Homotopy Reconstruction via the Cech Complex and the Vietoris-Rips Complex

Jisu Kim

Inria Saclay – Île-de-France, Palaiseau, France
http://pages.saclay.inria.fr/jisu.kim/
jisu.kim@inria.fr

Jaehyeok Shin

Department of Statistics & Data Science, Carnegie Mellon University, Pittsburgh, PA, USA
http://www.stat.cmu.edu/~jaehyeos/
shinjaehyeok@cmu.edu

Frédéric Chazal

Inria Saclay – Île-de-France, Palaiseau, France
https://geometrica.saclay.inria.fr/team/Fred.Chazal/
frederic.chazal@inria.fr

Alessandro Rinaldo

Department of Statistics & Data Science, Carnegie Mellon University, Pittsburgh, PA, USA
http://www.stat.cmu.edu/~arinaldo/
arinaldo@cmu.edu

Larry Wasserman

Department of Statistics & Data Science, Machine Learning Department, Carnegie Mellon University, Pittsburgh, PA, USA
http://www.stat.cmu.edu/~larry/
larry@stat.cmu.edu

Abstract

We derive conditions under which the reconstruction of a target space is topologically correct via the Čech complex or the Vietoris-Rips complex obtained from possibly noisy point cloud data. We provide two novel theoretical results. First, we describe sufficient conditions under which any non-empty intersection of finitely many Euclidean balls intersected with a positive reach set is contractible, so that the Nerve theorem applies for the restricted Čech complex. Second, we demonstrate the homotopy equivalence of a positive μ-reach set and its offsets. Applying these results to the restricted Čech complex and using the interleaving relations with the Čech complex (or the Vietoris-Rips complex), we formulate conditions guaranteeing that the target space is homotopy equivalent to the Čech complex (or the Vietoris-Rips complex), in terms of the μ-reach. Our results sharpen existing results.

2012 ACM Subject Classification Mathematics of computing → Algebraic topology; Theory of computation → Computational geometry

Keywords and phrases Computational topology, Homotopy reconstruction, Homotopy Equivalence, Vietoris-Rips complex, Čech complex, Reach, μ-reach, Nerve Theorem, Offset, Double offset, Consistency

Digital Object Identifier 10.4230/LIPIcs.SoCG.2020.54

Related Version The full version of the paper is available at https://arxiv.org/abs/1903.06955 and https://hal.archives-ouvertes.fr/hal-02425686.

Supplementary Material The code is available at https://github.com/jisuk1/nerveshape.

Funding *Jisu Kim*: Partially supported by Samsung Scholarship.

Acknowledgements We want to thank André Lieutier and Henry Adams for the thoughtful discussions and comments.

1 Introduction

A fundamental task in topological data analysis, geometric inference, and computational geometry is that of estimating the topology of a set $\mathbb{X} \subset \mathbb{R}^d$ based on a finite collection of data points $\mathcal{X}$ that lie in it or in its proximity. This problem naturally occurs in many applications area, such as cosmology [30], time series data [28], machine learning [17], and so on.

A natural way to approximate the target space is to consider an r-offset of the data points, that is, to take the union of the open balls of radius $r > 0$ centered at the data points. Under appropriate conditions, by the Nerve theorem [5] this offset is topologically equivalent to the target space $\mathbb{X}$ via the Čech complex [7, 23]. For computational reasons, the Alpha shape complex may be used instead, which is homotopy equivalent to the Čech complex [20]. To further speed up calculations, and in particular if the data are high dimensional, the Vietoris-Rips complex may be preferable as only the pairwise distances between the data points are used.

To guarantee that the topological approximation based on the data points recovers correctly the homotopy type of $\mathbb{X}$, it is necessary that the data points are dense and close to the target space, and that the radius parameter used for constructing the Čech complex or the Vietoris-Rips complex be of appropriate size.

The conditions require the offset r to be lower bounded by a constant times the Hausdorff distance between the target space and the data points, and upper bounded by another constant times a measure of the size of the topological features of the target space. Originally, the topological feature size was described as a sufficiently small number, for the Vietoris-Rips complex in [24, 25]. Then, the topological feature size was expressed in terms of the reach of $\mathbb{X}$: see, for the Čech complex, in [12, 27]. Subsequently, the notion of μ-reach was put forward to allow for more general target spaces: the condition for the Čech complex is studied in [6, 8], and the condition for the Vietoris-Rips complex is studied in [6]. Also, the radii parameters are allowed to vary across the data points in [12]. For the case when the target space equals the data points, the conditions for the Čech complex or the Vietoris-Rips complex is studied in [3, 4]. When the offset r is beyond the topological feature size so that the homotopy equivalence does not hold, the homotopy type of the Vietoris-Rips complex was studied for the circle in [2].

In this paper, we derive conditions under which the homotopy type of the target space is correctly recovered via the Čech complex or the Vietoris-Rips complex, in terms of the Hausdorff distance and the μ-reach of the target space. To tackle this problem, we provide two novel theoretical results. First, we describe sufficient conditions under which any non-empty intersection of finitely many Euclidean balls intersected with a set of positive reach is contractible, so that the Nerve theorem applies for the restricted Čech complex. Second, we demonstrate the homotopy equivalence of a positive μ-reach set and its offsets. These results are new and of independent interest.

Overall, our new bounds offer significant improvements over the previous results in [27, 6] and are sharp: in particular, they achieve the optimal upper bound for the parameter of the Čech complex and the Vietoris-Rips complex under a positive reach condition. We will provide a detailed comparison of our results with existing ones in Section 6.

