.1, but we give a different and slightly more direct argument.

Lemma 3.7 (doubling lemma) *Assume that we are given $0 < r < R$, $0 < N$ and l and that X is a doubling metric space with doubling constant D. Fix a point $p \in X$. If for every point $q \in B_p(R)$ there is an l–bi-Lipschitz embedding $f_q\colon B_q(r) \to \mathbb{R}^N$, then there is an L–bi-Lipschitz embedding $G\colon B_p(R) \to \mathbb{R}^M$ for some $M > 0$. Further, we can bound the distortion of the bi-Lipschitz embedding G by $L \le L(N, l, D, R/r)$ and the target dimension by $M \le M(N, l, D, R/r)$.*

Proof Construct a $\frac{1}{8}r$–net $Q = \{q_i\}$ in $B_p(R)$ and let $f_i\colon B_{q_i}(r) \to \mathbb{R}^N$ be the bi-Lipschitz mappings assumed. The size of the net is $|Q| = H \le D^{\log_2(R/r)+4}$. Translate the mappings so that $f_i(q_i) = 0$, and scale so that the Lipschitz-constant is 1. This will ensure that for any $a, b \in B_{q_i}(r)$ we have $|f_i(a) - f_i(b)| \ge (1/l^2) d(a,b)$ and $|f_i(a)| \le r$. Further, let $F\colon x \to (d(x, q_i))_i \in \mathbb{R}^H$ be a distance embedding from the net. Group the q_i into K groups J_k such that if $q_i, q_j \in J_k$ then $d(q_i, q_j) > 4r$. The number of points in the net Q within distance at most $4r$ from any q_i is at most D^4. Thus, a standard graph coloring argument such as in [27] will furnish the partition J_k with K sets with $K \le D^4$.

For each J_k, construct a map $g_k\colon \bigcup_{q_i \in J_k} B_{q_i}(r) \to \mathbb{R}^N$ by setting it equal to f_i on $B_{q_i}(r)$. Because the distance between the balls is at least $2r$, the Lipschitz constant will not increase. To see this take arbitrary $a \in B_{q_i}(r)$ and $b \in B_{q_j}(r)$, where $i \ne j$ and $q_i, q_j \in J_k$. Then $d(a,b) \ge 2r$, and

$$|g_k(a) - g_k(b)| \le |f_i(a) - f_i(q_i)| + |f_j(b) - f_j(q_j)| \le 2r \le d(a,b).$$

Extend g_k using Theorem 3.1 to give a $\sqrt{N}$–Lipschitz map from the entire ball $B_p(R)$, and denote it by the same name. For the embedding, combine these all into one vector

$$G = (g_1, \dots, g_K, F).$$

Clearly, this is $(2\sqrt{N}D^4 + 2D^{\log_2(R/r)+4})$–Lipschitz. It is also a map $G\colon B(p,R) \to \mathbb{R}^M$ with $M = KN + H$. For the lower Lipschitz bound in (1.2), take arbitrary $a, b \in B_p(R)$. First assume $d(a,b) < \frac{1}{2}r$. Then there is a q_i such that $d(a,q_i) < \frac{1}{8}r$ and $d(b,q_i) < r$. Both belong to the same ball $B_{q_i}(r)$ and for the k such that $q_i \in J_k$ we have $|G(a) - G(b)| \geq |g_k(a) - g_k(b)| = |f_i(a) - f_i(b)| \geq (1/l^2)d(a,b)$. Next assume $d(a,b) \geq \frac{1}{2}r$. Then there is a $q_i \in Q$ such that $d(a,q_i) < \frac{1}{8}r \leq \frac{1}{4}d(a,b)$ and thus $d(b,q_i) \geq d(a,b) - \frac{1}{4}d(a,b) \geq \frac{3}{4}d(a,b)$. Thus, $|d(a,q_i) - d(b,q_i)| > \frac{1}{2}d(a,b)$ and we get $|G(a) - G(b)| \geq |F(a) - F(b)| \geq \frac{1}{2}d(a,b)$. □

Definition 3.8 Let X, Y be metric spaces and let $s > 0$. A mapping $f\colon X \to Y$ is called an (s–)quasi-isometry if for all $x, y \in X$,

$$d_X(x,y) - s \leq d_Y(f(x), f(y)) \leq d_X(x,y) + s,$$

and $Y \subset N_s(f(X))$.

Remark It is easy to see that an s–quasi-isometry will be a 2–bi-Lipschitz map when restricted onto a $2s$–net.

Using quasi-isometries, one can define a distance between compact metric spaces and an associated notion of convergence. For more detailed discussion on convergence of metric and metric measure spaces, we refer the reader to [56]. See also the excellent references [7; 24].

Definition 3.9 Let X, Y be compact metric spaces. We denote by $d_{GH}(X,Y)$ the infimum of numbers $s > 0$ such that there are maps $f\colon X \to Y$ and $g\colon Y \to X$ which are s–quasi-isometries. Further, we say that a sequence of compact metric spaces X_i Gromov–Hausdorff converges to another metric space X if

$$\lim_{i \to \infty} d_{GH}(X_i, X) = 0.$$

Lemma 3.10 (Gromov–Hausdorff lemma) *Let X be a doubling metric space with doubling constant D, and let $\epsilon, \epsilon', l', l, m, N > 0$ be constants. Assume that Y is a metric space admitting a bi-Lipschitz map $h\colon Y \to \mathbb{R}^m$ with distortion l and* *that $d_{GH}(X,Y) \le \epsilon'$. If for every $p \in X$, there is an l'–bi-Lipschitz embedding $f_p\colon B_p(\epsilon) \to \mathbb{R}^N$, then there is an L–bi-Lipschitz map $G\colon X \to \mathbb{R}^K$ with distortion $L \le L(D,l,l',m,\epsilon'/\epsilon,N)$ and target dimension $K \le K(D,l,l',m,\epsilon'/\epsilon,N)$.*

Proof Fix $M = \epsilon'/\epsilon$. Let $g\colon X \to Y$ be an ϵ'–quasi-isometry. By Lemma 3.7, we can first construct for every $p \in X$ an embedding $f'_p\colon B_p(100l\sqrt{m}\epsilon') \to \mathbb{R}^{N'}$. By a scaling and dilation, we can assume that f'_p are 1–Lipschitz and $f'_p(p) = 0$. The distortion and the dimension N' will depend on l, l', D, N, m and M. Grouping f_p similarly to the proof in Lemma 3.7, we can define corresponding maps $g_k\colon X \to \mathbb{R}^N$ for $k = 1,\dots,H'$ for some H' depending on the parameters of the problem. Then we can construct a map $F\colon X \to \mathbb{R}^{K'}$ with $F = (g_1,\dots,g_H)$ with $K' = H'N$. Moreover, this construction guarantees that for every $x,y \in X$ with $d(x,y) \le 10l\epsilon'\sqrt{m}$, we have a $p \in X$ and an index i such that $x,y \in B_p(100l\sqrt{m}\epsilon')$ for some p and $g_i|_{B_p(100l\sqrt{m}\epsilon')} = f_p$. This can be used to give the lower bi-Lipschitz bound for (1.2) for any pair $x,y \in X$ such that $d(x,y) \le 10l\epsilon'\sqrt{m}$.

Next, take a δ–net N_δ for $\delta = 4\epsilon'$. The map $h \circ g|_{N_\delta}\colon N_\delta \to \mathbb{R}^m$ is a $2l$–bi-Lipschitz map because g is 2–bi-Lipschitz on a $4\epsilon'$–net. Extend this to a map $H\colon X \to \mathbb{R}^m$ which is $2l$–bi-Lipschitz on N_δ and $2l\sqrt{m}$–Lipschitz on X. The map H is used to give a lower Lipschitz bound for pairs $x,y \in X$ with $d(x,y) > 10l\sqrt{m}\epsilon'$. The proof can now be completed by similar estimates as in Lemma 3.7 and defining $G(x) = (H(x), F(x))$. This gives a map $G\colon X \to \mathbb{R}^{K'+m}$. □

The main source of quasi-isometries will be via quotients, so we state the following:

Lemma 3.11 *Assume (X,d_X) and (Y,d_Y) are metric spaces and $\delta > 0$. If $\pi\colon X \to Y$ is a quotient map and if we define $d_Y(a,b) = d_X(\pi^{-1}(a),\pi^{-1}(b))$ for all $a,b \in Y$ and $\mathrm{diam}(\pi^{-1}(a)) \le \delta$ for all $a \in X$, then $d_{GH}(X,Y) \le 2\delta$.*

3.2 Group actions and quotients

Assume that X is a proper metric space, ie that $B_p(r)$ is precompact for every $p \in X$ and $r > 0$. Further, assume that Γ is a discrete Lie group acting on X by isometries. For any group, its identity element is denoted by either e or 0, depending on if it is abelian. For any $p \in X$, its *isotropy* group is denoted by $\Gamma_p = \{\gamma \in \Gamma \mid \gamma p = p\}$. We say that the action is *properly discontinuous* if for every $p \in X$, there exists $r > 0$ such that $\{\gamma \in \Gamma \mid \gamma(B_p(r)) \cap B_p(r) \ne \varnothing\} = \Gamma_p$ and Γ_p is a finite group. We will study quotient spaces, which will be denoted by X/Γ, where the action of Γ is properly discontinuous. These spaces are also in some cases referred to as orbit spaces.

For $a \in X$, we will denote its orbit or equivalence class in X/Γ by $[a]$. Since the action is by isometries, we can define a quotient metric by

$$d([a],[b]) = \inf_{\gamma,\gamma' \in \Gamma} d(\gamma a, \gamma' b).$$

This distance makes X/Γ a metric space. For more terminology, see [49, Chapter 5]. The global group action may be complicated, but the local action can often be greatly simplified. The following lemma is used to make this precise. First recall, that for any group G and any set of elements $S \subset G$ the smallest subgroup containing them is denoted by $\langle S \rangle$. We say that S generates $\langle S \rangle$. For a group G, we say that $\gamma_1, \dots, \gamma_n$ are generators of the group, if $G = \langle \gamma_1, \dots, \gamma_n \rangle$.

Lemma 3.12 *Let X be a metric space and Γ a discrete Lie group of isometries acting properly discontinuously on X. Take a point $p \in X$. Let $\Gamma_p(r) = \langle g \in \Gamma \mid d(gp,p) \leq 8r \rangle$ be the subgroup generated by elements such that $d(p,gp) \leq 8r$. Then $B_{[p]}(r) \subset Y = X/\Gamma$ is isometric to $B_{[p]'}(r) \subset X/\Gamma_p(r)$ with the quotient metric.*

Proof We will denote the cosets, functions and points in $Y' = X/\Gamma_p(r)$ using primes and corresponding objects in Y without primes. For example, elements of $X/\Gamma_p(r)$ can be represented by $\Gamma_p(r)$–cosets $[x]'$, for $x \in X$, and elements of $Y = X/\Gamma$ will be represented as $[x]$. The action is still properly discontinuous and the distance function on the orbit space is given by $d'([x]',[y]') = \inf_{\gamma,\gamma' \in \Gamma_p(r)} d(\gamma x, \gamma' y)$. Since $\Gamma_p(r) \subset \Gamma$, we have a continuous map $F\colon (Y', d') \to (Y, d)$ which sends a $\Gamma_p(r)$–coset to the Γ–coset that contains it, ie $[x]' \mapsto [x]$. We will next show that F maps $B_{[p]'}(r) = \{y' \in Y' \mid d'([p]', y') < r\}$ isometrically onto $B_{[p]}(r)$, from which the conclusion follows.

Let $[a]', [b]' \in B_{[p]'}(r) \subset X/\Gamma_p(r)$, and as before, $F([a]') = [a]$ and $F([b]') = [b]$. Clearly, $d([a]',[b]') \geq d([a],[b])$ as the Γ–cosets are supersets of $\Gamma_p(r)$–cosets. We can choose the coset representatives a, b such that $a, b \in B_p(r) \subset X$. By the triangle inequality, $d([a],[b]) < 2r$, and for any small $\epsilon > 0$, there is a $\gamma_\epsilon \in \Gamma$ such that $d(a, \gamma_\epsilon b) < d([a],[b]) + \epsilon < 2r$. But then $\gamma_\epsilon b \in B_p(3r) \subset X$ and $d(\gamma_\epsilon b, b) \leq 4r$. Therefore, $d(\gamma_\epsilon p, p) \leq d(\gamma_\epsilon p, \gamma_\epsilon b) + d(\gamma_\epsilon b, b) + d(b,p) \leq r + r + 4r \leq 6r$. Thus $\gamma_\epsilon \in \Gamma_p(r)$, and therefore, $d([a]',[b]') \leq d(a, \gamma_\epsilon b) \leq d([a],[b]) + \epsilon$. The result follows since ϵ is arbitrary.

Finally, the map is onto, because if $d([p],[a]) < r$, then we can choose the representatives such that $d(p,a) < r$. In particular, $[a]' \in B_{[p]'}(r)$, and $F([a]') = [a]$ by definition. □

Motivated by the previous lemma, we define a notion of a local group.

Definition 3.13 Let Γ be a discrete Lie group acting properly discontinuously on X. We denote by $\Gamma_p(r) = \langle g \in \Gamma \mid d(gp, p) \le 8r \rangle$, and call it the local group at scale r (and location p).

In collapsing theory, it is often useful to find invariant points and submanifolds. However, often it is easier to first construct almost invariant points/submanifolds, and then to apply averaging to such a construction to give an invariant one. For this reason, we will recall Grove's and Karcher's center of mass technique. See [25; 32] for a more detailed discussion and proofs.

Lemma 3.14 (center of mass lemma) *Let M be a complete Riemannian manifold with a two-sided sectional-curvature bound $|K| \le \kappa$, and let $p \in M$ be a point. Further, let μ be a probability measure supported on a ball $B_p(r)$ of radius $r \le \min\{\pi/(2\sqrt{\kappa}), \frac{1}{2}\mathrm{inj}(M,p)\}$. There exists a unique point $C_\mu \in B_p(r)$ which minimizes the functional*

$$F(q) = \int d(q,x)^2 \, d\mu_x.$$

The minimizer C_μ is called the center of mass. The center of mass is invariant under isometric transformations of the measure. If $f\colon M \to M$ is an isometry, and $f_*(\mu)$ is the pushforward measure of μ, then

$$f(C_\mu) = C_{f_*(\mu)}.$$

As a corollary, one obtains the existence of fixed points for groups with small orbits. If G is any group acting on a manifold M, a point $p \in M$ is called a fixed point of a group action if $gp = g$ for all $g \in G$. A more traditional form of this lemma appears in [49, Lemma 5.9].

Corollary 3.15 (fixed-point lemma) *Let M satisfy the same assumptions as in Lemma 3.14. Let a compact Lie group G act by isometries on M. If*

$$\mathrm{diam}(G_p) < \min\left\{\frac{\pi}{2\sqrt{\kappa}}, \frac{1}{2}\mathrm{inj}(M,p)\right\},$$

then there exists a fixed point q of the G–action such that $d(q,p) \le \mathrm{diam}(G_p)$.

In fact, one can use the Haar measure μ on G, and consider the pushforward measure $f_*(\mu)$ by the map $f\colon G \to M$ which is given by $f(g) = gp$. The center of mass $C_{f_*(\mu)}$ will be invariant under the group action by G.

We will repeatedly use an argument that allows us to derive metric conclusions from a stratification of a finite-index subgroup. The idea of the statement is that a semidirect product with the stratified normal subgroup will itself admit a stratification. Such stratifications arise from Lemma 3.27 below. Assume Γ is a group. Then we call $|\cdot|\colon \Gamma \to \mathbb{R}$ a subadditive norm if the following hold:

- for all $g \in \Gamma$, $|g| \geq 0$ and $|e| = 0$;
- $|ab| \leq |a| + |b|$;
- $|a| = |a^{-1}|$.

For groups acting on spaces, one can choose a natural class of subadditive norms.

Definition 3.16 Let Γ act on a metric space X, and let $p \in X$ be fixed. We define the subadditive norm at p for $g \in \Gamma$ as

$$|g|_p = d(gp, p).$$

If $g \in \mathrm{Isom}(X)$ is an isometry, we let $|g|_p = d(gp, p)$. Further, if Γ acts on a Lie group N, then we will often choose $p = e$ and let $|\gamma| = |\gamma|_e$.

Lemma 3.17 (local group argument) *Fix an arbitrary scale parameter $l > 0$. Let Γ be a group with a subadditive norm $|\cdot|$, and assume it admits a short exact sequence*

$$0 \to \Lambda \to \Gamma \to H \to 0,$$

where H is a finite group of size $|H| < \infty$. Assume $\Lambda_0 < \Lambda$ is a normal subgroup in Γ generated by all $g \in \Lambda$ of length $|g| \leq l$, and further assume that every element $g \in \Lambda \setminus \Lambda_0$ has length $|g| \geq 10|H|l$. Then the subgroup $\Gamma_0 < \Gamma$ generated by all $g \in \Gamma$ with $|g| \leq l$ admits a short exact sequence

$$0 \to \Lambda_0 \to \Gamma_0 \to H_0 \to 0,$$

where $H_0 < H$. Further, any element $g \in \Gamma_0$ can be expressed as

$$g = \prod_{i=1}^{|H|+1} g_i \lambda,$$

where $|g_i| \leq l$ and $\lambda \in \Lambda_0$.

Proof Define the group by generation: $\Gamma_0 = \langle g \in \Gamma : |g| \leq l \rangle$. Define S' to be the collection of left cosets of Λ in Γ that contain a representative $s \in \Gamma$ with $|s| \leq l$.

Define S to consist of a representative s for each left coset in S' with $|s| \le l$. Clearly, $|S| = |S'| \le |H|$. Also, the cosets in S' coincide with Λ_0–cosets, because if $s = \lambda s'$ for some $\lambda \in \Lambda$ and $|s|, |s'| \le l$, then $|\lambda| \le |s| + |s'| \le 2l$, and thus $\lambda \in \Lambda_0$.

Construct W using all the "words" w of length at most $|H|+1$ formed by multiplying elements s and s^{-1} for $s \in S$, including the empty word which represents the identity element of Γ. None of these words represent elements of $\Lambda \setminus \Lambda_0$ since for any $w \in W$ we have $|w| \le (|H|+1)l$. Denote by $W' = \{[w]_{\Lambda_0} \mid w \in W\}$ the left cosets of Λ_0 represented by elements in W. Also, let $\bar{\Gamma} = \bigcup_{A \in W'} A \subset \Gamma$. We will show that $\bar{\Gamma} = \Gamma_0$. In other words, we show that W exhausts the left cosets of Λ_0 in Γ_0, and that these cosets have a product given by the product of their representatives. To do so, it is sufficient to show that any product of two elements in W is equivalent modulo Λ_0 to an element in W. This shows that $\bar{\Gamma}$ is closed under products. Since also $\bar{\Gamma} < \Gamma_0$ as it is generated by $S \cup \Lambda_0$, we have $\Gamma_0 = \bar{\Gamma}$.

