THE DATA MANIFOLD UNDER THE MICROSCOPE

Anonymous authors
Paper under double-blind review

ABSTRACT

A significant gap exists between theory and practice in deep learning. While generalization and approximation error bounds have been proposed, they are often restricted to overly simplified models or result in loose guarantees. Many of these bounds rely on the manifold hypothesis and depend on geometric regularity properties such as intrinsic dimension, curvature, or reach of the data manifold or target functions. However, evaluations of such bounds typically fall into two extremes: either synthetic, analytically defined manifolds where geometric properties are precisely known, or real-world datasets where the bounds are judged solely by downstream performance. Neither approach adequately reveals how data geometry affects the tightness or applicability of the theoretical results.

We propose a general-purpose framework for studying data geometry by creating dense, controllable versions of dSprites and COIL-20 with additional transformation dimensions and fine sampling resolution. This setup enables accurate finite-difference estimates of geometric measures such as curvature, reach, and volume, offering a flexible benchmark for analyzing manifold learning methods. As illustrative applications, we evaluate two established manifold learning bounds by Genovese et al. and Fefferman et al., and we examine how manifold geometry evolves across network layers in β-VAEs. Our results highlight both the limitations of existing bounds and the potential of such controlled benchmarks to guide future theoretical developments.

1 INTRODUCTION

Deep learning has become the dominant paradigm for a broad range of tasks, and in recent years generative models in particular such as variational autoencoders (VAEs) (Kingma & Welling, 2013), diffusion models (Ho et al., 2020), and modern masked or autoencoding architectures (He et al., 2022) have seen striking empirical success. A common way to interpret this success is through the manifold hypothesis (Cayton et al., 2005): high-dimensional data often concentrate near low-dimensional manifolds, and learning amounts to finding useful parameterizations of them. Many modern models can thus be viewed as procedures for fitting or approximating data manifolds, whether explicitly, as in latent-variable models(Arvanitidis et al., 2017), or implicitly, as in denoisers and score-based flows(Horvat & Pfister, 2021).

While appealing, it remains unclear how networks fit manifolds in practice. Classical results such as Genovese et al. (2012) on minimax rates for manifold estimation, Fefferman et al. (2016) on testing the manifold hypothesis , or more recent work on approximation of Sobolev classes on manifolds (Tan et al., 2024) highlight the role of geometric quantities like curvature, reach, or sampling density in learning. Yet in realistic data these quantities are rarely observable: the data-generating process is unknown, the intrinsic dimension is uncertain(Campadelli et al., 2015), and sampling can be irregular(Shroff et al., 2011; Sedghi et al., 2020). As a result, theoretical guarantees often rely on constants that cannot be directly checked or measured.

This creates two complementary needs. On the theoretical side, sharper and more data-adaptive bounds are desirable, or at least a clearer picture of when existing bounds are informative. On the empirical side, there is a lack of benchmarks that balance realism with geometric control: synthetic manifolds are too simple, while real-world datasets obscure the geometry entirely.

To narrow this gap, we introduce a framework that constructs low-dimensional image families sampled densely along controlled transformation axes (e.g., rotation, translation, scale), and provides

efficient finite-difference estimators for geometric measures such as curvature, reach, and volume. Coupled with an experimental pipeline for probing manifold fitting methods (e.g., β-VAEs), our framework enables systematic tests of how theory and practice align. While in this paper we illustrate its use through generative models and manifold-fitting bounds, the framework is general and can support investigations in other settings, such as discriminative learning, where geometry plays a role. Our contributions are the following:

- A reproducible framework of adapted low-dimensional datasets with dense, axis-aligned sampling, plus an experimental pipeline for probing modern generative and representation models.
- A suite of efficient finite-difference estimators for pointwise and global geometric quantities such as the curvature, reach, or the volume of manifolds.
- Empirical studies that demonstrate how the framework can be used to test theoretical bounds and trace how learned representations reshape data geometry.

2 Related Work

Manifold Analysis: Following the manifold hypothesis (Cayton et al., 2005), many works analyze the geometry of data and learned representations. Studies on VAEs and β-VAEs (Kingma et al., 2019; Higgins et al., 2017) report that latent manifolds are often relatively flat (Arvanitidis et al., 2017; Shao et al., 2018). Geometric invariants such as curvature, tangent spaces, and reach have been used to probe robustness and entanglement (Aamari et al., 2019; Berenfeld et al., 2022; Birdal et al., 2021), though results vary strongly with dataset, architecture, and estimator (Brahma et al., 2015; Kaufman & Azencot, 2023). Our work continues this line of research, but focuses on controllable datasets where geometric measures can be estimated more accurately.

Manifold Fitting Bounds: Genovese et al. (2012) established the minimax rate $O(n^{-2/(2+d)})$ depending only on intrinsic dimension d, proven optimal by Kim & Zhou (2015), though their estimator requires intractable non-convex optimization. Aamari & Levrard (2019) achieved $\tilde{O}(n^{-k/(k+d)})$ rates for k-smooth manifolds with noise, while Yao & Xia (2019) handled unbounded noise via projection, achieving $O(\sigma^2 \log(1/\sigma))$ Hausdorff error. Neural approaches (Yao et al., 2023; 2024) trade guarantees for efficiency. The geometric approach of Fefferman et al. (2016) preserves structural properties via reach, with extensions to noisy data (Fefferman et al., 2018; 2020) improving complexity from double- to single-exponential. A key challenge across these methods is their dependence on unknown geometric parameters. Bounds typically scale with quantities such as reach τ or volume V, which must be specified a priori; e.g., the sample complexity in Yao & Xia (2019) includes a factor τ^{-2} hidden in the constant. Since such quantities are rarely known in practice and can be very small, this creates a circular dependency: estimating the manifold requires geometric information that is only available from the manifold itself.

Geometric Property Estimation: To address this, recent work develops estimators for reach and related quantities (Aamari et al., 2019; Berenfeld et al., 2022), though validation without ground truth remains difficult. Hardness results (Kiani et al., 2024) and positive findings such as the "blessing of dimensionality" for Sobolev approximation (Tan et al., 2024) further underscore the subtle role of geometry in learning. In deep models, learned manifolds often deviate from intrinsic geometry: e.g., flat latent spaces (Shao et al., 2018), model-specific distortions (Grementieri & Fioresi, 2022), or global warping in GANs (Choi et al., 2021). Estimator choice itself can drive contradictory conclusions (Kaufman & Azencot, 2023).

Overall, theory provides good asymptotic rates, but empirical understanding is limited: synthetic data offer tractability but lack realism, while real datasets obscure geometry. Our framework extends prior manifold analyses by combining dense, axis-aligned sampling with finite-difference estimators, enabling accurate ground-truth evaluation of geometric quantities and systematic tests of how theoretical bounds behave under realistic approximations.

3 BACKGROUND AND DEFINITIONS

Datasets as Manifolds: The manifold hypothesis suggests that high-dimensional datasets often concentrate near low-dimensional manifolds. In this paper, we focus on datasets where the number of intrinsic factors of variation d is small and fixed, and where these factors are explicitly known. Each dataset is modeled as a union of smooth d-dimensional manifolds embedded in $\mathbb{R}^D$, possibly with boundary. Typically, each of the k semantic classes in the dataset corresponds to a separate connected component of this union.