2 Background

This section provides a brief introduction to simplicial complex, Nerve theorem, reach, and μ-reach. We refer to Appendix A and [23, 19, 21, 1, 8, 13, 26, 18] for further definitions and details. Throughout the paper, we let $\mathbb{X}$ and $\mathcal{X}$ be subsets of $\mathbb{R}^d$. For $x, y \in \mathbb{R}^d$, we let $d(x, y) := \|x - y\|$ be the Euclidean distance with $\|\cdot\|$ being the Euclidean norm. Let

$d(x,\mathbb{X}) = \inf_{y\in\mathbb{X}} d(x,y)$ denotes the distance from a point x to a set $\mathbb{X}$, and let $d_{\mathbb{X}} : \mathbb{R}^d \to \mathbb{R}$ be the distance function $x \mapsto d(x,\mathbb{X})$. For $r > 0$, we let $\mathbb{B}_{\mathbb{X}}(x,r) := \{y \in \mathbb{X} : d(x,y) < r\}$ be the open restricted ball centered at $x \in \mathbb{R}^d$ of radius $r > 0$. For $r > 0$, we let $\mathbb{X}^r$ be an r-offset of a set $\mathbb{X}$ defined by the collection of all points that are within r distance to $\mathbb{X}$, that is, $\mathbb{X}^r := \bigcup_{x\in\mathbb{X}} \mathbb{B}_{\mathbb{R}^d}(x,r)$. Finally, for two sets $X, Y \subset \mathbb{R}^d$, we let $d_H(X,Y) := \inf\{r > 0 : X \subset Y^r \text{ and } Y \subset X^r\}$ be the Hausdorff distance between X and Y.

2.1 Simplicial complex and Nerve theorem

A natural way to approximate the target space $\mathbb{X}$ with the data points $\mathcal{X}$ is to take the union of open balls centered at the data points. In detail, let $r = \{r_x, x \in \mathcal{X}\} \in \mathbb{R}_+^{\mathcal{X}}$ be pre-specified radii and consider the union of restricted balls

$$\bigcup_{x\in\mathcal{X}} \mathbb{B}_{\mathbb{X}}(x,r_x). \tag{1}$$

Though we allow for the points in $\mathcal{X}$ to lie outside $\mathbb{X}$, we will assume throughout that $\mathbb{B}_{\mathbb{X}}(x,r_x) \neq \emptyset$ for all $x \in \mathcal{X}$.

To infer the topological properties of the union of balls in (1), we rely on a simplicial complex, which can be seen as a high dimensional generalization of a graph. Given a set V, an *(abstract) simplicial complex* is a collection K of finite subsets of V such that $\alpha \in K$ and $\beta \subset \alpha$ implies $\beta \in K$. Each set $\alpha \in K$ is called its *simplex*, and each element of α is called a *vertex* of α.

A simplicial complex encoding the topological properties of the union of balls in (1) is the Čech complex.

▶ **Definition 1** (Čech complex). *Let $\mathcal{X}$, $\mathbb{X}$ be two subsets and $r \in \mathbb{R}_+^{\mathcal{X}}$. The (weighted)* Čech complex *Čech$_{\mathbb{X}}(\mathcal{X},r)$ is the simplicial complex*

$$\check{C}ech_{\mathbb{X}}(\mathcal{X},r) := \left\{ \sigma = \{x_1,\ldots,x_k\} \subset \mathcal{X} : \bigcap_{j=1}^{k} \mathbb{B}_{\mathbb{X}}(x_j, r_{x_j}) \neq \emptyset \right\}. \tag{2}$$

Computing the Čech complex requires computing all possible intersections of the balls. To further speed up the calculation, we only check the pairwise distances between the data points and instead build the Vietoris-Rips complex.

▶ **Definition 2** (Vietoris-Rips complex). *Let $\mathcal{X}$, $\mathbb{X}$ be two subsets and $r \in \mathbb{R}_+^{\mathcal{X}}$. The weighted* Vietoris-Rips complex *Rips$(\mathcal{X},r)$ is the simplicial complex defined as*

$$Rips(\mathcal{X},r) := \left\{\sigma \subset \mathcal{X} : d(x_i,x_j) < r_{x_i} + r_{x_j},\ \textit{for all } x_i,x_j \in \sigma\right\}. \tag{3}$$

The ambient Čech complex in (2) (that is, $\mathbb{X} = \mathbb{R}^d$) and the Vietoris-Rips complex in (3) have the following interleaving relationship when all radii are equal (e.g., see Theorem 2.5 in [16]). That is, when $r_x = r > 0$ for all $x \in \mathcal{X}$, then

$$\check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}(\mathcal{X},r) \subset \mathrm{Rips}(\mathcal{X},r) \subset \check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}\left(\mathcal{X}, \sqrt{\frac{2d}{d+1}}r\right). \tag{4}$$

This interleaving relation is extended to the case of different radii in Lemma 16.

The union of balls in (1) and the Čech complex in (2) are homotopy equivalent under appropriate conditions. This remarkable result is precisely the renowned nerve theorem [5, 7, 23], which we recall next. We first introduce the *nerve*, which is a more abstract notion of the Čech complex.

▶ **Definition 3** (Nerve). *Let $\mathcal{U} = \{U_\alpha\}$ be an open cover of a given topological space $\mathbb{X}$. The nerve of $\mathcal{U}$, denoted by $\mathcal{N}(\mathcal{U})$, is the abstract simplicial complex defined as*

$$\mathcal{N}(\mathcal{U}) = \left\{ \{U_1, \ldots, U_k\} \subset \mathcal{U} : \bigcap_{j=1}^{k} U_j \neq \emptyset \right\}.$$

The nerve theorem prescribes conditions under which the nerve of an open cover of $\mathbb{X}$ is homotopy equivalent to $\mathbb{X}$ itself.