Take any word w' with length longer than $|H|+1$. Consider the subwords $w'_1, \dots, w'_{|H|}$ formed by the first $1, \dots, |H|+1$ letters of w', respectively. Because there are only $|H|$ cosets, two of them will be in the same Λ–coset in Γ. Thus, assume w_i and w_j with $i < j$ belong to the same coset. Let $w' = w_j v_j$, where v_j is the product of the remaining letters. We have $w_i w_j^{-1} \in \Lambda$. But also, $|w_i w_j^{-1}| \le (2|H|+1)l$, so $w_i w_j^{-1} \in \Lambda_0$. Thus, we have for the Λ_0–cosets $[w_i]_{\Lambda_0} = [w_j]_{\Lambda_0}$, and thus $[w']_{\Lambda_0} = [w_j]_{\Lambda_0}[v_j]_{\Lambda_0} = [w_i]_{\Lambda_0}[v_j]_{\Lambda_0} = [w_i v_j]_{\Lambda_0}$. The word $w_i v_j$ is shorter than w'. This process can be continued as long as w' is longer than $|H|+1$. Thus, any word longer than $|H|+1$ is equivalent modulo elements in Λ_0 to an element in W. In particular, the cosets formed by products of cosets corresponding to elements in W can be represented by elements contained in W. This allows one to define a group structure for the cosets represented by W, and therefore a subgroup $H_0 < H$. It is immediate, that H_0 contains all the cosets generated by elements in S.

By the previous argument, Γ_0 is a finite extension of Λ_0 of index at most $|H_0|$. Also, Λ_0 is normal in Γ_0 since it is normal in Γ. The final statement about representing elements of g as products follows from the above, since g belongs to a coset represented by an element in W. □

3.3 Nilpotent Lie groups, lattices and collapsing theory

Nilpotent geometries occur as model spaces for collapsing phenomena. Thus, we start with an estimate relating to the curvature of a general simply connected nilpotent Lie group N equipped with a left-invariant metric g. For definitions of nilpotent

Lie groups as well as the computations involved here, see [14]. Associate to the Lie group N its nilpotent Lie algebra $\mathfrak{n}$. We construct a triangular basis for $\mathfrak{n}$ as follows. Let $\mathfrak{n}_0 = \mathfrak{n}$, $\mathfrak{n}_1 = [\mathfrak{n}, \mathfrak{n}]$, $\dots$, $\mathfrak{n}_{k+1} = [\mathfrak{n}, \mathfrak{n}_k]$. For k large enough, $\mathfrak{n}_k = \{0\}$. Let X_i be an orthonormal basis such that there are integers k_j and $X_i \in \mathfrak{n}_j$ for $i \geq k_j$ and $X_i \perp \mathfrak{n}_j$ for $i < k_j$. Let c_{ij}^k be the Maurer–Cartan structural constants with respect to this basis; ie

$$[X_i, X_j] = \sum_k c_{ij}^k X_k.$$

By construction of the basis, $c_{ij}^k \neq 0$ only if $k > j$. Computing the sectional curvatures $K(X_i, K_j)$ using [42], we get for $j \geq i$,

$$K(X_i, X_j) = \frac{1}{4} \sum_{k<j} (c_{ik}^j)^2 - \frac{3}{4} \sum_{k>j} (c_{ij}^k)^2.$$

As such, bounding the curvature is, up to a factor, the same as bounding the structural constants. Most of the time we will scale spaces so that sectional curvature satisfies $|K| \leq 2$. Therefore, we will assume throughout this chapter that $\|[\cdot,\cdot]\|_g \leq C'(n)$, where $C'(n)$ depends only on the dimension n. For comparison and later use, we will mention the following theorem.

Theorem 3.18 (Bieberbach [9; 49]) *There exists a dimension-dependent constant $c(n)$ such that the following holds. If M is a complete (not necessarily compact) connected flat Riemannian orbifold of dimension n, there exists a discrete group of isometries Γ acting on $\mathbb{R}^n$ such that $M = \mathbb{R}^n/\Gamma$. Also, there exists an affine k–dimensional subspace $O \subset \mathbb{R}^n$ which is invariant under Γ, and on which the action of Γ is properly discontinuous and cocompact. Further, there exists a subgroup $\Lambda < \Gamma$ with index $[\Gamma : \Lambda] \leq c(n)$ which acts on O by translations. Moreover, M is isometric to a locally flat Riemannian $(n-k)$–dimensional vector bundle over O* (see Definition 3.30 below).

Collapsing theory involves working with certain classes of spaces, which we define here. All notions in this chapter are defined for orbifolds. The versions for manifolds are obtained by a slight variation of terminology.

Definition 3.19 A compact Riemannian orbifold S is called $\epsilon-$almost flat if $|K| \leq 1$ and $\operatorname{diam}(S) \leq \epsilon$.

Throughout this chapter ϵ will be small enough; ie we will only consider $\epsilon(n)'$–almost flat n–dimensional manifolds for small enough $\epsilon(n)'$. The choice is dictated by Theorem 3.20. For such manifolds, a detailed structure theory was developed by Gromov and Ruh [23; 52; 10; 9]. Later, Ding observed [15] that with minor modifications the proof as presented in [52; 10] generalizes to orbifolds.[4] We state this structure theory here.

For definitions of orbifolds and some of the other terminology, we refer to the appendix and [58; 49; 34]. For any Lie group G, we denote by $\mathrm{Aff}(G)$ the Lie group of affine transformations of G, which preserve the flat left-invariant connection ∇_{can}. The left-invariant connection ∇_{can} is defined as the flat connection making left-invariant fields parallel. By $\mathrm{Aut}(G)$, we denote the Lie group of its automorphisms. Throughout, N will denote a simply connected nilpotent Lie group.

Theorem 3.20 (Gromov–Ruh almost flat theorem[5] [15; 52; 10]) *For every n, there are constants $\epsilon'(n)$, $c(n)$ such that for any $\epsilon'(n)$–almost flat orbifold S with metric g, there exists a simply connected nilpotent Lie group N, a left-invariant metric g_N on N, and a subgroup $\Gamma < \mathrm{Aff}(N)$ with the following properties:*

- *Γ acts on (N, g_N) isometrically, cocompactly and properly discontinuously. In particular, N/Γ is a Riemannian orbifold.*
- *We have the bound $|K| \leq 2$ for the sectional curvature of N and the bound $\|[\cdot,\cdot]\|_{g_N} \leq C'(n)$ for the norm of the bracket.*[6]
- *There is an orbifold diffeomorphism $f\colon S \to N/\Gamma$ such that $\|f_*g - g_N\|_g < \frac{1}{2}$. (The bound here could be replaced by any $\delta < 1$, but we fix a choice to reduce the number of parameters in the theorem.)*
- *There exists a normal subgroup $\Lambda < \Gamma$ such that $[\Gamma : \Lambda] \leq c(n)$ and Λ acts on N by left-translations. In other words, by slight abuse of notation, $\Lambda < N$.*

We remark that we consistently use N/Γ to indicate a quotient under a group action even though here Γ is acting on N on the left and the translational subgroup Λ acts on N by *left* multiplication. Again, $\|\cdot\|_g$ denotes the operator norm of the various tensors. In the work of [15] and [52], no explicit bounds are derived for the Lie bracket and curvature, but these can be obtained by using the tools from [51] and curvature bounds at the beginning of this section.

[4] The metric closeness part is not explicit in [15], but a direct consequence of applying [52].

[5] We also recommend consulting the preprint version of the paper by Ding which contains a more complete proof [16].

[6] For the precise sectional-curvature bound, it is necessary to use [51, Theorem 2.1 and Proposition 2.5].

In the course of the proof of the main theorem, we will need certain special classes of orbifolds that we will call quasiflat. These will be certain quotients of a nilpotent Lie group N by a group $\Gamma < \mathrm{Aff}(N)$. Each element $\gamma \in \mathrm{Aff}(N)$ can be represented by $\gamma = (a, A)$, where $a \in N$ and $A \in \mathrm{Aut}(N)$. The action of (a, A) on N is given by

$$(a, A)m = aAm,$$

where $m \in N$. The component A is called *the holonomy* of γ. If A is reduced to the identity, the operator can be expressed as (a, I) and is called translational. Translational elements $\gamma \in \Gamma$ can be canonically identified by an element in N, and we will frequently abuse notation and say $\gamma \in N$. There is a holonomy homomorphism

$$h\colon \Gamma \to \mathrm{Aff}(N)$$

given by $h(\gamma) = A$.

The Lie group N will always be equipped with a left-invariant metric g. As in the previous section, if $p \in N$ is given and $g \in \mathrm{Isom}(N)$, then we let

$$|g|_p = d(gp, p).$$

Definition 3.21 A Riemannian orbifold S is called a (ϵ–)*quasiflat* orbifold if S is isometric to a quotient N/Γ of a simply connected nilpotent Lie group (N, g) equipped with a left-invariant metric g by a discrete cocompact group of isometries Γ satisfying the following:

- For the left-invariant metric g, the sectional curvature is bounded by $|K| \leq 2$ and $\|[\cdot,\cdot]\|_g \leq C'(n)$.
- The group Γ acts on the space N isometrically, properly discontinuously and cocompactly.
- The quotient space N/Γ is a Riemannian orbifold with

$$\mathrm{diam}(S) = \mathrm{diam}(N/\Gamma) < \epsilon.$$

The space N is referred to as the (orbifold) universal cover, and Γ is called the orbifold fundamental group.

We also define a flat orbifold as follows. Another definition, used in Theorem 3.18, requires that the space be complete and that the sectional curvatures vanish. This is, however, equivalent by the same theorem to the following.

Definition 3.22 If Γ is a group acting cocompactly, properly discontinuously and isometrically on $\mathbb{R}^n$, then we call $S = \mathbb{R}^n/\Gamma$ a flat orbifold.

The manifold versions of the previous definitions only differ by assuming that the actions are free. Recall that a group action of Γ on a manifold M is *free* if the isotropy group is trivial for any $p \in M$, ie $\Gamma_p = \{e\}$, where e is the identity element of Γ.

By Theorem 3.20, any $\epsilon(n)'$–almost flat manifold is bi-Lipschitz to an $\epsilon(n)''$–quasiflat manifold with a slightly different $\epsilon(n)''$. Also, by Theorem 3.18 every compact flat manifold is also a flat orbifold in the previous sense. In the following, we will choose $\epsilon(n)$ so small that any $\epsilon(n)$–quasiflat manifold is $\epsilon(n)'$–almost flat, and so that (3.28) holds. Further, we will assume, by possibly making $\epsilon(n)$ smaller, that $\epsilon(n) < \epsilon(m)$ for $n > m$.

We first prove a result concerning the algebraic structure of Γ. This can be thought of as a generalization of the Bieberbach theorem; see [9].

Lemma 3.23 *Let $c(n)$ be the constant in the Gromov–Ruh almost flat theorem* (Theorem 3.20). *Let $S = N/\Gamma$ be an $\epsilon(n)$–quasiflat n–dimensional orbifold. Then Γ has a finite-index subgroup Λ with $[\Gamma : \Lambda] \le c(n)$, and $\Lambda = N \cap \Gamma$; ie Λ acts on N by translations.*

Proof Consider the left-invariant metric g of N. Since, up to a dilation, an $\epsilon(n)$–quasiflat space is also $\epsilon(n)'$–almost flat, by Theorem 3.20 we can construct a new group Γ', and a nilpotent Lie group N' such that S is orbifold-diffeomorphic to N'/Γ'. The proof also gives that Γ' acts by affine transformations on N' and has a finite-index normal subgroup $\Lambda' \lhd \Gamma'$, with $[\Gamma' : \Lambda'] \le c(n)$, and Λ' acts on N' by translations.

Both N' and N constitute orbifold universal covers in the sense of [58, Chapter 13]. Thus, there is a diffeomorphism $f\colon N' \to N$ conjugating the action of Γ' to that of Γ. Denote the induced homomorphism by $f_*\colon \Gamma' \to \Gamma$.

By [4, Theorem 2], there is a finite index $\Lambda < \Gamma$ such that Λ acts by translations on N'. In particular, Λ is a cocompact subgroup of N. We still need to show that it is possible to choose Λ with a universal bound on the index $c(n)$, since [4] does not bound the index. We wish to translate the index bound for Λ' to one for Λ. Applying [4, Proposition 2] to the nilpotent group $f_*^{-1}(\Lambda')$, we get $f_*^{-1}(\Lambda') < \Lambda$. Thus, $[\Gamma : \Lambda] \le c(n)$. □

In the previous section, we defined generators for groups. For lattices in nilpotent Lie groups, there are special classes of generators that we introduce here.

Definition 3.24 Let N be a simply connected nilpotent Lie group. A cocompact discrete subgroup $\Lambda < N$ is called a lattice.

Definition 3.25 Let N be an n–dimensional simply connected nilpotent Lie group and $\Lambda < N$ a lattice. Then a set $\{\gamma_1, \dots, \gamma_n\} \subset \Lambda$ is called a triangular basis for Λ if the following properties hold:

- $\gamma_1, \dots, \gamma_n$ generate Λ.
- $\gamma_i = e^{X_i}$, and the X_i form a vector space basis for the Lie algebra $\mathfrak{n} \cong \mathbb{R}^n$ of N.
- For $i < j$, we have $[\gamma_i, \gamma_j] \in \langle \gamma_1, \dots, \gamma_{i-1} \rangle$.
- For $i < j$, we have $[X_i, X_j] \in \operatorname{span}(X_1, \dots, X_{i-1})$.

Here, e^X is the Lie group exponential map. In the case of $N = \mathbb{R}^n$, the previous definition reduces to the standard definition of a triangular basis for a Lattice. An immediate consequence of the previous definition is that the group N can be given coordinates by $(t_1, \dots, t_n) \mapsto e^{t_1 X_1} \cdots e^{t_n X_n}$. This will be used in the appendix.

Lemma 3.26 *Let $S = N/\Gamma$ be an $\epsilon(n)$–quasiflat manifold, and $\Lambda = \Gamma \cap N$. For any $p \in N$, we can generate Γ and Λ by elements γ with $|\gamma|_p \le 16c(n)\epsilon(n)$.*

Proof Consider the action of $\Lambda = \Gamma \cap N$. Then N/Λ is a Riemannian manifold. Further by Lemma 3.23, the index of Λ in Γ is at most $c(n)$, so there is an m–fold orbifold covering map $\pi\colon N/\Lambda \to N/\Gamma$ with $m \le c(n)$. Since $\operatorname{diam}(N/\Gamma) < \epsilon(n)$ and since N is connected, we have $\operatorname{diam}(N/\Lambda) < 2c(n)\epsilon(n)$.

Let $\Gamma' = \langle g \in \Gamma : |g|_p \le 16\epsilon(n)c(n) \rangle$ and $\Lambda' = \langle g \in \Lambda : |g|_p \le 16c(n)\epsilon(n) \rangle$. By Lemma 3.12, the space N/Γ' is isometric to N/Γ, and thus $\Gamma = \Gamma'$. Similarly, we can show $\Lambda = \Lambda'$. This completes the proof. □

The main conclusion of the following statement is the existence of a basis $\gamma_1, \dots, \gamma_n$ with several technical properties. The latter part of the statement concludes the existence of associated subgroups Λ_k, L_k and scalars l_k with certain geometric properties. These parameters l_k are the "collapsing scales", and part of the conclusion is that they are independent of the chosen base point in a specific sense.

Lemma 3.27 (stratification lemma) *Fix arbitrary integers $n \ge 1$ and $l > 4$, and let $L = 2l^n$. For every $\epsilon(n)$–quasiflat orbifold $S = N/\Gamma$, there exist the following objects with certain desired properties described below:*

(1) *a triangular basis* $\gamma_1, \dots, \gamma_n$ *for* Λ*;*

(2) *a natural number* s *("the number of collapsing scales");*

(3) *for every* $k = 1, \dots, s$*, subsets* $I_k, J_k \subset \{1, \dots, n\}$*;*

(4) *subgroups* $\Lambda_k = \langle \gamma_i \mid \gamma_i \in J_k \rangle \subset N$ *corresponding to* J_k *for* $k = 1, \dots, s$*;*

(5) *connected normal subgroups* $L_k \lhd N$ *corresponding to* Λ_k *for* $k = 1, \dots, s$*;*

(6) *scalars* $l_k = 2|\gamma_{i_k}|_e > 0$*, where* $i_k = \max I_k$ *("collapsing scales") for each* $k = 1, \dots, s$*.*

The basis γ_i *and the groups* $I_k \subset \{1, \dots, n\}$ *for* $k = 1, \dots, s$ *satisfy the following:*

- $|\gamma_i|_e < |\gamma_j|_e$ *for* $i < j$*;*
- $|\gamma_i|_e \leq L|\gamma_j|_e$ *for all* k *and* $\gamma_i, \gamma_j \in I_k$*;*
- $l|\gamma_i|_e \leq |\gamma_j|_e$ *for all* $\gamma_i \in I_s$*,* $\gamma_j \in I_t$ *and* $s < t$*.*

We can set $J_k = \bigcup_{i=1}^k I_k$*. The subgroups* Λ_k *are normal in* Γ*, and* $\Lambda_k < L_k$ *is a cocompact lattice in* L_k*. Further, the subgroups* L_k *are invariant with respect to the action of the image of the holonomy* $h\colon \Gamma \to \mathrm{Aut}(N)$*.*

Finally, the scalars l_k *satisfy for every* $p \in N$ *and every* $k = 1, \dots, s$*, that*

$$\Lambda_k = \langle \lambda \in \Lambda : |\lambda|_p < l_k \rangle$$

and that

$$|\lambda|_p > l \cdot l_k / 4$$

for every $\lambda \in \Lambda \setminus \Lambda_k$*. Also, for any nontrivial* $\lambda \in \Lambda$*, we have* $|\lambda|_p > l_1/(2L)$*.*

Remark The properties of l_k are needed to be able to apply Lemma 3.17.

Proof The basis is constructed similarly to [10]. Let γ_1 minimize $|\cdot|_e = |\cdot|$ in Λ. Set $G_0 = \{e\}$ to be the trivial group. Define the subgroup G_1 generated by γ_1. Proceed to choose $\gamma_2 \in \Lambda \setminus G_1$ as the shortest with respect to $|\cdot|_e$. Define $G_2 = \langle \gamma_1, \gamma_2 \rangle$. We proceed inductively to define $\gamma_1, \dots, \gamma_n$. The element γ_i is the shortest element with respect to $|\cdot|_e$ in $\Lambda \setminus G_{i-1}$ and $G_i = \langle \gamma_1, \dots, \gamma_i \rangle$. A priori, this process could last more than n steps, but by an argument below the number of γ_i agrees with the dimension n. Thus, we abuse notation slightly by using the same index.