For the purposes of this work, we restrict attention to simple topologies in which each manifold factors into cyclic and non-cyclic directions. Concretely, every manifold is assumed to be homeomorphic to

$$[0,1]^r \times (S^1)^s, \quad r+s=d, \tag{3.1}$$

so that some coordinates vary over a compact interval while others wrap around a circle. This setting captures many commonly used synthetic datasets. A canonical example is *dSprites*, a collection of 64×64 grayscale images of objects (square, ellipse, heart) undergoing controlled transformations such as scaling, rotation, and translation. The manifold of a class of this dataset is homeomorphic to $[0,1]^3 \times S^1$ and exists in $\mathbb{R}^{4096}$

To obtain discrete datasets from this continuous geometric picture, we impose a grid structure. Specifically, we consider

$$G = \left\{ \left(\tfrac{j_1}{n_1}, \dots, \tfrac{j_r}{n_r}, 2\pi \tfrac{j_{r+1}}{n_{r+1}}, \dots, 2\pi \tfrac{j_d}{n_d} \right) \,\middle|\, 0 \le j_\ell < n_\ell \right\}, \tag{3.2}$$

where the first r coordinates sample $[0,1]$ uniformly and the last s coordinates sample S^1 uniformly. For each class $i \le k$, we define a mapping

$$u_i : G \to M_i \tag{3.3}$$

that associates each grid point with a dataset element obtained by applying the corresponding transformations. The image $u_i[G]$ provides a discrete parametrization of the manifold M_i, and by cutting along cyclic directions one obtains discrete patches of M_i. The complete discretized dataset is then

$$X_G = \bigcup_{i \le k} u_i[G]. \tag{3.4}$$

Geometric Measures: Critical to the analysis of a manifold are a few geometric quantities that describe its local and global structure. In this work, we focus on the *volume*, the *scalar curvature*, and the *reach*. For completeness, we first recall the definitions of the Riemannian metric and related objects needed to introduce these quantities.

Definition 3.1 (Riemannian metric)**.** A Riemannian metric g on a d-dimensional differentiable manifold M assigns to each point $p \in M$ an inner product g_p on the tangent space T_pM. If M is embedded in $\mathbb{R}^D$, the ambient Euclidean metric induces a Riemannian metric by restriction. In local coordinates $u : U \to M$, the metric matrix is

$$g_{ij}(x) = \langle \partial_i u(x), \partial_j u(x) \rangle_{\mathbb{R}^D}. \tag{3.5}$$

Definition 3.2 (Volume element and volume)**.** On a Riemannian manifold (M, g) with local coordinates $(u^1, \dots, u^d)$, the natural volume element is

$$d\,\mathrm{Vol} = \sqrt{\det(g)}\, du^1 \wedge \cdots \wedge du^d. \tag{3.6}$$

The volume of a region $R \subset M$ is then

$$\mathrm{Vol}(R) = \int_{u^{-1}(R)} \sqrt{\det(g)}\, dx^1 \cdots dx^d. \tag{3.7}$$

Definition 3.3 (Notions of curvature)**.** The curvature of (M, g) is encoded in the Riemann tensor

$$R^i{}_{jkl} = \partial_k \Gamma^i_{jl} - \partial_l \Gamma^i_{jk} + \Gamma^i_{kr}\Gamma^r_{jl} - \Gamma^i_{lr}\Gamma^r_{jk}, \tag{3.8}$$

where the Christoffel symbols are

$$\Gamma^i_{jk} = \tfrac{1}{2} g^{ir}(\partial_j g_{rk} + \partial_k g_{rj} - \partial_r g_{jk}). \tag{3.9}$$

Contracting yields the Ricci tensor $Ric_{ij} = R^r{}_{irj}$ and finally the scalar curvature

$$R = g^{ij} Ric_{ij}. \tag{3.10}$$

Definition 3.4 (Reach)**.** For a closed set $A \subset \mathbb{R}^D$, the *medial axis* is the set of points with more than one nearest neighbor in A. The *reach* of A is the distance to its medial axis:

$$\tau_A = \inf_{p \in A} d(p, Med(A)). \tag{3.11}$$

Intuitively, τ_A is the largest radius such that every point within distance $< \tau_A$ of A has a unique nearest neighbor in A. For a compact d-dimensional submanifold $M \subset \mathbb{R}^D$, the reach can be expressed as

$$\tau_M = \inf_{p \neq q \in M} \frac{\|p - q\|^2}{2\, d(q - p, T_p M)}, \tag{3.12}$$

where $d(v, T_p M)$ is the distance of a vector v to the tangent space at p.

In our applications we consider both *global* quantities (total volume, scalar curvature integral, global reach) and their *local* counterparts (volume element, pointwise scalar curvature, and local reach estimators as in Aamari et al. (2019)). These quantities will also serve as the basis for efficient finite-difference approximations introduced later.

Definition 3.5 (Hausdorff distance)**.** Let $A, B \subset \mathbb{R}^D$ be two non-empty subsets. The *Hausdorff distance* between A and B is defined as

$$H(A, B) \;=\; \max\Big\{ \sup_{a \in A} \inf_{b \in B} \|a - b\|\,,\; \sup_{b \in B} \inf_{a \in A} \|b - a\| \Big\}. \tag{3.13}$$

In words, $H(A, B)$ measures how far two sets are from coinciding by quantifying the worst-case pointwise deviation: every point of A must lie within distance $H(A, B)$ of some point of B, and vice versa.

This metric is particularly natural for manifold fitting. If M denotes the ground-truth data manifold and $\hat{M}$ a learned approximation (e.g. from an autoencoder), then $H(M, \hat{M})$ expresses the largest geometric error: the maximum distance that a single point on either manifold can have from the other. Thus, a small Hausdorff distance guarantees that the two manifolds are close everywhere, not just on average.

Manifold Fitting and Bounds: Manifold learning is closely connected with generative modeling: in both cases, one assumes that high-dimensional data lie near a low-dimensional manifold. Generative models typically learn a mapping between the data manifold in $\mathbb{R}^D$ and a latent space of lower dimension, where semantic factors of variation are disentangled and sampling is easier. If the encoder–decoder network is sufficiently regular (locally $\mathcal{C}^1$ or $\mathcal{C}^2$), then composing the encoder with the manifold charts u_i yields a parametrization of the latent manifold. In practice, each data point is mapped to its latent representation, and decoding corresponds to projecting back onto the estimated data manifold.

We focus on two theoretical results that provide bounds on the error of manifold fitting, expressed in terms of the sample size, intrinsic dimension, and geometric quantities of the manifold.

Minimax manifold estimation bound: Genovese et al. (2012) establish minimax rates for estimating a manifold from samples corrupted by small normal noise. The result shows that the intrinsic dimension d alone governs the difficulty of estimation.

