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 $\Gamma$ 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; 53B20, 53C21, 20H15

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 [24]. Recall the definition of a Bi-Lipschitz map.

Definition 1–1 Let $L > 0$. A map $f : X \to Y$ between metric spaces is called a $(L)$-bi-Lipschitz function (also embedding) if

$$\frac{1}{L}d(a, b) \leq d(f(a), f(b)) \leq Ld(a, b)$$

for all $a, b \in X$. The constant $L$ in the definition, while non-unique, is referred to as the distortion of $f$.

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.
We are generally not concerned with optimal distortions of a bi-Lipschitz map, and for a given map $f$ any constant $L$ such that the above definition is satisfied, is called the distortion of such a mapping. When $Y = \mathbb{R}^N$ a necessary condition for embeddability is that $(X, d)$ is metric doubling (see Definition 3–6). As a partial converse Assouad proved that any doubling metric space admits an $\alpha$−bi-Hölder embedding into Euclidean space for any $0 < \alpha < 1$ [2] (also see [41]). But due to counter-examples by Pansu [44] (as observed by Semmes [51]) and Laakso [32] 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 in this paper 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 control the distortion we will need to resort to very different techniques. Also, some of our embedding results will apply to complete non-compact manifolds.

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. Examples of approximating graphs by manifolds, as well as considering the hyperbolic plane, show that indeed these assumptions are necessary. Our results require the somewhat stronger assumption of bounding the sectional curvature in absolute value. Thus we are lead to the following theorem.

**Theorem 1–2** Every bounded subset $A$ with $\text{Diam}(A) \leq D$ in an $n$-dimensional complete Riemannian manifold $M^n$ with sectional curvature $|K| \leq 1$ admits a bi-Lipschitz embedding $f: A \to \mathbb{R}^N$ with distortion less than $C(D, n)$ and dimension of the image $N \leq N(D, n)$. 
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 [20]). While somewhat unexpected, we are also able to prove a version of the previous theorem for orbifolds.

**Theorem 1–3** Every bounded subset $A$ with $\text{Diam}(A) \leq D$ in an $n$-dimensional complete Riemannian orbifold $O^n$ with sectional curvature $|K| \leq 1$ admits a bi-Lipschitz embedding $f : A \to \mathbb{R}^N$ with distortion less than $C(D,n)$ and dimension of the image $N \leq N(D,n)$.

In particular we have the following non-trivial result.

**Theorem 1–4** Every complete flat and elliptic orbifold $O$ of dimension $n$ admits a bi-Lipschitz mapping $f : O \to \mathbb{R}^N$ with distortion less than $D(n)$ and dimension of the image $N \leq N(n)$. The constants $D(n)$ and $N(n)$ depend only on the dimension.

**Remark:** The bounds for the distortion and dimension can be made explicit for the case of flat and elliptic orbifolds, but for a general bounded curvature orbifolds we can not give explicit bounds. This is because the proof of Fukaya’s fibration theorem in [10] uses compactness in a few steps and thus does not give an explicit bound on the parameters $\delta(n), \rho(n)$. Some related work in [19] when combined with techniques from [10] 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 in [48, 8]). 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 flat spaces to the context of quotients of nilpotent Lie groups. They correspond to model spaces for collapsing phenomena. The precise definition is in Definition 3–19.

**Theorem 1–5** For sufficiently small $\epsilon(n)$, every $\epsilon(n)$-quasiflat $n$-dimensional manifold $M$ admits a bi-Lipschitz embedding into Euclidean space with distortion and dimension depending only on $n$. Further every locally flat Riemannian vector 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 non-negative sectional curvature. Yet they admit global bi-Lipschitz embeddings. Similarly, we are able to prove a version of this theorem for orbifolds \(^2\).

Finally we remark that the results above can trivially be generalized to Gromov-Hausdorff limits of bounded curvature orbifolds, which themselves may not be orbifolds. We remark, that the Gromov-Hausdorff limits of bounded curvature Riemannian manifolds have been intrinsically described using weak Alexandrov-type curvature bounds by Nikolaev in [43].

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 [34, 42, 52]. Constructing global embeddings may be difficult, but by assuming a diameter bound and lower curvature bound we can reduce it to embedding a certain fine \(\epsilon\)-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 [16] (see also [10]). 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 1–6** (Fukaya, [10, 15, 47]) Let \((M, g)\) be a complete Riemannian manifold of dimension \(n\) and sectional curvature \(|K| \leq 1\). For every \(\epsilon > 0\) there exists a universal \(\rho(n, \epsilon) > 0\) and \(\delta(n) > 0\) such that for any point \(p \in M\) if \(\text{inj}_p(M) < \delta(n)\) there exists a metric \(g'\) on the ball \(B_p(\rho(n))\) and a complete Riemannian manifold \(M'\) with the following properties.

- \(||g - g'||_g < \epsilon\)
- \((B_p(5\rho(n)), g')\) is isometric\(^3\) to a subset of a complete Riemannian manifold \(M'\)

\(^2\)For terminology related to orbifolds we recommend to consult for example [54, 45, 13, 31]. Some of this terminology is covered in the appendix to this paper.

\(^3\)By isometry we mean Riemannian isometry. However, what we really use is that the ball \(B_p(\rho(n), g')\) will have a distance preserving mapping.
\begin{itemize}
  \item $M'$ is either a $\epsilon(n)$-quasi flat manifold or a locally flat orbivector bundle over a $\epsilon(n)$-quasiflat manifold.
\end{itemize}

We remark that this theorem isn’t stated as such in the references. Our terminology is also slightly different and is introduced in Definitions 3–19 and 3–27. In [15] a topological version is stated and in [10] the main results concern a fibration structure. The statement of this theorem is contained in a similar form in the appendix of [10, Appendix 1], and the proof of our version can be derived from it. The statement could also be strengthened a bit, but for us this version is sufficient. The main difference is that in [10] the role of $M'$ is played by the normal bundle of a singular orbit (to the $N$-structure), $\nu(O_q)$, with a natural metric. Here we have merely explicated that such a normal bundle can be described as a $\epsilon(n)$-quasiflat manifold. In the appendix of this paper we outline a proof of the analogous statement for orbifolds. We also use the results in [47] 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”).

We next give a rough description of our approach to embedding the vector bundles $M$ that arise. We first use an approximation argument from Lemmas 4–6 and Lemma 3–8 to reduce the problem to embedding locally flat Riemannian vector bundles over flat manifolds. Next the flat manifold $M$ is decomposed into pieces which are patched together using Lipschitz extension theorems. These decompositions are similar in spirit but were developed independently of arguments by Seo and Romney in [46, 52]. Assume the diameter of the base of the vector bundle is unity. Denote by $S$ the zero-section of the vector bundle with its induced metric, and for $i \geq 1$ let $T_i$ be the points at distance roughly $2^i$ from the zero section. Additionally define $T_0$ to be the points $x \in O$ with $d(x, S) \leq 4$. By using the radial function $r(x) = d(x, S)$ we are able to reduce the problem to embedding each $T_i$ individually.

To embed each $T_i$ we use again that at small but definite scales they possess a description in terms of simpler spaces. The difficult aspect of the proof here is showing that these local descriptions exist at a scale comparable to the diameter of $T_i$, which is necessary for applying a doubling Lemma 3–7 \footnote{If $T'$ 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 1–10, 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.}. Further, the nature in which the scaling produces a simpler space is somewhat delicate. For $i = 0$, the resulting spaces are of the same dimension, but have “fewer collapsing scales”, i.e. 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, one can assume by induction (either on dimension, or simplicity) that each local model can be embedded. To construct the local models on the sets $T_i$, we use a quantitative version of the local group arguments applied in collapsing theory [16]. Many of these arguments were inspired by proofs in [45, 55].

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 [10, Appendix 1], and [13], but we provide a rough outline in the Appendix to this paper.

**Theorem 1–7** Let $(O, g)$ be a complete Riemannian orbifold of dimension $n$ and sectional curvature $|K| \leq 1$. For every $\epsilon > 0$ there exists a universal $\rho(n, \epsilon) > 0$ and $\delta(n) > 0$ such that for any point $p \in O$ if $\text{Vol}_p(O) < \delta(n)$ there exists a metric $g'$ on the ball $B_p(\rho(n))$ and a complete Riemannian orbifold $O'$ with the following properties.

1. $||g - g'||_g < \epsilon$

2. $(B_p(5\rho(n)), g')$ is isometric to a subset of $O'$

3. $O'$ is either a $\epsilon(n)$-quasiflat orbifold or a locally flat Riemannian vector bundle over a $\epsilon(n)$-quasiflat orbifold $S$.

The methods to embed bounded curvature manifolds 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 [5, 27, 38]. 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 (e.g., 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 [40].
For non-finite metric spaces the research is somewhat more limited. We mention only some of contexts within which embedding problems have been studied: subsets of $\mathbb{R}^n$ with intrinsic metrics by Tatiana Toro [57, 56], ultrametric spaces [35, 36], certain weighted Euclidean spaces by David and Semmes [50, 51]; and an embedding result for certain geodesic doubling spaces with bicombings [34] by Lang and Plaut. Some of these works concern bi-Lipschitz parametrizations, which is a stronger problem. Concurrently with this paper Matthew Romney [46] has considered related embedding problems on Grushin type spaces and applied methods from [52] as well as the current paper. The different abstract question of existence of any bi-Lipschitz parametrization for non-smooth manifolds has been discussed in [26]. 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 closest results to this work are from Naor and Khot, who construct embeddings for flat tori and estimate the worst possible distortion [30]. Improvements to their results were obtained in [23]. In comparison, our work provides embeddings for any compact or non-compact complete flat orbifold. Our bounds are arguably weaker due to the higher generality. Also noteworthy is the paper of Bonk and Lang [4], where they prove a bi-Lipschitz result for Alexandrov-surfaces with bounded integral curvature. Related questions on bi-Lipschitz embeddability have been discussed in [1], where one considers 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 [44] (as observed by Semmes [51]) and Laakso [32]. These examples are related to a large class of examples that are covered by Lipschitz differentiation theory, which was initially developed by Cheeger [9].

### 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 [6, 58]. The conjecture is also interesting within the context of Riemannian manifolds.

**Conjecture 1–8** Every bounded subset $A$ with $\text{Diam}(A) \leq D$ in an $n$-dimensional complete Alexandrov space $X^n$ with curvature $K \geq -1$ admits a bi-Lipschitz em-
bedding \( f : X^n \rightarrow \mathbb{R}^N \) with distortion less than \( C(D,n) \) and dimension of the image \( N \leq N(D,n) \).

If we assume in addition that the space \( X^n \) is volume non-collapsed the theorem follows from an argument in [49]. One of the main obstacles in proving such a theorem 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 [28]. Perelman claimed a proof without publishing it and thus it is appropriately referred to as a conjecture. See the discussion preceding Lemma 3–8 for the definition of the Gromov-Hausdorff-distance \( d_{GH} \).

**Conjecture 1–9** Let \( X^n \) be a fixed \( n \)-dimensional Alexandrov space with curvature \( K \geq -1 \). Then there exists an \( 0 < \epsilon_0 \) (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 \( f : X^n \rightarrow Y^n \) with distortion at most \( L \) (which may depend on \( X^n \) ).

Both of these conjectures seem hard. But mostly for lack of counter examples, 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 1–10** Let \( A \) be a subset with \( \text{Diam}(A) \leq D \) in an \( n \)-dimensional complete Riemannian manifold \( (M^n, g) \) with curvature \( \text{Ric}(g) \geq -(n - 1)g \). Assume \( \text{Vol}(B_1(p)) > v \) for some \( p \in A \). Then there exists a bi-Lipschitz embedding \( f : A \rightarrow \mathbb{R}^N \) with distortion less than \( C(D,n,v) \) and dimension of the image \( N \leq N(D,n,v) \).

Finally we remark on the problem of optimal bounds for our embeddings in the Theorem for flat and elliptic orbifolds 1–4. As remarked we obtain a bound of \( D(n) \) of the order \( O(e^{Cn \ln(n)}) \). 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 improve the super-exponential upper bound from [30] and obtain \( O(n^{\sqrt{\ln(n)}}) \) distortion for \( n \)-dimensional flat tori [23]. 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 [1] because certain finite approximations to Wasserstein spaces arise as quotients of permutation groups.
1.5 Outline

In the next section 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. Following this 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 [34, 42]. Finally in the fourth section we give a 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 also thank the referee for many detailed comments that helped us improve the paper tremendously. This research was supported by a NSF grant DGE 1342536.

2 Embedding some key examples

We will exhibit nearly explicit embeddings for some concrete examples. The examples were chosen to reflect some of the general techniques used to prove the main theorems. We will skip over some details that will be contained in proofs later.

2.1 Lens spaces

Take two distinct co-prime 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 $t(z_1, z_2) \rightarrow (e^{2\pi t/p}z_1, e^{2\pi tq/p}z_2)$ for $t \in \mathbb{Z}_p$. The action is free and isometric, wherefore $L(p, q)$ inherits a constant curvature metric. For an insightful discussion of these spaces see [55]. We ask for a bi-Lipschitz embedding $f : L(p, q) \rightarrow \mathbb{R}^N$ for some $N$ 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
\[ U_1 = \{ |(z_1, z_2)| \in L(p, q): |z_2| < \sqrt{3}|z_1| \} \text{ and } U_2 = \{ |(z_1, z_2)| \in L(p, q): |z_1| < \sqrt{3}|z_2| \}. \]
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 co-ordinates \((\alpha, \theta, \phi)\) by \((\alpha, \theta, \phi) \to (\cos(\alpha) e^{i\theta}, \sin(\alpha) e^{i\phi})\)
for \( \alpha \in [0, \pi/3) \) and \((\theta, \phi) \in S^1 \times S^1 \). The induced metric, becomes

\[
g = |\sin(\alpha)|^2 |d\alpha|^2 + |\cos(\alpha)|^2 |d\theta|^2 + |\cos(\alpha)|^2 |d\alpha|^2 + |\sin(\alpha)|^2 |d\phi|^2
\]
\[
= |\cos(\alpha)|^2 |d\theta|^2 + |d\alpha|^2 + |\sin(\alpha)|^2 |d\phi|^2.
\]

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

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

Note that the norms induced by \( g \) and \( g_f \) differ by a factor at most 4. 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
push-forward of the original metric \( g \). In fact \((U_1, g_f')\) is isometric to a subset of the
holonomy bundle, given by the \( \mathbb{R}^2 \)-vector bundle on \( \mathbb{R}^2 \frac{S^1}{p} \) (the circle of radius \( 2\pi/p \)),
with holonomy generated by a \( \frac{2\pi}{p} \)-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 \( i_1: (U_1, g) \to (U_1, g_f') \) is bi-Lipschitz with factor 4. Similar analysis
can be performed for \( U_2 \), where the bundle has holonomy generated by a \( 2\pi s/p \),
where \( sq \equiv 1(\text{mod}(p)) \). One is thus led to consider flat vector bundles with non-trivial
holonomy. These spaces will be discussed in the next subsection where we indicate
how to construct of bi-Lipschitz maps \( F_j: (U_j, g_f') \to \mathbb{R}^3 \).