Take any $a, b \in N$ such that $|a|_e, |b|_e \leq 16c(n)\epsilon(n)$. We have

$$|[a,b]|_e \leq 16C'(n)c(n)\epsilon(n)\min\{|a|_e, |b|_e\}. \tag{3.28}$$

This estimate follows either from curvature estimates such as in [10], or directly from the bounds on the structural constants on N. Choose $\epsilon(n) < 1/(10^2 c(n)C'(n))$, so that $\|[a,b]\|_e < \min\{|a|_e, |b|_e\}$.

Because of Lemma 3.26 we can choose generators α_i for the translational subgroup $\Lambda < \Gamma$ such that $|\alpha_i|_e < 16c(n)\epsilon(n)$. Moreover, by construction we get then that $|\gamma_i| < 16c(n)\epsilon(n)$. By estimate (3.28), we have $\|[\alpha_i, \gamma_i]\|_e < |\gamma_i|_e$, and thus $[\alpha_i, \gamma_i] \in G_{i-1}$ by construction. Since α_i generate Λ, we get for any $\lambda \in \Lambda$ that $[\lambda, \gamma_i] \in G_{i-1}$. In particular, for $i < j$ we have $[\gamma_i, \gamma_j] \in G_{i-1}$.

We define a map $f\colon (a_1, \dots, a_n) \to \gamma_1^{a_1} \cdots \gamma_n^{a_n}$. By induction, and the product relation below, we can show that the map is injective. Using the commutator relations from above and reordering terms, we can show that f is bijective onto Λ. Let $a = (a_1, \dots, a_n)$, $b = (b_1, \dots, b_n)$. We can define polynomials p_j by the defining relation

$$\gamma_1^{a_1} \cdots \gamma_n^{a_n} \times \gamma_1^{b_1} \cdots \gamma_n^{b_n} = \gamma_1^{a_1 + b_1 + p_1(a,b)} \cdots \gamma_n^{a_n + b_n + p_n(a,b)}.$$

The polynomials are defined by appropriately applying commutator relations. Also, p_i depends only on a_j, b_j for $j < i$. This defines a product on $\mathbb{Z}^n$, which can be extended to $\mathbb{R}^n$ by the previous relation. This defines a nilpotent Lie group N' with a cocompact lattice $\mathbb{Z}^n$. The lattice has a standard basis $e_i = f^{-1}(\gamma_i) \in \mathbb{Z}^n$, and $e_i = e^{X_i}$, where $X_i' = (0, \dots, 1, \dots, 0) \in T_0\mathbb{R}^n$ with the 1 in the i^{th} position. By a theorem of Malcev (see [40] and [12]), we can extend the homomorphism $f\colon \mathbb{Z}^n \to \Lambda$ to a bijective homomorphism $\overline{f}\colon N' \to N$. Since the map is bijective and since X_i' form a basis at $T_e N$, then $d\overline{f}(X_i') = X_i$ will form a basis for $T_e N$. Further $\gamma_i = e^{X_i}$, and $[X_i, X_j] = d\overline{f}([X_i', X_j'])$ it is direct to verify that $[X_i, X_j] \in \operatorname{span}(X_1, \dots, X_{i-1})$ for $j > i$. In particular, the basis is triangular. Moreover, the number of basis elements has to agree the dimension of the space.

Next, we define the sets I_j. Define $i_0 = 1$. Let i_1 be the smallest index such that $|\gamma_{i_1+1}| > l|\gamma_{i_1}|$, or if there is none $i_1 = n$. Continue inductively defining i_k to be the smallest index bigger than i_{k-1} such that $|\gamma_{i_k+1}| > l|\gamma_{i_k}|$, or if there is no such index $i_k = n$ and we stop. Define s such that $i_s = n$, and $I_k = \{\gamma_i \mid i_{k-1} < i \le i_k\}$. The properties in the statement are easy to verify. The k most collapsed directions are collected into $J_k = \bigcup_{i=1}^k I_i$, and we define $\Lambda_k = \langle \gamma_i \mid \gamma_i \in J_k \rangle$. Define also $\Lambda_0 = \{e\}$

By choosing $\epsilon(n) < 1/(10^2 c(n)C'(n)L)$ and applying (3.28), we can prove that, if $\gamma_i \in I_s$, $\gamma_j \in I_t$ and $s < t$, then $[\gamma_i, \gamma_j] \in \Lambda_{s-1}$. We remark that by virtue of the construction, if $\gamma_i \in \Lambda_k$, and $\gamma \in \Lambda$ such that $|\gamma| < l|\gamma_i|$, then also $\gamma \in \Lambda_k$.

Since γ_i form a basis, we can show that Λ_k is normal in Λ. Next, we want to show that it is also normal in Γ. Take an element $\sigma \in \Gamma$ with nontrivial holonomy, and a $\gamma_i \in \Lambda_k$ from the basis. Further, represent $\sigma = (t, A)$, where $A \in \operatorname{Aut}(N)$ is in the holonomy and $t \in N$. By left-multiplying with an element in Λ, we can assume $|t|_e = |\sigma|_e \leq 16c(n)\epsilon(n)$. This follows from the diameter bound for N/Λ in Lemma 3.26. Then $\sigma\gamma_i\sigma^{-1} = t(A\gamma_i)t^{-1} \in \Lambda$ by the normality of Λ. Since A is an isometry, $|A\gamma_i| = |\gamma_i|$. This combined with the commutator estimate (3.28) gives $d(t(A\gamma_i)t^{-1}, A\gamma_i) \leq \frac{1}{10}|\gamma_i|$, and thus $|t(A\gamma_i)t^{-1}| \leq \bigl(1 + \frac{1}{10}\bigr)|\gamma_i|$. By the iterative definition of Λ_k, we must have $t(A\gamma_i)t^{-1} \in \Lambda_k$. Note that we are assuming $l > 2$.

The connected Lie subgroups L_k can be generated by $X_1, \dots, X_{i_k}$, and normality in Λ follows since the basis is triangular. It is also easy to see that L_k is normal in N. Further, cocompactness of Λ_k follows by considering the map $(t_1, \dots, t_{i+k}) \mapsto e^{t_1 X_1} \cdots e^{t_{i_k} X_k}$. By the argument in the previous paragraph, for any element $\sigma = (t, A) \in \Gamma$ with $|\sigma|_e < 16\epsilon(n)c(n)$ and $\gamma \in \Lambda_k$, we have $tA\gamma t^{-1} \in \Lambda_k \in L_k$. Since L_k is normal in N, we have $A\gamma \in L_k$. From this, we can conclude that $AL_k = L_k$, ie that L_k are invariant under the holonomy action of Γ.

Finally, define $l_k = 2|\gamma_{i_k}|$. Fix $p \in N$. We will show that $\Lambda_k = \langle \lambda \in \Lambda \mid d(\lambda p, p) < l_k \rangle$ and $|\lambda|_p = d(\lambda p, p) > \frac{1}{4} l \cdot l_k$ for every $\lambda \in \Lambda \setminus \Lambda_k$. First of all, we can assume that $d(p, e) < 16c(n)\epsilon(n)$. This follows again from the diameter bound for N/Λ in Lemma 3.26. By assumption, for any $\gamma_i \in \Lambda_k$, we have $|\gamma_i|_e \leq |\gamma_{i_k}|_e < \frac{1}{2} l_k$. Thus by the commutator estimate (3.28),

$$d(\gamma_i p, p) = d(p^{-1}\gamma_i p, e) \leq d(p^{-1}\gamma_i p, \gamma_i) + d(\gamma_i, e) \leq \bigl(1 + \tfrac{1}{10}\bigr)|\gamma_{i_k}|_e < l_k.$$

Thus, $\gamma_i \in \langle \lambda \in \Lambda \mid d(\lambda p, p) < l_k \rangle$, and further by varying γ_i and generation we get $\Lambda_k < \langle \lambda \in \Lambda \mid d(\lambda p, p) < l_k \rangle$. Next if $\gamma \notin \Lambda_k$, then $|\gamma|_e = d(\gamma, e) > \frac{1}{2} l \cdot l_k$ by the construction of Λ_k. Simple estimates give

$$d(\gamma p, p) = d(p^{-1}\gamma p, e) \geq d(\gamma, e) - d(p^{-1}\gamma p, \gamma) \geq \bigl(1 - \tfrac{1}{10}\bigr)|\gamma|_e \geq \tfrac{1}{4} l \cdot l_k.$$

In particular, for every element such that $d(\lambda p, p) < l_k < \frac{1}{4} l \cdot l_k$ we must have $\lambda \in \Lambda_k$. Thus, $\Lambda_k = \langle \lambda \in \Lambda \mid d(\lambda p, p) < l_k \rangle$ and for any $\gamma \notin \Lambda_k$ we have $|\lambda|_p > \frac{1}{4} l \cdot l_k$. The final conclusion, that $d(\lambda p, p) > l_1/(2L)$ for all $\lambda \in \Lambda$, follows from $|\lambda|_e > l_1/L$ which follows from the construction of γ_i. □

It follows from the previous lemma that the scales l_k and subgroups Λ_k, L_k are essentially independent of the chosen base point. Further, they correspond to "local groups" at scales l_k.

Definition 3.29 Let $S = N/\Gamma$ be an $\epsilon(n)$–quasiflat orbifold. If $l = 400c(n)$ and $L = 2l^n$, then call a basis $\gamma_1, \dots, \gamma_n$ satisfying the previous lemma a short basis for Λ. The grouping $I_1, \dots, I_s$ is called canonical grouping and the number s is called the number of collapsing scales. If such a basis exists, we say that S has s collapsing scales. For $k = 1, \dots, s$, the sets J_k are called the k^{th} most collapsed scales.

Definition 3.30 Let S be a complete Riemannian orbifold and $L = V(S)$ an orbivector bundle over S. Then L is called a (k–dimensional) *locally flat (Riemannian) orbivector bundle* if the orbivector bundle $V(S)$ admits a bundle-metric and a metric connection such that it is locally flat (and such that the fibers have dimension k). In the case where V and S are complete Riemannian manifolds, we call V a locally flat Riemannian vector bundle over S.

A bundle metric is a metric on the fibers of S which is parallel with respect to the connection. For the notion of an orbivector bundle, one may consult [34] and the appendix. The notions of a bundle metric and metric connection are naturally extended from vector bundles to orbivector bundles by considering local charts. Note that we permit the case when $S = \{\text{pt}\}$ is a single point (ie a zero-dimensional orbifold), and $L = \mathbb{R}^k/\Gamma$, where $\Gamma < O(k)$ is a finite group.

Because the monodromy action[7] is by isometries, and the bundle is locally trivial, there is a canonical Riemannian metric and distance function on L (and not just on the fibers). To understand the geometry of L we are reduced to understanding the geometry of S and that of the monodromy action. These metrics are somewhat easier to visualize in the case where S is a compact flat manifold, and V is a locally flat vector bundle over such a base. In this case, L is given by a quotient of Euclidean space by an action of a free, discrete and isometric group action, which may fail to be cocompact.

In our case, all the orbifolds $L = V(S)$ will arise as group quotients. By assumption, S will be $\epsilon(n)$–quasiflat. Thus, S has an orbifold universal cover N, which is a simply connected nilpotent Lie group. Further, S is given by the quotient N/Γ, where Γ is the orbifold fundamental group, which acts discretely, cocompactly and by isometries on N. The orbivector bundle L has a pullback vector bundle L' on N. Since N is simply connected, and L was assumed locally flat, we have a flat pullback connection on L' making it trivial. In particular, we have, by choosing an orthonormal frame on

[7] Monodromy may also be defined for orbifold. A similar discussion is used to prove that flat orbifolds are good in [49].

some fiber, a natural trivialization $L' = N \times \mathbb{R}^k$, where k is the dimension of the fiber. The pullback metric on L' is simply a product metric, and the distance is given by the product structure. From this perspective, the distance on L is given by a quotient metric on L'/Γ, where Γ acts on the $\mathbb{R}^k$ factor via a monodromy action. Note, all these vector bundles are quotients and will be studied as such. However, they are not *compact*, because the action of Γ on the space $N \times \mathbb{R}^k$ is not cocompact.

The monodromy action is given by a homomorphism $\Gamma \to O(n)$, where Γ corresponds to the orbifold fundamental group. The rest of this section collects some necessary facts about these group representations. Note that in our case Γ is "virtually" nilpotent, and its action on $\mathbb{R}^n$ will turn out to be "virtually" abelian. By "virtually", we mean that there is a finite-index subgroup with the desired property, and that the index can be controlled in terms of the dimension n.

Definition 3.31 Let H be an abelian subgroup of $O(n)$. We call a decomposition $\mathbb{R}^n = \bigoplus V_i$ the *canonical decomposition*, and the subspaces $V_0, \dots, V_k$ are called *canonical invariant subspaces*, if they satisfy the following properties:

- Each subspace has positive dimension except possibly V_0.
- For every $h \in H$ and every j, we have $hV_j = V_j$.
- There is one V_j such that for all $h \in H$, we have $hv = v$ for all $v \in V_j$ (ie each $h \in H$ acts trivially on this subspace). We will always denote this subset by V_0, possibly allowing $V_0 = \{0\}$.
- On each other subspace V_j, exactly one of the following holds:
  (1) for every $h \in H$ and every $v \in V_j$, either $hv = v$ or $hv = -v$;
  (2) V_j is even-dimensional, and we can associate isometrically complex coordinates $(z_1, \dots, z_n)$ such that for any $h \in H$, the action is given by $h((z_1, \dots, z_n)) = (e^{\theta_j(h)i} z_1, \dots, e^{\theta_j(h)i} z_n)$ for some $\theta_j(h) > 0$.
- The subspaces are maximal with respect to these properties and pairwise orthogonal.

If $A \in O(n)$ is an isometry, the canonical decomposition associated to it is the one associated to the cyclic subgroup generated by A.

Every abelian isometric action has a unique canonical decomposition. For an individual matrix, this should be thought of as a decomposition corresponding to the factorization of its minimal polynomial. Using this terminology we prove a minor modification of the Jordan lemma.

Lemma 3.32 (modified Jordan theorem) *There is a constant $c''(n)$ depending on the dimension with the following property: Let Γ be a group with a nilpotent normal subgroup Λ with index $[\Gamma : \Lambda] < \infty$, and let $h\colon \Gamma \to O(n)$ be a homomorphism. Then $h(\Gamma)$ has a normal abelian subgroup of index at most $c''(n)$.*

Proof Let $L = \overline{h(\Gamma)}$ be the closure of the image of h. The subgroup L is Lie subgroup of $O(n)$. It is either finite or has nontrivial identity component T. In the first case, we may proceed as in the standard Jordan theorem [57]. In the latter case, since Γ has a finite-index nilpotent subgroup, $T \subset O(n)$ must be a compact nilpotent group. This is only possible if it is abelian, ie a torus. The torus T corresponds to a decomposition of $\mathbb{R}^n$ into canonical invariant subspaces V_j. Since T is a continuous group, except possibly for one V_0, every V_j for $j = 1, \dots, s$ will be even-dimensional, and the action will correspond to a Hopf-action $\theta(z_1, \dots, z_k) = (e^{i\theta} z_1, \dots, e^{i\theta} z_k)$. On V_0 the action is trivial.

Since T is normal in L, the decomposition to the V_j corresponds to a homomorphism

$$\phi = (\phi_1, \dots, \phi_k)\colon L \to \prod_{j=1}^{s} O(V_j).$$

Also, the action of L on each subspace conjugates the Hopf-action on each V_i to itself (which agrees with the action of T). The infinitesimal generator in $\mathfrak{o}(V_j)$ corresponding to the Hopf-action is J which maps $J(z_1, \dots, z_k) = (iz_1, \dots, iz_k)$, if $\dim(V_j) = 2k$. In particular, for any $g \in \phi(L)$ we have $gJg^{-1} = J$ or $gJg^{-1} = -J$ (because it preserves the action), which means that g either preserves the complex structure, or it composed with the conjugate map does so. In other words, g is either symplectic or a product of a symplectic matrix with a matrix B giving the conjugation on V_j. Thus, $\phi_j(L) \subset U(V_j) \cup \overline{U(V_j)}$, where $U(V_j)$ is the symplectic group of isometries and is isomorphic to the unitary group of k-by-k matrices, and $\overline{U(V_j)}$ is obtained by composing such a matrix with B.

Consider $\phi(L) \cap \prod_j \bigl(\mathrm{SU}(V_j) \cup \overline{\mathrm{SU}(V_j)}\bigr) = K$, where $\mathrm{SU}(V)$ and $\overline{\mathrm{SU}(N)}$ are those matrices in $U(V_j)$ with unit determinant (using the k-by-k complex coordinates), or their product with B. This group must be finite because every left coset of T contains a finite number of elements in $\prod_j \mathrm{SU}(V_j)$. Here, we used that the group T acts on V_j via a Hopf-action $\theta(z_1, \dots, z_n) = (e^{i\theta} z_1, \dots, e^{i\theta} z_n)$, which has determinant $e^{i\dim(V_j)\theta}$. Also, $e^{i\dim(V_j)\theta} = 1$ only when $\theta = 2\pi n/\dim(V_j)$. Since K is finite, conjugation by T will leave it fixed, and so any element of T commutes with any element in K. Further, $\phi(L) = KT$ by the previous discussion.

By the standard Jordan theorem, the group K has a finite-index abelian normal subgroup $H \triangleleft K$ of index at most $d(n) \le O((n+1)!)$. Define the abelian group $G = HT$, which is generated by H and T. We have

$$[\phi(L) : G] = [KT : HT] \le [K : H].$$

Since we have an effective bound for the index $[K : H] \le d(n)$, we get a bound for $[\phi(L) : G] \le d(n)$. It is also easy to see that G is normal in $\phi(L)$. □

Remark The optimal bound for the index in the classical Jordan theorem is $(n+1)!$, which is valid for $n \ge 71$ [13]. Thus, from the previous argument we get the optimal bound $[\Gamma : \Lambda] \le c''(n) \le (n+1)!$ for $n \ge 71$. Below, we will set $c(n)$ to be the larger of $c(n)''$ in Lemma 3.32 and $c(n)$ in Theorem 3.20.

4 Construction of embeddings

The proof will proceed by increasing generalities. First we discuss the case of flat manifolds, then the case of flat orbifolds, and finally the case of quasiflat orbifolds. At the end, we use Theorems 1.7 and 1.8 to obtain the proof for Theorem 1.3.

4.1 Flat manifolds

In this section, we construct embeddings for flat manifolds by proving the following:

Theorem 4.1 *Every connected complete flat manifold M of dimension n admits a bi-Lipschitz embedding $f\colon M \to \mathbb{R}^N$ with distortion less than $D(n)$ and dimension of the image N at most $N(n)$.*

We will prove this theorem by induction. The somewhat complicated induction argument is easier to understand through the examples in Section 2. While it is true that the proof could be directly done for orbifolds, the author believes it is easier to start with the manifold context and later highlight the small differences with the orbifold case.