Theorem 1 (Genovese et al. (2012))**.** Let $\mathcal{M}$ be the class of compact, smooth, boundaryless d-dimensional manifolds embedded in $\mathbb{R}^D$. For each $M \in \mathcal{M}$, let Q_M be the distribution of $Y = \xi + Z$, where ξ is uniformly distributed on M and Z is uniform on a normal fiber of radius σ at ξ. For an n-sample estimator $\hat{M} : \mathbb{R}^n \to \mathcal{M}$, define the minimax risk

$$R_n(\mathcal{Q}) = \inf_{\hat{M}} \sup_{Q \in \mathcal{Q}} \mathbb{E}_Q[H(\hat{M}, M)], \tag{3.14}$$

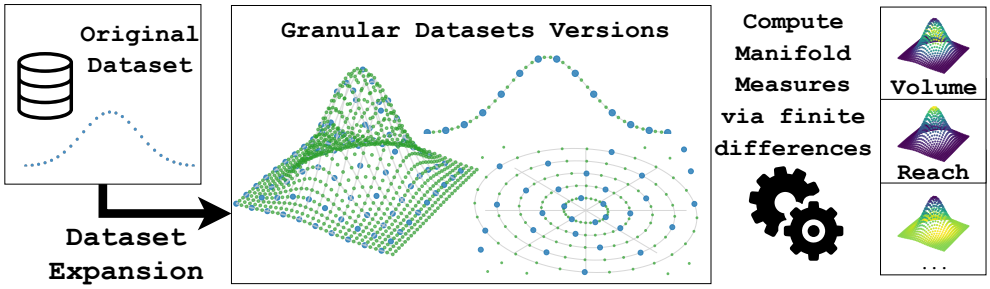

Figure 1: From a low-dimensional seed we produce dense, regular grids via analytic parametrizations or systematic transforms; central finite differences recover geometric estimators such as reach and curvature enabling validation of manifold-fitting bounds.

where $H(\cdot,\cdot)$ is the Hausdorff distance. Then there exist constants $C_1, C_2 > 0$ such that

$$C_1 \left(\frac{1}{n}\right)^{\frac{2}{2+d}} \;\leq\; R_n(\mathcal{Q}) \;\leq\; C_2 \left(\frac{\log n}{n}\right)^{\frac{2}{2+d}}. \tag{3.15}$$

This result shows that the sample complexity depends exponentially on the intrinsic dimension d, but not directly on the ambient dimension D. Intuitively, the distributions Q_M correspond to sampling a point on M and perturbing it orthogonally within the reach of the manifold. In our context, an estimator $\hat{M}$ can be thought of as a neural network (e.g., an autoencoder) that outputs a fitted manifold.

Testing the manifold hypothesis: Fefferman et al. (2016) provide a testing framework for the manifold hypothesis. Their result connects the number of samples to geometric parameters of the manifold, namely its volume and reach. Assuming that a manifold M which satisfies the hypothesis exists, the following upper bound on the Hausdorff distance between the true and estimated manifolds emerges:

$$R_n(\mathcal{Q}) \;\leq\; C_1 \frac{V^{\frac{1}{d}}}{\tau} \left(\frac{1}{n}\right)^{\frac{1}{d}}. \tag{3.16}$$

A more comprehensive version of the background can be found in the appendix A.

4 Methods and Framework

We construct controlled synthetic manifolds of low intrinsic dimension ($d = 1$–4) sampled on regularly spaced grids. Dense sampling enables stable finite-difference approximations of partial derivatives, allowing accurate computation of the induced metric, volume element, curvature tensors, and reach. This framework provides a setting to empirically assess theoretical manifold-fitting bounds and a reproducible dataset suite for validating geometric estimators.

Datasets with a dense grid structure: We consider two complementary approaches to construct datasets of low intrinsic dimension with a grid structure.

For manifolds with an explicit mathematical description, we use known parametrizations to generate a regular grid of points. For image-based or domain-specific datasets, we extend existing data by systematically applying transformations (e.g., translations, rotations, scalings). Sampling all possible combinations of transforms yields a structured grid analogous to the toy manifold case.

While grid sampling provides dense coverage, it is not necessarily uniform with respect to the underlying manifold geometry. To obtain more uniform subsets, we employ two strategies:

- Iterative farthest-point sampling on the grid, ensuring that selected points are maximally separated.
- Sampling with respect to the volume form: either by weighting grid points and sampling proportionally, or—when analytic formulas are available—by using inverse transform sam-

pling iteratively across dimensions with numerical integration and inversion. The latter, however, becomes computationally expensive as dimension grows.

For experiments, we fix a test set by selecting a uniformly spread fraction of grid points. The remaining points are used for training, with varying training set sizes. This procedure ensures a clear train/test separation while keeping the test set representative enough to approximate Hausdorff and average distances between manifolds.

Computation of Geometric Measures: To estimate geometric quantities such as the volume element, scalar curvature, and reach, we rely on finite difference approximations of partial derivatives on a dense grid.

Given a parametrization $u : \mathbb{R}^d \to \mathbb{R}^D$, the central difference approximation

$$f'(x) \;=\; \frac{f(x+h) - f(x-h)}{2h} + \mathcal{O}(h^2) \tag{4.1}$$

extends naturally to each coordinate direction. For instance, we define

$$D^c_{h,i}u(x) = \frac{u(x_1, \ldots, x_i + h, \ldots, x_d) - u(x_1, \ldots, x_i - h, \ldots, x_d)}{2h}, \tag{4.2}$$

which provides second-order accurate estimates of the partial derivatives. Products of such approximations yield

$$g_{ij} = \langle u_{,i}, u_{,j} \rangle + \mathcal{O}(h^2), \tag{4.3}$$

ensuring that all quantities depending smoothly on the metric (volume element, Christoffel symbols, Riemann and Ricci curvature tensors, scalar curvature) inherit $\mathcal{O}(h^2)$ error. These intermediate tensors are also made available in the framework for further study.

For the volume element and scalar curvature, standard Taylor expansion arguments imply that the relative error decreases quadratically with grid spacing:

$$\hat{V} - V = \mathcal{O}(h^2), \tag{4.4}$$

$$\hat{R} - R = \mathcal{O}(h^2). \tag{4.5}$$

For the reach, we adopt the estimator of Aamari et al. (2019), which, when applied to an equidistant grid of spacing h, also yields

$$|\hat{\tau} - \tau| = \mathcal{O}(h^2). \tag{4.6}$$

All three estimators have been tested on families of manifolds with known closed-form quantities, including 2- and 3-dimensional ellipsoids, hyperboloids, and 4-spheres embedded in various ambient spaces. Across these cases, the relative error of our estimates is typically below 10^{-2}, confirming that finite differences on dense grids provide highly accurate approximations.

5 Applications

Datasets: Our experimental setup employs two types of datasets: (i) toy manifolds with explicit mathematical parametrizations for analytical validation, and (ii) adapted versions of established image datasets (dSprites Matthey et al. (2017) and COIL-20 Nene et al. (1996)) that provide controlled settings for empirical evaluation. Table 1 summarizes the key characteristics of each dataset.

The toy manifolds serve as analytically tractable benchmarks with known closed-form geometric properties. For the image datasets, we modified the original sampling schemes to better suit geometric computations: dSprites was regenerated with reduced resolution but improved anti-aliasing, while COIL-20 was extended with additional transformations including scale and orientation. Both datasets employ dense grid sampling with margin oversampling on non-cyclic dimensions to enable finite difference computations of geometric measures. Detailed specifications are in Appendix B.

Manifold fitting methods: We consider two complementary approaches to fit manifolds to the datasets: a geometric method, *Manifold Moving Least Squares* (MMLS) (Sober & Levin, 2020) , and a deep learning method, the β-VAE autoencoder(Higgins et al., 2017).