Finally 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 especially the reference [34, Theorem 3.2],
where a similar embedding is considered. We remark, that there are many ways in
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.
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 $E^3_\theta$, which is a flat $\mathbb{R}^2$-bundle over $S^1$ with holonomy $\theta$. One has $E^3_\theta = \mathbb{R}^3/\mathbb{Z}$, where the $\mathbb{Z}$-action is defined as $t(x,z) = (x+2\pi t, e^{i\theta} z)$ for $t \in \mathbb{Z}$. Let $r$ be the distance function to the zero section $r(x,z) = |z|$. 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 $f_j : T_j \to \mathbb{R}^n$. Then we can collect the maps $f_j$ with disjoint domains by defining four functions $F_s$ for $s = 0, \ldots, 3$ as follows. Consider the sets $D_s = \bigcup_{k=0}^{\infty} T_{s+4k}$, and define $F_s(x) = f_{s+4k}$ on $T_{s+4k}$. The resulting functions will be Lipschitz on their respective domains $D_s$. By an application of McSchane extension Theorem 3–1 we can extend them to $E^3_\theta$. The embedding we consider is $F = (r,F_0,F_1,F_2,F_3)$. By argument similar to the proof of the Theorem 4–1 below this map can be seen to be bi-Lipschitz. We can also control its distortion.

Next we discuss the construction of the embeddings $f_j$. On $T_0$ the injectivity radius is bounded from below. Thus we can find $f_0$ using Lemma 3–7. For $T_j$ when $j \geq 1$ we will first modify the metric. Use co-ordinates $(x,r,\theta) \to (x, re^{i\theta}) \in \mathbb{R}^3$. The metric on $T_i$ (or its lift in $\mathbb{R}^3$) space can be expressed as

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

We change the metric to $g_f = |dx|^2 + |dr|^2 + 2^j|d\theta|^2 = g_1$. 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 $(2^{j-1}, 2^{j+1}) \times T^j_\theta$, where $T^j_\theta$ 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 $(2^{j-1}, 2^{j+1}) \times T^j_\theta$ 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 [30].

We comment briefly on the general case of an arbitrary flat vector bundle. Consider a flat $\mathbb{R}^d$-bundle over $S^1$. The argument remains unchanged for $T_0$, but for $T_j$ when $j \geq 1$ 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.

The key argument in the proof of the main theorems below is indicated in these two examples. First the space is decomposed into sets (“charts”) on which we have in some sense a simpler geometry. These charts can be patched together to give an embedding of the space. In the Lens space the simpler geometry was that of a vector bundle. Further 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 non-collapsed space. 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 Heisenberg Group and curvature of Nilpotent groups

The Heisenberg group may be described as the simply connected Lie group of upper triangular $3 \times 3$-matrices with diagonal entries equal to one.

\[
\mathbb{H} = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \mid a, b, c \in \mathbb{R} \right\}
\]

Define a lattice

\[
\Gamma = \left\{ \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix} \mid a, b, c \in \mathbb{Z} \right\} \subset \mathbb{H}
\]

and a basis for the Lie algebra

\[
X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, Y = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, Z = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\]

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 $|Z| = \epsilon$. A direct computation shows that the sectional curvatures
of this space lie in
\[ \left[ -\frac{3\epsilon^2}{4}, \frac{\epsilon^2}{4} \right]. \]

As \( \epsilon \to 0 \) the spaces \( \mathbb{H}_\epsilon \) Gromov-Hausdorff converge to a torus \( T^2 \). The Gromov-Hausdorff approximation is given by the \( S^1 \)-fibration map
\[
\pi: \mathbb{H}_\epsilon \to ([a], [b]) \in \frac{1}{2\pi} S^1 \times \frac{1}{2\pi} S^1 = T^2,
\]
where \([a]\) is the fractional part of \( a \in \mathbb{R} \). By \( \frac{1}{2\pi} S^1 \) we mean the circle of unit length. This map is easily seen to be well-defined on \( \mathbb{H}_\epsilon \). An embedding for \( \mathbb{H}_\epsilon \) can now be visualized using a doubling Lemma 3–7. Consider a small ball \( B_p(\delta) \subset T^2 \). Since \( T^2 \) is locally simply connected, \( \pi^{-1}(B_p(\delta)) \) 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 bi-Lipschitz map with small distortion. On the other hand, the space \( B_p(\delta) \times S^1 \) is is easy to embed since it splits as a product of two simple spaces. The result can then be deduced from the doubling Lemma 3–7.

### 3 Frequently used results

As a general remark. Below there are a number of constants determined from specific theorems. In order to keep track of these different constants and functions throughout the paper, we fix them. The special constants are: \( C'(n) \) from Lemma 3–18 and Definition 3–19, \( c(n) \) is the maximum of the constants necessary for Lemmas 3–29, 3–18 and 3–16, \( \epsilon(n) \) is fixed via Lemmas 3–25, 4–5 and 4–6, \( \delta(n) \) from Lemma 5–5, and \( C(n) \) from Lemma 5–4. Also, we will choose \( c(n) \) increasing and \( \epsilon(n) \) decreasing. Their value will be considered fixed throughout the paper, although the precise value is left implicit. In general, we will use the convention that a constant \( M(a_1, a_2, \ldots, a_n) \) is a quantity depending on \( a_1, \ldots, a_n \) only.

#### 3.1 Embedding lemmas

Consider any metric space \( X \). Throughout this paper we will denote open balls by \( B_x(r) = \{ y \in X | d(y, x) < r \} \). For any subset \( A \subset X \) we call \( A(\delta) = \{ x \in A | d(x, A) < \delta \} \) its \( \delta \)-tubular neighborhood. The metric spaces in this paper will be complete 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, [25]) Let X be a metric space and \( A \subset X \). Then any \( L \)-Lipschitz function \( f: A \to \mathbb{R}^n \) has an \( \sqrt{nL} \)-Lipschitz extension \( \tilde{f}: X \to \mathbb{R}^n \) s.t \( \tilde{f}|_A = f \).

**Remark:** McShane 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 Kirzbraun’s theorem [33]. For most of the paper Kirzbraun 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 \( f: X \to \mathbb{R}^N \) and \( g: Y \to \mathbb{R}^M \) with distortion \( L \), then \( X \times Y \) admits a bi-Lipschitz embedding \((f, g): X \times Y \to \mathbb{R}^{N+M}\) with distortion at most \( \sqrt{2}L \).

The product space could also be equipped with different bi-Lipschitz equivalent metrics, but the one in this lemma arises in the applications. In other cases the structure admits a conical splitting. Thus we define cones over metric spaces.

**Definition 3–3** Let \( X \) be a compact metric space with \( \text{diam}(X) \leq \pi \). 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 this lemma are a priori finite.

**Lemma 3–4** Let \( \Gamma \) be a group acting properly discontinuously and by isometries on \( \mathbb{R}^n \) which fixes \( 0 \in \mathbb{R}^n \). Then the group \( \Gamma \) acts discontinuously 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 $\text{diam}(X) \leq \pi$, and $Y = C(X)$. Further assume that $X$ 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:** First consider $Y$. Let $r$ be the radial function. Assume that $f : X \to \mathbb{R}^n$ is bi-Lipschitz with distortion $L$ and $0 \in \text{Im}(f)$. Define $g(r, x) = (Lr, rf(x))$ on $Y$. If $(s, x)$ and $(t, y)$ are points in $Y$ and $s \leq t$, then

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

Note that $|f(y)| \leq L\pi$ (because $\text{diam}(X) \leq \pi$). Further

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

Using $s|f(x) - f(y)| \leq Lsd(x, y)$ we get

$$|g(s, x) - g(t, y)| \leq (L + L + 2L/\pi)d((s, x), (t, y)) \leq 10Ld((s, x), (t, y)).$$

Next we derive the necessary lower bound. Take any $(s, x)$ and $(t, y)$ are points in $Y$ and $s \leq t$. Either $|s - t| \geq \frac{1}{20L^2}d((s, x), (t, y))$ or not. In the first case

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

In the latter case $|s - t| \leq \frac{1}{20L^2}d((s, x), (t, y))$, and

$$|g(s, x) - g(t, y)| \geq |g(t, x) - g(t, y)| - |s - t||f(x)| \geq \frac{1}{L}d((t, x), (t, y)) - \frac{\pi}{20L}d((s, x), (t, y)) \geq \frac{1}{2L}d((s, x), (t, y)).$$

These estimates complete the proof.

\[\square\]

**Remark:** Similar constructions would also apply for a spherical suspension of a metric space (see [58]) but we do not need that result in this paper.

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”.
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 $p_1, \ldots, p_D$ such that

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

We emphasize that it is necessary for a metric space to be doubling in order to possessing a bi-Lipschitz embedding into $\mathbb{R}^N$.

Lemma 3–7  (Doubling Lemma) Assume $0 < r < R, N$ are given 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 a $1$-bi-Lipschitz embedding $f_q : B_q(r) \to \mathbb{R}^N$, then there is a $L$-bi-Lipschitz embedding of $B_p(R)$ to $\mathbb{R}^M$. Further we can bound the distortion of the bi-Lipschitz embedding by $L \leq L(N,l,D,R/r)$ and the target dimension by $M \leq M(N,l,D,R/r)$.

Proof: Let $q_i$ be an $r/8$-net in $B_p(R)$ and $f_i : B_{q_i}(r) \to \mathbb{R}^N$ the bi-Lipschitz mappings assumed. The size of the net is bounded by a function of $D$ and $R/r$. Translate the mappings so that $f_i(q_i) = 0$, and scale so that the Lipschitz constant is 1. Further let $F : x \to (d(x,q_i))_{i \in I}$ be a distance embedding from the net. The size of the net is $H \leq D \log^2 (R/r) + 4$. 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 within distance at most $4r$ of $q_i$ is at most $D^4$. Thus a standard graph coloring argument such as in [24] will furnish the partition $J_k$ with $K$ sets with $K \leq D^4$.

For each $J_k$ construct a map $g_k : \bigcup_{q_i \in J_k} B_{q_i}(r) \to \mathbb{R}^N$ 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 $i \neq j$ and $q_i, q_j \in J_k$. Then $d(a,b) \geq 2r$ and $|f(a) - f(b)| \leq |f(a) - f(q_i)| + |f(b) - f(q_j)| \leq 2r \leq d(a,b)$. Extend $g_k$ 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

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

Clearly this is $2\sqrt{N}D^4 + 2D \log^2 (R/r) / 2 + 2$-Lipschitz. For the lower Lipschitz bound take arbitrary $a, b \in B_p(R)$. First assume $d(a,b) < r/2$. Then there is a $q_i$ such that $d(a,q_i) < r/8$ and $d(b,q_i) < r$. So both belong to the same ball $B_{q_i}(r)$ and for the $k$ such that $q_i \in J_k$ we have $|g_k(a) - g_k(b)| = |f_i(a) - f_i(b)| \geq 1/2d(a,b)$. On
the other hand assume \(d(a, b) \geq r/2\), then there is an index \(q_i\) such that \(d(a, q_i) < r/8 \leq d(a, b)/4\) and thus \(d(b, q_i) \geq d(a, b) - d(a, b)/4 \geq 3/4d(a, b)\). Thus \(|d(a, q_i) - d(b, q_i)| > d(a, b)/2\), and we get \(|F(a) - F(b)| \geq d(a, b)/2\).

\[
\square
\]

For certain spaces it is easier to construct embeddings for nearby spaces, and thus we need a result concerning Gromov-Hausdorff distance. Recall, that a map between two metric spaces \(f: X \to Y\) is an \(\epsilon\)-approximation, if \(\text{Im}(f)\) is \(\epsilon\)-dense in \(Y\) and

\[
|d(a, b) - d(f(a), f(b))| \leq \epsilon.
\]

A map \(f: X \to Y\) is called \(\epsilon\)-dense if for every \(y \in Y\)

there is a \(x \in X\), such that \(d(y, f(x)) \leq \epsilon\). We say that

\[
d_{GH}(X, Y) = \inf\{\text{there are } \epsilon\text{-approximations } f: X \to Y \text{ and } g: Y \to X\}.
\]

Lemma 3–8 (Gromov-Hausdorff-lemma) Let \(X\) be a doubling metric space with doubling constant \(D\) and let \(\epsilon, \epsilon', l > 0\) be constants. Further assume that \(Y\) is a metric space admitting a bi-Lipschitz map \(h: Y \to \mathbb{R}^m\) with distortion \(l\). Assume \(d_{GH}(X, Y) \leq \epsilon'\) and that for every \(p \in X\) there is a bi-Lipschitz embedding \(f_p: B_p(\epsilon) \to \mathbb{R}^N\), then there is a \(L\)-bi-Lipschitz map \(f: X \to \mathbb{R}^M\), with distortion \(L \leq L(D, l, m, \epsilon'/\epsilon, N)\) and target dimension \(M \leq M(D, l, m, \epsilon/\epsilon', N)\).

**Proof:** Fix \(M = \epsilon'/\epsilon\). Let \(g: X \to Y\) be a Gromov-Hausdorff approximation. By Lemma 3–7 we can first construct for every \(p \in X\) an embedding \(f_p: B_p(100l\epsilon')\). This distortion will depend on \(l\) and \(M\). Similar to the proof in Lemma 3–7 we can construct a map \(F: X \to \mathbb{R}^K\) and \(F = (f_1, \ldots, f_n)\) such that for every \(x, y \in X\) with \(d(x, y) \leq 10l\epsilon\) we have a \(p \in X\) such that \(x, y \in B_p(100l\epsilon)\) and a \(f_i\) with \(f_i|_{B_p(100l\epsilon)} = f_p\).