The goal is to use embeddings at different scales and reveal ever simpler vector bundles at smaller scales. Simplicity is measured in the total dimensionality of the space or the dimensionality of the base space of a vector bundle. This is also a frequent idea of collapsing theory, and as such the proof can be seen as a refinement of collapsing theory to a family of explicit spaces. To avoid lengthy phrases, we use $V-(n,d)$ to abbreviate the induction statement, where $d \geq 0$ and $n > 0$. The case $V-(n,0)$ covers *compact* cases of the theorem, and $V-(n,d)$ for $d > 0$ covers the *noncompact* cases, which are vector bundles. Bieberbach's Theorem 3.18 can be used to show that any connected complete flat manifold is covered by either of these two cases. The notation $V-(n,0)$ is justified by imagining compact manifolds as 0–dimensional vector bundles over themselves.

Next, we define these statements more carefully (the case where $d = 0$ is treated in a special way).

- $\boldsymbol{V-(n,0)}$ There exist constants $L(n,0)$ and $N(n,0)$ such that the following holds: Assume M is an arbitrary n–dimensional connected compact flat manifold. Then there exists a bi-Lipschitz embedding $f\colon M \to \mathbb{R}^N$ with distortion $L \leq L(n,0)$ and target dimension $N \leq N(n,0)$ depending only on n.
- $\boldsymbol{V-(n,d)}$ There exist constants $L(n,d)$ and $N(n,d)$ such that the following holds: Assume V is an arbitrary d–dimensional connected locally flat Riemannian vector bundle over a compact flat manifold M of dimension n. Then there exists a bi-Lipschitz embedding $f\colon V \to \mathbb{R}^N$ with distortion $L \leq L(n,d)$ and target dimension $N \leq N(n,d)$ depending only on n, d.

In the proof below, we will perform a double induction to prove $V-(n,d)$, which is based on reducing the embedding problem for $V-(n,d)$ to embedding instances of $V-(n-k, k+d)$ for some $k > 0$ or of $V-(a,b)$ with $a+b < n+d$. For $V-(n,0)$, the reduction is similar to that of a lens space, and for $V-(n,d)$ the reduction is analogous to that for holonomy-bundles; see Section 2.

The main steps of these reductions are choosing a length scale $\delta > 0$, applying a stratification of a lattice from Lemma 3.27, followed by an application of the local group Lemma 3.17 to expose the structure of a local group at the scale δ, and then applying Lemma 3.12 to describe the metric structure at those scales. The metric structure at a small but definite scale is always that of a vector bundle. The base of such a vector bundle is an embedded manifold of smaller dimension. This embedded manifold is found by finding submanifolds $\mathbb{R}^k$ in the universal cover (that is, $\mathbb{R}^n$) which remain invariant under the action of a local group. Here, the averaging from Corollary 3.15 plays a crucial role. Once the embeddings are constructed at scale δ, using the induction hypothesis, the embedding of the space is obtained by using Lemma 3.7.

Proof of Theorem 4.1 We first give a description of the induction argument, and then move onto defining the relevant notation and proving the induction claims.

Summary of induction argument We need to embed all complete connected flat manifolds. By Bieberbach's Theorem 3.18, every complete connected flat manifold can be represented as a quotient by a discrete group of isometries. In particular, any such manifold is either isometric to $\mathbb{R}^n$, a quotient of $\mathbb{R}^n$ by a cocompact group action or a locally flat vector bundle over such a base. In the first case, the embedding is given by the identity mapping. Otherwise, either the space can be represented as $M = \mathbb{R}^n/\Gamma$, or $V = \mathbb{R}^n \times \mathbb{R}^d/\Gamma$, where Γ is some discrete Lie group which acts cocompactly, properly discontinuously and by isometries on $\mathbb{R}^n$. Additionally, for V, the action of Γ leaves $\mathbb{R}^n \times \{0\}$ invariant, and acts by orthogonal transformations on the fibers $\mathbb{R}^d$. In the latter case, the action is *not* cocompact on all of $\mathbb{R}^n \times \mathbb{R}^d$, and the quotient will be a noncompact space. The case $V-(n,0)$ covers the problem of embedding M, and $V-(n,d)$ covers the problem of embedding V.

Thus, we need to show $V-(n,d)$ for $n \geq 1$ and $d \geq 0$. First, we induct on $n+d$. If $n+d=1$, then we only have the case $V-(1,0)$, which is the base case for our induction.

Next, assume the statement has been shown for $n+d \leq m$, and we prove it for $V-(n,d)$ with $n+d=m+1$, that is $V-(k,m+1-k)$ for $k=1,\dots,m+1$. We show this by induction on k. The base case is $V-(1,m)$, which is reduced to $V-(a,b)$ with $a+b \leq m$.

Assume now that we have shown $V-(k,m+1-k)$ for all $k=1,\dots,k'-1$ with $k'<m+1$ and show it for $k=k'$. If $k'<m+1$, then the third reduction below reduces the embedding problem to instances of $V-(a,b)$ with either $a<k'$ or $a+b \leq m$. In both cases, we have reduced the embedding problem to cases we have assumed to be true. For $k'=m+1$, we want to prove $V-(m+1,0)$, which is reduced to embedding instance of $V-(a,b)$ with $a<m+1$. By the induction principle, we can assume these are already embedded, and thus the statement follows.

Defining variables We will use Γ to denote the group acting on $\mathbb{R}^n \times \mathbb{R}^d$ and $\mathbb{R}^n$ as described above. By Bieberbach's theorem, since the action of Γ on $\mathbb{R}^n$ is cocompact, there is a finite-index lattice $\Lambda = \Gamma \cap \mathbb{R}^n$ acting by translations on $\mathbb{R}^n$ with index $[\Gamma : \Lambda] \leq c(n)$. Denote by 0 the identity element of $\mathbb{R}^n$ and use the subadditive norm $|\gamma|_0 = |\gamma| = d(\gamma 0, \gamma)$ on Γ. Denote by e the identity element of Γ. Denote by $D = \mathrm{diam}(M) = \mathrm{diam}(\mathbb{R}^n/\Gamma)$. Apply Lemma 3.27 to choose the following objects. Choose a short basis $\gamma_1,\dots,\gamma_n$ for Λ, and the canonical grouping $I_1,\dots,I_S$ with parameters $l = 400c(n)$ and $L = 2l^n$; see Definition 3.29. Let S be the number of collapsed scales. By J_k for $k = 1, \dots, S$, we denote the k^{th} most collapsed scales. By the proof for Lemma 3.27, these correspond to normal subgroups $\Lambda_k < \Lambda$, which are normal in Γ. Also, they correspond to connected normal Lie-subgroups $L_k < \mathbb{R}^n$ which are invariant under the holonomy action of Γ, and such that $\Lambda_k < L_k$ are cocompact. Denote by l_k the collapsing scales given by Lemma 3.27.

Points in $\mathbb{R}^n$ are denoted by lower case letters such as p and points in the quotient space by $[p]$. Cosets with respect to local groups will be denoted by $[p]_r$. We will now proceed by induction on n and $n+d$ to show $V-(n,d)$ for all n.

Base case $V-(1,0)$ In this case, the flat manifold is one-dimensional and isometric to S^1. Such a manifold is trivial to embed.

Case $V-(n,0)$ when $S = 1$ From the definition of $l_1 = 2|\gamma_n|_0$ in Lemma 3.27, and Lemma A.4 we obtain $\operatorname{diam}(M) \le C(n)l_1$. Let $p \in \mathbb{R}^n$ be arbitrary. Set $\delta = l_1/(320c(n)L)$. Define the group $\Gamma_\delta = \langle g \in \Gamma \mid d(gp, p) < 8\delta \rangle$. By Lemma 3.12, for any $[p] \in M$ we have $B_{[p]}(\delta) \subset M$ is isometric to $B_{[p]_\delta}(\delta) \subset \mathbb{R}^n / \Gamma_\delta(p)$. We next describe the structure of the group $\Gamma_\delta(p)$.

Apply Lemma 3.17 with $l = 8\delta$. When we define $\Lambda_0 = \langle \lambda \in \Lambda \mid d(\lambda p, p) < 8\delta \rangle = \{e\}$, and observe that by Lemma 3.27 for any $\lambda \in \Lambda$ we have $|\lambda|_e > l_1/(2L) \ge 80c(n)\delta$, we obtain a short exact sequence

$$0 \to 0 \to \Gamma_\delta \to H_0 \to 0.$$

The group H_0 is finite and thus also Γ_δ is finite. Since M is a manifold, Γ_δ acts freely on $\mathbb{R}^n$. Applying Corollary 3.15, we get that, if Γ_δ were nontrivial, then it would have a fixed point. Since this would contradict that the action of Γ_δ is free, we must have Γ_δ is trivial. Thus, $B_{[p]_\delta}(\delta) \subset \mathbb{R}^n / \Gamma_\delta(p)$ is isometric to a Euclidean ball of the same radius $B_0(\delta) \subset \mathbb{R}^n$. Thus, each ball $B_{[p]}(\delta)$ is isometric to a subset of $\mathbb{R}^n$, and obviously admits a bi-Lipschitz embedding to $\mathbb{R}^n$. Also, $\operatorname{diam}(M)/\delta \le 320C(n)c(n)L$. Since p was arbitrary, Lemma 3.7 implies that M also possesses an embedding.

Reduction $V-(n,0)$ to $V-(n-k,k)$ for some $n > k \ge 1$ when $S > 1$ The gain here is that the dimension of the new base space $n-k$ is smaller than n. Consider $\delta = \max\bigl(l_S/(200Lc(n)), 2l_{S-1}\bigr)$. Note that $\operatorname{diam}(M) < C(n)l_S \le 200Lc(n)C(n)\delta$ by Lemma A.4. Thus, similarly to the previous case, it is sufficient by Lemma 3.7 to embed each ball $B_{[p]}(\delta)$. Next fix $[p] \in M$, and its representative $p \in \mathbb{R}^n$. We will show that $B_{[p]}(\delta)$ is isometric to a subset of a vector bundle over a $(n-k)$–dimensional flat manifold for some k.

Define the local group $\Gamma_\delta = \langle g \in \Gamma \mid d(gp, p) < 8\delta \rangle$. By Lemma 3.12, $B_{[p]}(\delta) \subset M$ is isometric to $B_{[p]_\delta}(\delta) \subset \mathbb{R}^n / \Gamma_\delta$. Define the subgroup $\Lambda_\delta = \langle \lambda \in \Lambda \mid d(\lambda p, p) < 8\delta \rangle$. By Lemma 3.27, for each $\lambda \notin \Lambda_{S-1}$ we have $d(\lambda p, p) > \frac{1}{4} l \cdot l_{S-1} > 80c(n)\delta$, and thus we can see that $\Lambda_\delta = \Lambda_{S-1}$. Further, by Lemma 3.17 applied with $l = 8\delta$, there is a finite group H_0 and a short exact sequence

$$0 \to \Lambda_{S-1} \to \Gamma_\delta \to H_0 \to 0.$$

Consider the connected Lie-subgroup L_{S-1} corresponding to Λ_{S-1} via Lemma 3.27. The subgroup L_{S-1} is invariant under Γ_δ and the holonomy. Thus, the action of Γ_δ descends to an action of H_0 on the orbit-space $E = \mathbb{R}^n / L_{S_1}$. Represent elements in E as orbits $[x]_{L_{S-1}}$ for $x \in \mathbb{R}^n$. Thus by Corollary 3.15, there is an orbit $[o]_{L_{S-1}} \in E$ fixed by the action of H_0. Let $O = [o]_{L_{S-1}} \subset \mathbb{R}^n$ be the orbit. By construction, O is invariant under the group action Γ_δ. Since O is invariant under Γ_δ we have $\mathbb{R}^n / \Gamma_\delta = V'$, where V' is a locally flat Riemannian vector bundle over O / Γ_δ of dimension $n-k$. Also, the action of Γ_δ is cocompact on O, because $\Lambda_{S-1} < \Gamma_\delta$. In particular, $B_{[p]_\delta}(\delta)$ is isometric to a subset of a k–dimensional locally flat Riemannian vector bundle over a $(n-k)$–dimensional compact connected flat manifold. Thus, the ball $B_{[p]}(\delta)$ can be embedded by $V-(n-k, k)$.

Reduction $V-(n, d)$ to $V-(n-k, d+k)$ when $S > 1$, and cases of the form $V-(a, b)$ with $a+b < n+d$ Denote by $V = \mathbb{R}^n \times \mathbb{R}^d / \Gamma$ the locally flat vector bundle over $M = \mathbb{R}^n \times \{0\} / \Gamma \subset V$. This reduction guarantees that either the dimension of the base space drops, or the total dimension drops. The proof is similar to the holonomy bundle example discussed in Section 2.

The space V is divided into sets defined by their distance to the zero-section M. Take the sets $T_0 = \{p \in V \mid d(p, M) \leq 2D\}$ and $T_k = \{p \in V \mid 2^{k+1} D \geq d(p, M) \geq 2^{k-1} D\}$ for $k \in \mathbb{N}$. Recall $D = \operatorname{diam}(M)$. We first show that it is sufficient to construct an embedding f_k for each T_k of bounded dimension and distortion. Suppose such embeddings exist. First scale and translate f_k so that $0 \in \operatorname{Im}(f_k)$ and f_k is 1–Lipschitz. Consider the sets $D_s = \bigcup_{k=0}^{\infty} T_{s+4k}$, and define $F_s(x) = f_{s+4k}(x)$ on T_{s+4k}. The resulting functions will be Lipschitz on their respective domains D_s. By an application of McShane extension Theorem 3.1, we can extend them to all of V. A bi-Lipschitz embedding will result from

$$G(x) = (r(x), F_0(x), F_1(x), F_2(x), F_3(x)),$$

where $r(x) = d(x, M)$. To see that this is an embedding, repeat the argument from Section 2. Next, construct embeddings f_i for each T_i.

First, we consider T_i for $i \geq 1$. Define $K = \{r(x) = 2^i D\}$. It is sufficient to find an embedding for the set K with its induced metric. The metric space T_i is bi-Lipschitz equivalent to $[2^{i-1}, 2^{i+1}] \times K$, and we can use the product embedding from Lemma 3.2 to embed it. Therefore, we can focus on embedding K. If $d = 1, 2$, then K will be compact and flat, and we can apply the case $V-(n+d-1, 0)$ or the case $V-(n+d-2, 0)$ to embed K. If $d > 2$, then K is not flat. But it is lower-dimensional and has the structure of a sphere bundle. For each $q \in K$, we find a ball with radius proportional to $2^i D$ and which is bi-Lipschitz to a lower-dimensional vector bundle. This combined with the doubling Lemma 3.7 gives an embedding for T_i.

For a real number t, denote by tS^{d-1} the $(d-1)$–dimensional sphere of radius t. The space K can be represented as $K = \mathbb{R}^n \times D2^i S^{d-1}/\Gamma$, where Γ acts on S^{d-1} by isometries via the holonomy action $h\colon \Gamma \to O(d)$. Any point in K can be represented as an orbit $[(q, v)] \in K$ for an element $(q, v) \in \mathbb{R}^n \times D2^i S^{d-1}$. Fix $\delta = D2^i/(10^5 c(d)\sqrt{d})$. We will show that $B_{[(q,v)]}(\delta) \subset K$ is bi-Lipschitz to a subset of locally flat vector bundle V of total dimension $m = n-1$. This will be done by studying the local group $\Gamma_\delta = \langle g \in \Gamma \mid d(g(q, v), (q, v)) \leq 8\delta \rangle$, and observing by Lemma 3.12 that $B_{[q]}(\delta)$ is isometric to $B_{[(q,v)]_\delta}(\delta) \subset \mathbb{R}^n \times D2^i S^{d-1}/\Gamma_\delta$ and describing in detail the action of Γ_δ.

By the modified Jordan theorem (Lemma 3.32), the group $h(\Gamma)$ has a normal abelian subgroup L of index at most $c(d)$, and $h^{-1}(L) = \Lambda_L$ is also normal and has finite index at most $c(d)$. Decompose $\mathbb{R}^d$ into canonical invariant subspaces V_j corresponding to L. By uniqueness, the subspaces V_j are invariant under the action of Γ. Decompose v into its components $v_j \in V_j$. Consider the induced norm $\|\cdot\|$ on $\mathbb{R}^d$. There is an index j such that $\|v_j\| \geq 2^i D/(3\sqrt{d})$. To reduce clutter, reindex so that $j = 1$ and let $\xi_1 = D2^i v_1/\|v_1\|$. Denote by $h_1\colon \Gamma \to O(V_1)$ the action of Γ on V_1. Denote by $d_{S^{d-1}}$ the metric on the sphere $2^i D S^{d-1}$. Then

$$d_{S^{d-1}}(v, \xi_1) \leq (2^i D) \cos^{-1}\Big(\frac{1}{3\sqrt{d}}\Big) \leq 2^i D\Big(\frac{\pi}{2} - \frac{1}{3\sqrt{d}}\Big).$$

For each generator g of Γ_δ, we have $d(g(q, v), (q, v)) \leq 8\delta$. Then also

$$d_{S^{d-1}}(g\xi_1, \xi_1) \leq 24\delta\sqrt{d}. \tag{4.2}$$

Let $\Lambda_0 = \Gamma_\delta \cap \Lambda_L$. Clearly, $[\Gamma_\delta : \Lambda_0] \leq c(d)$. Also, define

$$\Lambda_\delta = \langle g \in \Lambda_L \mid d_{S^{d-1}}(g\xi_1, \xi_1) \leq 24\delta\sqrt{d} \rangle.$$

Either the action of L, and of Λ_L, on the subspace V_1 will consist of reflections and identity transformations, or it will be a Hopf-type action. Consider the first case. We have $24c(d)\delta\sqrt{d} < D2^i\pi/4$, and thus the action of Λ_δ on V_1 is trivial. Consider the seminorm $|g|'_{\xi_1} = d_{S^{d-1}}(g\xi_1, \xi_1)$, and the action of the groups $h_1(\Lambda_L)$ and $h_1(\Gamma)$ on $O(V_1)$ with the parameter $l = 24\delta\sqrt{d}$. Applying Lemma 3.17, and using (4.2) we get a short exact sequence

$$0 \to h_1(\Lambda_\delta) \to h_1(\Gamma_\delta) \to H_0 \to 0,$$

where H_0 is a finite group of order at most $c(d)$.