MMLS: MMLS is a simple and robust method for locally projecting points onto an approximated manifold. Intuitively, it can be seen as averaging tangent spaces in local neighborhoods and projecting points onto these planes. In our experiments, we use six nearest neighbors and a Gaussian kernel with bandwidth $\sigma = 1.0$. MMLS is applied to all datasets, and for the toy datasets it serves as the primary fitting method.

β-VAE: The β-VAE provides a complementary, learning-based approach. It maps input data into a low-dimensional latent space and reconstructs them back, with the reconstruction lying on a learned manifold. The reconstruction distance can thus be interpreted as the distance to this manifold. As the input dimensionality is too small for meaningful bottlenecks in toy datasets, we employ β-VAE only for image datasets (dSprites, COIL-20). The main hyperparameters are latent dimension $z = 10$ and $\beta = 4$, with models trained using batch size 64 and learning rate 5×10^{-4}.

Sampling protocol: For both methods, a fixed test set of 500 uniformly spread points is used. On the toy datasets, training sets range from 5–500 points, and each experiment is repeated 20 times. On dSprites and COIL-20, data ratios range from 0.01 up to 1.0, with three repetitions per setup.

The motivation for considering both methods is that MMLS directly fits a manifold in the original data space, whereas β-VAE learns a latent manifold with richer semantic structure. Together, they provide complementary views of manifold fitting performance.

Error Bounds evaluation: We now turn to the first application of manifold fitting: the empirical evaluation of theoretical bounds on approximation error. Recall from sub-section 3 that Genovese et al. provide an upper and lower bound depending only on the intrinsic dimension, while Fefferman et al. derive a sharper upper bound that additionally incorporates the reach and volume of the manifold.

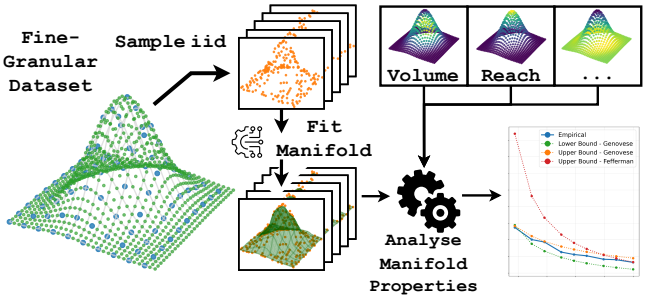

Figure 2: Experimental setup for error bound estimation.

Experimental setup: For the fitting methods (MMLS, β-VAE), each dataset (toy, dSprites, COIL-20), and a range of intrinsic dimensions, we fit manifolds for multiple training set sizes. As distance metric we use the Hausdorff distance between the fitted and the true manifold. Concretely, given a test point on $\mathcal{M}$, its projection is defined as the weighted projection under MMLS or as the reconstruction under β-VAE. For each size, we thus obtain results of distance versus sample size.

Comparing to bounds: To compare with theoretical rates, we regress $\log(\text{dist})$ against $\log(n)$, motivated by the fact that both theoretical bounds decay exponentially in $1/n$ or $\log n/n$, which are asymptotically close. This yields an empirical exponential fit, around which we place the bounds: the upper bound constant is set just above the fitted curve (99th quantile) and the lower bound just below (1st quantile). For Fefferman's upper bound, volume and reach are used to refine the constant, while in the β-VAE case these quantities are estimated on the latent manifold.

Results: We present three representative comparisons:

Table 1: Dataset characteristics and parametrizations used in our experiments:

Dataset	d	D	**Factors**	**Range**	**Components**
Circle	1	2	θ	$[0, 2\pi]$	1
Two moons	1	2	θ	$[0, 2\pi]$	2
Sphere	2	3	θ, ϕ	$[0, 2\pi] \times [0, \pi]$	1
Torus	2	3	θ, ϕ	$[0, 2\pi] \times [0, 2\pi]$	1
dSprites	4	4096	scale, orientation, pos. x, pos. y	see Appendix B	3
COIL-20	3	4096	horiz. orient., scale, img. orient.	see Appendix B	20

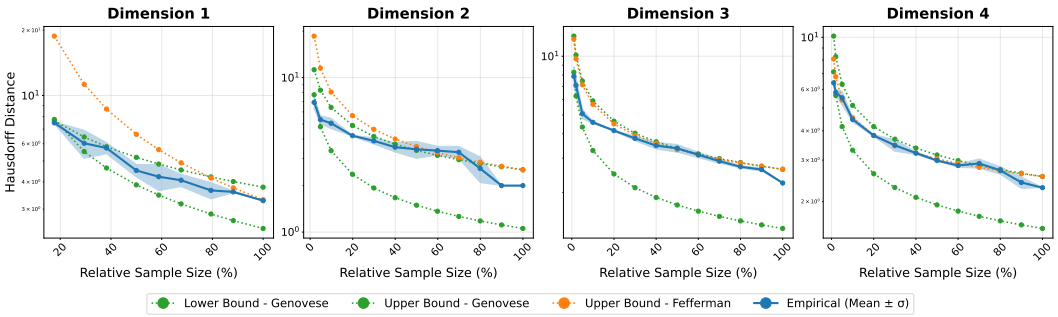

Figure 3: Fitting bounds for different datasets utilizing MMLS.

- **Cross-dataset** (Figure 5): curves for the sphere, torus, 3D COIL-20 and 4D dSprites using MMLS.
- **Cross-dimension** (Figure 3): curves for dSprites with intrinsic dimensions $d = 1, 2, 3, 4$, again using MMLS.
- **Cross-model** (Figure 4): comparison of MMLS and β-VAE on 4D dSprites.

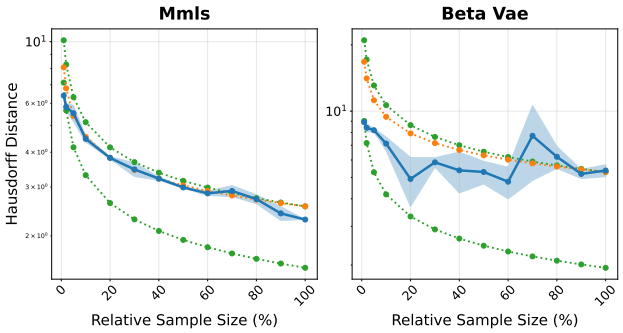

Figure 4: Bounds for MMLS & β-VAE on dSprites (4D)

The main observations are: (i) Fefferman's bounds, which explicitly exploit reach and volume, yield significantly tighter upper envelopes than dimension-only bounds. (ii) For toy datasets, the empirical curves initially follow the upper rate $\left(\frac{\log n}{n}\right)^{1/d}$ but then approach the lower rate $\left(\frac{1}{n}\right)^{1/d}$ as sample size grows. This suggests that an intermediate function, interpolating between $\log n$ and 1, might more faithfully capture the transition. (iii) For image datasets (COIL-20, dSprites), the behavior is reversed: curves start closer to the lower bound and move upward with larger n.

Overall, these results demonstrate that empirical manifold fitting can recover the qualitative behavior of theoretical approximation bounds, while highlighting that constants—and sometimes even the effective regime—depend sensitively on manifold geometry and the choice of fitting method.

Manifold analysis: In the second application, we analyze how geometric properties of data manifolds evolve through the layers of a β-VAE using the dSprites (4D) and COIL-20 (3D) datasets. For each layer we compute the manifold volume, integrated scalar curvature (with positive and negative parts separated), reach, and the average distance between class manifolds (Figure 6).