Next take a \(\delta\)-net \(N_\delta\) for \(\delta = 4\epsilon\). The map \(h \circ g|_{N_\delta}: N_\delta \to \mathbb{R}^n\) is an \(2l\)-bi-Lipschitz map because \(g\) is \(2\)-Lipschitz on a \(4\epsilon\)-net. This map is \(1\)-Lipschitz Extend this to a map \(H: X \to \mathbb{R}^n\) which is \(2l\)-bi-Lipschitz on \(N_\delta\), and \(2l\)-Lipschitz on \(X\). The proof can now be completed by similar estimates as in Lemma 3–7. The embedding is given by \(G(x) = (H(x), F(x))\).

\[
\square
\]

The main source of Gromov-Hausdorff approximants will be via quotients so we state

the following Lemma.

Lemma 3–9 Assume \((X, d_X), (Y, d_Y)\) are metric spaces and \(\delta > 0\). Let \(\pi: X \to Y\) be a quotient map with \(d_Y(a, b) = d_X(\pi^{-1}(a), \pi^{-1}(b))\) for all \(a, b \in Y\) and \(\text{diam}(\pi^{-1}(a)) \leq \delta\). Then \(d_{GH}(X, Y) \leq 2\delta\).
3.2 Group actions and quotients

Assume that $X$ is a proper metric space, i.e. that $B_p(r)$ is precompact for every $p \in X$ and $r > 0$. Further assume that $\Gamma$ is a discrete Lie group acting on $X$ by isometries. For any group we will denote 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 $\Gamma_p = \{ \gamma \in \Gamma | \gamma p = p \}$. We say that the action is properly discontinuous if for every $p \in X$ there exists an $r > 0$, such that $\{ \gamma \in \Gamma | \gamma (B_p(r)) \cap B_p(r) \neq \emptyset \} = \Gamma_p$ and $\Gamma_p$ is a finite group. We will study quotient spaces which will be denoted by $X/\Gamma$. 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 $X/\Gamma$ 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 $X/\Gamma$ a metric space. For more terminology see [45, 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 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 $\gamma_1, \ldots, \gamma_n$ are generators of the group, if $G = \langle \gamma_1, \ldots, \gamma_n \rangle$.

**Lemma 3–10** Let $X$ be a metric space and $\Gamma$ a discrete Lie group of isometries acting properly discontinuously on $X$. Take a point $p \in X$. Let $\Gamma_p(r) = \{ g \in \Gamma | d(gp,p) \leq 8r \}$ be the subgroup generated by elements such that $d(p, gp) \leq 8r$. Then $B_{\{p\}}(r) \subset Y$ 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, such as $[x]'$, and corresponding objects in $Y$ without primes. Elements of $X/\Gamma_p$ can be represented by a $\Gamma_p(r)$-cosets $[x]'$, for $x \in X$. The action is still properly discontinuous and the distance function on the orbit space is given by $d'([x]',[y]') = \inf_{\gamma,\gamma' \in \Gamma_p} d(\gamma x, \gamma' y)$. Since $\Gamma_p(r) \subset \Gamma$, we have a continuous map $F: (Y',d') \to (Y,d)$ which sends a $\Gamma_p(r)$-coset to the $\Gamma$-coset that contains it, i.e. $[x]' \to [x]$. We will next show that $F$ maps $B_{\{p\}}(r) = \{ y' \in Y' | d'([p]'), y' < r \}$ isometrically onto $B_{\{p\}}(r)$ from which the conclusion follows.

Let $[a]', [b]' \in B_{\{p\}}(r) \subset X/\Gamma(r)$ and as before $F([a]') = [a], F([b]') = [b]$. Clearly $d([a]',[b]') \geq d([a],[b])$, as the $\Gamma$-cosets are super-sets of $\Gamma_p(r)$-cosets. We can choose the representatives of the cosets $a,b$ such that $a,b \in B_p(r) \subset X$. By 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 $d(\gamma_\epsilon p, p) \leq d(\gamma_\epsilon p, \gamma_\epsilon b) + d(b, p) \leq r + r + 4r \leq 6r$. Thus $\gamma_\epsilon \in \Gamma(r)$ and therefore $d([a]', [b]') \leq d(a, \gamma_\epsilon b) \leq d([a], [b]) + \epsilon$. The result follows since $\epsilon$ is arbitrary.

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

**Definition 3–11** Let $\Gamma$ be a discrete Lie group acting properly discontinuously on $X$. We denote by $\Gamma_p(r) = \{ g \in \Gamma : d(gp, p) \leq 8r \}$, 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 we will recall Grove’s and Karcher’s center of mass technique. See [22, 29] for a more detailed discussion and proofs.

**Lemma 3–12** (Center of mass Lemma) Let $M$ be a complete Riemannian manifold with sectional curvature bounded by $|K| \leq \kappa$ and $p \in M$ a point. Further let $\mu$ be a probability mass supported on a ball $B_p(r)$ with $r \leq \min \left\{ \frac{\pi}{2\sqrt{\kappa}}, \text{inj}(M, p) \right\}$. 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 $C_\mu$ is called the center of mass. The center of mass is invariant under isometric transformations of the measure. If $f : M \to M$ is an isometry, and $f_* (\mu)$ is the push-forward measure of $\mu$, 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 [45, Lemma 5.9].
Corollary 3–13 (Fixed point Lemma) Let $M$ satisfy the same assumptions as in Lemma 3–12. Let a compact Lie group $G$ act by isometries on $M$. If
\[ \text{diam}(G_p) < \min \left\{ \frac{\pi}{2\sqrt{\kappa}}, \text{inj}(M, p) \right\}, \]
then there exists a fixed point $q$ of the $G$-action such that $d(q, p) \leq \text{diam}(G_p)$.

In fact one can use the Haar measure $\mu$ on $G$, and consider the push-forward measure $f_\ast(\mu)$ by the map $f : G \to M$ which is given by $f(g) = gp$. The center of mass $C_{f_\ast(\mu)}$ 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 3–25 below. Assume $\Gamma$ is a group. Then we call $| \cdot | : \Gamma \to \mathbb{R}$ 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 3–14 Let $\Gamma$ 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
\[ |g|_p = d(gp, p). \]
In particular if $g \in \text{Isom}(X)$ is an isometry we denote $|g|_p = d(gp, p)$.

Lemma 3–15 (Local Group Argument) Fix an arbitrary scale parameter $l > 0$. Let $\Gamma$ 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 $|H| < \infty$. Assume $\Lambda_0 < \Lambda$ is a normal subgroup in $\Gamma$ 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 $H_0 < H$. 
Proof: Define $S$ to consist of a single representative $s$ for each left cosets of $\Lambda$ in $\Gamma$ with $|s| < l$. Clearly $|S| \leq |H|$. 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 $\Gamma$. None of these words represent elements of $\Lambda \setminus \Lambda_0$ since for any $w \in W$ we have $|w| \leq (|H| + 1)l$. Take any word $w'$ with length longer than $|H| + 1$. Consider the subwords $w'_1, \ldots, w'_{|H|}$ formed by the first $1, \ldots, |H| + 1$ letters of $w'$, respectively. Because there are only $|H|$ cosets, two of them will be in the same $\Lambda$-coset in $\Gamma$. Thus assume $w_i$ and $w_j$ with $j < i$ belong to the same coset. Let $w' = w_jv_j$, where $v_j$ is the product of the remaining letters. We have $w_jw_j^{-1} \in \Lambda$. But also $|w_jw_j^{-1}| \leq (2|H| + 1)l$ so $w_jw_j^{-1} \in \Lambda_0$. Thus we have for the $\Lambda_0$-cosets $[w_i]_{\Lambda_0} = [w_j]_{\Lambda_0}$, and thus $[w_i]_{\Lambda_0} = [w_jw_j^{-1}]_{\Lambda_0} = [w_j]_{\Lambda_0}v_j = [w_j]_{\Lambda_0}$. The word on the right is shorter. 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 $\Lambda_0$ to an element in $W$. In particular products of cosets corresponding to elements in $W$ are 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 $S$.

Define the group by generation: $\Gamma_0 = \langle g \in \Gamma ||g| \leq l \rangle$. By the previous argument $\Gamma_0$ is a finite extension of $\Lambda_0$ of index at most $|H_0|$. Also $\Lambda_0$ is normal in $\Gamma_0$ since it is normal in $\Gamma$.

3.3 Nilpotent Lie groups, lattices and collapsing theory

Nilpotent geometries occur as model spaces for collapsing phenomena. Thus we start by an estimate relating to the curvature of a general simply connected nilpotent Lie group $N$. For definitions of nilpotent Lie groups as well as the computations involved here see [12]. Associate to the Lie group $N$ its nilpotent Lie algebra $\mathfrak{n}$. For $\mathfrak{n}$ we may give it a triangular basis as follows. Let $\mathfrak{n}_0 = \mathfrak{n}, \mathfrak{n}_1 = [\mathfrak{n}, \mathfrak{n}], \ldots, \mathfrak{n}_{k+1} = [\mathfrak{n}, \mathfrak{n}_k]$. For $k$ large enough $\mathfrak{n}_k = \{0\}$. Further 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 Maurier-Cartan structural constants with respect to this basis, i.e.

$$[X_i, X_j] = \sum_k c_{ij}^k X_k.$$
By construction of the basis $c^k_{ij} \neq 0$ only if $k > j$. Computing the sectional curvatures $K(X_i, K_j)$ using [39] we get for $j \geq i$:

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

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 paper that $|c^k_{ij}| \leq C'(n)$, where $C'(n)$ depends only on the dimension $n$. For comparison and later use we will mention the following theorem.

**Theorem 3–16** (Bieberbach, [7, 45]) If $M$ is a complete (not necessarily compact) flat Riemannian orbifold of dimension $n$, there exists a discrete group of isometries $\Gamma$ 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 $\Gamma$, and on which the action of $\Gamma$ is co-compact. Further there exists a subgroup $\Lambda < \Gamma$ with index $[\Gamma : \Lambda] \leq c(n)$ which acts on $O$ by translations.

Collapsing theory involves working with certain classes of spaces, which we define here. All notions in this chapter are directly defined for orbifolds, and their manifold analogues are obvious. While the following notion is not directly used we give it as a motivational example.

**Definition 3–17** A compact Riemannian orbifold $S$ is called $\epsilon$—almost flat if $|K| \leq 1$ and $\text{diam}(S) \leq \epsilon$.

Throughout this paper $\epsilon$ will be small enough, i.e we will only consider $\epsilon(n)\epsilon$-almost flat $n$-dimensional manifolds for small enough $\epsilon(n)\epsilon$. The choice is dictated by Theorem 3–18 and Lemma 3–25. For such manifolds a detailed structure theory was developed by Gromov and Ruh ([21, 48, 8, 7]). Later Ding observed that with minor modifications the proof as presented in [48, 8] generalizes to orbifolds. We state this structure theory here. For definitions of orbifolds and some of the other other terminology we refer to the appendix and [54, 45, 31]. For any Lie group $G$ we denote by $\text{Aff}(G)$ the Lie group of affine transformations of $G$, which preserve the flat left-invariant connection $\nabla_{\text{can}}$ with respect to which the left-invariant vector fields are parallel. By $\text{Aut}(G)$ we denote the Lie group of its automorphisms. Throughout $N$ will denote a connected nilpotent Lie group.

\footnote{The metric closeness part is not explicit in [13], but a direct consequence of applying [48].}
Theorem 3–18 ([13, 48, 8] Gromov-Ruh Almost Flat Theorem\textsuperscript{6}) For every $n$ there are constants $\varepsilon'(n), c(n)$ such that for any $\varepsilon'(n)$-almost flat orbifold $S$ with metric $g$ there exists a connected nilpotent Lie group $N$, a left-invariant metric $g_N$ on $N$, group $\Gamma < \text{Aff}(N)$ with the following properties.

- $\Gamma$ acts on $(N, g_N)$ isometrically, co-compactly and properly discontinuously. In particular $N/\Gamma$ is a Riemannian orbifold.
- We have the bound $|K| \leq 2^7$ for the sectional curvature of $N$, and $\|\cdot\|_g \leq C'(n)$ for the.
- There is a orbifold diffeomorphism $f : S \to N/\Gamma$, such that $\|f_*g - g_N\|_g < \frac{1}{2}$\textsuperscript{8}.
- There exists a normal subgroup $\Lambda < \Gamma$, such that $[\Gamma : \Lambda] \leq c(n)$ and $\Lambda$ acts on $N$ by left-translations. In other words, by slight abuse of notation, $\Lambda < N$.

We remark that we consistently use $N/\Gamma$ to indicate a quotient under a group action even though here $\Gamma$ is acting on $N$ on the left and the translational subgroup $\Lambda$ acts on $N$ by left multiplication. Again $\|\cdot\|_g$ denotes the operator norm of the various tensors. In the work of [13] and [48] no explicit bounds are derived for the Lie bracket and curvature, but as stated in [10] these follow from the methods and the curvature bounds at the beginning of this section.

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 $\Gamma < \text{Aff}(N)$. Each element $\gamma \in \text{Aff}(N)$ can be represented by $\gamma = (a, A)$, where $a \in N$ and $A \in \text{Aut}(N)$. 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 $\gamma$. If $A$ is reduced to 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 $h : \Gamma \to \text{Aff}(N)$ given by $h(\gamma) = A$.

\textsuperscript{6}We also recommend consulting the preprint version of the paper by Ding which contains a more complete proof, [14].

\textsuperscript{7}For the precise sectional curvature bound it is necessary to also use a remark in ??.

\textsuperscript{8}The bound here could be replaced by any $\delta < 1$, but we fix a choice to reduce the number of parameters in the theorem.
The Lie group $N$ will always be equipped with a left-invariant metric $g$. As in the previous section, if $p$, and $g \in \text{Isom}(N)$, then we denote

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

**Definition 3–19** A Riemannian orbifold $S$ is called a $(\epsilon)$-quasiflat orbifold if $S$ is isometric to a quotient of a connected nilpotent Lie group $(N, g)$ equipped with a left-invariant metric $g$ by a discrete co-compact group of isometries $\Gamma$ 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 $\Gamma$ acts on $N$ isometrically, properly discontinuously and co-compactly.
- The quotient space $N/\Gamma$ is a Riemannian orbifold with $\text{diam}(N/\Gamma) < \epsilon$. The space $N$ is referred to as the (orbifold) universal cover and $\Gamma$ is called the orbifold fundamental group.

We also define a flat orbifold as follows.