▶ **Theorem 4** (Nerve theorem). *Let $\mathbb{X}$ be a paracompact space and $\mathcal{U}$ be an open cover of $\mathbb{X}$. If every nonempty intersection of finitely many sets in $\mathcal{U}$ is contractible, then $\mathbb{X}$ is homotopy equivalent to the nerve $\mathcal{N}(\mathcal{U})$.*

Thus, in order to conclude that the $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{X}}(\mathcal{X}, r)$ complex in (2) has the same homotopy type as $\mathbb{X}$, it is enough to show, by the nerve theorem, that the union of restricted balls $\bigcup_{x \in \mathcal{X}} \mathbb{B}_{\mathbb{X}}(x, r_x)$ covers the target space $\mathbb{X}$ and that any arbitrary non-empty intersection of restricted balls is contractible. The difficulty in establishing the latter, more technical, condition lies in the fact that it is not clear a priori what properties of $\mathbb{X}$ will imply it. If $\mathbb{X}$ is a convex set, then the nerve theorem applies straightforwardly. But for more general spaces, such as smooth lower-dimensional manifolds, it is not obvious how contractibility may be guaranteed. One of the main results of this paper, given below in Theorem 9, asserts that if $\mathbb{X}$ has positive reach and the radii of the restricted balls are small compared to the reach, then any non-empty intersection of restricted balls is contractible.

2.2 The reach

First introduced by [21], the reach is a quantity expressing the degree of geometric regularity of a set. In detail, given a closed subset $\mathbb{X} \subset \mathbb{R}^d$, the medial axis of $\mathbb{X}$, denoted by $\mathrm{Med}(\mathbb{X})$, is the subset of $\mathbb{R}^d$ consisting of all the points that have at least two nearest neighbors in $\mathbb{X}$. Formally,

$$\mathrm{Med}(\mathbb{X}) = \left\{ x \in \mathbb{R}^d \setminus \mathbb{X} \colon \text{there exist } q_1 \neq q_2 \in \mathbb{X}, ||q_1 - x|| = ||q_2 - x|| = d(x, \mathbb{X}) \right\}, \tag{5}$$

The reach of $\mathbb{X}$ is then defined as the minimal distance from $\mathbb{X}$ to $\mathrm{Med}(\mathbb{X})$.

▶ **Definition 5.** *The reach of a closed subset $\mathbb{X} \subset \mathbb{R}^d$ is defined as*

$$\tau_{\mathbb{X}} = \inf_{q \in \mathbb{X}} d\left(q, \mathrm{Med}(\mathbb{X})\right) = \inf_{q \in \mathbb{X}, x \in \mathrm{Med}(\mathbb{X})} ||q - x||. \tag{6}$$

Some authors [see, e.g. 27, 29] refer to $\tau_{\mathbb{X}}^{-1}$ as the *condition number*. From the definition of the medial axis in (5), the projection $\pi_{\mathbb{X}}(x) = \arg\min_{p \in \mathbb{X}} \|p - x\|$ onto $\mathbb{X}$ is well defined (i.e. unique) outside $Med(\mathbb{X})$. In fact, the reach is the largest distance $\rho \geq 0$ such that $\pi_{\mathbb{X}}$ is well defined on the ρ-offset $\left\{x \in \mathbb{R}^d : d(x, \mathbb{X}) < \rho\right\}$. Hence, assuming the set $\mathbb{X}$ has positive reach can be seen as a generalization or weakening of convexity, since a set $\mathbb{X} \subset \mathbb{R}^d$ is convex if and only if $\tau_{\mathbb{X}} = \infty$. In the next section, we describe how to use the reach condition to ensure that the union of restricted balls is contractible, which in turn allows us to apply the Nerve theorem to recover the homotopy type of the target space $\mathbb{X}$.

For a non-smooth target space, the reach of the space can be zero. In this case, we can deploy a more general notion of feature size, called μ-reach, introduced by [8]. For any point $x \in \mathbb{R}^d \setminus \mathbb{X}$, let $\Gamma_{\mathbb{X}}(x)$ be the set of points in $\mathbb{X}$ closest to x. Let $\Theta_{\mathbb{X}}(x)$ be the center of the

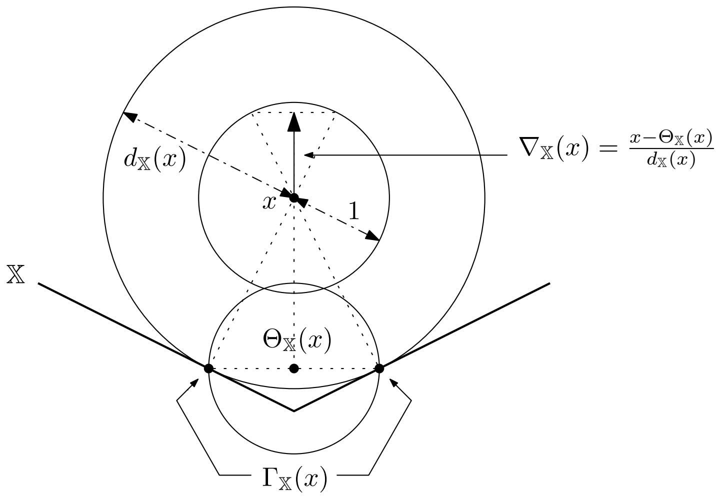

Figure 1 The graphical illustration for the generalized gradient $\nabla_{\mathbb{X}}(x)$, from [9, 8].