We observe that curvature systematically increases while reach decreases in deeper layers, indicating that the intermediate manifolds become progressively more intricate and approach self-intersections. At the same time, average distances between class manifolds grow closer to the latent space, suggest-

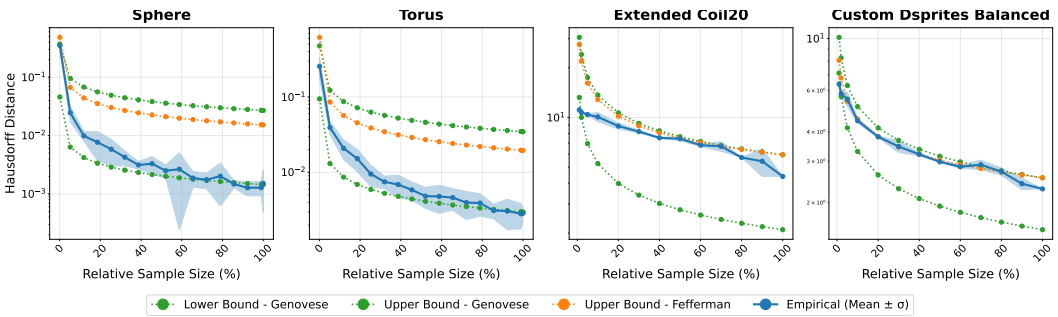

Figure 5: Fitting bounds for MMLS on from left to right Sphere, Torus, COIL-20 and dSprites.

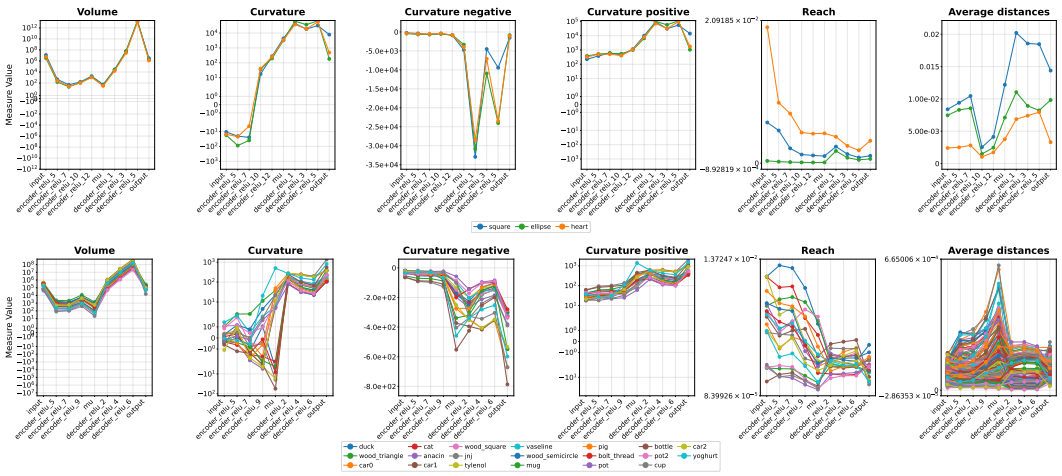

Figure 6: Evolution of volume, scalar curvature, and reach of the manifold across layers of a β-VAE, shown for both dSprites and COIL-20:

ing that as semantic information strengthens, the network prioritizes class separation over preserving low-level visual transformations. This shift may explain why the β-VAE struggles to converge: its latent manifold emphasizes semantic clustering at the cost of geometric distortion, whereas methods such as MMLS, which operate directly in data space, do not exhibit the same trade-off.

6 Discussion

We presented a framework for constructing dense synthetic manifolds and accurately estimating their geometric properties, enabling controlled studies of manifold fitting and learning. Our applications illustrated two use cases: assessing the behavior of existing manifold learning bounds via log–log scaling fits, and analyzing how data geometry evolves across layers of a β-VAE. Both revealed insights into how geometry influences learning, such as the trade-off between increasing intra-class complexity and improved inter-class separation.

The main advantage of this framework lies in its ability to provide ground truth geometric quantities that are otherwise inaccessible. In contrast to real datasets, where assumptions cannot be verified, or simple analytic manifolds, which lack representational richness, our synthetic constructions strike a balance between realism and control. This makes them particularly suitable for probing theoretical assumptions, validating estimators, and stress-testing bounds under controlled perturbations.

At the same time, several limitations should be emphasized. The framework currently handles only manifolds with simple topology and relatively low intrinsic dimension. Dense grid sampling and finite-difference estimation do not easily extend to higher dimensions or to manifolds with complex connectivity. Moreover, the present experiments focus primarily on generative models in unsupervised settings, and do not yet capture the full range of discriminative or semi-supervised tasks.

Future work could address these limitations in multiple ways. Expanding the dataset suite to cover richer transformations and more diverse topologies would broaden applicability. Another direction is to evaluate methods that approximate geometric measures, for example by comparing local PCA- or kernel-based estimators of curvature and reach against ground truth. Likewise, the framework could be used to study discriminative bounds, where classifiers effectively fit functions on manifolds, and to quantify how geometric fidelity interacts with generalization. Beyond bounds, one could also apply it to study robustness, transfer learning, or the geometry of learned latent spaces in larger-scale architectures.

Finally, we note that large language models (LLMs) were used to polish the writing of this paper. All scientific ideas, results, and analyses are original; the LLMs contributed only to improving clarity and readability of the text.

REFERENCES

Eddie Aamari and Clément Levrard. Stability and minimax optimality of tangential delaunay complexes for manifold reconstruction. *Discrete & Computational Geometry*, 59:923–971, 2018.

Eddie Aamari and Clément Levrard. Nonasymptotic rates for manifold, tangent space and curvature estimation. 2019.

Eddie Aamari, Jisu Kim, Frédéric Chazal, Bertrand Michel, Alessandro Rinaldo, and Larry Wasserman. Estimating the reach of a manifold. 2019.

Antonio G Ache and Micah W Warren. Ricci curvature and the manifold learning problem. *Advances in Mathematics*, 342:14–66, 2019.

Alessio Ansuini, Alessandro Laio, Jakob H Macke, and Davide Zoccolan. Intrinsic dimension of data representations in deep neural networks. *Advances in Neural Information Processing Systems*, 32, 2019.

Georgios Arvanitidis, Lars Kai Hansen, and Søren Hauberg. Latent space oddity: on the curvature of deep generative models. *arXiv preprint arXiv:1710.11379*, 2017.

Iskander Azangulov, George Deligiannidis, and Judith Rousseau. Convergence of diffusion models under the manifold hypothesis in high-dimensions. *arXiv preprint arXiv:2409.18804*, 2024.

Randall Balestriero, Sebastien Paris, and Richard Baraniuk. Max-affine spline insights into deep generative networks. *arXiv preprint arXiv:2002.11912*, 2020.

Ronen Basri and David Jacobs. Efficient representation of low-dimensional manifolds using deep networks. *arXiv preprint arXiv:1602.04723*, 2016.

Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6):1373–1396, 2003.

Clément Berenfeld, John Harvey, Marc Hoffmann, and Krishnan Shankar. Estimating the reach of a manifold via its convexity defect function. *Discrete & Computational Geometry*, 67(2):403–438, 2022.