**Definition 3–20** If $\Gamma$ is a group acting co-compactly, properly discontinuously and isometrically on $\mathbb{R}^n$, then we call $S = \mathbb{R}^n/\Gamma$ a flat orbifold.

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

By Theorem 3–18 any $\epsilon(n)'$-almost flat manifold is bi-Lipschitz to an $\epsilon(n)$-quasiflat manifold with a slightly different $\epsilon(n)$. Also by Theorem 3–16 every compact flat manifold is also a flat orbifold in the previous sense. We will assume $\epsilon(n)$ small enough, and choose them so that $\epsilon(n) < \epsilon(m)$ for $n > m$. We first prove a result concerning the algebraic structure of $\Gamma$. This can be thought of as a generalization of Bieberbach theorem (see [7]).

**Lemma 3–21** Let $c(n)$ be the constant in Gromov-Ruh Almost Flat theorem 3–18. Let $S = N/\Gamma$ be a $\epsilon(n)$-quasiflat $n$-dimensional orbifold, then $\Gamma$ has a finite index subgroup $\Lambda$ with $[\Gamma : \Lambda] \leq c(n)$, and $\Lambda = N \cap \Gamma$, i.e. $\Lambda$ acts on $N$ by translations.

\textsuperscript{9}The only reason we assume a diameter bound is to attain a commutator estimate below. Thus the diameter bound could be somewhat weakened as well.
Proof: Consider the left invariant metric $g$ of $N$. Since a dilation of a $\epsilon(n)$-quasiflat space is also $\epsilon(n)''$ almost flat for some $\epsilon(n)''$ by Theorem 3–18 in [13] and [8] we can construct a new group $\Gamma'$, and a nilpotent Lie group $N'$ such that $S$ is orbifold-diffeomorphic to $N'/\Gamma'$. The proof also gives that $\Gamma'$ acts by affine transformations on $N'$ and has a finite index normal subgroup $\Lambda' \triangleleft \Gamma'$, with $[\Gamma' : \Lambda'] \leq \epsilon(n)$, and $\Lambda'$ acts on $N'$ by translations.

Both $N'$ and $N$ constitute orbifold universal covers in the sense of [54, Chapter 13]. Thus there is a diffeomorphism $f : N' \to N \times \mathbb{R}^k$ conjugating the action of $\Gamma'$ to that of $\Gamma$. Denote the induced homomorphism by $f_* : \Gamma' \to \Gamma$.

By [3, Theorem 2] there is a finite index $\Lambda < \Gamma$, such that $\Lambda$ acts by translations on $N'$. In particular $\Lambda$ is a co-compact subgroup of $N \times \mathbb{R}^k$. Note that we still need to show that it is possible to choose $\Lambda$ with finite index, since [3] does not bound the index. We wish to translate the index bound for $\Lambda'$ to one for $\Lambda$. Applying [3, Proposition 2] to the nilpotent group $f_*^{-1}(\Lambda')$ we get $f_*^{-1}(\Lambda') < \Lambda$. Thus $[\Gamma : \Lambda] \leq \epsilon(n)$.

□

Fix $p \in S$ a point in a $n$-dimensional $\epsilon(n)$-quasiflat orbifold $S$ for some small enough $\epsilon(n)$. We have $S = N/\Gamma$, where $\Gamma$ corresponds to what is in sometimes called the orbifold fundamental group. 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 3–22 Let connected $N$ be a nilpotent Lie group. A co-compact discrete subgroup $\Lambda < N$ is called a lattice.

Definition 3–23 Let $N$ be a $n$-dimensional connected nilpotent Lie group and $\Lambda < N$ a lattice. Then a set $\{\gamma_1, \ldots, \gamma_n\} \subset \Lambda$ is called a triangular basis for $\Lambda$ if the following properties hold.

- $\gamma_i$ generate $\Lambda$.
- $\gamma_i = e^{X_i}$, and $X_i$ form a basis for $n$.
- For $i < j$ we have $[\gamma_i, \gamma_j] \in \langle \gamma_1, \ldots, \gamma_{i-1} \rangle$.
- For $i < j$ $[X_i, X_j] \in \text{span}(X_1, \ldots, X_{i-1})$.

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 co-ordinates by $(t_1, \ldots, t_n) \to e^{t_1 X_1} \cdots e^{t_n X_n}$. This will be used in the appendix.
Lemma 3–24  Let $S = N/\Gamma$ be an $\epsilon(n)$-quasiflat manifold, and $\Lambda = \Gamma \cap N$. For any $p \in N$ we can generate $\Gamma$ and $\Lambda$ by elements $\gamma$ with $|\gamma|_e \leq 16c(n)e(n)$.

Proof: Consider the action of $\Lambda = \Gamma \cap N$. Then $N/\Lambda$ is a Riemannian manifold. Further by Lemma 3–21 the index of $\Lambda$ in $\Gamma$ is at most $c(n)$, so there is at most $c(n)$-fold map $\pi: N/\Lambda \to N/\Gamma$. Since $\text{diam}(N/\Gamma) < \epsilon(n)$ we have $\text{diam}(N/\Lambda) < 2c(n)e(n)$.

Let $\Gamma' = \{g \in \Gamma ||g|_e \leq 16c(n)\}$, and $\Lambda' = \{g \in \Lambda ||g|_e \leq 16c(n)e(n)\}$. By Lemma 3–10 the space $N/\Gamma'$ is isometric to $N/\Gamma$, and thus $h_\Gamma = \Gamma'$. Similarly we can show $\Lambda = \Lambda'$. This completes the proof.

Proof: The basis is constructed similarly to [8]. Let $\gamma_1$ be minimize $| \cdot |$ in $\Lambda$. Set $G_0 = \{ e \}$ the trivial group. Define subgroup $G_1$ generated by $\gamma_1$. Proceed to choose $\gamma_2 \in \Lambda \setminus G_1$ as the shortest with respect to $| \cdot |$. Define $G_2 = \langle \gamma_1, \gamma_2 \rangle$. We proceed inductively to define $\gamma_1, \ldots, \gamma_n$. A priori this process could last more than $n$ steps, but by an argument below the number of $\gamma_i$ agrees with the dimension. Thus we abuse notation slightly by using the same index.

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

$$(a, b)|_e \leq 16C(n)c(n)e(n) \min \{|a|_e, |b|_e\}.$$ 

This estimate follows either from curvature estimates such as in [8], or directly from the bounds on the structural constants on $N$. Choose $\epsilon(n) < \frac{1}{10^d c(n)C(n)}$, so that $|(a, b)|_e < \min \{|a|_e, |b|_e\}$.
By Lemma 3–24 we can choose generators $\alpha_i$ for the translational subgroup $\Lambda < \Gamma$ such that $|\alpha_i|_e < 16c(n)e(n)$, and generators $\gamma_i$ for $\Gamma$ with $|\gamma_i|_e < 16c(n)e(n)$. Further for any $\gamma_i$ we have then $||\alpha_i, \gamma_i||_e < |\gamma_i|_e$, and thus $[\alpha_i, \gamma_i] \in G_{i-1}$ by construction. Since $\alpha_i$ generate $\Lambda$ 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 $f: (a_1, \ldots, a_n) \rightarrow \gamma_1^{a_1} \ldots \gamma_n^{a_n}$. By induction, and the product relation below, we can show that the map is injective. Further using the commutator relations from above and reordering terms we can show that $f$ is bijective onto $\Lambda$. Let $a = (a_1, \ldots, a_n), b = b_1, \ldots, b_n$. We can define polynomials $p_j$ by the defining relation

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

The polynomials are defined by appropriately applying commutator relations. Also, $p_j$ depends only on $a_i, b_j$ 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 co-compact 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 $X_i = (0, \ldots, 1, \ldots, 0) \in T_0\mathbb{R}^n$ with the 1 in the $i$'th position. By a Theorem of Malcev (see [37] and [10]) we can extend the homomorphism $f: \mathbb{Z}^n \rightarrow \Lambda$ to a bijective homomorphism $\overline{f}: N' \rightarrow N$. Since the map is bijective and since $X'_i$ form a basis at $T_eN$, then $d\overline{f}X'_i = X_i$ will form a basis for $T_eN$. Further $\gamma_i = e^{X_i}$, and $[X_i, X_j] = d\overline{f}([X'_i, X'_j])$ it is direct to verify that $[X_i, X_j] \in \text{span}(X_1, \ldots, X_{i-1})$ for $j > i$. In particular the basis is triangular.

Next we define the sets $I_j$. We need that $L > l^n$. Define $i_0 = 1$. Let $i_1$ be the smallest index such that $|\gamma_i|_e > l|\gamma_1|$, 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 $|\gamma_i|_e > l|\gamma_k|$, or if there is no such index $i_k = n$ and we stop. Define $s$ such that $i_s = n$, and $I_k = \{\gamma_i|_k < i \leq i_k\}$. 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 $\Lambda_k = \langle \gamma_i | \gamma_i \in J_k \rangle$. Define also $\Lambda_0 = \{e\}$.

By choosing $e(n) < \frac{1}{10c(n)e(n)}$ 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 $\gamma_i$ are a basis we can show that $\Lambda_k$ is normal in $\Lambda$. Next we want to show that it is also normal in $\Gamma$. Take an element $\sigma \in \Gamma$ with non-trivial holonomy, and a $\gamma_i \in \Lambda_k$ from the basis. Further represent $\sigma = (t, A)$, where $A \in \text{Aut}(N)$ is in the holonomy and $t \in N$. By left-multiplying with a element in $\Lambda$ we can assume $|t|_e \leq 16c(n)e(n)$. Then $\sigma \gamma_i \sigma^{-1} = t(A\gamma_i)t^{-1} \in \Lambda$. Since $A$ is an isometry $|A\gamma_i| = |\gamma_i|$. Also using the
diameter estimate we can assume $|σ|_e < 4ε(n)$. This combined with the commutator estimate gives $d(t(Aγ_i)^{-1}, Aγ_i) ≤ \frac{1}{10} |γ_i|$, and thus $|t(Λγ_i)^{-1}| ≤ (1 + \frac{1}{10}) |γ_i|$, and thus by the iterative definition of $Λ_k$ we must have $t(Λγ_i)^{-1} ∈ Λ_k$. Note that we are assuming $l > 2$.

The connected Lie subgroups $L_k$ can be generated by $X_1, \ldots , X_k$, and normality in $Λ$ follows since the basis is triangular. It is also easy to see that $L_k$ is normal in $N$. Further $κ$-co-compactness of $Λ_k$ follows by considering the map $(t_1, \ldots , t_{i+κ}) → e^{t_1X_1} \cdots e^{t_iX_i}$. By the previous paragraph, for any element $σ = (t, A) ∈ Γ$ with $|σ|_e < 4ε(n)$ and $γ ∈ Λ_k$, we have $tAγ^{-1} ∈ Λ_k ∈ L_k$. Since $L_k$ is normal in $N$ we have $Λγ ∈ L_k$. From this we can conclude that $AL_k = L_k$, i.e. that $L_k$ are invariant under the holonomy action of $Γ$.

Next define $l_k = 2|γ_i|$. Fix $p ∈ N$. We will show that for $Λ_k = \langle λ ∈ Λ|d(λp, p) < l_k \rangle$ and $d(λp, p) > l/4l_k$ for every $λ ∈ Λ \setminus Λ_k$. First of all we can assume that $d(p, e) < 16c(n)ε(n)$. By assumption for any $γ_i ∈ Λ_k$ we have $|γ_i| ≤ |γ_i|_e < \frac{1}{10} l_k$. Then by commutator estimates

$$d(γ_i p, p) = d(p^{-1}γ_i p, e) ≤ d(p^{-1}γ_i p, γ_i) + d(γ_i, e) ≤ (1 + \frac{1}{10}) |γ_i| ≤ l_k.$$ 

Thus $γ_i ∈ \langle λ ∈ Λ|d(λp, p) < l_k \rangle$, and further by varying $γ_i$ and generation we get $Λ_k < \langle λ ∈ Λ|d(λp, p) < l_k \rangle$. Next if $γ ∉ Λ_k$, then $|γ|_e = d(γ, e) > l/2l_k$ by the construction of $Λ_k$. Thus

$$d(γ p, p) = d(p^{-1}γ p, e) ≥ d(γ, e) − d(p^{-1}γ p, γ) ≥ (1 − \frac{1}{10}) |γ|_e ≥ l/4l_k.$$ 

In particular for every element such that $d(λp, p) < l_k < l/4l_k$ we have $λ ∈ Λ_k$. Thus $Λ_k = \langle λ ∈ Λ|d(λp, p) < l_k \rangle$, and for any $γ ∉ Λ_k$ we have $|λ|_p > l/4l_k$. The final conclusion $d(λp, p) > l_1/(2L)$ for all $λ ∈ Λ$ follows from $|λ|_e > l_1/L$ by 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 $l_k$. 

Sylvester Eriksson-Bique
Definition 3–26 Let $S = N/\Gamma$ be an $\epsilon(n)$-quasiflat orbifold. If $l = 400c(n)$ and $L = 2^p$, then call a basis $\gamma_1, \ldots, \gamma_{n}$ satisfying the previous lemma a short basis for $\Lambda$. The grouping $I_1, \ldots, I_s$ 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 $k = 1, \ldots, s$, the sets $J_k$ are called the $k$'th most collapsed scales.

Definition 3–27 Let $S$ be a complete Riemannian orbifold and $L = V(S)$ a vector 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 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 [31] 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 $S = \{pt\}$ is a single point (i.e. a zero-dimensional orbifold), and $S = \mathbb{R}^k/\Gamma$, where $\Gamma < O(k)$ is a finite group.

Because the monodromy action 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. The monodromy action is given by a homomorphism $\Gamma \to O(n)$, where $\Gamma$ corresponds to the orbifold fundamental group. The rest of this section collects some necessary facts about these group representations. Note that in our case $\Gamma$ 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.

Definition 3–28 Let $H$ be an abelian group subgroup of $O(n)$. A decomposition $\mathbb{R}^n = \bigoplus V_j$ is called the canonical decomposition and the subspaces $V_1, \ldots, V_k$ 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 $hV_j = V_j$.

---

Monodromy may also be defined for orbifold. A similar discussion is used to prove that flat orbifolds are good in [45].
• There is one $V_j$ such that for all $h \in H$ we have $hv = v$ for all $v \in V_j$ (i.e. each $h \in H$ acts trivially on this subspace). We will always denote this subset $V_0$ possibly allowing $V_0 = \{0\}$.