unique smallest closed ball enclosing $\Gamma_{\mathbb{X}}(x)$. Then, for $x \in \mathbb{R}^d \setminus \mathbb{X}$, the generalized gradient of the distance function $d_{\mathbb{X}}$ is defined as

$$\nabla_{\mathbb{X}}(x) = \frac{x - \Theta_{\mathbb{X}}(x)}{d_{\mathbb{X}}(x)}, \tag{7}$$

and set $\nabla_{\mathbb{X}}(x) = 0$ for $x \in \mathbb{X}$. See Figure 1 for a graphical illustration. Then, for $\mu \in (0, 1]$, the μ-medial axis of $\mathbb{X}$ is defined as

$$\mathrm{Med}_{\mu}(\mathbb{X}) = \left\{x \in \mathbb{R}^d \setminus \mathbb{X} : \|\nabla_{\mathbb{X}}(x)\| < \mu\right\}, \tag{8}$$

Finally, the μ-reach of $\mathbb{X}$ is defined as the minimal distance from $\mathbb{X}$ to $\mathrm{Med}_{\mu}(\mathbb{X})$.

▶ **Definition 6.** *The μ-reach of a closed subset $\mathbb{X} \subset \mathbb{R}^d$ is defined as*

$$\tau_{\mathbb{X}}^{\mu} = \inf_{q \in \mathbb{X}} d\left(q, \mathrm{Med}_{\mu}(\mathbb{X})\right) = \inf_{q \in \mathbb{X}, x \in \mathrm{Med}_{\mu}(\mathbb{X})} ||q - x||. \tag{9}$$

Note that if $\mu = 1$, the corresponding μ-reach equals to the reach of $\mathbb{X}$.

Two offsets $\mathbb{X}^r$ and $\mathbb{X}^s$ of the target space $\mathbb{X}$ are topologically equivalent if they are free of critical points of the distance function $d_{\mathbb{X}}$ in the sense specified below (see e.g., [22] or Proposition 3.4 in [11]).

▶ **Lemma 7** (Isotopy Lemma). *Let $\mathbb{X} \subset \mathbb{R}^d$ be a set, and for $r, s > 0$ with $s \leq r$, let $\mathbb{X}^r$ and $\mathbb{X}^s$ be two offsets of $\mathbb{X}$. Suppose the distance function $d_{\mathbb{X}}$ does not have a critical point on $\overline{\mathbb{X}^r} \backslash \mathbb{X}^s$, that is, $\nabla_{\mathbb{X}}(x) \neq 0$ for all $x \in \overline{\mathbb{X}^r} \backslash \mathbb{X}^s$ where $\nabla_{\mathbb{X}}$ is from* (7)*. Then $\mathbb{X}^r$ and $\mathbb{X}^s$ are homeomorphic.*

Note that requiring $\nabla_{\mathbb{X}}(x) \neq 0$ for all $x \in \overline{\mathbb{X}^r} \backslash \mathbb{X}^s$ is weaker than the μ-reach condition $\tau_{\mathbb{X}}^{\mu} > r$ for any $\mu \in (0, 1]$. One of the main results of the paper, given in Theorem 12, generalizes this topological relation to the relation between the target space and its offset under a stronger positive μ-reach condition.

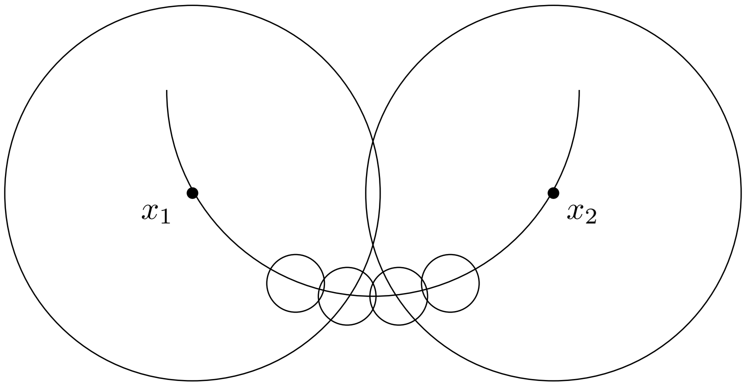

Figure 2 *An example in which the union of balls is different from the underlying space in terms of the homotopy. In the figure, the union of balls deformation retracts to a circle, hence its homotopy is different from the underlying semicircle.*

2.3 Restricted versus Ambient balls

It is important to point out that the nerve theorem needs not to be applied to the Čech complex built using ambient, as opposed, to restricted balls. In particular, the homotopy type of $\mathbb{X}$, may not be correctly recovered using unions of ambient balls even if the point cloud is dense in $\mathbb{X}$ and the radii of the balls all vanish. We elucidate this point in the next example. Below, $\mathbb{B}_{\mathbb{R}^d}(x, r)$ denotes the open ambient ball in $\mathbb{R}^d$ centered at x and of radius $r > 0$.