Since the action of Λ_δ on V_1 is trivial, we get that $h_1(\Gamma_\delta)$ is finite. By (4.2) and the product bound of Lemma 3.17, we get for any $g \in h_1(\Gamma_\delta)$ that $d_{S^{d-1}}(g\xi_1, \xi_1) < 40c(d)\delta\sqrt{d} < D2^i\pi/(100\sqrt{d})$. By the fixed-point lemma (Corollary 3.15) applied to the sphere $D2^iS^{d-1}$, we find a $w \in V_1 \cap 2^iDS^{d-1}$ fixed by the action of $h_1(\Gamma_\delta)$. Moreover, this vector is fixed by the action of $h(\Gamma_\delta)$. Also, $d_{S^{d-1}}(\xi_1, w) \le D2^i\pi/(100\sqrt{d})$. Thus, we have

$$\begin{aligned} d_{S^{d-1}}(w, v) &\le d_{S^{d-1}}(v, \xi_1) + d_{S^{d-1}}(\xi_1, w) \\ &\le D2^i\Big(\frac{\pi}{100\sqrt{d}} + \frac{\pi}{2} - \frac{1}{3\sqrt{d}}\Big) \\ &\le D2^i\Big(\frac{\pi}{2} - \frac{1}{10\sqrt{d}}\Big). \end{aligned} \tag{4.3}$$

The submanifold $\mathbb{R}^n \times \{w\}$ is invariant under the action of Γ_δ. Since the action of Γ_δ on $\mathbb{R}^n$ is also cocompact, the manifold $M_\delta = \mathbb{R}^n \times \{w\}/\Gamma_\delta$ is a compact flat manifold. Fix $R = D2^i\big(\frac{\pi}{2} - 1/(20\sqrt{d})\big)$. The normal exponential map gives a map of the R–ball B' in the tangent space of $D2^iS^{d-1}$ at w onto $B_w(R) \subset D2^iS^{d-1}$. Call this normal exponential map $e_w\colon \mathbb{R}^{d-1} \to 2^iDS^{d-1}$. When B' is equipped with the induced Euclidean metric, this map becomes a bi-Lipschitz map $B' \to B_w(R)$. (See the calculation for lens spaces in Section 2, which can be generalized to this setting using polar coordinates). Taking products with $\mathbb{R}^n$ we get a bi-Lipschitz map $F\colon \mathbb{R}^n \times B' \to N_R(\mathbb{R}^n \times \{w\})$, where $N_R(\mathbb{R}^n \times \{w\})$ is the R–tubular neighborhood of $\mathbb{R}^n \times \{w\}$ in $\mathbb{R}^n \times 2^iDS^{d-1}$.

The action of Γ_δ induces an action on the normal bundle of the subset $\mathbb{R}^n \times \{w\}$, and thus on $\mathbb{R}^n \times B'$. Further, this action of Γ_δ on $\mathbb{R}^n \times B' \subset \mathbb{R}^n \times \mathbb{R}^{d-1}$ commutes with the action on $\mathbb{R}^n \times 2^iDS^{d-1}$ via the map F. Thus, F induced a bi-Lipschitz map

$$f\colon \mathbb{R}^n \times B'/\Gamma_\delta \to N_R(\mathbb{R}^n \times \{w\}).$$

Use (4.3) to get that $B_{[p,v]_r}(\delta)$ is contained in $N_R(\mathbb{R}^n\times\{w\})/\,\Gamma_\delta$. Thus, the restricted inverse $f^{-1}\colon B_{[(q,v)]_\delta}(\delta)\to\mathbb{R}^n\times B'/\,\Gamma_\delta$ is a bi-Lipschitz map. The image $\mathbb{R}^n\times B'/\,\Gamma_\delta$ is an isometric subset of $\mathbb{R}^n\times\mathbb{R}^{d-1}/\,\Gamma_\delta$ which is a $(d-1)$–dimensional locally flat vector bundle over M_δ, which is a compact flat manifold of dimension n. Thus, it can be embedded using the statement $V-(n,d-1)$, which has smaller total dimension.

The other case is when V_1 is a $2k$–dimensional rotational subspace where the action of L is given by the Hopf-action. We can define Λ_L and $\Lambda_0=\Lambda_L\cap\Gamma_\delta$ similarly to before. This time the action of Λ_0 on $\mathbb{R}^n$ will not necessary be trivial, but will be a Hopf-type action. For any point $\xi\in 2^iDS^{2k-1}\cap V_1$, its Λ_0–orbit is contained in a unique Hopf-fiber $S^1_\xi\subset 2^iDS^{2k-1}$. The action of Γ_δ conjugates the Hopf-action to itself, and thus any orbit S^1_ξ is mapped to another orbit S^1_ψ. We realize this orbit space as the complex projective space of diameter 2^iD, ie $2^iD\mathbb{CP}^{(k-2)/2}$, and the action of Γ_δ descends to $2^iD\mathbb{CP}^{(k-2)/2}$. Denote by $S^1_{\xi_1}$ the orbit through ξ_1.

Apply the argument in Lemma 3.17 with $\Lambda=\Lambda_0$, $\Gamma=\Gamma_\delta$, $l=24\delta\sqrt{d}$, and the seminorm $|g|'=d(gS^1_{\xi_1},S^1_{\xi_1})$. This gives that $\Lambda_0\cap\Gamma_\delta$ has finite index in Γ_δ and any coset of Λ_0 in Γ_δ can be represented by an element $g\in\Gamma_\delta$ such that

$$d(gS^1_{\xi_1},S^1_{\xi_1})\le d(g\xi_1,\xi_1)\le 24\delta\sqrt{d}\,c(d). \tag{4.4}$$

Thus, the action of Γ_δ on $2^iD\mathbb{CP}^{(k-2)/2}$ is by a finite group $H_0=\Gamma_0/\Lambda_0$. By (4.4), we have for any $h\in H_0$ that for the orbit $S^1_{\xi_1}$ corresponding to ξ_1,

$$d(hS^1_{\xi_1},S^1_{\xi_1})\le d(g\xi_1,\xi_1)\le 24c(d)\delta\sqrt{d}<\tfrac14 D2^i\pi.$$

Corollary 3.15 can be used to give an orbit S^1_w for some $w\in 2^iDS^{d-1}$, which is fixed under the action of H_0. In particular, the orbit as a subset of S^{d-1} is invariant under Γ_δ. Again, by Corollary 3.15 and (4.4) we get

$$d_{S^{d-1}}(S^1_w,v)\le D2^i\Big(\frac{\pi}{2}-\frac{1}{10\sqrt{d}}\Big). \tag{4.5}$$

This time define $M_\delta=\mathbb{R}^n\times S^1_w/\,\Gamma_\delta$ and consider it as a subset of K. Define $R=D2^i\big(\frac{\pi}{2}-1/(20\sqrt{d})\big)$. The ball $B_{(q,v)}(\delta)\subset K$ is contained in the R–tubular neighborhood $N_R(\mathbb{R}^n\times S^1_w/\,\Gamma_\delta)$ by (4.5). See Figure 1 for a simplified image of this normal neighborhood. The R–tubular neighborhood can be identified in a bi-Lipschitz fashion with $\mathbb{R}^n\times S^1_w\times B'/\,\Gamma_\delta$ for B' a Euclidean ball of radius R and dimension $d-2$. The identification is via the normal exponential map of the submanifold $\mathbb{R}^n\times S^1_w/\,\Gamma_\delta\subset K$. The quotient $\mathbb{R}^n\times S^1_w\times B'/\,\Gamma_\delta$ is isometric to a subset of a $(d-2)$–dimensional locally

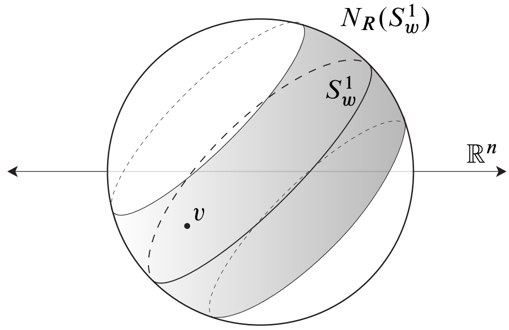

Figure 1: The normal neighborhood of an orbit of the Hopf-action is presented in this simplified image, where S^1_w is the invariant fiber and the shaded region its normal neighborhood. We have simplified the setting by reducing the sphere to a 2–dimensional sphere and by indicating the $\mathbb{R}^n$ factor abstractly as a line. The point v is depicted as belonging to the normal neighborhood.

flat vector bundle V'' over M_δ. In particular, we get a map $F\colon B_{(q,v)}(\delta) \to V''$. The conclusion thus follows from the statement $V-(n+1, d-2)$.

The previous paragraphs give bi-Lipschitz maps for each T_i when $i \geq 1$. For $i = 0$, we need a different argument. First assume $S = 1$. Similar to the case $V-(n,0)$ when $S = 1$, we can use Lemma 3.7 and embed each ball $B_{[p]}(\delta)$ since for $\delta = l_1/(320c(n)L) \sim D$ we have $\Gamma_\delta = \{e\}$. In other words, the space has a lower bound on its injectivity radius. Also, note $\operatorname{diam}(T_0) \leq 6D$.

Assume instead that $S > 1$. Consider an arbitrary $B_{[p]}(\delta)$ with $[p] \in T_0$ for $\delta = \max(l_S/(200Lc(n)), 2l_{S-1})$. Here $[p]$ is an orbit of the Γ–action for $p = (p^*, v) \in \mathbb{R}^n \times \mathbb{R}^d$. Since $\operatorname{diam}(T_0) \leq 6D \leq 200Lc(n)C(n)\delta$, we can use the Lemma 3.7 to construct embeddings once we can embed all of the balls $B_{[p]}(\delta)$. Again define $\Gamma_\delta = \langle g \in \Gamma \mid d(gp, p) \leq 8\delta\rangle$. Also, define $\Gamma'_\delta = \langle g \in \Gamma \mid d(gp^*, p) \leq 8\delta\rangle$. We have $\Gamma_\delta < \Gamma'_\delta$. Denote their orbits containing x by $[x]_\delta$ and $[x]'_\delta$ respectively. Then we have $B_{[p]}(\delta) \subset V$ is isometric to $B_{[p]_\delta}(\delta) \subset \mathbb{R}^n \times \mathbb{R}^d/\,\Gamma_\delta$ and $B_{[p]'_\delta}(\delta) \subset \mathbb{R}^n \times \mathbb{R}^d/\,\Gamma'_\delta$. Here we used Lemma 3.12.

By the second case in the proof, there is an invariant affine subspace $O \subset \mathbb{R}^n$ of dimension $n-k$ such that $\Gamma'_\delta O \subset O$. In particular, $\mathbb{R}^n \times \mathbb{R}^d/\Gamma'_\delta$ is isometric to $O \times \mathbb{R}^{d+k}/\,\Gamma'_\delta$. The space $O \times \mathbb{R}^{d+k}/\,\Gamma'_\delta$ is a locally flat vector bundle over $O/\,\Gamma_\delta$, and since $B_{[p]}(\delta)$ is isometric to a subset of it, we can embed it by the statement $V-(n-k, d+k)$ for some $k > 0$. $\square$

4.2 Flat orbifolds

We next modify the proof slightly to allow us to embed complete flat orbifolds.

Theorem 4.6 *Every complete connected flat orbifold O of dimension n admits a bi-Lipschitz embedding into Euclidean space with distortion and dimension depending only on the dimension n. Further, every locally flat orbivector bundle over such a base with its natural metric admits such an embedding with distortion and target dimension depending only on the total dimension.*

Proof We first give a description of the induction argument, and then move onto defining the relevant notation and proving the induction claims.

Summary of induction argument The arguments are sufficiently similar to the proof of Theorem 4.1 that we only indicate the main differences. We will need to modify the induction statements slightly to add the case $V-(0,d)$, which corresponds to an orbivector bundle over a point, ie $\mathbb{R}^n/\Gamma$, where Γ is a finite group. We assume the same notational conventions as used in the beginning of Theorem 4.1:

- $\boldsymbol{V-(n,0)}$ There exist constants $L(n,0)$ and $N(n,0)$ such that the following holds: Assume M is an arbitrary connected compact flat orbifold with dimension n. Then there exists a bi-Lipschitz embedding $f\colon M \to \mathbb{R}^N$ with distortion $L \le L(n,0)$ and target dimension $N \le N(n,0)$ depending only on n.
- $\boldsymbol{V-(n,d),\ n \ge 1,\ d > 1}$ There exist constants $L(n,d)$ and $N(n,d)$ such that the following holds: Assume V is an arbitrary d–dimensional locally flat Riemannian orbivector bundle over a compact flat orbifold M of dimension n. Then there exists a bi-Lipschitz embedding $f\colon V \to \mathbb{R}^N$ with distortion $L \le L(n,d)$ and target dimension $M \le N(n,d)$ depending only on n, d.
- $\boldsymbol{V-(0,d)}$ There exist constants $L(0,d)$ and $N(0,d)$ such that the following holds: Assume $V = \mathbb{R}^d/\Gamma$ where $\Gamma < O(d)$ is a finite group. Then there exists a bi-Lipschitz embedding $f\colon V \to \mathbb{R}^N$ with distortion $L \le L(0,d)$ and target dimension $N \le N(0,d)$ depending only on d.

Overview of induction We need to embed all complete connected flat orbifolds. Again, by Bieberbach's theorem, every complete connected flat orbifold can be represented as a quotient by a discrete group of isometries (Theorem 3.18). In particular, any such manifold is either isometric to $\mathbb{R}^n$, a quotient of $\mathbb{R}^n$ by a cocompact quotient or a locally flat vector bundle over such a base. In the first case, the embedding is given by the identity mapping. Otherwise, either the space can be represented as $M = \mathbb{R}^n/\Gamma$, or $V = \mathbb{R}^n \times \mathbb{R}^d/\Gamma$, where Γ acts cocompactly, properly discontinuously and by isometries on $\mathbb{R}^n$ and by isometries on $\mathbb{R}^d$. The case $V-(n,0)$ covers the problem of embedding M, and $V-(n,d)$ covers the problem of embedding V. We allow for $n=0$ in these statement.

The induction works in the same way as for Theorem 4.1, except we have new base cases for the induction $V-(0,1)$ and $V-(0,2)$.

Remark We will assume the same notation as in Theorem 4.1, except for the slight variation that the group actions are not assumed to be free.

Special cases $V-(0,1)$, $V-(0,2)$ and $V-(1,0)$ The spaces that result can all be enumerated and checked individually: a ray, two-dimensional cone over a circle, a circle, or a cone over an interval. Their bi-Lipschitz embeddings can be explicitly constructed.

The rest of the proof covers various inductive steps similar to Theorem 4.1, but including the degenerate one where the base is zero-dimensional.

Reduction $V-(0,n)$ to $V-(n-k,k-1)$ Let $\Gamma < O(n)$ be a finite subgroup. The space $M = \mathbb{R}^n/\Gamma$ is isometric to $C(S^{n-1}/\Gamma)$, ie the cone over S^{n-1}/Γ. By Lemma 3.5, it is sufficient to embed $S = S^{n-1}/\Gamma$. By the Jordan theorem,[8] Lemma 3.32, we have a finite-index normal subgroup $K \lhd \Gamma$ which is abelian with index bounded by $[\Gamma : K] \le c(n)$. Decompose $\mathbb{R}^n$ to canonical invariant subspaces V_i for K.

Let $\delta = 1/(10^5 c(n)\sqrt{n})$, and fix an arbitrary $[v] \in S^{n-1}/\Gamma$ which is represented by an element $v \in S^{n-1}$. By Lemma 3.7, it will be sufficient to embed $B_{[v]}(\delta) \subset S^{n-1}/\Gamma$. Define $\Gamma_\delta = \langle g \in \Gamma \mid d(gv, v) \le 8\delta \rangle$, $K_\delta = \langle g \in K \mid d(gv, v) \le 8\delta \rangle$ and $K'_\delta = \Gamma_\delta \cap K$. Decompose v into its components $v_i \in V_i$. There is an index i such that $\|v_i\| \ge 1/(3\sqrt{n})$. To reduce clutter, reindex so that $i = 1$, and define $\xi_1 = v_1/\|v_1\|$. But then $d(g\xi_1, \xi_1) \le 24\sqrt{n}\delta$. Repeating the argument in the proof of Theorem 4.1 we get either an element $w \in S^{n-1}$ fixed by Γ_δ, or a Hopf-orbit $S^1_w \subset V_1 \cap S^{n-1}$ through w which is invariant under Γ_δ. By using the normal exponential map, we get that $B_{[v]}(\delta) \subset S^{n-1}$ is bi-Lipschitz to a subset of either a $(n-1)$–dimensional locally flat orbivector bundle over a point (ie w), or a $(n-2)$–dimensional orbivector bundle

[8] In this case, Λ is the trivial group, and the statement is the classical Jordan theorem [57].

over a one-dimensional compact flat orbifold (ie S^1_w/Γ_δ). These can be embedded using the statements $V-(0,n-1)$ or $V-(1,n-2)$.

Reduction $V-(n,0)$ to $V-(0,n)$ when $S=1$ This is very similar to the $V-(n,0)$ case in Theorem 4.1. We still have the estimate $\operatorname{diam}(M) < C(n)|l_1|$. Let $p \in \mathbb{R}^n$ be arbitrary. Set $\delta = l_1/(320c(n)L)$. Define the group $\Gamma_\delta = \langle g \in \Gamma \mid d(gp,p) < 8\delta \rangle$. On the other hand, by Lemma 3.17 for any $[p] \in M$ we have $B_{[p]}(\delta) \subset M$ is isometric to $B_{[p]_\delta}(\delta) \subset \mathbb{R}^n/\Gamma_\delta(p)$. Thus, by Lemma 3.7 we can reduce the embedding problem to finding embeddings for $B_{[p]_\delta}(\delta)$. Define $\Lambda_0 = \langle \lambda \in \Lambda \mid d(\lambda p, p) < 8\delta \rangle = \{e\}$, and similarly to before we get

$$0 \to 0 \to \Gamma_\delta \to H_0 \to 0,$$

where H_0 and thus Γ_δ is finite. The group Γ_δ has a fixed point o by Corollary 3.15, and thus we can identify $\mathbb{R}^n/\Gamma_\delta$ by a quotient of isometries fixing o. Thus, $\mathbb{R}^n/\Gamma_\delta$ can be embedded by reduction to the $V-(0,n)$ case.

Reduction $V-(n,0)$ to $V-(n-k,k)$ when $S>1$ The proof for manifolds translates almost verbatim. Consider $\delta = \max(l_S/(200Lc(n)), 2l_{S-1})$. Again, $\operatorname{diam}(M) < C(n)l_S \le 200Lc(n)C(n)\delta$ by Lemma A.4. Thus, it is sufficient by Lemma 3.7 to embed each ball $B_{[p]}(\delta)$. Next fix $[p] \in M$, and its representative $p \in \mathbb{R}^n$. Repeating the arguments from the proof of Theorem 4.1 we can show that $B_{[p]}(\delta)$ is isometric to an orbivector bundle over a $(n-k)$–dimensional flat orbifold O. The only difference is that the action is not free.