Tolga Birdal, Aaron Lou, Leonidas J Guibas, and Umut Simsekli. Intrinsic dimension, persistent homology and generalization in neural networks. *Advances in neural information processing systems*, 34:6776–6789, 2021.

Paul S Bradley and Olvi L Mangasarian. K-plane clustering. *Journal of Global optimization*, 16: 23–32, 2000.

Pratik Prabhanjan Brahma, Dapeng Wu, and Yiyuan She. Why deep learning works: A manifold disentanglement perspective. *IEEE transactions on neural networks and learning systems*, 27 (10):1997–2008, 2015.

Matthew Brand. Charting a manifold. *Advances in neural information processing systems*, 15, 2002.

Paola Campadelli, Elena Casiraghi, Claudio Ceruti, and Alessandro Rozza. Intrinsic dimension estimation: Relevant techniques and a benchmark framework. *Mathematical Problems in Engineering*, 2015(1):759567, 2015.

Gunnar Carlsson. Topology and data. *Bulletin of the American Mathematical Society*, 46(2):255–308, 2009.

Lawrence Cayton et al. Algorithms for manifold learning. *Univ. of California at San Diego Tech. Rep*, 12(1-17):1, 2005.

Jaewoong Choi, Junho Lee, Changyeon Yoon, Jung Ho Park, Geonho Hwang, and Myungjoo Kang. Do not escape from the manifold: Discovering the local coordinates on the latent space of gans. *arXiv preprint arXiv:2106.06959*, 2021.

Manfredo Perdigao Do Carmo and J Flaherty Francis. *Riemannian geometry*, volume 2. Springer, 1992.

Piotr Dollár, Vincent Rabaud, and Serge Belongie. Learning to traverse image manifolds. *Advances in neural information processing systems*, 19, 2006.

Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

Herbert Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93(3): 418–491, 1959.

Charles Fefferman, Sanjoy Mitter, and Hariharan Narayanan. Testing the manifold hypothesis. *Journal of the American Mathematical Society*, 29(4):983–1049, 2016.

Charles Fefferman, Sergei Ivanov, Yaroslav Kurylev, Matti Lassas, and Hariharan Narayanan. Fitting a putative manifold to noisy data. In *Conference On Learning Theory*, pp. 688–720. PMLR, 2018.

Charles Fefferman, Sergei Ivanov, Yaroslav Kurylev, Matti Lassas, and Hariharan Narayanan. Reconstruction and interpolation of manifolds. i: The geometric whitney problem. *Foundations of Computational Mathematics*, 20(5):1035–1133, 2020.

Christopher R Genovese, Marco Perone-Pacifico, Isabella Verdinelli, and Larry Wasserman. Minimax manifold estimation. *The Journal of Machine Learning Research*, 13(1):1263–1291, 2012.

Luca Grementieri and Rita Fioresi. Model-centric data manifold: The data through the eyes of the model. *SIAM Journal on Imaging Sciences*, 15(3):1140–1156, 2022.

Quanquan Gu, Zhenhui Li, and Jiawei Han. Linear discriminant dimensionality reduction. In *Machine Learning and Knowledge Discovery in Databases: European Conference, ECML PKDD 2011, Athens, Greece, September 5-9, 2011. Proceedings, Part I 11*, pp. 549–564. Springer, 2011.

Kaiming He, Xinlei Chen, Saining Xie, Yanghao Li, Piotr Dollár, and Ross Girshick. Masked autoencoders are scalable vision learners. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 16000–16009, 2022.

Irina Higgins, Loic Matthey, Arka Pal, Christopher Burgess, Xavier Glorot, Matthew Botvinick, Shakir Mohamed, and Alexander Lerchner. beta-vae: Learning basic visual concepts with a constrained variational framework. In *International conference on learning representations*, 2017.

Jeffrey Ho, Ming-Husang Yang, Jongwoo Lim, Kuang-Chih Lee, and David Kriegman. Clustering appearances of objects under varying illumination conditions. In *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, volume 1, pp. I–I. IEEE, 2003.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Christian Horvat and Jean-Pascal Pfister. Denoising normalizing flow. *Advances in neural information processing systems*, 34:9099–9111, 2021.

Daniel N Kaslovsky and François G Meyer. Non-asymptotic analysis of tangent space perturbation. *Information and Inference: a Journal of the IMA*, 3(2):134–187, 2014.

Ilya Kaufman and Omri Azencot. Data representations' study of latent image manifolds. In *International Conference on Machine Learning*, pp. 15928–15945. PMLR, 2023.

Bobak Kiani, Jason Wang, and Melanie Weber. Hardness of learning neural networks under the manifold hypothesis. *Advances in Neural Information Processing Systems*, 37:5661–5696, 2024.

Arlene KH Kim and Harrison H Zhou. Tight minimax rates for manifold estimation under hausdorff loss. 2015.

Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

Diederik P Kingma, Max Welling, et al. An introduction to variational autoencoders. *Foundations and Trends® in Machine Learning*, 12(4):307–392, 2019.

Joseph B Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29(1):1–27, 1964.

David Levin. The approximation power of moving least-squares. *Mathematics of computation*, 67 (224):1517–1531, 1998.

Chunyuan Li, Heerad Farkhoor, Rosanne Liu, and Jason Yosinski. Measuring the intrinsic dimension of objective landscapes. *arXiv preprint arXiv:1804.08838*, 2018.

Didong Li, Minerva Mukhopadhyay, and David B Dunson. Efficient manifold approximation with spherelets. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 84(4):1129–1149, 2022.

Tong Lin and Hongbin Zha. Riemannian manifold learning. *IEEE transactions on pattern analysis and machine intelligence*, 30(5):796–809, 2008.

Yaron Lipman, Daniel Cohen-Or, and David Levin. Error bounds and optimal neighborhoods for mls approximation. In *Proceedings of the fourth Eurographics symposium on Geometry processing*, pp. 71–80, 2006.

Gabriel Loaiza-Ganem, Brendan Leigh Ross, Rasa Hosseinzadeh, Anthony L Caterini, and Jesse C Cresswell. Deep generative models through the lens of the manifold hypothesis: A survey and new connections. *arXiv preprint arXiv:2404.02954*, 2024.

Xingjun Ma, Bo Li, Yisen Wang, Sarah M Erfani, Sudanthi Wijewickrema, Grant Schoenebeck, Dawn Song, Michael E Houle, and James Bailey. Characterizing adversarial subspaces using local intrinsic dimensionality. *arXiv preprint arXiv:1801.02613*, 2018.

Loic Matthey, Irina Higgins, Demis Hassabis, and Alexander Lerchner. dsprites: Disentanglement testing sprites dataset. https://github.com/deepmind/dsprites-dataset/, 2017.

Leland McInnes, John Healy, and James Melville. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.

Kitty Mohammed and Hariharan Narayanan. Manifold learning using kernel density estimation and local principal components analysis. *arXiv preprint arXiv:1709.03615*, 2017.

Sameer A Nene, Shree K Nayar, Hiroshi Murase, et al. Columbia object image library (coil-20). 1996.

Umut Ozertem and Deniz Erdogmus. Locally defined principal curves and surfaces. *The Journal of Machine Learning Research*, 12:1249–1286, 2011.