▶ **Example 8.** Let $\mathbb{X} = (\partial\mathbb{B}_{\mathbb{R}^2}(0, 1)) \cap \{x \in \mathbb{R}^2 : x_2 \geq 0\}$ be a semicircle in $\mathbb{R}^2$. Let $\epsilon \in (0, 1)$ be fixed, and x_1, x_2 be points on $\mathbb{X}$ satisfying $\|x_1 - x_2\| \in \left(\epsilon\sqrt{4-\epsilon^2}, 2\epsilon\right)$. Then, $\mathbb{B}_{\mathbb{R}^2}(x_1, \epsilon) \cap \mathbb{B}_{\mathbb{R}^2}(x_2, \epsilon)$ is nonempty but has an empty intersection with $\mathbb{X}$. Now, choose $\rho < d(\mathbb{X}, \mathbb{B}_{\mathbb{R}^2}(x_1, \epsilon) \cap \mathbb{B}_{\mathbb{R}^2}(x_2, \epsilon))$ and choose $\mathcal{X}_0 \subset \mathbb{X}$ be dense enough so that $\bigcup_{x\in\mathcal{X}_0} \mathbb{B}_{\mathbb{R}^2}(x, \rho)$ covers $\mathbb{X}$. Now, consider the union of ambient balls

$$\left(\mathbb{B}_{\mathbb{R}^2}(x_1, \epsilon) \bigcup \mathbb{B}_{\mathbb{R}^2}(x_2, \epsilon)\right) \bigcup \left(\bigcup_{x\in\mathcal{X}_0} \mathbb{B}_{\mathbb{R}^2}(x, \rho)\right). \tag{10}$$

Then from the fact $\rho < d\left(\mathbb{X}, \mathbb{B}_{\mathbb{R}^2}(x_1, \epsilon) \cap \mathbb{B}_{\mathbb{R}^2}(x_2, \epsilon)\right)$ and $\bigcup_{x\in\mathcal{X}_0} \mathbb{B}_{\mathbb{R}^2}(x, \rho)$ is a covering of $\mathbb{X}$, we have that the union of balls in (10) is homotopy equivalent to a circle, hence its homotopy is different from the semicircle $\mathbb{X}$. Note that the above construction holds for all choices of $\epsilon \in (0, 1)$. Since $\rho \to 0$ as $\epsilon \to 0$, $\mathcal{X}_0$ can be arbitrary dense in $\mathbb{X}$. See Figure 2.

3 The nerve theorem for Euclidean sets of positive reach

In order to apply the nerve theorem to the Čech complex built on restricted balls, it is enough to check whether any finite intersection of the restricted balls $\bigcap_{j=1}^{k} \mathbb{B}_{\mathbb{X}}(x_j, r_{x_j})$ is contractible (since $\mathbb{X}$ is a subset of $\mathbb{R}^d$ and is endowed with the subspace topology, it is paracompact.).

Theorem 9 is one of the main statements of this paper and shows that, if a subset $\mathbb{X} \subset \mathbb{R}^d$ has a positive reach $\tau > 0$, any non-empty intersection of restricted balls is contractible if the radii are small enough compared to τ.

▶ **Theorem 9.** *Let $\mathbb{X} \subset \mathbb{R}^d$ be a subset with reach $\tau > 0$ and let $\mathcal{X} \subset \mathbb{R}^d$ be a set of points. Let $\{r_x > 0 : x \in \mathcal{X}\}$ be a set of radii indexed by $x \in \mathcal{X}$. Then, if $r_x \leq \sqrt{\tau^2 + (\tau - d_{\mathbb{X}}(x))^2}$ for all $x \in \mathcal{X}$, any nonempty intersection of restricted balls $\bigcap_{x\in I} \mathbb{B}_{\mathbb{X}}(x, r_x)$ for $I \subset \mathcal{X}$ is contractible.*

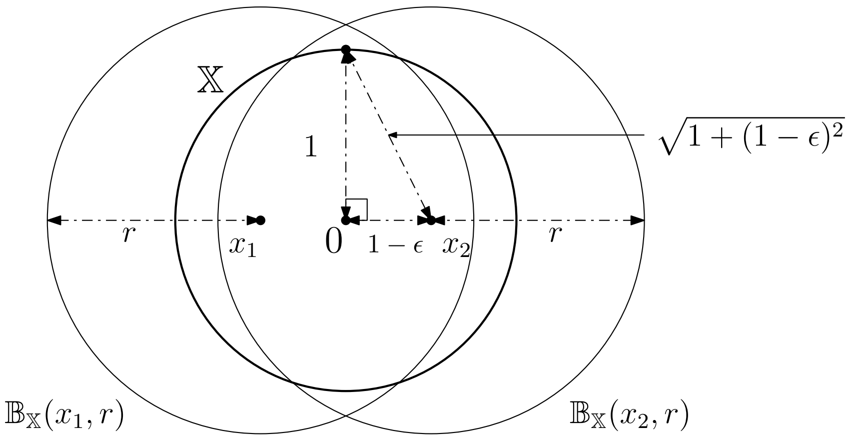

Figure 3 *An example in which $\mathbb{B}_{\mathbb{X}}(x_1, r) \bigcup \mathbb{B}_{\mathbb{X}}(x_2, r)$ is not homotopy equivalent to $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{X}}(\mathcal{X}, r)$ where $\mathbb{X} = \left\{x \in \mathbb{R}^2 : \|x\|_2 = 1\right\}$, $x_1 = (-1+\epsilon, 0)$, $x_2 = (1-\epsilon, 0)$, $\mathcal{X} = \{x_1, x_2\}$, and $r > \sqrt{1+(1-\epsilon)^2}$, for any $\epsilon > 0$.*

Therefore, by combining Theorem 9 and the Nerve Theorem (Theorem 4), we can establish that the topology of the subspace $\mathbb{X}$ can be recovered by the corresponding restricted Čech complex $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{X}}(\mathcal{X}, r)$, provided the radii of the balls are not too large with respect to the reach. This result is summarized in the following corollary.