• On each other subspace $V_j$ either 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 $(z_1, \ldots, z_n)$, such that for any $h \in H$ the action is given by $h((z_1, \ldots, z_n)) = (e^{\theta_j(h)}z_1, \ldots, e^{\theta_j(h)}z_n)$ 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 the minimal polynomial. Using this terminology we prove a minor modification of the Jordan lemma.

**Lemma 3–29** *(Modified Jordan Theorem)* Let $\Gamma$ be a group with a normal subgroup $\Lambda$ with index $[\Gamma : \Lambda] < \infty$ and $h: \Gamma \to O(n)$ 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. The subgroup $L$ is Lie subgroup of $O(n)$. It is either finite or has non-trivial identity component $T$. In the first case, we may proceed as in the standard Jordan Theorem [53]. In the latter case, since $\Gamma$ has a finite index nilpotent subgroup, $T \subset O(n)$ must be a compact nilpotent group. This is only possible if it is abelian, i.e. 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 $j = 1, \ldots, s$ will be even-dimensional, and the action will correspond to a Hopf-action $\theta(z_1, \ldots, z_n) \to (e^{i\theta}z_1, \ldots, e^{i\theta}z_n)$. On $V_0$ the action is trivial.

The decomposition to $V_j$’s corresponds to a homomorphisms $\phi = (\phi_1, \ldots, \phi_k): O(n) \to \prod_{j=1}^k O(V_j)$. Because $T$ is normal in $L$, we have that it conjugates the Hopf-action on each $V_i$ to its self (which agrees with the action of $T$). The infinitesimal generator in $\mathfrak{g}(V_j)$ corresponding to the Hopf-action is $J$ which maps $J(z_1, \ldots, z_n) = (iz_1, \ldots, iz_n)$. 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 $g$ is holomorphic or anti-holomorphic. In other
words $\phi_j(L) \subset U(V_j) \cup \overline{U(V_j)}$, where $U(V_j)$ and $\overline{U(V_j)}$ consist of the holomorphic and anti-holomorphic isometries on $V_j$. Consider $\phi(L) \cap \prod_j(SU(V_j) \cup SU(V_j)) = K$, where $SU(V)$ and $SU(N)$ are the holomorphic and anti-holomorphic special unitary groups of $V$. This group must be finite because every left coset of $T$ contains a finite number of elements in $\prod_j SU(V_j)$. Note that $t$ acts on $SU(V_j)$ via a Hopf-action $	heta(z_1, \ldots, z_n) \rightarrow (e^{i\theta z_1}, \ldots, 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 $K$.

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

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

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

$\square$

**Remark:** The optimal bound for the classical Jordan theorem is $c(n) \leq (n+1)!$, which is valid for $n \geq 71$ [11]. Thus from the previous argument we get the optimal bound $(n+1)!$ for $n \geq 71$ for $c(n)$.

### 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 1–6 and 1–7 to conclude the proof for general bounded curvature manifolds and orbifolds.

#### 4.1 Flat Manifolds

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

**Theorem 4–1** Every complete flat manifold $M$ of dimension $n$ admits a bi-Lipschitz embedding $f : M \rightarrow \mathbb{R}^N$ with distortion less than $D(n)$ and dimension of the image $N \leq N(n)$.
The somewhat complicated induction argument is easier to understand through examples for which we refer to 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 number of dimensions or different scales at which collapsing occurs. 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 two abbreviations for induction statements: \( C - (n, S) \) and \( V - (n, k, S) \). These statements are defined here. For the notion of collapsing scales see Definition 3–26.

- **\( C - (n, S) \)**: There exists constants \( L(n, S) \) and \( N(n, S) \) such that the following holds. Assume \( M \) is an arbitrary \( n \)-dimensional compact flat manifold with \( S \) collapsing scales. Then there exists a bi-Lipschitz embedding \( f: M \to \mathbb{R}^N \) with distortion \( L \leq L(n, S) \) and target dimension \( N \leq N(n, S) \) depending only on \( n \) and \( S \).

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

In the proof below we will do an induction argument by first showing \( C - (n, 1) \), and then reducing the other cases to simpler ones with either fewer collapsing scales, or a smaller dimension. The reductions are based on splitting and doubling arguments (Lemmas 3–2 and 3–7). For \( C - (n, S) \) the reduction is similar to that of a Lens space, and for \( V - (n, d, S) \) the reduction is analogous to that for holonomy-bundles (see Section 2).

The core aspects of the argument are to use a stratification of a lattice from Lemma 3–25 combined with the local group Lemma 3–15 to expose the structure of a local group at a smaller scale and then apply Lemma 3–10 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. This embedded manifold is found by finding submanifolds in the universal cover which remain invariant under the action of a local group. Here the averaging from Lemma 3–13 plays a crucial role. Some of the proofs in this section could also be done directly for \( \epsilon(n) \)-quasiflat spaces, but we avoid a number of technicalities with this approach.
**Proof or Theorem 4–1:** Consistently $M = \mathbb{R}^n / \Gamma$ will be a complete flat manifold, and $V = \mathbb{R}^n \times \mathbb{R}^d / \Gamma$ will be a locally flat Riemannian vector bundle over such a base. Denote the corresponding lattice by $\Lambda = \Gamma \cap \mathbb{R}^n$. For a base point $p$ use the natural subadditive norm $|\cdot|_p$ on $\Gamma$. Denote by $D = \text{diam}(M)$. Choose a short basis $\gamma_1, \ldots, \gamma_n$ for $\Lambda$, and the canonical grouping $I_1, \ldots, I_S$ with parameters $l = 400c(n)$ and $L = 2^m$ (see Definition 3–26). The number $S$ is used to denote the number of collapsed scales. By $J_k$ for $k = 1, \ldots, s$ we denote the $k$'th most collapsed scales. By the proof for Lemma 3–25 these correspond to normal subgroups $\Lambda^k < \Lambda$, which are normal in $\Gamma$. Also they correspond to connected normal Lie-subgroups $L_k < N$ which are invariant under the holonomy action of $\Gamma$. Call $l_k$ the collapsing scales given by Lemma 3–25.

Points in $\mathbb{R}^n$ are denoted by lower case letters such as $p$ and points in the quotient by $[p]$. Cosets with respect to local groups will be denoted by $[p]_r$.

**Case $C – (n, 1):** By the definition of $l_1 = |\gamma_i|_e = |\gamma_n|_e$ in 3–25, and Lemma 5–4 in the appendix we have that $\text{diam}(M) < C(n)|l_1|$. Let $p \in \mathbb{R}^n$ be arbitrary. Set $\delta = l_1/(320c(n)L)$. Define the group $\Gamma_\delta = \langle g \in \Gamma | d(gp, p) < 8\delta \rangle$. By Lemma 3–15 for any $[p] \in M$ we have $B_{[p]}(\delta) \subset M$ is isometric to $B_{[p]_\delta}(\delta) \subset \mathbb{R}^n / \Gamma_\delta(p)$. Since $\Lambda_0 = \langle \lambda \in \Lambda | d(\lambda p, p) < 8\delta \rangle = \{e\}$, and for any $\lambda \in \Lambda$ we have $|\lambda|_e > 10c(n)\delta$, we have by Lemma 3–10 an exact sequence

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

The group $H_0$ is finite and thus also $\Gamma_\delta$ is finite. Since $M$ is a manifold $\Gamma_\delta$ acts freely on $M$. Applying Lemma 3–13 we get that if $\Gamma_\delta$ were non-trivial then it would have a fixed point. Since this would contradict the freeness of the action of $\Gamma_\delta$, we must have that $\Gamma_\delta = \{e\}$. Thus $B_{[p]_\delta}(\delta) \subset \mathbb{R}^n / \Gamma_\delta(p)$ is isometric to $B_{e}(\delta) \subset \mathbb{R}^n$. 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 $\text{diam}(M)/\delta \leq 320C(n)c(n)L$. Thus by Lemma 3–7 also $M$ possesses an embedding.

**Reduction $C – (n, S)$ to $V – (n – k, k, S – 1)$ for some $n > k \geq 1$ and $S > 1$:** Consider $\delta = \max \left( \frac{L}{200Lc(n)}, \frac{2L_{S-1}}{S-1} \right)$. Note that $\text{diam}(M) < C(n)L_S \leq 200Lc(n)C(n)\delta$ by Lemma 5–4. 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 an $(n – k)$-dimensional flat manifold.

Define the local group $\Gamma_\delta = \langle g \in \Gamma | d(gp, p) < 8\delta \rangle$, and using Lemma 3–15 we get $B_{[p]}(\delta) \subset M$ is isometric to $B_{[p]_\delta}(\delta) \subset \mathbb{R}^n / \Gamma_\delta$. Define the subgroup $\Lambda_\delta = \langle \lambda \in \Lambda$...
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 $L_{S-1}$ via Lemma 3–25. The subgroup $L_{S-1}$ is invariant under $\Gamma_\delta$ and the holonomy. Thus the action of $\Gamma_\delta$ 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 Lemma 3–13 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 $\Gamma_r$. Since $O$ is invariant under $\Gamma_\delta$ we have $\mathbb{R}^n/\Gamma_\delta = V'$, where $V'$ is a locally flat Riemannian vector bundle over $O/\Gamma_\delta$ of dimension $n - k$ with $S - 1$ collapsing scales (since $\Lambda_{S-1}$ acts on $O$ as a lattice and due to Lemma 3–25). In particular $B_{[p]}(\delta)$ is isometric to a subset of a $k$-dimensional locally flat Riemannian vector bundle over an $n - 1$-dimensional base with $S - 1$-collapsing scales. Thus the ball $B_{[p]}(\delta)$ can be embedded by $V - (n - k, k, S - 1)$.

**Reduction** $V - (n, d, S)$ to $V - (n - c, d + c, S - 1), V - (a, b, c)$ and/or $C - (n + d - 1, c)$: 

Call $V$ the locally flat vector bundle over $M$. We will also establish $a + b < n + d$, which means that the dimension is smaller. 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 $T_0 = \{ p \in V | d(p, M) \leq 2D \}$ and $T_k = \{ p \in V | 2^{k+1}D \geq d(p, M) \geq 2^kD \}$ for $k \in \mathbb{N}$. 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 $0 \in \text{Im}(f_k)$ 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 McSchane extension Theorem 3–1 we can extend them to $V$. A bi-Lipschitz embedding will result from

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

where $r(x) = d(x, M)$. 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 to embed it -
as long as we know how to embed $K$. If $d = 1, 2$, $K$ will be compact and flat, and we can apply the case $C - (n + d - 1, c)$. 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 unit sphere of radius $t$. The space $K = \mathbb{R}^n \times D2^i S^{d - 1}/\Gamma$, where $\Gamma$ acts on $S^{d - 1}$ by isometries via $h: \Gamma \to O(d)$. Any point in $K$ can be represented as an orbit $[(q, v)] \in K$ for an element $(q, v) \in \mathbb{R}^n \times S^{d - 1}$. Fix $\delta = \frac{D2^i}{\cos(\pi/2\sqrt{d})}$. We will show that $B_{[(q, v)]}(\delta) \subset K$ is bi-Lipschitz to a subset of vector bundle with $V$ of total dimension $m = n - 1$. This will be done by studying the local group $\Gamma_\delta = \langle g \in \Gamma | d(g(q, v), (q, v)) \leq 8\delta \rangle$, and observing by Lemma 3–10 that $B_{[q]}(\delta)$ is isometric to $B_{[q]}(\delta) \subset \mathbb{R}^n \times D2^i S^{d - 1}/\Gamma_\delta$ and describing in detail the action of $\Gamma_\delta$.

By the modified Jordan Theorem 3–29 $h(\Gamma)$ has a normal abelian subgroup $L$ of index at most $c(d)$, and $h^{-1}(L) = \Lambda$ is also normal and has finite index at most $c(d)$. Decompose $\mathbb{R}^d$ into canonical invariant subspaces $V_i$ corresponding to $L$. By uniqueness the subspaces $V_i$ are invariant under the action of $\Gamma$. Decompose $v$ into its components $v_i \in V_i$. Consider the induced norm $|| \cdot ||$ on $\mathbb{R}^d$. There is an index $i$ such that $||v_i|| \geq \frac{2^i D}{\sqrt{d}}$. To reduce clutter, re-index so that $i = 1$ and denote $\xi_1 = D^i v_i/||v_i||$. Denote by $d_{g^i - 1}$ the metric on the sphere $2^i DS^{d - 1}$. Then

$$d_{g^i - 1}(v, \xi_1) \leq (2^i D) \cos^{-1} \left( \frac{1}{3\sqrt{d}} \right) \leq \frac{\pi}{2} - \frac{1}{3\sqrt{d}}$$

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

$$d(g(q, \xi), (q, \xi)) \leq 24\delta\sqrt{d}.$$ 

Denote by $\Lambda_0 = \Gamma_\delta \cap \Lambda$. Clearly $[\Gamma_\delta : \Lambda_0] \leq c(d)$. Also define $\Lambda_\delta = \langle g \in \Lambda | d(g(q, v), (q, v)) \leq 8\delta \rangle$.

Either action of $\Lambda$ on the sub-space $V_i$ will consist of reflections and identity transformations, or it will be a Hopf-type action. Consider the first case. We have $24c(d)\delta\sqrt{d} < 2D^i \pi/4$. By the argument in Lemma 3–10 we can show that $\Lambda_\delta$ on $V_1$ is in fact trivial, and that $\Gamma_\delta$ is a finite extension of $\Lambda_\delta$ by a group $H_0$ with $[H_0] \leq c(d)$. Thus the group $h(\Gamma_\delta)$ is a finite group acting on $\mathbb{R}^d$ and for any $g \in h(\Gamma_\delta)$ we have $d_{g^i - 1}(g(2^i D\xi_1), 2^i D\xi_1) < 40c(d)\delta\sqrt{d} < 2D^i \pi/(100\sqrt{d})$. By the fixed point Lemma 3–13 we find a $w \in 2^i DS^{d - 1}$ fixed by the action of $h(\Gamma_\delta)$. Also $d_{g^i - 1}(2^i D\xi_1, w) \leq 2D^i \pi/(100\sqrt{d})$. Thus
The submanifold \( \mathbb{R}^n \times \{w\} \) is invariant under \( \Gamma_\delta \), and thus the manifold \( M_\delta = \mathbb{R}^n \times \{w\} / \Gamma_\delta \) is a compact flat manifold. Fix \( R = D2^i(\pi/2 - \frac{1}{10\sqrt{d}}) \). The normal exponential map gives a map of the \( R \)-ball \( B' \) in the tangent space of \( D2^iS^{d-1} \) at \( w \) onto \( B_w(R) \subset D2^iS^{d-1} \): \( e_w : \mathbb{R}^{d-1} \rightarrow 2iDS^{d-1} \). When \( B' \) is equipped with the induced Euclidean metric this map becomes a bi-Lipschitz map of \( \mathbb{R}^n \times \{w\} \) in \( \mathbb{R}^n \times 2iDS^{d-1} \). Recall that if \( A \) is a subset of a metric space \( X \) then its \( r \)-tubular neighborhood is defined as \( A_r = \{ x \in X | d(x, A) < r \} \).