Reduction $V-(n,d)$ to $V-(n-k,k+d)$, $V-(a,b)$ with $k>0$ or $a+b<n+d$ There is no major change to the proof in Theorem 4.1 as we didn't use the fact that the action is free in any substantive way. The same decomposition into the T_i is used, followed by a consideration of the action of Γ_δ on S^{d-1}, and finding a fixed point w or a fixed Hopf-orbit S^1_ξ. □

Remark (on bounds) The methods are likely not optimal, so we only give a rough estimate of the size of the dimension and distortion in the previous construction. Similar bounds could be obtained for the $\epsilon(n)$–quasiflat spaces considered below. Let $D(n)$ be the worst case distortion for a flat orbifold of dimension n, and let $N(n)$ be the worst case dimension needed. In particular, $D(n) \le \max_{k=0,\dots,n} L(n-k,k)$, and $N(n) \le \max_{k=0,\dots,n} N(n-k,k)$, where L, N are defined in the statement of $V-(n,k)$ after Theorem 4.6. For dimension $n=1$, we can bound the distortion by constants $L(1,0)$, and $L(0,1)$, and they can be embedded in three dimensions.

Next, $V-(n,d)$ was reduced to $V-(n-k,d+k)$ and $V-(a,b)$ with $a+b<n+d$. Each such reduction is done by using Lemma 3.7 at the scale $R/r \sim n!^{n+1} \sim O(e^{Cn^2 \ln(n)})$, which increases distortion by a multiplicative factor of $O(e^{Cn^2 \ln(n)})$. In order to reduce $V-(n,d)$ to lower-dimensional spaces, this reduction might need to be repeated up to n times. Thus, $D(n) = O(e^{Cn^3 \ln(n)} D(n-1))$, which gives $D(n) = O(e^{Cn^4 \ln(n)})$. By a similar analysis, we get that the dimension of the embedding can be bounded by $N(n) = O(e^{Cn \ln(n)})$. Our arguments are not optimal and probably some exponential factor could be removed, but the methods don't seem to yield better than superexponential distortion.

4.3 Quasiflat orbifolds

Below, the proof of Theorem 1.6 will proceed by approximating the orbifold by a flat orbifold. This will involve quotienting away the commutator subgroup $[N,N]$ and using Lemma 3.10. Thus, we begin by studying lattices and their commutators.

Lemma 4.7 *Let $\Lambda < N$ be a lattice. Then $[\Lambda,\Lambda]$ is a lattice in $[N,N]$.*

Proof $\Lambda' = [\Lambda,\Lambda]$ is a torsion-free nilpotent Lie group and by Malcev it corresponds to a connected Lie group N' such that N'/Λ' is compact; see [40] and [12]. Further, by Malcev there is a unique homomorphism $f\colon N' \to N$ such that $f_*(\Lambda') = \Lambda'$. We want to show that $f(N') = [N,N]$. First of all, since $\Lambda' < [N,N]$, we get that $f(N') < [N,N]$.

The image $f(N')$ is normal in N, and $N/f(N')$ is a nilpotent Lie group. The normality follows since Λ' is normal in Λ, and thus $\lambda f(N')\lambda^{-1} < N'$ for all $\lambda \in \Lambda$. From this we can conclude, using the nilpotency of N' and the fact that Λ is a lattice, that $Nf(N')N^{-1} < N'$. This follows since every element can be represented as $n = \prod_i e^{t_i X_i}$. The derivative of $\mathrm{Ad}_n(x) = nxn^{-1}$ at the identity is $\mathrm{ad}_n(t_i)$, which is a matrix-polynomial in t_i, which has the tangent space of $f(N')$ as an invariant subspace whenever t_i is integer. Thus, the same tangent space is left invariant by $\mathrm{ad}_n(t_i)$ for any $t_i \in \mathbb{R}$.

The projection $\pi\colon N/\Lambda' \to N/f(N')$ has compact fibers and thus we can show that the action of $\Lambda/[\Lambda,\Lambda]$ on $N/f(N')$ is properly discontinuous and cocompact. However, $\Lambda/[\Lambda,\Lambda]$ is abelian, and thus $N/f(N')$ is also abelian (by another application of Malcev's theorems). Therefore, $[N,N] < f(N')$, which completes the proof. □

Lemma 4.8 *Let n be an integer, $\epsilon(n)$ sufficiently small,*[9] *and $l \le \epsilon(n)$ fixed. There exists a dimension dependent constant $C(n)$ such that the following holds. If N is an n–dimensional simply connected nilpotent Lie group with a left-invariant metric g and sectional curvature $|K| \le 2$, and if $\Lambda < N$ is a lattice which is generated by elements of length $|l|_e \le l \le \epsilon(n)$, then $\operatorname{diam}(N/\Lambda) \le C(n)l$.*

Proof As in Lemma 3.27, we choose a short basis γ_i for the lattice Λ. By construction, $|\gamma_i|_e \le l$ for all $i = 1, \dots, n$. The result then follows from Lemma A.4, with the same choice of $C(n)$. □

Lemma 4.9 *Let S be an $\epsilon(n)$–quasiflat n–dimensional orbifold. Then there exist constants $0 < \delta < 1$ depending only on the dimension n and a constant $A(n)$ such that the following two properties hold:*

- *For every $[p] \in S$, there exists a flat Riemannian orbivector bundle $V = V(S')$ over a k–dimensional $\epsilon(k)$–quasiflat orbifold S' with $k < n$, with dimensions of the fibers $n-k$, and an $A(n)$–bi-Lipschitz map $f_{[p]}\colon B_{[p]}(\delta \operatorname{diam}(S)) \to V(S')$.*
- *There exists a flat orbifold S'' such that $d_{GH}(S, S') < \delta \operatorname{diam}(S)$.*

Proof Represent S as a quotient $S = N/\Gamma$, where Γ is the discrete group guaranteed by Definition 3.21. Further, take the finite-index subgroup $\Lambda = N \cap \Gamma$ with $[\Gamma : \Lambda] < c(n)$, which is guaranteed by Lemma 3.23. Let $l = 400c(n)$, $L = 2l^n$, I_j, J_k, γ_j, Λ_k, L_k and s be as in Lemma 3.27 and Definition 3.29. As in Lemma 3.27, we need to assume $\epsilon(n)$ small enough. For this proof, $\epsilon(n) \le \delta(n)/(C(n)10^3 c(n)C'(n)L)$ is enough. See Section 3 for the definitions of the various constants. We will first prove the existence of a nearby flat orbifold.

We can generate Λ by elements α_i such that $|\alpha_i|_e \le 4c(n) \operatorname{diam}(S) \le 2c(n)\epsilon(n)$ as in Lemma 3.27. As such, from nilpotency we see that we can generate $[\Lambda, \Lambda] = \{[\mu, \nu] \mid \mu, \nu \in \Lambda\} < \Lambda$ by elements $\beta_{i,j} = [\alpha_i, \alpha_j] \in N$. From (3.28) we get $|\beta_{i,j}|_e < \operatorname{diam}(S)/(10C(n)L^2) < \operatorname{diam}(S)/(8LC(n)c(n))$.

Choose the largest index a such that $|\gamma_a|_e < \operatorname{diam}(S)/(8LC(n)c(n))$, and let $a \in I_b$ for some b. By construction, $l_b \le L|\gamma_a|_e \le \operatorname{diam}(S)/(8C(n)c(n))$. Also, $l_b < l_s$ because $\operatorname{diam}(S) \le C(n)l_s$ by Lemma A.4. Thus, $L_b \ne N$.

[9] See the remark at the beginning of Section 3.

We have that $\beta_{i,j} \in \Lambda_b$, because if $\beta_{i,j} \notin \Lambda_b$, then, by construction, $i_b + 1 \notin J_b$ and we have (by construction of the basis in Lemma 3.27)

$$|\gamma_{i_b+1}|_e \leq |\beta_{i,j}|_e < \operatorname{diam}(S)/(8Lc(n)C(n)),$$

which would contradict the choice of a. Thus, we see that $[\Lambda, \Lambda] < L_b$ and that $l_b \leq \operatorname{diam}(S)/(8C(n)c(n))$. Since L_b is invariant under conjugation by elements in Γ and the holonomy of Γ, we see that the action of Γ descends to an action of a discrete group Γ' on $N/L_b = N'$, where $\Gamma' = \Gamma/\Gamma_b$ and Γ_b is the isotropy group of L_b. Define $N'/\Gamma' = S'$. Note that Γ' has a translational subgroup $\Lambda' = \Lambda/\Lambda_b$ which is abelian and acts cocompactly on N'. Therefore, N' is in fact flat and S' is a flat orbifold. We will next show that $d_{GH}(S', S) < \delta \operatorname{diam}(S)$ for $\delta = 1/(2c(n))$.

The orbits of the L_b–action on N descend to orbits on S, and S' can be identified by the orbit-space of the L_b–orbits on S. This gives a map between orbifolds $\pi\colon S \to S'$, which lifts to a Riemannian submersion on their orbifold universal covers, which are N and N' respectively. We have $\operatorname{diam}(S') \leq \operatorname{diam}(S) \leq \epsilon(n)$. Choose a coset $[p]' \in S'$ represented by an element $p \in N$ with

$$|p| \leq 4c(n)\epsilon(n). \tag{4.10}$$

Denote by $[p] = \pi^{-1}([p]')$ the coset in S.

Thus, it will be sufficient to prove that $\operatorname{diam}(\pi^{-1}([p])) < \frac{1}{2}\delta \operatorname{diam}(S)$. We see that $\pi^{-1}([p]') = O_{[p]}$, where $O_{[p]} \subset S$ is the orbit of L_b passing through $[p] \in S$. The set $O_{[p]}$ is isometric to L_b/Γ_p where $\Gamma_p = \{p^{-1}\gamma p \mid \gamma \in \Gamma\}$ with a translational subgroup $\Lambda_p = p^{-1}\Lambda'_b p$. This translational subgroup has a basis given by $p^{-1}\gamma_i p$ for $i = 1, \ldots, i_{b+1} - 1$. We have, by (3.28) and (4.10), that

$$|p^{-1}\gamma_i p|_e \leq 2|\gamma_i|_e \leq \operatorname{diam}(S)/(4C(n)c(n)).$$

By Lemma 4.8, we have $\operatorname{diam}(L_b/\Gamma_p) \leq \operatorname{diam}(S)/(4c(n))$. Thus, $d_{GH}(S', S) \leq \delta \operatorname{diam}(S)$ follows by our choice $\delta = 1/(2c(n))$ and using Lemma 3.11.

Next, we prove the first part, that at the scale of $\delta' = \delta \operatorname{diam}(S)$ the space has a description in terms of a simpler vector bundle. Define $\Gamma_{\delta'}(p) = \langle \gamma \in \Gamma \mid d(\gamma p, p) \leq 8\delta' \rangle$. By Lemma 3.12, we know that $B_{[p]}(\delta') \subset S$ is isometric to $B_{[p]_{\delta'}}(\delta') \subset N/\Gamma_{\delta'}(p)$. Further by Lemma 3.17, and observing that $\Lambda_{\delta'} = \langle \gamma \in \Lambda \mid d(\gamma p, p) \leq 8\delta' \rangle = \Lambda_b$ we have a finite group H_0 and a short exact sequence

$$0 \to \Lambda_b \to \Gamma_{\delta'} \to H_0 \to 0.$$

The action of H_0 descends to an action on N/L_b by $h[x]_{L_b} = [hx]_{L_b}$. Further, by Lemma 3.17, for every element $h \in H_0$ we can choose a representative $\gamma_h \in \Gamma_{\delta'}$ with $d(\gamma_h p, p) \le 8(c(n)+1)\delta'$. In particular, $d(h[p]_{L_b}, [p]_{L_b}) \le 8(c(n)+1)\delta'$. Thus, by Corollary 3.15 used on N/L_b (which is flat) there is an orbit $[q]_{L_b}$ which is fixed by the action of H_0 and $d([q]_{L_b}, p) \le 8(c(n)+1)\delta'$. Further, the orbit $[q]_{L_b} \subset N$ is invariant under the action of $\Gamma_{\delta'}$.

Consider the normal bundle $\nu([q]_{L_b})$ of $[q]_{L_b}$ in N. The bundle has a trivialization and can be identified with $\mathbb{R}^t \times L_b$ for some t. This trivialization can be obtained similarly to Lemma A.5. By Lemma A.5, there is a map $f\colon N_{2\delta(n)}([q]_{L_b}) \to \mathbb{R}^t \times L_b$ from the $2\delta(n)$–tubular neighborhood $N_{2\delta(n)}([q]_{L_b})$ of $[q]_{L_b}$. This map is also $A(n)$–bi-Lipschitz on the ball $B_q(2\delta(n))$. Further, the group $\Gamma_{\delta'}$ acts on $N_{2\delta(n)}([q]_{L_b})$ and $\mathbb{R}^t \times L_b$ by isometries, and the action commutes with f. We can thus show that restricted to the ball $B = B_{[q]_{\delta'}}(\delta(n)) \subset N/\Gamma_{\delta'}$ the induced map $\overline{f}\colon N_{2\delta(n)}([q]_{L_b})/\Gamma_{\delta'} \to \mathbb{R}^t \times L_b/\Gamma_{\delta'}$ is also $A(n)$–bi-Lipschitz.

Denote the orbifold by $L_b/\Gamma_\delta = S'$, and its dimension by $k < n$. The orbifold S' is lower-dimensional because $L_b < N$ strictly. The space $\mathbb{R}^t \times L_b/\Gamma_\delta = V(S')$ is a locally flat Riemannian orbivector bundle over the k–dimensional $\epsilon(k)$–quasiflat orbifold S'.[10] Also, $B_{[p]_{\delta'}}(\delta') \subset B_{[q]_{\delta'}}(\delta(n))$, so further restricting we get an $A(n)$–bi-Lipschitz map $\overline{f}\colon B_{[p]_{\delta'}}(\delta') \to V(S')$, which is our desired map. □

A similar statement is needed for orbivector bundles.

Lemma 4.11 *For a small enough $\epsilon(n)$, and for every d and n, there is a constant $A(n,d)$ such that the following holds: Let $V = V(S)$ be a locally flat Riemannian d–dimensional orbivector bundle over an $\epsilon(n)$–quasiflat n–dimensional orbifold S. Then there exist constants $0 < \delta, \delta' < 1$, which depend only on the dimensions n and d such that the following two properties hold:*

- *For every $p \in V(S)$, there exists a flat Riemannian orbivector $V = V(S')$ bundle over a k–dimensional $\epsilon(k)$–quasiflat orbifold S' with $k < n$ and an $A(n,d)$–bi-Lipschitz map $f_p\colon B_p(\delta \operatorname{diam}(S)) \to V(S')$.*
- *There exists a locally flat Riemannian orbivector bundle $V'' = V(S'')$ over a flat basis S'' such that $d_{GH}(V, V'') < \delta' \operatorname{diam}(S)$.*

[10] The orbifold S' will have the same bound on its curvature and diameter as S. We use here the assumption that $\epsilon(k) \ge \epsilon(n)$ for $k \le n$.

Proof For this proof, $\epsilon(n) \leq \delta(n)/(C(n)10^3 c(n)C'(n)L)$ is enough; for the definition of L, see Definition 3.29. We can represent the spaces as $S = N/\Gamma$ and $V = N \times \mathbb{R}^d/\Gamma$, where Γ acts by rotations on $\mathbb{R}^d$ and cocompactly on N so that S is $\epsilon(n)$–quasiflat. By Lemma 3.23, there is a lattice $\Lambda = \Gamma \cap N$ of N with $[\Gamma : \Lambda] \leq c(n)$. Further, there is an induced action $h\colon \Lambda \to O(d)$. By Lemma 3.32, we have an index $c(d)$ subgroup $\Lambda' < \Lambda$ with $[\Lambda : \Lambda'] \leq c(d)$ and $h(\Lambda')$ abelian. In particular, $[\Lambda', \Lambda'] = \Lambda'_c < \Lambda'$. Also define $[N, N] = N_c$

The group Λ'_c acts trivially on $\mathbb{R}^d$ and by translations on N. Consider the left action of N_c on $N \times \mathbb{R}^d$, which acts trivially on $\mathbb{R}^d$ and by left multiplication on N. The group Γ maps orbits of N_c to similar orbits, and thus $\Gamma_c = \Gamma/\Lambda'_c$ is a discrete group which acts properly discontinuously on $N/N_c \times \mathbb{R}^d$. Further, the orbits of N_c descend to V. Denote by $V_c = (N/N_c \times \mathbb{R}^d)/\Gamma_c$ the group quotient, and $\pi\colon V \to V_c$ the induced projection map. The group action of Γ_c is cocompact on N/N_c, and by rotations on $\mathbb{R}^d$. Further, N/N_c is abelian, and thus isomorphic and isometric to $\mathbb{R}^l$ for some l. In other words, up to an isometry $V_c = \mathbb{R}^l \times \mathbb{R}^d/\Gamma_c$, which is a d–dimensional locally flat orbivector bundle over the flat orbifold $\mathbb{R}^l/\Gamma_c$.

Next, we give a diameter bound needed to apply Lemma 3.11. Take any $q = [(p, v)] \in V$ and the N_c orbit O_q passing through that point. We can assume, as in Lemma 4.9, that $|p|_e \leq 2c(n)\epsilon(n)$. Since Λ' is a lattice of N by Lemma 4.7, also Λ'_c is a lattice in $[N, N] = N_c$. Also $O_q = N_c/p^{-1}\Lambda_c p$. The compact manifold O_q is a subset of $N/p^{-1}\Lambda' p$, and $N_c/p^{-1}[\Lambda, \Lambda]p$ is a compact manifold with $\operatorname{diam}(N_c/[\Lambda, \Lambda]) \leq 20C(n)\operatorname{diam}(S)c(n)$ by computations from Lemma 4.9. Further, $N_c/p^{-1}[\Lambda, \Lambda]p$ covers O_q with index at most $c(d)$. Thus,

$$\operatorname{diam}(O_q) \leq 50C(n)\operatorname{diam}(S)c(n)c(d). \tag{4.12}$$

Note that O_q can be covered by $N \times \{(p, v)\}/\Lambda_c$, where (p, v) is a representative of q in $N \times \mathbb{R}^d$. Also, the action of Λ_c is trivial on $\mathbb{R}^d$. Define $\delta' = 100C(n)c(n)c(d)$ and $\Delta = 100C(n)c(n)c(d)\operatorname{diam}(S)$. As before, we can show that π is a Δ–quasi-isometry. In particular, $d_{GH}(V_c, V) \leq \Delta$. This shows the second part.

We will next show the first statement by a similar argument as in Lemma 4.9. Fix an arbitrary $p = (q, v) \in N \times \mathbb{R}^d$. Let $l = 400c(n)$, $L = 2l^n$, I_j, J_k, γ_j, Λ_k, L_k and s be as in Lemma 3.27 and Definition 3.29. As in Lemma 3.27, we need to assume $\epsilon(n)$ small enough.

As in Lemma 3.27, choose the largest index a with $|\gamma_a|_e < \mathrm{diam}(S)/(8Lc(n)C(n))$, and let $a \in I_b$ for some b. By construction, $l_b \leq L|\gamma_a|_e \leq \mathrm{diam}(S)/(8c(n)C(n))$ and $l_b < l_s$.