▶ **Corollary 10** (Nerve Theorem on the restricted balls)**.** *Under the same condition of Theorem 9, suppose $r_x \leq \sqrt{\tau^2 + (\tau - d_{\mathbb{X}}(x))^2}$ for all $x \in \mathcal{X}$, then the union of restricted balls $\bigcup_{x \in \mathcal{X}} \mathbb{B}_{\mathbb{X}}(x, r_x)$ is homotopy equivalent to the restricted Čech complex $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{X}}(\mathcal{X}, r)$. If, in addition, the union of restricted balls covers the target space $\mathbb{X}$, that is,*

$$\mathbb{X} \subset \bigcup_{x \in \mathcal{X}} \mathbb{B}_{\mathbb{X}}(x, r_x), \tag{11}$$

then $\mathbb{X}$ is homotopy equivalent to the restricted Čech complex $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{X}}(\mathcal{X}, r)$.

The reach condition $r_x \leq \sqrt{\tau^2 + (\tau - d_{\mathbb{X}}(x))^2}$ is tight as the following example shows.

▶ **Example 11.** Let $\mathbb{X}$ be the unit Euclidean sphere in $\mathbb{R}^d$, and fix $\epsilon > 0$. Let $x_1 := (1-\epsilon, 0, \ldots, 0), x_2 := (-1+\epsilon, 0, \ldots, 0) \in \mathbb{R}^d$, and set $\mathcal{X} := \{x_1, x_2\}$. For a unit Euclidean sphere, the reach is equal to its radius 1. Therefore, if $r = (r_1, r_2) \in \left(0, \sqrt{1+(1-\epsilon)^2}\right]^2$ then $\mathbb{B}_{\mathbb{X}}(x_1, r_1) \bigcup \mathbb{B}_{\mathbb{X}}(x_2, r_2)$ is homotopy equivalent to $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{X}}(\mathcal{X}, r)$ by Corollary 10. However, if $r_1, r_2 > \sqrt{1+(1-\epsilon)^2}$, $\mathbb{B}_{\mathbb{X}}(x_1, r_1) \bigcup \mathbb{B}_{\mathbb{X}}(x_2, r_2) \simeq \mathbb{X}$ but $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{X}}(\mathcal{X}, r) \simeq 0$. Figure 3 illustrates the 2-dimensional case.

4 Deformation retraction on positive μ-reach

The positive reach condition is critical for the nerve theorem on the restricted Čech complex. However, it is not easily generalized to the positive μ-reach condition. Instead, we find a positive reach set that approximates the positive μ-reach set. And to show their homotopy equivalence, we discover the topological relation between the positive μ-reach set and its offset.

The homeomorphic relation between two offsets $\mathbb{X}^r$ and $\mathbb{X}^s$ of the target space $\mathbb{X}$ in Lemma 7 does not hold in general between the target space and its offset, but a weakened topological relation holds under a stronger condition on the target space. Theorem 12, which

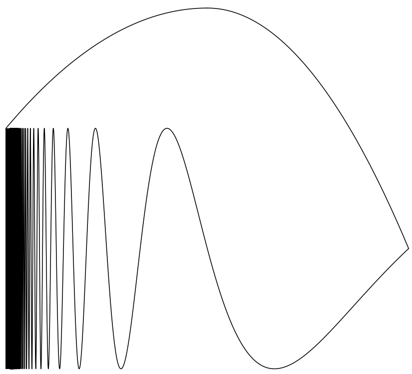

Figure 4 *An example where $\mathbb{X}^r$ does not deformation retracts to $\mathbb{X}$. $\mathbb{X}$ is a topologist's sine circle, that is, $\mathbb{X} = \mathbb{X}_0 \cup \mathbb{X}_1 \cup \mathbb{X}_2$, with $\mathbb{X}_0 = \left\{\left(x, \sin\frac{\pi}{x}\right) \in \mathbb{R}^2 : x \in [0,1]\right\}$, $\mathbb{X}_1 = \{0\} \times [-1,1]$, and $\mathbb{X}_2$ is a sufficiently smooth curve joining $(0,1)$ and $(1,0)$ and meeting $\mathbb{X}_0 \cup \mathbb{X}_1$ only at its endpoints.*

is one of the main results in our paper, asserts that if the target space $\mathbb{X}$ has a positive μ-reach, then the offset $\mathbb{X}^r$ deformation retracts to $\mathbb{X}$ when the offset size is not large, and in particular, they are homotopy equivalent.

▶ **Theorem 12.** *Let $\mathbb{X} \subset \mathbb{R}^d$ be a subset with positive μ-reach $\tau^\mu > 0$. For $r \leq \tau^\mu$, the r-offset $\mathbb{X}^r$ deformation retracts to $\mathbb{X}$. In particular, $\mathbb{X}$ and $\mathbb{X}^r$ are homotopy equivalent.*

The positive μ-reach condition $r \leq \tau^\mu$ in Theorem 12 is critical and cannot be weakened to $\nabla_{\mathbb{X}}(x) \neq 0$ for all $x \in \overline{\mathbb{X}^r} \backslash \mathbb{X}$ as in Lemma 7. Indeed, Example 13 shows that the offset does not deformation retract to the target space although $\nabla_{\mathbb{X}}(x) \neq 0$ for all $x \in \mathbb{R}^d$.