The action of \( \Gamma_\delta \) 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 \( \Gamma_\delta \) on \( \mathbb{R}^n \times B' \subset \mathbb{R}^n \times \mathbb{R}^{d-1} \) commutes with the action on \( \mathbb{R}^n \times 2iDS^{d-1} \) via the map \( F \). Thus \( F \) induced a bi-Lipschitz map of \( \mathbb{R}^n \times \{w\} \times B'/\Gamma_\delta \rightarrow M_\delta \). Further note that \( B_{[p,v,\infty]}(\delta) \) is contained in \( \nu_\delta(\mathbb{R}^n \times \{w\}) \). Thus we have that the restricted inverse \( f^{-1} : B_{[p,v,\infty]}(\delta) \rightarrow \mathbb{R}^n \times \{w\} \times B'/\Gamma_\delta \) 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 \( M_\delta \) which is a flat manifold of dimension \( n \). Thus, it can be embedded using the statement \( V = (n, d-1, s) \) for some \( s \).

The other case is when \( V_1 \) is a \( 2k \)-dimensional rotational subspace where the action of \( L \) given by the Hopf-action. This time the action of \( \Lambda_0 \) on \( \mathbb{R}^n \) will not necessary be trivial, but will be a Hopf-type action. For any point \( \xi \in 2iDS^{2k-1} \cap V_1 \) its \( \Lambda_0 \)-orbit is contained in a unique Hopf-fiber \( S^1_\xi \subset 2iDS^{2k-1} \). The action of \( \Gamma_\delta \) conjugates the Hopf-action to its self, and thus any orbit \( S^1_\xi \) is mapped to another orbit \( S^1_\eta \). Apply the argument in Lemma 3–10 with \( \Lambda = \Lambda_0 \) and thus any coset of \( \Lambda_0 \) in \( \Gamma_\delta \) can be represented by an element \( g \in \Gamma_\delta \) such that

\[
d(gS^1_\xi, S^1_\xi) \leq d(g\xi_1, \xi_1) \leq 24\delta \sqrt{d}c(d)
\]

The action of \( \Gamma_\delta \) on \( 2iDS^{2k-1} \) descends to an action of the finite group \( H_0 = \Gamma_\delta / \Lambda_0 \) on the orbit-space of the Hopf-action. We realize this orbit space as the complex projective...
space of diameter $2^i D$, i.e. $2^i D \mathbb{C} P^{(k-2)/2}$. By the previous we have for any $h \in H_0$ that for the orbit $S^1_{\xi_1}$ corresponding to $\xi_1$:

$$d(hS^1_{\xi_1}, S^1_{\xi_1}) \leq d(g\xi_1, \xi_1) \leq 2c(d(\delta\sqrt{d}) < D2^i \pi/4.$$ 

Thus Lemma 3–13 can be used to give an orbit $S^1_w$ for some $w \in 2^i D\mathbb{C}^2$. Similar to above we can get $d_{S^1_w} (S^1_w, v) \leq D2^i (\pi/2 - \frac{1}{10\sqrt{d}})$). This time define $M_\delta = \mathbb{R}^n \times S^1_w / \Gamma_\delta$ and consider it as a subset of $K$. Define $R$ as before. The ball $B(q, \nu)(\delta)$ is contained in the $R$-tubular neighborhood $\nu_R(S^1_w)$. The $R$-tubular neighborhood can be identified in a bi-Lipschitz fashion with $\mathbb{R}^n \times S^1_w \times B'$ for $B'$ a Euclidean ball of radius $R$ and dimension $d - 2$. Similar to the previous the action $\Gamma_\delta$ induces an action on $\mathbb{R}^n \times S^1_w \times B'$, and $\mathbb{R}^n \times S^1_w \times B' \Gamma_\delta$ is isometric to a subset of a $d - 2$-dimensional locally flat vector bundle $V''$ over $M_\delta$. In particular we get a map $F$: $B(q, \nu)(\delta) \rightarrow V''$. The conclusion thus follows from the statement $V - (n + 1, d - 2, s)$ for some $s$.

The previous paragraphs gives 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 $C(n, 1)$ we can use Lemma 3–7 and embed each ball $B(p)(\delta)$ since for $\delta = l_1/(320c(n)L)$ we have $\Gamma_\delta = \{e\}$. Also note diam($T_0) \leq 6D$.

Assume instead that $S > 1$. Consider an arbitrary $B(p)(\delta)$ with $[p] \in T_0$ for $\delta = \max(\frac{l_s}{200c(n)L}, 2l_{S-1})$. Here $[p]$ is an orbit of the $\Gamma$ action for $p = (p^*, v) \in \mathbb{R}^n \times \mathbb{R}^d$. Since diam($T_0) \leq 6D \leq 200c(n)L \delta$ we can use the Lemma 3–7 to construct embeddings once we can embed all of these balls. Again define $\Gamma_\delta = \langle g \in \Gamma | d(gp^*, p) \leq 8\delta \rangle$. Also define $\Gamma'_\delta = \langle g \in \Gamma | d(gp^*, p) \leq 8\delta \rangle$. We have $\Gamma_\delta \subset \Gamma'_\delta$. 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 $B(p)'_\delta(\delta) \subset \mathbb{R}^n \times \mathbb{R}^d / \Gamma'_\delta$.

By the second case in the proof there is an invariant affine subspace $O \subset \mathbb{R}^n$ of dimension $t$ such that $\Gamma'_\delta O \subset O$. In particular $\mathbb{R}^n \times \mathbb{R}^d / \Gamma'_\delta = \mathbb{R} \times \mathbb{R}^{d-n-t} / \Gamma'_\delta$. Also the flat manifold $O/\Gamma'_\delta$ has a translational subgroup $A_{S-1}$ with at most $S - 1$ collapsed scales, and thus $O/\Gamma'_\delta$ has $S - 1$ collapsing scales. The space $O \times \mathbb{R}^{k+n-t} / \Gamma'_\delta$ 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 $V - (t, k + n - t, S - 1)$.

□

4.2 Flat orbifolds

We next modify the proof slightly to allow us to embed complete flat orbifolds.
**Theorem 4–2** Every complete flat orbifold $O$ admits a bi-Lipschitz embedding into Euclidean space with distortion and dimension depending only on $n$. Further every locally flat orbivector bundle over such a base with its natural metric admits such an embedding.

**Proof:**

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 and add a new case that didn’t occur previously. We assume the same notational conventions as in Theorem 4–1.

- $C - (n, S)$: There exists constants $L(n, S)$ and $N(n, S)$ such that the following holds. Assume $M$ is an arbitrary compact flat orbifold with dimension $n$ and $S$ collapsing scales. Then there exists a bi-Lipschitz embedding $f: M \rightarrow \mathbb{R}^N$ with distortion $L \leq L(n, S)$ and target dimension $N \leq N(n, S)$ depending only on $n$ and $S$.

- $V - (n, d, S), n \geq 1$: There exists constants $L(n, d, S)$ and $N(n, d, S)$ 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 $n$ with $S$ collapsing scales. Then there exists a bi-Lipschitz embedding $f: V \rightarrow \mathbb{R}^N$ with distortion $L \leq L(n, d, S)$ and target dimension $M \leq N(n, d, S)$ depending only on $n, d$ and $S$.

- $V - (0, d, 0)$: There exists constants $L(0, d, 0)$ and $N(0, d, 0)$ 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 $f: V \rightarrow \mathbb{R}^N$ with distortion $L \leq L(0, d, 0)$ and target dimension $N \leq N(0, d, 0)$ depending only on $n$.

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

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

**Reduction** $V - (0, n, 0)$ to $V - (n - k, k - 1, s)$: Let $\Gamma < O(n)$ be a finite subgroup. The space $M = \mathbb{R}^n/\Gamma$ is isometric to $C(S^{n-1}/\Gamma)$, i.e. the cone over $S^{n-1}/\Gamma$. By Lemma 3–5, it is sufficient to embed $S = S^{n-1}/\Gamma$. The proof is similar to the vector bundle reductions considered in the previous proof. By the Jordan theorem 3–29 we
have a finite index normal subgroup $K \triangleleft \Gamma$ which is abelian with index bounded by $[\Gamma : K] \leq c(n)$. Decompose $\mathbb{R}^n$ to canonical invariant subspaces $V_i$ for $K$.

Let $\delta = \frac{1}{10^6 c(n) \sqrt{d}}$, 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 $B_{\{v\}}(\delta)$. Define $\Gamma_\delta = \langle g \in \Gamma | d(gp, p) < 8\delta \rangle$, $K_\delta = \langle g \in K | d(kv, v) < 8\delta \rangle$ and $K'_\delta = \Gamma_\delta \cap K$. Decompose $v$ into itâ€™s components $v_i \in V_i$. There is an index $i$ such that $||v_i|| \geq \frac{1}{3\sqrt{d}}$. To reduce clutter, reindex so that $i = 1$, and define $\xi_1 = \frac{v_1}{||v_1||}$. But then $d(g\xi_1, \xi_1) \leq 24\sqrt{\delta}$. Repeating the argument in the proof of Theorem 4–1 we get either an element $w \in S^{n-1}$ fixed by $\Gamma_\delta$, or a Hopf-orbit $S^1 \subset V_1 \cap S^{n-1}$ which is invariant under $\Gamma_\delta$. By using the normal exponential map we get that $B_{\{v\}}(\delta) \subset S^{n-1}$ is bi-Lipschitz to either a $n - 1$ dimensional locally flat orbivector bundle over a point (i.e. $w$), or a $n - 2$ dimensional orbivector bundle over a one dimensional compact flat orbifold. These can be embedded using the statements $V - (0, n - 1, 0)$ or $V - (1, n - 2, 1)$.

**Reduction** $C - (n, 1)$ to $V - (0, n, 0)$: This is very similar to the $C - (n, 1)$ case above. We still have the estimate $\text{diam}(M) < C(n) |l_1|$. Let $p \in \mathbb{R}^n$ be arbitrary. Set $\delta = \frac{l_1}{(320c(n)L)}$. Define the group $\Gamma_\delta = \langle g \in \Gamma | d(gp, p) < 8\delta \rangle$. On the other hand by Lemma 3–15 for any $[p] \in M$ we have $B_{[p]}(\delta) \subset M$ is isometric to $B_{[p]}(\delta) \subset \mathbb{R}^n / \Gamma_\delta(p)$. Define $\Lambda_0 = \langle \lambda \in \Lambda | d(\lambda p, p) < 8\delta \rangle = \{e\}$, and similarly to before we get

$$0 \rightarrow 0 \rightarrow \Gamma_\delta \rightarrow H_0 \rightarrow 0,$$

where $H_0$ and thus $\Gamma_\delta$ is finite. The group $\Gamma_\delta$ has a fixed point $o$ by Lemma 3–13, 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 $V - (0, n, 0)$-case above. Finally the proof of embedding $M$ is completed by applying Lemma 3–7.

**Reduction** $C - (n, S)$ to $V - (n - k, k, S - 1)$: The proof for manifolds translates almost verbatim. Consider $\delta = \max(\frac{b_{l_1}}{200L^{n-1}}, 2l_{S-1})$. Again $\text{diam}(M) < C(n) |l_1| \leq 200Lc(n)C(n)\delta$ by Lemma 5–4. 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 bi-Lipschitz to an orbivector bundle over an $(n - k)$-dimensional flat orbifold $O$. The only difference is that the action is not free.

**Reduction** $V - (n, d, S)$ to $V - (n - d, k + d, S - 1)$, $V - (a, b, c)$ and/or $C - (n + d \text{â€šŠ} 1, c)$: There is no major change to the manifold case as we didn’t use the fact that the action
is free in any substantive way. The same decomposition into $T_i$’s is used, followed by a consideration of the action of $\Gamma_\delta$ on $S^{d-1}$, and finding a fixed point $w$ or a fixed Hopf-orbit $S^1_\xi$.

□

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 $n$. For dimensions 1, 2 we can bound the distortion by constants $D(1)$ and $D(2)$, and they can be embedded in 3 dimensions. In general, for $n$-dimensional compact spaces with only one collapsing scale we have to employ a doubling Lemma 3–7 at the scale $n! \sim O(e^{Cn \ln(n)})$, which gives a distortion of $O(e^{Cn \ln(n)})$. For $S > 1$ collapsing scales, the model on the smaller scale is a vector bundle over a base with $S - 1$ collapsed scales. Another decomposition and scaling argument is performed. For the sets $T_i$ we get additional distortion $O(e^{Cn^2 \ln(n)}D(n-1))$. However decompositions near the base will result in vector bundles with fewer collapsed scales but of dimension $n$, so the argument needs to be repeated up to $S$ times to give a worst case distortion of the order $O((e^{Cn^2 \ln(n)})^SD(n-1))$. For vector bundles of dimension $n$, the bound is similar. As $S \leq n$, we get a recursion relation $D(n) = O(e^{Cn \ln(n)}D(n-1))$, which gives $D(n) = O(e^{Cn^2 \ln(n)})$. The dimension of the embedding is of the order $O(e^{Cn^2 \ln(n)})$. Our arguments are not optimal and probably some exponential factor could be removed, but the methods don’t seem to yield better than super exponential distortion.

4.3 Quasiflat orbifolds

Below the proof of Theorem 1–5 will proceed by approximating the orbifold by a flat orbifold. This will involve quotienting away the commutator subgroup $[N, N]$ and using Lemma 3–8. Thus we begin by studying lattices and their commutators.