Let $\delta = 1/(8c(n))$ and redefine $\Delta = \delta\,\mathrm{diam}(S)$. Define $\Gamma_\Delta(q) = \{\gamma \in \Gamma \mid d(\gamma q, q) \leq 8\Delta\}$ and $\Gamma_\Delta(p) = \{\gamma \in \Gamma \mid d(\gamma p, p) \leq 8\Delta\}$. Since $\Gamma_\Delta(p) < \Gamma_\Delta(q)$, by Lemma 3.12 we know that $B_{[p]}(\Delta) \subset V$ is isometric to $B_{[p]_\Delta}(\Delta) \subset N \times \mathbb{R}^d / \Gamma_\Delta(q)$. Note, we first show that it is isometric to a ball in $N/\Gamma_\Delta(p)$, but by the inclusion $\Gamma_\Delta(p) < \Gamma_\Delta(q)$ it is easy to see that this induces the desired isometry. Further, by Lemma 3.17 we have a finite group H_0 and a short exact sequence

$$0 \to \Lambda_b \to \Gamma_\Delta(q) \to H_0 \to 0.$$

Here, $\Lambda_b < N$ is defined similarly to Lemma 4.9. The group Λ_b is a lattice in the connected Lie subgroup $L_b < N$ associated to it; see Lemma 3.26. By the same argument as in Lemma 4.9, we can find an L_b–orbit $[x]_{L_b}$ passing through $x \in N$ with $d(x,q) < 8(c(n)+1)\delta'$ in N which is invariant under the action of Γ. By Lemma A.5, we find for the $\delta(n)$–tubular neighborhood $N_{\delta(n)}([x]_{L_b})$ of $[x]_{L_b}$ in N a map $f\colon N_{\delta(n)}([x]_{L_b}) \times \mathbb{R}^d \to L_b \times \mathbb{R}^t \times \mathbb{R}^d$, which is identity on the $\mathbb{R}^d$ factor. Further, f is $A(n)$–bi-Lipschitz when restricted to $B_x(2\delta(n)) \times \mathbb{R}^d$.

The action of the group $\Gamma_\Delta(q)$ induces canonically an action on $L_b \times \mathbb{R}^t \times \mathbb{R}^d$. Quotienting by this action induces a map

$$\bar{f}\colon N_{\delta(n)}(([x]_{L_b})) \times \mathbb{R}^d / \Gamma_\Delta(p) \to L_b \times \mathbb{R}^t \times \mathbb{R}^d / \Gamma_\Delta(p).$$

Let $B(\delta(n)) = N_{\delta(n)}(([x]_{L_b})) \times \mathbb{R}^d / \Gamma_\Delta(p)$. The space $V'' = L_b \times \mathbb{R}^t \times \mathbb{R}^d / \Gamma_\Delta(p)$ is a locally flat Riemannian orbivector bundle over $L_b / \Gamma_\Delta(p)$, which is an $\epsilon(n{-}l)$–quasiflat $(n{-}l)$–dimensional orbifold for some $n \geq l > 0$. Note that $\epsilon(n-l) \geq \epsilon(n)$ by assumption.

Finally, consider the ball $B_{[p]_\Delta}(3\Delta) \subset B(\delta(n))$. This inclusion can be seen by considering lifted balls and noting that $B_p(2\Delta) \subset A(\delta(n))$. Note that we have chosen $\epsilon(n)$ so small that $8(c(n)+2)\Delta < \frac{1}{10}\delta(n)$. The restricted map $\bar{f}|_{B_{[p]_\Delta}(\Delta)}$ is also $A(n)$–bi-Lipschitz. □

Finally, we can prove the embedding theorem for quasiflat manifolds and vector bundles.

Proof of Theorem 1.6 The proof will proceed by a reverse induction argument initiated by the flat case of the theorem. We will define the induction statements $E(d,n)$ as follows:

- $\boldsymbol{E(0,n)}$ There exist $N(0,n)$ and $L(0,n)$ such that every $\epsilon(n)$–quasiflat n–dimensional orbifold S admits a bi-Lipschitz map $f\colon S \to \mathbb{R}^{N(0,n)}$ with distortion $L(0,n)$.
- $\boldsymbol{E(d,n)}$ **for** $\boldsymbol{n \geq d \geq 1}$ There exist $N(d,n)$ and $L(d,n)$ such that every d–dimensional locally flat orbivector bundle $V = V(S)$ over an $\epsilon(n-d)$–quasiflat $(n-d)$–dimensional orbifold S admits a bi-Lipschitz map $f\colon V \to \mathbb{R}^{N(d,n)}$ with distortion $L(d,n)$.

The statement will be proved by an induction on n. The case $n=1$ is completed since the cases included in $E(0,1)$ and $E(1,1)$ are flat orbifolds and can be embedded using Theorem 4.6. Next, fix $n>1$ and assume we have proved $E(t,s)$ for all $1 \leq s \leq n$ and $0 \leq t \leq s$.

We proceed to show $E(d,n)$ by reverse induction on d. The base cases are $d=n$, $n-1$. The statement $E(n,n)$ corresponds to a flat orbifold and is covered in Theorem 4.6. Further, the case $E(n-1,n)$ corresponds to a locally flat orbivector bundle over a flat base, so is also covered in Theorem 4.6.

Next, fix $0 \leq d \leq n-2$. Assume the statement $E(k,n)$ has been proved for $k>d$. There are two slightly different cases $d=0$ and $d \geq 1$. Assume the first. In this case, by Lemma 4.9, there is a δ such that $B_p(\delta \operatorname{diam}(S))$ is bi-Lipschitz to a case covered by $E(s,n)$ for some $s>0$. Thus, $B_p(\delta \operatorname{diam}(S))$ admits a bi-Lipschitz embedding by induction. Also, $d_{GH}(S,S') \leq \delta \operatorname{diam}(S)$ where S' is a flat orbifold of lower dimension. Also, S' admits a bi-Lipschitz embedding by Theorem 4.6. The statement then follows from Lemma 3.10. The case $d \geq 1$ is similar, but one uses Lemma 4.11 instead and concludes that $B_p(\delta \operatorname{diam}(S))$ is bi-Lipschitz to a case covered by $E(d+s,n)$ for some $s>0$, and $d_{GH}(V,V') \leq \delta' \operatorname{diam}(S)$ for a locally flat orbivector bundle V' over a flat base. The constants δ, δ' here depend only on the dimension.

We remark that to apply Lemma 3.10 we technically need to show that every $\epsilon(n)$–quasiflat space is doubling. This could be shown by combining techniques from [62, IV.5.8] (and [43]). Also, Lemmas 4.11 and 4.9 combined with induction actually can be used to prove doubling directly. However, in order to apply the proof of Lemma 3.10 we only need the doubling estimate on all balls $B_p(R) \subset V$ and $B_p(R) \subset R$, where $R>0$ has a fixed upper bound depending on $\epsilon(n)$, δ, δ' and the constants $N(s,t)$ and $L(s,t)$ for cases $E(s,t)$ assumed true by induction. This doubling estimate follows by the sectional-curvature estimate $|K| \leq 2$. □

4.4 Proofs of main theorems

With the previous special cases at hand we can prove the main theorems.

Proof of Theorem 1.3 Fix $\epsilon = \epsilon(n)$ and $\eta = \frac{1}{4}$. By Fukaya's fibration Theorem 1.7, every point $p \in M^n$ has a definite sized ball $B_p(\delta)$, where $\delta = 5\rho(n, \eta, \epsilon)$, equipped with a metric g' such that $(B_p(\delta), g')$ is isometric to a subset of a compact $\epsilon(n)$–quasiflat manifold M' or a locally flat vector bundle V over such a base. Since $\|g - g'\| < \frac{1}{4}$, then the ball $B_p\left(\frac{1}{2}\delta\right)$ is bi-Lipschitz to a subset of V, or M', with distortion at most $\left(1 + \frac{1}{4}\right)/\left(1 - \frac{1}{4}\right)$. Since the spaces M' and V can be embedded by Theorem 1.6, we can embed M^n using a doubling Lemma 3.7 with $R = \operatorname{diam}(A)$ and $r = \delta$. Note, by our choice of ϵ and η, the constant δ depends only on the dimension n. □

Proof of Theorems 1.4 and 1.5 The only difference from the proof of Theorem 1.3 is using Fukaya's fibration theorem for orbifolds (Theorem 1.8) with $\epsilon = \epsilon(n)$ and $\eta = \frac{1}{4}$. □

Appendix

A.1 Collapsing theory of orbifolds

Much of the terminology here is presented in [34; 58]. Recall that a *topological orbifold* is a second-countable Hausdorff space with an atlas of charts U_i, which are open sets satisfying the following:

- Each U_i has a covering by $\widehat{U}_i$ with continuous open embeddings $\phi_i \colon \widehat{U}_i \to \mathbb{R}^n$ and a finite group G_i acting continuously on $\widehat{U}_i$ such that U_i is homeomorphic to $\widehat{U}_i / G_i$. Denote the projection by $\pi_i \colon \widehat{U}_i \to U_i$.
- There are transition homeomorphisms $\phi_{ij} \colon \pi_i^{-1}(U_i \cap U_j) \to \pi_j^{-1}(U_i \cap U_j)$ such that $\pi_j \circ \phi_{ij} = \pi_i$.

A chart is denoted simply by U_i, without explicating the covering $\widehat{U}_i$ or other aspects. An *orbifold* is defined as one for which $\phi_j \circ \phi_{ij} \circ \phi_i^{-1}$ are smooth maps. Further, a *Riemannian orbifold* is an orbifold with a metric tensor g_i on each U_i for which $d\phi_{ij}^*(g_j) = g_i$. Any other notion like geodesics, sectional curvature or Ricci-curvature, is defined by considering the lifted metric on $\widehat{U}_i$. One may also define the notion of orbifold maps as those that lift to the covering charts $\widehat{U}_i$ and are compatible with the projections. A smooth orbifold map is also naturally defined by considering charts. Such a map is nonsingular if in charts it can be expressed as a local diffeomorphism. A nonsingular orbifold map which is surjective is called an orbifold covering map.

We also need a notion of an orbivector bundle and an orbiframe bundle. An orbivector bundle is a triple (V, O, π), where V, O are smooth orbifolds and $\pi\colon V \to O$ is a smooth map with the following properties:

- There is an atlas of O consisting of charts U_i such that for all of the charts $(U_i, \widehat{U}_i, G_i)$ of O there is a chart $V_i = \pi_i^{-1}(U_i)$.
- There are covering charts $\widehat{V}_i = \widehat{U}_i \times \mathbb{R}^k$ and a G_i–action such that $V_i = \widehat{V}_i / G_i$, and maps $\pi_i^V\colon \widehat{V}_i \to V_i$.
- The projections $\widehat{\pi}_i\colon \widehat{V}_i \to \widehat{U}_i$ satisfy $\widehat{\pi}_i = \pi \circ \pi_i^V$.
- The action of G_i on $\widehat{U}_i \times \mathbb{R}^k$ splits as $g(u, x) = (gu, g_u x)$, where G_i acts linearly on $\mathbb{R}^k$ (the action can depend on the base point u), and the action on the first component is the original action.
- The actions of G_i and G_j on overlapping charts commute.

If there is a bundle metric on each $\widehat{V}_i$ such that G_i acts by isometries on the fibers $\mathbb{R}^k$ and such that it is invariant under transition maps, then we call the orbivector bundle a Riemannian orbivector bundle. A connection on such bundle is defined as an affine connection on each $\widehat{V}_i$ which is invariant under G_i and the compatible via the transition maps. Such a connection is flat if it is flat on each chart. Such a connection together with a Riemannian structure on the base induces a metric on V making it into a Riemannian orbifold. This allows us to define a central notion in this chapter.

Definition A.1 (V, O, π) is called a locally flat orbivector bundle if it is a Riemannian orbivector bundle over a Riemannian orbifold with a flat connection.

To every smooth orbifold we can associate a tangent orbivector bundle TO, where the group actions on the fibers are given by differentials. The fiber of a point $\pi^{-1}(p) \subset TO$ is denoted by $C_p(O)$. For Riemannian orbifolds, this is easily seen to correspond with the metric tangent cone at p. In that case, $C_p(O) = \mathbb{R}^n / G$ for some group G acting by isometries on $\mathbb{R}^n$. One can also define orbifold principal bundles by considering a fiber modeled on a Lie group G and assuming transition maps are given by left-multiplication.

This allows us to define for Riemannian orbifolds an orbiframe bundle FO. The orbiframe bundle FO is always a manifold.

In order to do collapsing theory, we need to define a local orbifold cover. A connected Riemannian orbifold has a distance function, and we can define a complete Riemannian orbifold as one which is complete with respect to its distance function. On such orbifolds we may define an exponential map $\exp\colon TO \to O$ by considering geodesics and their lifts. Fix now a point $p \in O$ and take the exponential map at that point $\exp_p\colon C_p(O) \to O$, which is surjective. As $C_p(O) = \mathbb{R}^n/G$, we can lift the exponential map to a map $\widetilde{\exp}_p\colon \mathbb{R}^n \to O$. Next we state a crucial property of such an exponential map.

Lemma A.2 *If a complete Riemannian orbifold has sectional curvature bounded by $|K| \le 1$, then $\widetilde{\exp}_p$ is a nonsingular orbifold map on the ball $B_0(\pi) \subset \mathbb{R}^n$.*

The map $\exp_p\colon C_p(O) \to O$ will define locally a homeomorphism in the neighborhood of $p \in C_p(O)$. We call the largest δ such that $\exp_p|_{B(p,\delta)}$ is injective on $B(p,\delta) \subset C_p(O)$ the *injectivity radius* at p, and denote it by $\mathrm{inj}_p(O)$. If $O' \subset O$ is a suborbifold, then we can also define a normal bundle of it and a normal injectivity radius by the largest radius such that the normal exponential map is injective. These terms are used in the proof of Theorem 1.8 below.

Lemma A.2 is alluded to in the appendix of [15] and can be proved by using local coordinates and standard Jacobi-field estimates. Using this terminology, we can now state and outline the proof of Fukaya's fibration theorem for collapsed orbifolds

Theorem A.3 *Let (O,g) be a complete Riemannian orbifold of dimension n and sectional curvature $|K| \le 1$, and let $\epsilon > 0$ be an arbitrary constant. For every $\eta > 0$, there exists a universal $\rho(n,\eta,\epsilon) > 0$ such that for any point $p \in O$ there exists a metric g' on the ball $B_p(10\rho(n,\eta,\epsilon))$ and a complete Riemannian orbifold O' with the following properties:*

- $\|g - g'\|_g < \eta$.
- $(B_p(5\rho(n,\eta,\epsilon)), g')$ *is isometric to a subset of O'.*
- *O' is either an ϵ–quasiflat orbifold or a locally flat Riemannian orbivector bundle over an ϵ–quasiflat orbifold S.*

Proof Fix $\epsilon, \eta > 0$. Without loss of generality, we can assume ϵ is very small. If the result were not to hold, then it would have to fail for a sequence of pointed orbifolds (O_i, p_i). Thus, let O_i be a sequence of examples such that there doesn't exist any $\delta > 0$ such that for any i, the ball $B_{p_i}(\delta)$ has a metric g''_η with the following two properties: $\|g''_\eta - g\|_g < \eta$, and $(B_{p_i}(\delta), g''_\eta)$ is isometric to a subset of an ϵ–quasiflat orbifold, or an orbivector bundle over such a base. We will see a contradiction by establishing the existence of such a δ.

One should first regularize the orbifold slightly as to attain uniform bounds on all covariant derivatives. This is done in [34, second part, Section 5]. The action of the local group on the frame bundle is free, and thus Fukaya's proofs in [17; 18] (see also [12]) show that a subsequence of the orbifold frame bundles FO_i converges to a smooth manifold Y with bounds on all covariant derivatives. Assume without loss of generality that FO_i is already this sequence.

Using standard collapsing theory from Cheeger, Gromov and Fukaya [12], we obtain an equivariant fibration of FO_i by nilpotent fibers, and a metric g_η which is invariant under the local nilpotent action and the local $O(n)$–action. It is possible for those orbits to be finite. Further, we have bounds on their second fundamental forms and normal injectivity radii of the nilpotent fibers. (In the case of a point, the normal injectivity radius is just the injectivity radius.) By [51, Theorem 2.1 and Proposition 2.5], we can also control the sectional curvatures of g_η. By equivariance, this structure descends onto a fibration of O_i by orbits of nilpotent Lie groups. (Again, the fibers may be points, in which case the normal injectivity radius bound means the injectivity radius bound). Corresponding to this fibration we have, after possibly passing to another subsequence and reindexing, an invariant metric g'_η, which satisfies $\|g - g'_\eta\|_g < \eta'(i)$ for some $\eta'(i)$ going to zero with i. Call the fiber passing through p by $\mathcal{O}_p$.

By the argument in [12, Appendix 1], we can find (for large enough i) a constant δ and a fiber $\mathcal{O}_{q_i} \subset O_i$ with $d(q_i, p_i) < \delta$ such that the orbit $\mathcal{O}_{q_i}$ has normal injectivity radius at least 20δ. One can also bound the second fundamental form of $\mathcal{O}_{q_i}$. The normal bundle has a natural metric inherited from g'_η, and a locally flat connection inherited from the nilpotent action on the fiber $\mathcal{O}_{q_i}$. By the second fundamental form bounds, for i sufficiently large, the restricted metric g'_η makes $\mathcal{O}_{q_i}$ an ϵ–quasiflat orbifold. Further, the metric induced by the locally flat connection makes the normal bundle to $\mathcal{O}_{q_i}$ a locally flat Riemannian orbivector bundle over an ϵ–quasiflat base. Denote by g''_η the metric tensor which is induced on this normal bundle by the locally flat connection.

Denote by $N_{10\delta}(\mathcal{O}_{q_i})$ the vectors in the normal bundle of the fiber $\mathcal{O}_{q_i}$ with length at most 10δ. One can show, for i sufficiently large, that the normal-exponential map $f\colon N_{10\delta}(\mathcal{O}_{q_i}) \to O_i$ induces a map onto the 10δ–tubular neighborhood (with respect to g'_η of O_i) with $\|f^*g - g''_\eta\|_g < \eta''(i)$ for some $\eta''(i)$ tending to zero with i. Further, the image of f contains $B_p(\delta)$ and we can consider the restricted map $f^{-1}\colon B_p(\delta) \to N_{10\delta}(\mathcal{O}_{q_i})$. This gives a metric $f^{-1*}(g''_\eta) = g'$ on $B_p(\delta)$ with the desired properties. The model space O' is the normal bundle of $\mathcal{O}_{q_i}$ with the metric g''_η and the isometry of $B_p(\delta)$ is given by f. However, we assumed that no such structure could exist, and thus the statement follows by contradiction. □

A.2 Diameter estimate

Lemma A.4 *Let $M = N/\Gamma$ be an n–dimensional $\epsilon(n)$–quasiflat orbifold and $\Lambda = N \cap \Gamma < \Gamma$ the translational normal subgroup of Γ. Further, let γ_i be a short basis for M as in Lemma 3.27. Then we have* $\operatorname{diam}(M) \le C(n) \max_{i=1,\dots,n} |\gamma_i|$.