▶ **Example 13.** Let $\mathbb{X} \subset \mathbb{R}^2$ be a topologist's sine circle, that is, $\mathbb{X} = \mathbb{X}_0 \cup \mathbb{X}_1 \cup \mathbb{X}_2$, with $\mathbb{X}_0 = \left\{\left(x, \sin\frac{\pi}{x}\right) \in \mathbb{R}^2 : x \in [0,1]\right\}$, $\mathbb{X}_1 = \{0\} \times [-1,1]$, and $\mathbb{X}_2$ is a sufficiently smooth curve joining $(0,1)$ and $(1,0)$ and meets $\mathbb{X}_0 \cup \mathbb{X}_1$ only at its endpoints. See Figure 4. Then, $\tau_{\mathbb{X}}^\mu = 0$ for any $\mu \in (0,1]$ but $\nabla_{\mathbb{X}}$ is nonzero for all $x \in \mathbb{R}^2 \backslash \mathbb{X}$. Now, $H_1(\mathbb{X}) = 0$, but for any sufficiently small $r > 0$, $\mathbb{X}^r$ is homeomorphic to an annulus $\mathbb{B}_{\mathbb{R}^2}(0,2) \backslash \overline{\mathbb{B}_{\mathbb{R}^2}(0,1)}$ and hence $H_1(\mathbb{X}^r) = \mathbb{Z}$. Hence $\mathbb{X}^r$ cannot deformation retract to $\mathbb{X}$.

Using Theorem 12, we find a positive reach set that approximates the positive μ-reach set. The set we will use is the double offset [9]. Recall that, for $r > 0$, an r-offset $\mathbb{X}^r$ of a set $\mathbb{X}$ is the collection of all points that are within r distance to $\mathbb{X}$, that is, $\mathbb{X}^r := \bigcup_{x \in \mathbb{X}} \mathbb{B}_{\mathbb{R}^d}(x, r)$. The double offset is to take offset, take complement, take offset, and take complement, that is, for $s \geq t > 0$, $\mathbb{X}^{s,t} := (((\mathbb{X}^s)^{\complement})^t)^{\complement}$. Roughly speaking, it is to inflate your set first, and then deflate your set, so that sharp corners become smooth. See [9] for more details. To set up the homotopy equivalence of the positive μ-reach set and its double offset, we need another tool for finding the homotopy equivalence of the complement set. This is done in the next lemma.

▶ **Lemma 14.** *Let $\mathbb{X} \subset \mathbb{R}^d$ be a subset with positive reach $\tau > 0$. For $r \leq \tau$, $\mathbb{X}^{\complement}$ deformation retracts to $(\mathbb{X}^r)^{\complement}$. In particular, $\mathbb{X}^{\complement}$ and $(\mathbb{X}^r)^{\complement}$ are homotopy equivalent.*

Now, combining Theorem 12 and Lemma 14 gives the desired homotopy equivalence between the target set of positive μ-reach and its double offset, where the double offset has a positive reach.

▶ **Corollary 15.** *Let $\mathbb{X} \subset \mathbb{R}^d$ be a subset with positive μ-reach $\tau^\mu > 0$. For $s, t > 0$ with $t \leq s$, let $\mathbb{X}^{s,t} := (((\mathbb{X}^s)^\complement)^t)^\complement$ be the double offset of $\mathbb{X}$. If $s < \tau^\mu$ and $t < \mu s$, then $\mathbb{X}^{s,t}$ and $\mathbb{X}$ are homotopy equivalent, and the reach of $\mathbb{X}^{s,t}$ is greater than or equal to t, that is, $\tau_{\mathbb{X}^{s,t}} \geq t$.*

5 Homotopy Reconstruction via Cech complex and Vietoris-Rips complex

Next, we derive conditions under which the homotopy type of the target space is correctly recovered via the Čech complex and the Vietoris-Rips complex. We first extend the interleaving relationship of the ambient Čech complex and the Vietoris-Rips complex in (4) to the different radii case in Lemma 16.

▶ **Lemma 16.** *Let $\mathcal{X} \subset \mathbb{R}^d$ be a set of points and $r = \{r_x > 0 : x \in \mathcal{X}\}$ be a set of radii indexed by $x \in \mathcal{X}$. Then,*

$$\check{C}ech_{\mathbb{R}^d}(\mathcal{X}, r) \subset Rips(\mathcal{X}, r) \subset \check{C}ech_{\mathbb{R}^d}\left(\mathcal{X}, \sqrt{\frac{2d}{d+1}}r\right).$$

To recover the homotopy of the target set via the ambient Čech complex and the Vietoris-Rips complex, we utilize the restricted Čech complex. Hence, we set up the interleaving relationship between the restricted Čech complex and the ambient Čech complex in Lemma 17 and between the restricted Čech complex and the Vietoris-Rips complex in Corollary 18.

▶ **Lemma 17.** *Let $\mathbb{X} \subset \mathbb{R}^d$ be a subset with reach $\tau > 0$ and let $\mathcal{X} \subset \mathbb{R}^d$ be a set of points. Let $r = \{r_x > 0 : x \in \mathcal{X}\}$ be a set of radii indexed by $x \in \mathcal{X}$. Then,*

$$\check{C}ech_{\mathbb{X}}(\mathcal{X}, r) \subset \check{C}ech_{\mathbb{R}^d}(\mathcal{X}, r) \subset \check{C}ech_{\mathbb{X}}(\mathcal{X}, r'),$$

where $r' = \{r'_x > 0 : x \in \mathcal{X}\}$ with

$$r'_x = \sqrt{\frac{2\tau\left(r_x^2 + d_{\mathbb{X}}(x)\left(2\tau - d_{\mathbb{X}}(x)\right)\right)}{\tau + \sqrt{\tau^2 - \left(r_x^2 + d_{\mathbb{X}}(x)\left(2\tau - d_{\mathbb{X}}(x)\right)\right)}} - d_{\mathbb{X}}(x)\left(2\tau - d_{\mathbb{X}}(x)\right)}.$$