Lemma 4–3 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 $N'/\Lambda'$ is co-compact (see [37] and [10]). Further by Malcev there is a unique homomorphism $f: N' \to N$. We want to show that $f(N') = [N, N]$. First of all, since $\Lambda' < [N, N]$, we get that $f(N') < [N, N]$. The
Quantitative Bi-Lipschitz embeddings

image $f(N')$ is normal in $N$, and $N/f(N')$ is a nilpotent Lie group. The projection $\pi: N/\Lambda' \to N/f(N')$ has compact fibers and thus we can show that the action of $\Lambda/[\Lambda, \Lambda]$ on $N/f(N')$ is properly discontinuous and co-compact. 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 4–4** Let $\epsilon(n)$ be sufficiently small\(^{11}\) and $l \leq \epsilon(n)$ fixed. There exists a dimension dependent constant $C(n)$ such that the following holds. Assume that $N$ is an $n$-dimensional nilpotent Lie group with a left invariant metric $g$ and sectional curvature $|K| \leq 2$. If $\Lambda < N$ is a lattice which is generated by elements of length $|l| \leq l \leq \epsilon(n)$, then $\text{diam}(N/\Lambda) \leq C(n)l$.

**Proof:** As in Lemma 3–25 we choose a short basis $\gamma_i$ for the lattice $\Lambda$. By construction $|\gamma_i| \leq l$. The result then follows from Lemma 5–4 in the appendix, with the same constant. The bound obtained from Lemma 3–25 is

□

**Lemma 4–5** Let $S$ be an $\epsilon(n)$-quasiflat $n$-dimensional orbifold. Then there exists constants depending only on the dimension $\delta < 1$ and a constant $A(n)$, such that the following two properties hold.

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

**Proof:** Let $S = N/\Gamma$, and $\Lambda = N \cap \Gamma$ with $[\Gamma : \Lambda] < c(n)$. Let $l = 400\epsilon(n), L = 2^m, I_j, J_k, \gamma_j, \Lambda_k, L_k$ and $S$ be as in Lemma 3–25 and Definition 3–26. As in Lemma 3–25 we need to assume $\epsilon(n)$ small enough. For this proof $\epsilon(n) \leq \frac{\delta_{(n)}}{c(n)\log^2 c(n)c(n)\epsilon(n)L}$ is enough. See Section 3 for the definitions of the various constants.

As in Lemma 3–25 we can generate $\Lambda$ by elements $\alpha_i$ such that $|\alpha_i|e < 4\epsilon(n)\epsilon(n)$. As such $[\Lambda, \Lambda] = \{[\mu, \nu]|\mu, \nu \in \Lambda\} < N \times \mathbb{R}^k$ is generated by elements $\beta_{ij} = [\alpha_i, \alpha_j] \in N$ with $|\beta_{ij}|e < \epsilon(n)/(10L^2) < \epsilon(n)/(8Lc(n))$.

\(^{11}\)See the remark at the beginning of Section 3.
Choose the largest index $a$ such that $|\gamma_a|_e < \epsilon(n)/(8Lc(n))$, and let $a \in I_b$ for some $b$. By construction $I_b \leq L|\gamma_a|_e \leq \epsilon(n)/(8c(n))$.

Also for any $\beta_{i,j} \in \Lambda_b$ because if $\beta_{i,j} \notin \Lambda_b$, then by construction for the first $i_{b+1} \notin J_b$ we have (by construction of the basis)

$$|\gamma_{b+1}|_e \leq |\beta_{i,j}|_e < \epsilon(n)/(8Lc(n)),$$

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

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 $S$\footnote{The action does not descend due to the action of $\Gamma$.}. This gives a map between orbifolds $\pi: S' \to S$, which lifts to a Riemannian submersion on their orbifold universal covers. Choose a coset $[p] \in S$ represented by an element $p \in N$ with $|p| \leq 4c(n)e(n)$.

Thus it will be sufficient to prove that $\text{diam}(\pi^{-1}([p])) < \delta$. Choose an orbit $O_{[p]} \subset S$ passing through $[p] \in S$. The $O_{[p]}$ is isometric to $L_b/\Gamma_p$ where $\Gamma_p = \{p^{-1}\gamma_p | \gamma \in \Gamma\}$ 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 $|p^{-1}\gamma_i|_e \leq 2|\gamma_i|_e \leq \epsilon(n)/(4c(n)).$ By Lemma 4–4 we have $\text{diam}(L_b/\Gamma_p) \leq C(n)e(n)/(4c(n))$. Thus $d_{GH}(S', S) \leq \delta$ follows if we set $\delta = C(n)e(n)/(2c(n))$ and using Lemma 3–9.

Next define $\Gamma_{\delta}(p) = \langle \gamma \in \Gamma | d(\gamma p, p) \leq 8\delta \rangle$. By Lemma 3–10 we know that $B[p](\delta) \subset S$ is isometric to $B[p](\delta) \subset N/\Gamma_{\delta}(p)$. Further by Lemma 3–10, and observing that $\Lambda_{\delta} = \langle \gamma \in \Lambda | d(\gamma p, p) \leq 8\delta \rangle = \Lambda_b$ we have a finite group $H_0$ and a short exact sequence

$$0 \to \Lambda_b \to \Gamma_{\delta} \to H_0 \to 0.$$
there is an orbit \([q]_{L_\delta}\) which is fixed by the action of \(H_\delta\) and \(d([q]_{L_\delta}, p) \leq 8(c(n) + 1)\delta\). Further the orbit \([q]_{L_\delta}\) is invariant under the action of \(\Gamma_\delta\).

Consider the normal bundle \(\nu([q]_{L_\delta})\) in \(N\). The bundle has a trivialization and can be identified with \(\mathbb{R}^t \times L_\delta\) for some \(t\). By Lemma 5–5 there is a map \(f: A(\delta(n)) \rightarrow \mathbb{R}^t \times L_\delta\) from the \(\delta(n)\)-tubular neighborhood \(A(2\delta(n))\) of \([q]_{L_\delta}\). This map is also \(A(n)\)-bi-Lipschitz on the ball \(B_q(2\delta(n))\). Further the group \(\Gamma_\delta\) acts on \(A(\delta(n))\) and \(\mathbb{R}^t \times L_\delta\) by isometries, and the action commutes with \(f\). We can thus show that that restricted to the ball \(B = B_q(\delta(n)) \subset N/\Gamma_\delta\) the induced map \(\overline{f}: A(\delta(n))/\Gamma_\delta \rightarrow \mathbb{R}^t \times L_\delta/\Gamma_\delta\) is also \(A(n)\)-bi-Lipschitz. To see this one can consider lifted geodesics. The space \(\mathbb{R}^t \times L_\delta/\Gamma_\delta = V(S')\) is a locally flat Riemannian orbivector bundle over a flat basis \(S'\). Also \(B_{[p]_{L_\delta}}(\delta) \subset B_{[q]_{L_\delta}}(\delta(n))\), so further restricting we get a \(A(n)\)-bi-Lipschitz map \(\overline{f}: B_{[p]_{L_\delta}}(\delta) \rightarrow V(S)\) is the desired map.

\[\Box\]

A similar statement is needed for orbivector bundles.

**Lemma 4–6** Let \(\epsilon(n)\) be small enough. For every \(d\) and \(n\) there are constants \(A(n, d)\) 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 exists a \(\delta < 1\), 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 \(f_p: B_\delta(p) \rightarrow V(S')\).
- There exists a locally flat Riemannian orbivector bundle \(V'' = V(S'')\) over a flat basis \(S''\) such that \(d_{GH}(V, V'') < \delta\).

**Proof:** For this proof \(\epsilon(n) \leq \frac{\delta(n)}{c(n)10^{c(n)}L_{c(n)}\mu}\) is enough (see Definition 3–26 for the definition of \(L\)). We have \(S = N/\Gamma\) and \(V = N \times \mathbb{R}^d/\Gamma\), where \(\Gamma\) acts by rotations on \(\mathbb{R}^d\) and co-compactly on \(N\) so that \(S\) is \(\epsilon(n)\)-quasiflat. There is a lattice \(\Lambda < \Gamma\) of \(N\) with \([\Gamma : \Lambda] \leq c(n)\). Further there is an induced action \(h: \Lambda \rightarrow O(d)\). By Lemma 3–29 we have an index \(c(d)\) subgroup \(\Lambda' < \Lambda\) with \([\Lambda : \Lambda'] \leq c(d)\) and \(h(\Lambda')\) abelian. In particular \([\Lambda, \Lambda] = \Lambda_c < \Lambda'\).

Since \(\Lambda\) is a lattice of \(N\) by Lemma 4–3 we have that \(\Lambda_c\) is a lattice in \([N, N] = N_c\). Further by the arguments in the Lemma 4–5 that \(\Lambda_c\) is generated by elements of length at most \(\epsilon(n)/(8Lc(n))\). Consider the left action of \(N_c\) on \(N \times \mathbb{R}^d\), which acts trivially
on $\mathbb{R}^d$. The group $\Gamma$ maps orbits to orbits, and thus $\Gamma_c = \Gamma/\Lambda_c$ is a discrete group and acts properly discontinuously on $N/N_c \times \mathbb{R}^d$. 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 $\pi : V \to V_c$ the induced projection map. The group action of $\Gamma_c$ is co-compact on $N/N_c$, and by actions on $\mathbb{R}^d$. Further $N/N_c$ is abelian, and thus isomorphic to $\mathbb{R}^l$ for some $l$. In other words we can represent it as $V_c = \mathbb{R}^l \times \mathbb{R}^d/\Gamma_c$, which is a $d$-dimensional locally flat orbifold over the flat base $\mathbb{R}^l/\Gamma_c$.

Next take any $q \in V$ and the $N_c$ orbit $O_q$ passing through that point. By a computation from Lemma 4–3 and using Lemma 4–4 we have that $\text{diam}(O_q) \leq C(n)e(n)/(4c(n))e(n)$. Define $\delta = C(n)e(n)/(2c(n))$. As before we can show that $\pi$ is a $\delta$-Gromov-Hausdorff approximation. In particular $d_{GH}(V/N_c, V) \leq \delta$. This shows the second part.

We will next show the first statement. Next fix a $p = (q, v) \in N \times \mathbb{R}^d$. The argument is similar to Lemma 4–5. Define $\Gamma_\delta(q) = \{ \gamma \in \Gamma | d(\gamma q, q) \leq 8\delta \}$. Further define $\Gamma_\delta(p) = \{ \gamma \in \Gamma | d(\gamma p, p) \leq 8\delta \}$. Since $\Gamma_\delta(p) < \Gamma_\delta(q)$, by Lemma 3–10 we know that $B(p) \subset V$ is isometric to $B(p) \subset N \times \mathbb{R}^d/\Gamma_\delta(q)$. Further by Lemma 3–10 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 $\Lambda_b$ is defined similarly as in Lemma 4–5. Further $\Lambda_b$ is a lattice in the connected Lie subgroup $L_b \subset N$. By the same argument as before we can find an $L_b$ orbit $[x]_{L_b}$ passing through $x$ with $d(x, q) < 8(c(n) + 1)\delta$ in $N$ which is invariant under the action of $\Gamma$. By Lemma 5–5 that for the $\delta(n)$-tubular neighborhood $A(\delta(n))$ of $[x]_{L_b}$ in $N$ there is a map $f : A(\delta(n)) \times \mathbb{R}^d \to L_b \times \mathbb{R}^l \times \mathbb{R}^d$, which is identity on the $\mathbb{R}^d$ factor. Further $f$ when restricted to $B_4(2\delta(n)) \times \mathbb{R}^d$ is $A(n)$-bi-Lipschitz on each component. The action of the group $\Gamma_\delta(q)$ induces canonically an action on $L_b \times \mathbb{R}^l \times \mathbb{R}^d$ which leaves $A(\delta(n))$ invariant and which acts on $\mathbb{R}^d$ by rotations. Define an induced map $\tilde{f} : A(\delta(n)) \times \mathbb{R}^d/\Gamma_\delta(p) \to L_b \times \mathbb{R}^l \times \mathbb{R}^d/\Gamma_\delta(p)$. Let $B(\delta(n)) = A(\delta(n)) \times \mathbb{R}^d/\Gamma_\delta(p)$. The space $V'' = L_b \times \mathbb{R}^l \times \mathbb{R}^d/\Gamma_\delta(p)$ is a Locally flat Riemannian orbivector bundle over $L_b/\Gamma_\delta(p)$, which is an $\epsilon(n)$-quasiisiflat $n - l$-dimensional orbifold.

Finally consider the ball $B_{|p|_{\delta}}(3\delta) \subset B(\delta(n))$. That this inclusion holds 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 $8(c(n) + 2)\delta < \frac{\delta(n)}{10}$. The restricted map $\tilde{f}_{|B_{\delta(n)}(\delta)}$ is $A(n)$-bi-Lipschitz. This can be seen by noting that, since $f$ is component-wise $A(n)$-bi-Lipschitz, it distorts the distance of any curve by a factor at most $A(n)$. On the other hand any pair of points in
$B_{|p|\delta}(\delta)$ can be connected by a geodesic in $B_{|p|\delta}(2\delta)$, and any points in $\overline{f(B_{|p|\delta}(\delta))}$ can be connected by a geodesic within $\overline{f(B_{|p|\delta}(3\delta))}$. In conclusion $\overline{f|_{B_{|p|\delta}(\delta)}}: B_{|p|\delta}(\delta) \to V''$ is our desired map.

□

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

**Proof of Theorem 1–5:** 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.

- $E(0, n)$: Every $\epsilon(n)$-quasiflat $n$-dimensional orbifold $S$ admits a bi-Lipschitz map $f: S \to \mathbb{R}^{N(n)}$ with distortion $L(n)$.

- $E(d, n)$ for $n \geq d \geq 1$: 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 $f: V \to \mathbb{R}^{N(d, n)}$ with distortion $L(d, n)$.

By Theorems 4–1 and 4–2 we have the cases $E(0, 1)$ and $E(n - 1, n)$. The statement will be completed by an induction on $n$. The case $n = 1$ is completed since $E(0, 1)$ and $E(1, 1)$ are one-dimensional flat orbifolds and easy to embed. Next fix $n > 1$ and assume we have proved $E(d, s)$ for all $s < n$. The statement $E(n, n)$ corresponds to a flat orbifold and is covered in Theorem 4–2. Further the case $E(n - 1, n)$ corresponds to a locally flat orbivector bundle over a flat base, so is covered also in Theorem 4–2.