Phillip Pope, Chen Zhu, Ahmed Abdelkader, Micah Goldblum, and Tom Goldstein. The intrinsic dimension of images and its impact on learning. *arXiv preprint arXiv:2104.08894*, 2021.

Sam T Roweis and Lawrence K Saul. Nonlinear dimensionality reduction by locally linear embedding. *science*, 290(5500):2323–2326, 2000.

H. Murase S.A. Nene, S.K. Nayar. Columbia Object Image Library (COIL-20). In *Technical Report, Department of Computer Science, Columbia University CUCS-005-96*, Feb 1996a.

H. Murase S.A. Nene, S.K. Nayar. Columbia Object Image Library (COIL-100). In *Technical Report, Department of Computer Science, Columbia University CUCS-006-96*, Feb 1996b.

Mahlagha Sedghi, George Atia, and Michael Georgiopoulos. A multi-criteria approach for fast and outlier-aware representative selection from manifolds. *arXiv preprint arXiv:2003.05989*, 2020.

Ali Sekmen and Bahadir Bilgin. Manifold-based approach for neural network robustness analysis. *Communications Engineering*, 3(1):118, 2024.

Hang Shao, Abhishek Kumar, and P Thomas Fletcher. The riemannian geometry of deep generative models. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 315–323, 2018.

Nitesh Shroff, Pavan Turaga, and Rama Chellappa. Manifold precis: An annealing technique for diverse sampling of manifolds. *Advances in Neural Information Processing Systems*, 24, 2011.

Barak Sober and David Levin. Manifold approximation by moving least-squares projection (mmls). *Constructive Approximation*, 52(3):433–478, 2020.

Hong Ye Tan, Subhadip Mukherjee, Junqi Tang, and Carola-Bibiane Schönlieb. Blessing of dimensionality for approximating sobolev classes on manifolds. *arXiv preprint arXiv:2408.06996*, 2024.

Rong Tang and Yun Yang. Adaptivity of diffusion models to manifold structures. In *International Conference on Artificial Intelligence and Statistics*, pp. 1648–1656. PMLR, 2024.

Joshua B Tenenbaum, Vin de Silva, and John C Langford. A global geometric framework for nonlinear dimensionality reduction. *science*, 290(5500):2319–2323, 2000.

Zhigang Yao and Yuqing Xia. Manifold fitting under unbounded noise. *arXiv preprint arXiv:1909.10228*, 2019.

Zhigang Yao and Yuqing Xia. Manifold fitting under unbounded noise. *Journal of Machine Learning Research*, 26(45):1–55, 2025.

Zhigang Yao, Jiaji Su, Bingjie Li, and Shing-Tung Yau. Manifold fitting. *arXiv preprint arXiv:2304.07680*, 2023.

Zhigang Yao, Jiaji Su, and Shing-Tung Yau. Manifold fitting with cyclegan. *Proceedings of the National Academy of Sciences*, 121(5):e2311436121, 2024.

Zhenyue Zhang and Hongyuan Zha. Nonlinear dimension reduction via local tangent space alignment. In *International Conference on Intelligent Data Engineering and Automated Learning*, pp. 477–481. Springer, 2003.

A Expanded Background and Definitions

A.1 Datasets as Manifolds

The manifold hypothesis suggests that high-dimensional datasets often concentrate near low-dimensional manifolds. In this paper, we focus on datasets where the number of intrinsic factors of variation d is small and fixed, and where these factors are explicitly known. Each dataset is modeled as a union of smooth d-dimensional manifolds embedded in $\mathbb{R}^D$, possibly with boundary. Typically, each of the k semantic classes in the dataset corresponds to a separate connected component of this union.

For the purposes of this work, we restrict attention to simple topologies in which each manifold factors into cyclic and non-cyclic directions. Concretely, every manifold is assumed to be homeomorphic to

$$[0,1]^r \times (S^1)^s, \quad r+s=d, \tag{A.1}$$

so that some coordinates vary over a compact interval while others wrap around a circle. This setting captures many commonly used synthetic datasets. A canonical example is *dSprites*, a collection of 64×64 grayscale images of objects (square, ellipse, heart) undergoing controlled transformations

such as scaling, rotation, and translation. These transformations can be viewed as a restricted subset of the special similarity group

$$SSim(n) = \mathbb{R}^n \rtimes (\mathbb{R}^+ \times SO(n)). \tag{A.2}$$

To obtain discrete datasets from this continuous geometric picture, we impose a grid structure. Specifically, we consider

$$G = \left\{ \left(\tfrac{j_1}{n_1}, \dots, \tfrac{j_r}{n_r}, 2\pi \tfrac{j_{r+1}}{n_{r+1}}, \dots, 2\pi \tfrac{j_d}{n_d} \right) \,\middle|\, 0 \le j_\ell < n_\ell \right\}, \tag{A.3}$$

where the first r coordinates sample $[0, 1]$ uniformly and the last s coordinates sample S^1 uniformly. For each class $i \le k$, we define a mapping

$$u_i : G \to M_i \tag{A.4}$$

that associates each grid point with a dataset element obtained by applying the corresponding transformations. The image $u_i[G]$ provides a discrete parametrization of the manifold M_i, and by cutting along cyclic directions one obtains discrete patches of M_i. The complete discretized dataset is then

$$X_G = \bigcup_{i \le k} u_i[G]. \tag{A.5}$$

A.2 Geometric Measures

Critical to the analysis of a manifold are a few geometric quantities that describe its local and global structure. In this work, we focus on the *volume*, the *scalar curvature*, and the *reach*. For completeness, we first recall the definitions of the Riemannian metric and related objects needed to introduce these quantities.

Definition A.1 (Riemannian metric)**.** A Riemannian metric g on a d-dimensional differentiable manifold M assigns to each point $p \in M$ an inner product g_p on the tangent space T_pM. If M is embedded in $\mathbb{R}^D$, the ambient Euclidean metric induces a Riemannian metric by restriction. In local coordinates $u : U \to M$, the metric matrix is

$$g_{ij}(x) = \langle \partial_i u(x), \partial_j u(x) \rangle_{\mathbb{R}^D}. \tag{A.6}$$

Definition A.2 (Volume element and volume)**.** On a Riemannian manifold (M, g) with local coordinates $(u^1, \dots, u^d)$, the natural volume element is

$$d\,\mathrm{Vol} = \sqrt{\det(g)}\, du^1 \wedge \cdots \wedge du^d. \tag{A.7}$$

The volume of a region $R \subset M$ is then

$$\mathrm{Vol}(R) = \int_{u^{-1}(R)} \sqrt{\det(g)}\, dx^1 \cdots dx^d. \tag{A.8}$$

Definition A.3 (Notions of curvature)**.** The curvature of (M, g) is encoded in the Riemann tensor

$$R^i{}_{jkl} = \partial_k \Gamma^i_{jl} - \partial_l \Gamma^i_{jk} + \Gamma^i_{kr} \Gamma^r_{jl} - \Gamma^i_{lr} \Gamma^r_{jk}, \tag{A.9}$$

where the Christoffel symbols are

$$\Gamma^i_{jk} = \tfrac{1}{2} g^{ir} (\partial_j g_{rk} + \partial_k g_{rj} - \partial_r g_{jk}). \tag{A.10}$$