Here, we use the notation $|\gamma| = |\gamma|_e = d(\gamma, e)$, where e is the identity element of N, and d is the Riemannian distance function on N.

Proof Let $D = \max_{i=1,\dots,n} |\gamma_i|$. To prove the lemma we can assume $\Gamma = \Lambda$, because taking a finite cover will only increase the diameter. We will show a more general statement: if a cocompact lattice $\Lambda < N$ admits a triangular basis $\gamma_1, \dots, \gamma_n$ in the sense of Definition 3.25, then $\operatorname{diam}(M) \le C(n) \max_{i=1,\dots,n} |\gamma_i|$. The statement is obvious for $n = 1$, and we will proceed to assume that $n > 1$ and that the statement has been shown for $(n-1)$–dimensional simply connected nilpotent Lie groups.

Assume that $\Lambda \subset N$ is a discrete cocompact lattice of an n–dimensional nilpotent Lie group with a triangular basis γ_i and set $\gamma_i = e^{X_i}$ with $X_i \in T_eN$ a basis and $[X_i, X_j] \in \operatorname{span}(X_1, \dots, X_{i-1})$ for $i < j$. Since γ_i is assumed to be a basis, the vectors X_i form a basis for T_eN. (Recall Definition 3.25.) In particular, we can introduce coordinates $(t_1, \dots, t_n) \mapsto \prod_{i=1}^n e^{t_{n-i} X_{n-i}}$ such that the product can be expressed as

$$\prod_{i=1}^n e^{t_{n-i} X_{n-i}} \times \prod_{i=1}^n e^{s_{n-i} X_{n-i}} = \prod_{i=1}^n e^{(s_{n-i} + t_{n-i} + p_{n-i}(s,t)) X_i},$$

where $p_i(s,t)$ depends only on s_j, t_j for $j < i$. We will show that $\operatorname{diam}(N) \le C(n) D$.

Since γ_1 is in the center of N, we have that X_1 commutes with all left-invariant fields. Further, by the Koszul formula, $t \mapsto e^{tX_1}$ is a geodesic in N. Thus, $d(e^{sX_1}, e) = s\,d(e^{X_1}, e) = s|\gamma_1|$. Consider the Lie group $N' = N/L$, where $L = \{e^{tX_1}\}$, and a discrete subgroup $\Gamma' = \langle \gamma_i', i = 1, \dots, n-1 \rangle = \Gamma / \langle \gamma_1^n, n \in \mathbb{Z} \rangle$. Here, the generators correspond to equivalence classes of generators of Γ in $\Gamma/\langle \gamma_1^n, n \in \mathbb{Z}\rangle$: $\gamma_i' = [\gamma_{i+1}]$. If we let $e^{X_i'} = \gamma_i'$, we have coordinates $(t_1, \dots, t_{n-1}) \mapsto \prod_{i=1}^{n-1} e^{t_{n-i} X_{n-i}'}$. Here the exponential map is in N'. Take as the metric on N' the quotient metric, which is induced by the restricted Riemannian metric on a vector space perpendicular to $\operatorname{span}(X_1)$. With respect to this metric $|\gamma_i'| \le |\gamma_{i+1}| \le |\gamma_n| \le D$ for all $1 \le i \le n-1$. The norm $|\cdot|$ on N' is defined as $|n'| = d(n'e', e')$, where $n' \in N'$ and e' is the identity element of N'.

Define $M' = N'/\Gamma'$. The group L acts naturally on M by isometries and M/L is isometric to M'. Further, the L–orbits in M have length $|\gamma_1|$, so by Lemma 3.11, $d_{GH}(M, M') \le 3|\gamma_1|$. By the inductive hypothesis, $\operatorname{diam}(M') \le C(n-1)D$. Thus,

$$\begin{aligned}\operatorname{diam}(M) &\le \operatorname{diam}(M') + 2d_{GH}(M, M') \le C(n-1)D + 6D \\ &\le (6 + C(n-1))\, D.\end{aligned}$$

In particular, we can choose $C(n) = 6n$. □

Lemma A.5 *Let N^n be a simply connected nilpotent Lie group and $M^k \lhd N^n$ a connected normal subgroup thereof. Assume that N is equipped with a left-invariant metric which satisfies the sectional-curvature bound $|K| \le 2$. Then for some universal $\delta(n) > 0$ and the $4\delta(n)$–tubular neighborhood $N_{4\delta(n)}(M^k)$ of M^k in N^n, there is a map $F\colon N_{\delta(n)}(M^k) \to M^k \times \mathbb{R}^{n-k}$ whose restriction to a ball $F\colon B_e(2\delta(n)) \to M^k \times \mathbb{R}^{n-k}$ is $A(n)$–bi-Lipschitz. Also, F is an isometry when restricted to $F^{-1}(M^k \times \{0\})$. Furthermore, for any affine isometry I of N^n which preserves M^k, we also have that $F \circ I \circ F^{-1}$ is a restricted affine isometry of $M^k \times \mathbb{R}^{n-k}$.*

Proof Introduce an orthonormal triangular basis $X_1, \dots, X_k$ for M^k, and extend it to an orthonormal basis for N^n with vectors $X_{k+1}, \dots, X_n$. Consider the map $G\colon M^k \times \mathbb{R}^{n-k} \to N$, which is defined by

$$G(m, x) = m e^{\sum_{i=k+1}^{n} x_i X_i}.$$

This is smooth. The exponential map used here is the Lie group exponential. Clearly, any affine isometry I preserving M^k will conjugate to be an isometry of $M^k \times \mathbb{R}^{n-k}$

by considering its action on M^k and its normal bundle. Apply the Campbell–Hausdorff formulas to compute the quantities $\partial_{x_i} G$:

$$\partial_{x_i} G(m,x) = X_i + \sum_{k=1}^{n} \sum_{i_1,\dots,i_k} M^k_{i_1,\dots,i_k} x_{i_1} \cdot x_{i_2} \cdots x_{i_k} [X_{i_1}, [X_{i_2}, [\dots, [X_{i_k}, X_i]]]].$$

Denote the error terms in the sum by $Y_i = \partial_{x_i} G(m,x) - X_i$. Now assume $x_i \leq \delta$ for any i. The metric tensor can now be expressed in the given coordinates as follows:

$$\begin{aligned}\langle \partial_{x_i} G(m,x), \partial_{x_j} G(m,x) \rangle &= \langle X_i + Y_i, X_j + Y_j \rangle \\ &= \delta_{ij} + \langle Y_i, Y_j \rangle + \langle X_i, Y_j \rangle + \langle Y_i, X_j \rangle.\end{aligned}$$

Since $\|\mathrm{Ad}\| \leq C'(n)$, we get $\|Y_i\| \leq M(n)\delta^k C'(n)^k$. Here, δ_{ij} is the Dirac delta. Choose $\delta(n) < 1/(10^3 M(n) C'(n))$ and apply Cauchy's inequality to give

$$1 - \tfrac{3}{8} \leq \langle \partial_{x_i} G(m,x), \partial_{x_i} G(m,x) \rangle \leq 1 + \tfrac{3}{8} \tag{A.6}$$

whenever $x_i \leq 10\delta(n)$. Let $C(\delta(n)) = \{(m,x) \mid x_i \leq 10\delta(n)\} \subset M^k \times \mathbb{R}^{n-k}$. From its definition and (A.6), we can see that $G(C(\delta(n)))$ contains the $4\delta(n)$–tubular neighborhood $N_{4\delta(n)}(M^k)$ of M^k. Also, when restricting G^{-1} onto $B_e(2\delta(n)) \subset N$ we get a $\big((1+\frac{3}{8})/(1-\frac{3}{8})\big)$–bi-Lipschitz into $M^k \times \mathbb{R}^{n-k}$. Thus, the inverse G^{-1} is our desired map. We remark that the distortion can be made arbitrarily small by choosing a smaller $\delta(n)$. □

References

[1] **S Alexander**, **V Kapovitch**, **A Petrunin**, *Alexandrov geometry*, in preparation

[2] **A Andoni**, **A Naor**, **O Neiman**, *Snowflake universality of Wasserstein spaces*, preprint (2015) arXiv

[3] **P Assouad**, *Plongements lipschitziens dans* $\mathbf{R}^n$, Bull. Soc. Math. France 111 (1983) 429–448 MR

[4] **L Auslander**, *Bieberbach's theorems on space groups and discrete uniform subgroups of Lie groups*, Ann. of Math. 71 (1960) 579–590 MR

[5] **M Bonk**, **U Lang**, *Bi-Lipschitz parameterization of surfaces*, Math. Ann. 327 (2003) 135–169 MR

[6] **J Bourgain**, *On Lipschitz embedding of finite metric spaces in Hilbert space*, Israel J. Math. 52 (1985) 46–52 MR

[7] **D Burago**, **Y Burago**, **S Ivanov**, *A course in metric geometry*, Graduate Studies in Mathematics 33, Amer. Math. Soc., Providence, RI (2001) MR

[8] **Y Burago**, **M Gromov**, **G Perel'man**, *A D Aleksandrov spaces with curvatures bounded below*, Uspekhi Mat. Nauk 47 (1992) 3–51 MR In Russian; translated in Russian Math. Surveys 47 (1992) 1–58

[9] **P Buser**, **H Karcher**, *The Bieberbach case in Gromov's almost flat manifold theorem*, from "Global differential geometry and global analysis" (D Ferus, W Kühnel, U Simon, B Wegner, editors), Lecture Notes in Math. 838, Springer (1981) 82–93 MR

[10] **P Buser**, **H Karcher**, *Gromov's almost flat manifolds*, Astérisque 81, Soc. Math. France, Paris (1981) MR

[11] **J Cheeger**, *Differentiability of Lipschitz functions on metric measure spaces*, Geom. Funct. Anal. 9 (1999) 428–517 MR

[12] **J Cheeger**, **K Fukaya**, **M Gromov**, *Nilpotent structures and invariant metrics on collapsed manifolds*, J. Amer. Math. Soc. 5 (1992) 327–372 MR

[13] **M J Collins**, *On Jordan's theorem for complex linear groups*, J. Group Theory 10 (2007) 411–423 MR

[14] **L J Corwin**, **F P Greenleaf**, *Representations of nilpotent Lie groups and their applications, I: Basic theory and examples*, Cambridge Studies in Advanced Mathematics 18, Cambridge Univ. Press (1990) MR

[15] **Y Ding**, *A restriction for singularities on collapsing orbifolds*, ISRN Geom. (2011) art. id. 502814, 7 pages

[16] **Y Ding**, *A restriction for singularities on collapsing orbifolds*, preprint (2011) arXiv

[17] **K Fukaya**, *A boundary of the set of the Riemannian manifolds with bounded curvatures and diameters*, J. Differential Geom. 28 (1988) 1–21 MR

[18] **K Fukaya**, *Collapsing Riemannian manifolds to ones with lower dimension, II*, J. Math. Soc. Japan 41 (1989) 333–356 MR

[19] **K Fukaya**, *Hausdorff convergence of Riemannian manifolds and its applications*, from "Recent topics in differential and analytic geometry" (T Ochiai, editor), Adv. Stud. Pure Math. 18, Academic Press, Boston (1990) 143–238 MR

[20] **K Fukaya**, *Collapsing Riemannian manifolds and its applications*, from "Proceedings of the International Congress of Mathematicians, I" (I Satake, editor), Math. Soc. Japan, Tokyo (1991) 491–500 MR

[21] **P Ghanaat**, **M Min-Oo**, **E A Ruh**, *Local structure of Riemannian manifolds*, Indiana Univ. Math. J. 39 (1990) 1305–1312 MR

[22] **R E Greene**, **H Wu**, *Lipschitz convergence of Riemannian manifolds*, Pacific J. Math. 131 (1988) 119–141 MR

[23] **M Gromov**, *Almost flat manifolds*, J. Differential Geom. 13 (1978) 231–241 MR

[24] **M Gromov**, *Metric structures for Riemannian and non-Riemannian spaces*, Progress in Mathematics 152, Birkhäuser, Boston (1999) MR

[25] **K Grove**, **H Karcher**, *How to conjugate C^1–close group actions*, Math. Z. 132 (1973) 11–20 MR

[26] **I Haviv**, **O Regev**, *The Euclidean distortion of flat tori*, J. Topol. Anal. 5 (2013) 205–223 MR

[27] **J Heinonen**, *Lectures on analysis on metric spaces*, Springer (2001) MR

[28] **J Heinonen**, *Lectures on Lipschitz analysis*, Report, Matematiikan Laitos 100, Univ. of Jyväskylä (2005) MR

[29] **J Heinonen**, **S Keith**, *Flat forms, bi-Lipschitz parameterizations, and smoothability of manifolds*, Publ. Math. Inst. Hautes Études Sci. 113 (2011) 1–37 MR

[30] **P Indyk**, **J Matoušek**, *Low-distortion embeddings of finite metric spaces*, from "Handbook of discrete and computational geometry" (J E Goodman, J O'Rourke, editors), 2nd edition, Chapman & Hall/CRC, Boca Raton, FL (2004) 177–196

[31] **V Kapovitch**, *Perelman's stability theorem*, from "Surveys in differential geometry, XI" (J Cheeger, K Grove, editors), International Press, Somerville, MA (2007) 103–136 MR

[32] **H Karcher**, *Riemannian center of mass and mollifier smoothing*, Comm. Pure Appl. Math. 30 (1977) 509–541 MR

[33] **S Khot**, **A Naor**, *Nonembeddability theorems via Fourier analysis*, Math. Ann. 334 (2006) 821–852 MR

[34] **B Kleiner**, **J Lott**, *Locally collapsed* 3*–manifolds*, Astérisque 365, Soc. Math. France, Paris (2014) 7–99 MR

[35] **T J Laakso**, *Ahlfors Q–regular spaces with arbitrary $Q > 1$ admitting weak Poincaré inequality*, Geom. Funct. Anal. 10 (2000) 111–123 MR

[36] **U Lang**, **C Plaut**, *Bilipschitz embeddings of metric spaces into space forms*, Geom. Dedicata 87 (2001) 285–307 MR

[37] **U Lang**, **V Schroeder**, *Kirszbraun's theorem and metric spaces of bounded curvature*, Geom. Funct. Anal. 7 (1997) 535–560 MR

[38] **K Luosto**, *Ultrametric spaces bi-Lipschitz embeddable in* $\mathbf{R}^n$, Fund. Math. 150 (1996) 25–42 MR

[39] **J Luukkainen**, **H Movahedi-Lankarani**, *Minimal bi-Lipschitz embedding dimension of ultrametric spaces*, Fund. Math. 144 (1994) 181–193 MR

[40] **A I Mal'cev**, *On a class of homogeneous spaces*, Izvestiya Akad. Nauk. SSSR. Ser. Mat. 13 (1949) 9–32 MR In Russian; translated in Amer. Math. Soc. Trans. 39 (1951), 33 pages

[41] **J Matoušek**, *Bi-Lipschitz embeddings into low-dimensional Euclidean spaces*, Comment. Math. Univ. Carolin. 31 (1990) 589–600 MR

[42] **J Milnor**, *Curvatures of left invariant metrics on Lie groups*, Advances in Math. 21 (1976) 293–329 MR

[43] **A Nagel**, **E M Stein**, **S Wainger**, *Balls and metrics defined by vector fields, I: Basic properties*, Acta Math. 155 (1985) 103–147 MR

[44] **A Naor**, *An introduction to the Ribe program*, Jpn. J. Math. 7 (2012) 167–233 MR

[45] **A Naor**, **O Neiman**, *Assouad's theorem with dimension independent of the snowflaking*, Rev. Mat. Iberoam. 28 (2012) 1123–1142 MR

[46] **A Naor**, **Y Peres**, **O Schramm**, **S Sheffield**, *Markov chains in smooth Banach spaces and Gromov-hyperbolic metric spaces*, Duke Math. J. 134 (2006) 165–197 MR

[47] **I G Nikolaev**, *Bounded curvature closure of the set of compact Riemannian manifolds*, Bull. Amer. Math. Soc. 24 (1991) 171–177 MR

[48] **P Pansu**, *Métriques de Carnot–Carathéodory et quasiisométries des espaces symétriques de rang un*, Ann. of Math. 129 (1989) 1–60 MR

[49] **J G Ratcliffe**, *Foundations of hyperbolic manifolds*, 2nd edition, Graduate Texts in Mathematics 149, Springer (2006) MR

[50] **M Romney**, *Conformal Grushin spaces*, Conform. Geom. Dyn. 20 (2016) 97–115 MR

[51] **X Rong**, *On the fundamental groups of manifolds of positive sectional curvature*, Ann. of Math. 143 (1996) 397–411 MR

[52] **E A Ruh**, *Almost flat manifolds*, J. Differential Geom. 17 (1982) 1–14 MR

[53] **S Semmes**, *Bi-Lipschitz mappings and strong A_∞ weights*, Ann. Acad. Sci. Fenn. Ser. A I Math. 18 (1993) 211–248 MR

[54] **S Semmes**, *On the nonexistence of bi-Lipschitz parameterizations and geometric problems about A_∞–weights*, Rev. Mat. Iberoamericana 12 (1996) 337–410 MR

[55] **J Seo**, *A characterization of bi-Lipschitz embeddable metric spaces in terms of local bi-Lipschitz embeddability*, Math. Res. Lett. 18 (2011) 1179–1202 MR

[56] **C Sormani**, *How Riemannian manifolds converge*, from "Metric and differential geometry" (X Dai, X Rong, editors), Progr. Math. 297, Springer (2012) 91–117 MR

[57] **T Tao**, *Hilbert's fifth problem and related topics*, Graduate Studies in Mathematics 153, Amer. Math. Soc., Providence, RI (2014) MR

[58] **W P Thurston**, *The geometry and topology of three-manifolds*, lecture notes, Princeton University (1979) Available at `http://msri.org/publications/books/gt3m`

[59] **W P Thurston**, *Three-dimensional geometry and topology, I*, Princeton Mathematical Series 35, Princeton Univ. Press (1997) MR

[60] **T Toro**, *Surfaces with generalized second fundamental form in L^2 are Lipschitz manifolds*, J. Differential Geom. 39 (1994) 65–101 MR

[61] **T Toro**, *Geometric conditions and existence of bi-Lipschitz parameterizations*, Duke Math. J. 77 (1995) 193–227 MR

[62] **N T Varopoulos**, **L Saloff-Coste**, **T Coulhon**, *Analysis and geometry on groups*, Cambridge Tracts in Mathematics 100, Cambridge Univ. Press (1992) MR

Department of Mathematics, UCLA
Los Angeles, CA, United States

syerikss@math.ucla.edu

Proposed: Yasha Eliashberg
Seconded: Dmitri Burago, Bruce Kleiner

Received: 21 October 2015
Revised: 3 July 2017