Next we will prove $E(d, n)$ by reverse induction on $d$. The base cases $d = n, n - 1$ are included above. So fix $d < n - 2$. Assume the statement $E(k, n)$ has been proved for $k > d$. There are two slightly different cases $d = 0$ and $d > 1$. Assume the first. In this case by Lemmas 4–5 there is a $\delta$ such that $B_p(\delta)$ is bi-Lipschitz to a case covered by $E(s, n)$ for some $s > 0$. Thus $B_p(\delta)$ admits a bi-Lipschitz embedding by induction. Also $d_{GH}(S, S') \leq \delta$ where $S'$ is a flat orbifold of lower dimension. $S'$ admits a bi-Lipschitz embedding by Theorems 4–2. The statement then follows from Lemma 3–8. The case $d \geq 1$ is similar, but one uses Lemma 4–6 instead and concludes that $B_p(\delta)$ is bi-Lipschitz to a case covered by $E(d + s, n)$ for some $s > 0$, and $d_{GH}(V, V') \leq \delta$ for a locally flat orbivector bundle $V'$ over a flat base.

□
4.4 Proofs of main theorems

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

**Proof of Theorem 1–10:** By Fukaya’s fibration Theorem 1–6, every point \( p \in M^n \) has a definite sized ball \( B_p(\delta) \) equipped with a metric \( g_\epsilon \) such that \((B_p(\delta), g_\epsilon)\) is isometric to a subset of a compact \( \epsilon(n) \)-quasiflat space or a locally flat vector bundle \( V \) over such a base. Since \( ||g - g_\epsilon|| < \epsilon \), then the ball \( B_p(\delta/2) \) is bi-Lipschitz to a subset of \( V \) with distortion at most \( 1 + \epsilon \). Since the vector bundle can be embedded by Theorem 1–5, and we can embed \( M^n \) using a doubling Lemma 3–7.

\[ \square \]

**Proof of Theorems 1–3 and 1–4:** The only difference to the proof of Theorem 1–10 is using Fukaya’s fibration theorem for orbifolds 1–7. In the particular case of flat and elliptic orbifolds we can also use Theorem 3–16 to identify these spaces as quotients, and embed them using Theorem 4–2 and 4–1.

\[ \square \]

5 Appendix

5.1 Collapsing theory of orbifolds

Much of the terminology here is presented in \([31, 54]\). Recall that a *topological orbifold* is a second countable Hausdorff space with an atlas of charts \( U_i \) satisfying the following.

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

A chart is denoted simply by \( U_i \), without explicating the covering \( \hat{U}_i \) 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
Any other notion like geodesics, sectional curvature or Ricci-curvature, is defined by considering the lifted metric \( g_i \). One may also define the notion of orbifold maps as those that lift to the covering charts \( \hat{U}_i \) and are compatible with the projections. A smooth orbifold map is also naturally defined by considering charts. Such a map is non-singular if in charts it can be expressed as a local diffeomorphism.

A non-singular 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 \( (V, O, \pi) \), where \( V, O \) are smooth orbifolds and \( \pi: V \to O \) is a smooth map with the following properties.

- There is an atlas of \( O \) consisting of charts \( U_i \) such that for any chart \( (U_i, \hat{U}_i, G_i) \) of \( O \) there is a chart \( V_i = \pi_i^{-1}(U_i) \).
- There are covering charts \( \hat{V}_i = \hat{U}_i \times \mathbb{R}^k \) and an \( G_i \) action, such that \( V_i = \hat{V}_i / G_i \), and maps \( \pi_i^V: \hat{V}_i \to V_i \).
- The projections \( \hat{\pi}_i: \hat{V}_i \to \hat{U}_i \) satisfy \( \hat{\pi}_i = \pi \circ \pi_i^V \).
- The action of \( G_i \) on \( \hat{U}_i \times \mathbb{R}^k \) 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 \( \hat{V}_i \), such that \( G_i \) acts by isometries on the fibers \( \mathbb{R}^k \) and 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 \( \hat{V}_i \) 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 paper.

**Definition 5-1** \((V, O, \pi)\) 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 to it its tangent orbivector bundle \( TO \), where the group \( G_i \) actions on the fibers are given by differentials. The fiber of a point \( \pi^{-1}(p) \subset TO \) is denoted \( C_p(O) \). For Riemannian orbifolds this is easily seen to correspond with the metric tangent cone at \( p \). Note that 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 and corresponding local orbifold fundamental group. A connected Riemannian orbifold has a distance function and we can define a complete Riemannian manifold as one which is complete with respect to its distance function. On such orbifolds we may define an exponential map $\exp: T_O \to O$ by considering geodesics and their lifts. Fix now a point $p \in O$ and take the exponential co-ordinates $\mathbb{C}_p(O) = \mathbb{R}^n / G$, we can lift the exponential map to a map $\tilde{\exp}_p: \mathbb{R}^n \to O$. Next we state a crucial property of such an exponential map.

**Lemma 5–2** If a complete Riemannian orbifold has sectional curvature bounded by $|K| \leq 1$, then $\tilde{\exp}_p$ is a non-singular orbifold map on the ball $B_0(\pi) \subset \mathbb{R}^n$.

This result is alluded to in the appendix of [13] and can be proved by using local co-ordinates and standard Jacobi-field estimates. Let us now consider the orbifold local covering by $B_0(\pi)$. For a fixed $\delta > 0$, $C > 1$ we also define the group

$$\pi_1^{loc}(\delta, p, C) = \{ \gamma \in \text{Isom}(B_0(\delta), B_0(C\delta)) | \tilde{\exp}_p(\gamma x) = \tilde{\exp}_p(x), \forall x \in B_0(\delta) \},$$

where $\text{Isom}(B_0(\delta), B_0(C\delta))$ is the set of isometries mapping $B_0(\delta)$ into $B_0(C\delta)$.

Using this terminology we can now state and outline the proof of Fukaya’s fibration theorem for collapsed orbifolds

**Theorem 5–3** Fix $\epsilon(n)$ arbitrarily small. Let $(O, g)$ be a complete Riemannian orbifold of dimension $n$ and sectional curvature $|K| \leq 1$. For every $\epsilon > 0$ there exists a universal $\rho(n, \epsilon) > 0$ and $\delta(n) > 0$ such that for any point $p \in O$ if $\text{Vol}_p(O) < \delta(n)$ there exists a metric $g'$ on the ball $B_p(\rho(n))$ and a complete Riemannian orbifold $O'$ with the following properties.

- $\|g - g'\|_g < \epsilon$
- $(B_p(5\rho(n, \epsilon)), g')$ is isometric to a subset of $O'$
- $O'$ is either a $\epsilon(n)$-quasiflat orbifold or a locally flat Riemannian vector bundle over a $\epsilon(n)$-quasiflat orbifold $S$. 
Proof: Fix an $\epsilon$. The argument proceeds by contradiction. Let $O_i$ be a sequence of examples that have $\text{Vol}_{\rho_i}(O) < \frac{1}{i}$ where there doesn’t exist any $\rho(n, \epsilon)$ such that the conclusion holds. One should first regularize the manifold slightly as to attain uniform bounds on all co-variant derivatives. This is done in [31]. The action of the local group on the frame bundle is free and thus Fukaya’s proofs in [17, 18] (see also [10]) 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 [10] we obtain an equivariant fibration of $FO_i$ by nilpotent fibers, and a metric $g_\epsilon$ which is invariant under the local nilpotent action and the local $O(n)$ action. Further we have bounds on their second fundamental forms and normal injectivity radii of the nilpotent fibers. By a remark in [47] we can also control the sectional curvatures of $g_\epsilon$. By equivariance this structure descends onto a fibration of $O_i$. Call the fiber passing through $p$ by $O_p$.

By the argument in [10, Appendix 1] we can find (for large enough $i$) a constant $\delta$ and a fiber $O_{q_i} \subset O_i$ with $d(q_i, p_i) < \delta$, and the orbit $O_{q_i}$ is normal injectivity radius at least $20\delta$ footnotethe normal injectivity radius can be defined by considering the normal exponential map, which is defined similar to the above. The normal injectivity radius can be bounded by another compactness argument from [10] and using the local universal covers above. One can bound the second fundamental form of $O_{q_i}$. Denote by $\nu_{10\delta}(O_{q_i})$ the vectors in the normal bundle of the fiber with length at most $10\delta$, and $g'$ the bundle metric of this locally flat bundle. Then one can show that the normal-exponential map $f: \nu_{10\delta}(O_{q_i}) \to M$ induces a map onto the $10\delta$-tubular neighborhood of $O_{p_i}$ with $\|f^*g - g''\|_g < \epsilon'$. Further the image contains $B_{p}(\delta)$ and we can consider the restricted map $f^{-1}: B_{p}(\delta) \to \nu_{10\delta}(O_{q_i})$. The desired metric is $f^{-1*}(g'') = g'$.

\square

5.2 Diameter estimate

Lemma 5–4 Let $M = N/\Gamma$ be an $n$-dimensional $\epsilon(n)$-quasiflat orbifold and $\Lambda < \Gamma$ the translational normal subgroup in $N$. Further let $\gamma_i$ be a short basis for $M$ as in Lemma 3–25. Then we have $\text{diam}(M) \leq C(n) \max_{i=1,\ldots,n} |\gamma_i|$.

Proof: Denote $D = \max_{i=1,\ldots,n} |\gamma_i|$. To prove the lemma we can assume $\Gamma = \Lambda$, because taking a finite cover will only increase the diameter. We will show in fact
show a more general statement. The statement we prove is that if a cocompact lattice \( \Lambda < N \) admits a triangular basis \( \gamma_1, \ldots, \gamma_n \) in the sense of definition 3–24, then \( \text{diam}(M) \leq C(n) \max_{1 \leq i \leq n} |\gamma_i| \). 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 nilpotent Lie groups. Assume thus that \( \Lambda \subset N \) is a discrete co-compact lattice of an \( n \)-dimensional nilpotent Lie group with a triangular basis \( \gamma_i \) and set \( \gamma_i = e^{X_i} \) with \( X_i \in T_e N \) a basis and \([X_i, X_j] \in \text{span}(X_1, \ldots, X_{i-1})\) for \( i < j \). Since \( \gamma_i \) is assumed to be a basis, we have that \( X_i \) form a basis for \( T_e N \). In particular we can introduce coordinates \( (t_1, \ldots, t_n) \to \prod_{i=1}^n e^{t_{n-i} X_{n-i}} \), such that the product can be expressed as

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

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

Since \( \gamma_1 \) is in the center of \( N \), we have \( X_1 \) commutes with all left-invariant fields. Further by the Koszul formula \( t \to e^{tX_1} \) is seen to be a geodesic in \( N \). Thus \( d(e^{tX_1}, e) = \text{sd}(e^{tX_1}, e) = s|\gamma_1| \). Consider the Lie group \( N' = N/L \), where \( L = \{e^{tX_1}\} \), and a discrete subgroup \( \Gamma' = \langle \gamma'_i, i = 1, \ldots, n - 1 \rangle = \Gamma/\langle \gamma_i^n, n \in \mathbb{Z} \rangle \). Here the generators correspond to equivalence classes of generators of \( \Gamma \) in \( \Gamma/\langle \gamma_i^n, n \in \mathbb{Z} \rangle \): \( \gamma'_i = [\gamma_{i+1}] \), and if we denote \( e^{tX_i} = \gamma'_i \), we have co-ordinates \( (t_1, \ldots, t_{n-1}) \to \prod_{i=1}^{n-1} e^{t_{n-i} X_{n-i}} \). Here the exponential map is in \( N' \). The metric is a quotient metric, so that \( |\gamma'_i|_e \leq |\gamma_i|_e \leq D \) for all \( 1 \leq i \leq n - 1 \). Define \( M' = N' / \Gamma' \). The group \( L \) acts naturally on \( M \) by isometries and \( M/L \) is isometric to \( M' \). Further we have that the \( L \)-orbits on \( M \) have length \( |\gamma_1| \), so by Lemma 3–9 \( d_{GH}(M, M') \leq 3|\gamma_1| \). By the inductive hypothesis \( \text{diam}(M') \leq C(n - 1)D \). And thus

\[
\text{diam}(M) \leq \text{diam}(M') + 2d_{GH}(M, M') \leq C(n - 1)D + 6D \leq (6 + C(n - 1))D.
\]

In particular we can choose \( C(n) = 6n \).

\[\square\]

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

**Proof:** Introduce an orthonormal triangular basis $X_1, \ldots, X_k$ for $M^k$, and extend it to an orthonormal basis for $N^n$ with vectors $X_{k+1}, \ldots, X_n$. Motivated by the normal exponential map from Riemannian geometry consider the map $G: M^k \times \mathbb{R}^{n-k} \to M^k \ltimes N^n$ as

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

This is smooth. 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. Applying the Campbell-Hausdorff formulas to compute the quantities $\partial_{\delta} G$.

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

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

$$\langle \partial_{\delta} G(m,x), \partial_{\delta} G(m,x) \rangle = \langle X_i + Y_i, X_j + Y_j \rangle = g_{ij} + \langle Y_i, Y_j \rangle + \langle X_i, Y_j \rangle + \langle Y_i, X_j \rangle$$

Since $||Ad|| \leq C'(n)$, we get $||Y_i|| \leq M(n)\delta^k C'(n)$. Choosing $\delta(n) < \frac{1}{10M(n)C'(n)}$ and applying Cauchy’s inequality gives

$$1 - \frac{3}{8} \leq \langle \partial_{\delta} G(m,x), \partial_{\delta} G(m,x) \rangle \leq 1 + \frac{3}{8},$$

whenever $x_i \leq 10\delta(n)$. Let $C(\delta(n)) = \{(m,x)|x_i \leq 10\delta(n)\} \subset M^k \times \mathbb{R}^{n-k}$. The estimate shows that $F(C(\delta(n)))$ contains the $4\delta(n)$-tubular neighborhood $A(\delta(n))$ of $M^k$. Also when restricted onto $B_\rho(2\delta(n))$ we get a $\frac{1+\frac{3}{8}}{1-\frac{3}{8}}$ bi-Lipschitz into $N$ whose image contains the ball $B_c(\delta(n))$. Thus the inverse $F^{-1}$ is our desired map. We remark, that the distortion can be made arbitrarily small by choosing a smaller $\delta(n)$.

□
References


Quantitative Bi-Lipschitz embeddings


Quantitative Bi-Lipschitz embeddings


Courant Institute, New York University,
251 Mercer St,
New York, NY-10012, USA

ebs@cims.nyu.edu

http://www.cims.nyu.edu/~ebs