(Contracting yields the Ricci tensor $Ric_{ij} = R^r{}_{irj}$ and finally the scalar curvature

$$R = g^{ij} Ric_{ij}. \tag{A.11}$$

The scalar curvature governs the deviation of small geodesic balls from Euclidean volume:

$$\frac{\mathrm{Vol}(B_{\mathcal{M}}(p, \varepsilon))}{\mathrm{Vol}(B_{\mathbb{R}^d}(0, \varepsilon))} = 1 - \frac{R}{6(n+2)} \varepsilon^2 + \mathcal{O}(\varepsilon^3). \tag{A.12}$$

Definition A.4 (Reach)**.** For a closed set $A \subset \mathbb{R}^D$, the *medial axis* is the set of points with more than one nearest neighbor in A. The *reach* of A is the distance to its medial axis:

$$\tau_A = \inf_{p \in A} d(p, Med(A)). \tag{A.13}$$

Intuitively, τ_A is the largest radius such that every point within distance $< \tau_A$ of A has a unique nearest neighbor in A. For a compact d-dimensional submanifold $M \subset \mathbb{R}^D$, the reach can be expressed as

$$\tau_M = \inf_{p \neq q \in M} \frac{\|p - q\|^2}{2\, d(q - p, T_p M)}, \tag{A.14}$$

where $d(v, T_p M)$ is the distance of a vector v to the tangent space at p.

In our applications we consider both *global* quantities (total volume, scalar curvature integral, global reach) and their *local* counterparts (volume element, pointwise scalar curvature, and local reach estimators as in Aamari et al. Aamari et al. (2019)). These quantities will also serve as the basis for efficient finite-difference approximations introduced later.

Definition A.5 (Hausdorff distance)**.** Let $A, B \subset \mathbb{R}^D$ be two non-empty subsets. The *Hausdorff distance* between A and B is defined as

$$H(A, B) \;=\; \max\Big\{ \sup_{a \in A} \inf_{b \in B} \|a - b\|\,,\; \sup_{b \in B} \inf_{a \in A} \|b - a\| \Big\}. \tag{A.15}$$

In words, $H(A, B)$ measures how far two sets are from coinciding by quantifying the worst-case pointwise deviation: every point of A must lie within distance $H(A, B)$ of some point of B, and vice versa.

This metric is particularly natural for manifold fitting. If M denotes the ground-truth data manifold and $\hat{M}$ a learned approximation (e.g. from an autoencoder), then $H(M, \hat{M})$ expresses the largest geometric error: the maximum distance that a single point on either manifold can have from the other. Thus, a small Hausdorff distance guarantees that the two manifolds are close everywhere, not just on average.

A.3 Manifold Fitting and Bounds

Manifold learning is closely connected with generative modeling: in both cases, one assumes that high-dimensional data lie near a low-dimensional manifold. Generative models typically learn a mapping between the data manifold in $\mathbb{R}^D$ and a latent space of lower dimension, where semantic factors of variation are disentangled and sampling is easier. If the encoder–decoder network is sufficiently regular (locally $\mathcal{C}^1$ or $\mathcal{C}^2$), then composing the encoder with the manifold charts u_i yields a parametrization of the latent manifold. In practice, each data point is mapped to its latent representation, and decoding corresponds to projecting back onto the estimated data manifold.

We focus on two theoretical results that provide bounds on the error of manifold fitting, expressed in terms of the sample size, intrinsic dimension, and geometric quantities of the manifold.

Minimax manifold estimation bound Genovese et al. (2012)

Genovese et al. Genovese et al. (2012) establish minimax rates for estimating a manifold from samples corrupted by small normal noise. The result shows that the intrinsic dimension d alone governs the difficulty of estimation.

Theorem 2 (Genovese et al. Genovese et al. (2012))**.** Let $\mathcal{M}$ be the class of compact, smooth, boundaryless d-dimensional manifolds embedded in $\mathbb{R}^D$. For each $M \in \mathcal{M}$, let Q_M be the distribution of $Y = \xi + Z$, where ξ is uniformly distributed on M and Z is uniform on a normal fiber of radius σ at ξ. For an n-sample estimator $\hat{M} : \mathbb{R}^n \to \mathcal{M}$, define the minimax risk

$$R_n(\mathcal{Q}) = \inf_{\hat{M}} \sup_{Q \in \mathcal{Q}} \mathbb{E}_Q[H(\hat{M}, M)], \tag{A.16}$$

where $H(\cdot, \cdot)$ is the Hausdorff distance. Then there exist constants $C_1, C_2 > 0$ such that

$$C_1 \left(\frac{1}{n}\right)^{\frac{2}{2+d}} \;\leq\; R_n(\mathcal{Q}) \;\leq\; C_2 \left(\frac{\log n}{n}\right)^{\frac{2}{2+d}}. \tag{A.17}$$

This result shows that the sample complexity depends exponentially on the intrinsic dimension d, but not directly on the ambient dimension D. Intuitively, the distributions Q_M correspond to sampling a point on M and perturbing it orthogonally within the reach of the manifold. In our context, an estimator $\hat{M}$ can be thought of as a neural network (e.g., an autoencoder) that outputs a fitted manifold.

TESTING THE MANIFOLD HYPOTHESIS FEFFERMAN ET AL. (2016)

Fefferman et al. Fefferman et al. (2016) provide a testing framework for the manifold hypothesis. Their result connects the number of samples to geometric parameters of the manifold, namely its volume and reach.

Theorem 3 (Fefferman et al. Fefferman et al. (2016))**.** There exists an algorithm which, given samples from a distribution P and $\varepsilon > 0$, distinguishes with probability at least $1 - \delta$ between the following two cases:

- There exists $M \in G(d, CV, \tau/C)$ such that $\mathcal{L}(M, P) \leq C\varepsilon$.
- There does not exist $M \in G(d, V/C, \tau C)$ such that $\mathcal{L}(M, P) \leq \varepsilon/C$.

Here $G(d, V, \tau)$ denotes the class of d-dimensional manifolds of volume $\leq V$ and reach $\geq \tau$. The required sample size is

$$n \;=\; \frac{N_p \,\ln^4(N_p/\varepsilon) + \ln(\delta^{-1})}{\varepsilon^2}, \qquad N_p = V\left(\frac{1}{\tau^d} + \frac{1}{\tau^{d/2}\varepsilon^{d/2}}\right). \tag{A.18}$$

Assuming that such a manifold M exists, the result yields an upper bound on the Hausdorff distance between the true and estimated manifolds:

$$R_n(\mathcal{Q}) \;\leq\; \frac{C_1}{\tau}\left(\frac{V}{n}\right)^{1/d}. \tag{A.19}$$

B DATASET PARAMETER DETAILS

dSprites parameters:

- Image shape: 64×64
- Shape: 3 values (square, ellipse, heart)
- Scale: 16 values linearly spaced in $[0.5, 1]$
- Orientation: 16 values in $[0, 2\pi]$
- Position X: 16 values in $[0, 1]$
- Position Y: 16 values in $[0, 1]$
- Total: 196,608 images (3×16^4)

COIL-20 parameters:

- Image shape: 64×64
- Objects: 20 objects from original COIL-20
- Horizontal orientation: 18 values in $[0, 2\pi]$
- Scale: 16 values linearly spaced in $[0.5, 1]$
- Image orientation: 16 values in $[0, 2\pi]$
- Total: 92,160 images ($20 \times 18 \times 16^2$)