Equivalently,

$$\check{C}ech_{\mathbb{R}^d}(\mathcal{X}, r'') \subset \check{C}ech_{\mathbb{X}}(\mathcal{X}, r) \subset \check{C}ech_{\mathbb{R}^d}(\mathcal{X}, r),$$

where $r'' = \{r''_x > 0 : x \in \mathcal{X}\}$ with

$$r''_x = \sqrt{\tau^2 - d_{\mathbb{X}}(x)(2\tau - d_{\mathbb{X}}(x)) - \frac{(2\tau^2 - r_x^2 - d_{\mathbb{X}}(x)(2\tau - d_{\mathbb{X}}(x)))^2}{4\tau^2}}.$$

▶ **Corollary 18.** *Let $\mathbb{X} \subset \mathbb{R}^d$ be a subset with reach $\tau > 0$ and let $\mathcal{X} \subset \mathbb{R}^d$ be a set of points. Let $r = \{r_x > 0 : x \in \mathcal{X}\}$ be a set of radii indexed by $x \in \mathcal{X}$. Then,*

$$\check{C}ech_{\mathbb{X}}(\mathcal{X}, r) \subset Rips(\mathcal{X}, r) \subset \check{C}ech_{\mathbb{X}}(\mathcal{X}, r'''),$$

where $r''' = \{r'''_x > 0 : x \in \mathcal{X}\}$ with

$$r'''_x = \sqrt{\frac{2\tau\left(\frac{2d}{d+1}r_x^2 + d_{\mathbb{X}}(x)\left(2\tau - d_{\mathbb{X}}(x)\right)\right)}{\tau + \sqrt{\tau^2 - \left(\frac{2d}{d+1}r_x^2 + d_{\mathbb{X}}(x)\left(2\tau - d_{\mathbb{X}}(x)\right)\right)}} - d_{\mathbb{X}}(x)\left(2\tau - d_{\mathbb{X}}(x)\right)}.$$

Combining Nerve Theorem on the restricted balls (Corollary 10) with the covering condition (11) and Lemma 17 or Corollary 18 gives the following commutative diagram:

$$\begin{tikzcd}
& \mathbb{X} \arrow[dl] & \\
\check{\mathrm{C}}\mathrm{ech}_{\mathbb{X}}(\mathcal{X},r) \arrow[rr] \arrow[dr] & & \check{\mathrm{C}}\mathrm{ech}_{\mathbb{X}}(\mathcal{X},r'''') \arrow[ul] \\
& \mathcal{S} \arrow[ur] &
\end{tikzcd} , \tag{12}$$

where $\mathcal{S}$ is either the ambient Čech complex $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}(\mathcal{X},r)$ or the Vietoris-Rips complex $\mathrm{Rips}(\mathcal{X},r)$. Using this diagram, we develop the homotopy equivalence between the target space and either the ambient Čech complex or the Vietoris-Rips complex. First, Theorem 19 asserts that when the target space of positive reach is densely covered by the data points and if they are not too far apart, the ambient Čech complex can be used to recover the homotopy type.

▶ **Theorem 19.** *Let $\mathbb{X} \subset \mathbb{R}^d$ be a subset with reach $\tau > 0$ and let $\mathcal{X} \subset \mathbb{R}^d$ be a closed discrete set of points. Let $\{r_x > 0 : x \in \mathcal{X}\}$ be a set of radii indexed by $x \in \mathcal{X}$ with $r_{\min} := \min_{x\in\mathcal{X}}\{r_x\}$ and $r_{\max} := \max_{x\in\mathcal{X}}\{r_x\}$, and let $\epsilon := \max\{d_{\mathbb{X}}(x) : x \in \mathcal{X}\}$. Suppose $\mathbb{X}$ is covered by the union of balls centered at $x \in \mathcal{X}$ and radius δ as*

$$\mathbb{X} \subset \bigcup_{x\in\mathcal{X}} \mathbb{B}_{\mathbb{R}}(x,\delta). \tag{13}$$

Suppose that the maximum radius $r_{\max}$ is bounded as

$$r_{\max} \leq \tau - \epsilon. \tag{14}$$

Also, suppose δ satisfies the following condition:

$$\begin{aligned}
&\delta + \sqrt{r_{\max}^2 - \tilde{l}^2 + \epsilon(2\tau-\epsilon) - ((\tau-\epsilon)^2 - r_{\max}^2 + \tilde{l}^2 + (\tau-\epsilon_{\tilde{l}})^2)\left(\frac{\tau}{\sqrt{\tau^2-\tilde{r}_{\delta,c}}}-1\right)} \\
&\leq r_{\min}, \\
&\sqrt{\frac{d}{2(d+1)}}\frac{r_{\max}}{r_{\min}}\left(\sqrt{\tilde{r}_b^2 - (2\tau^2-\tilde{r}_b^2)\left(\frac{\tau}{\sqrt{\tau^2-\tilde{r}_{\delta,b}^2}}-1\right)}+2\delta\right) \leq r''_{\min},
\end{aligned} \tag{15}$$

$$\begin{aligned}
&\tilde{l} := \frac{1}{2}\left(r_{\min} - \tau + \sqrt{(\tau-\epsilon)^2 - r_{\max}^2} - \delta\right), \qquad \epsilon_{\tilde{l}} := \tau - \sqrt{(\tau-\epsilon)^2 - r_{\max}^2} + \tilde{l}, \\
&\tilde{r}_{\delta,c}^2 := \min\left\{\delta^2 + \epsilon(2\tau-\epsilon), \frac{1}{2}(r_{\max}^2 - \tilde{l}^2 + \epsilon(2\tau-\epsilon) + \epsilon_{\tilde{l}}(2\tau-\epsilon_{\tilde{l}}))\right\}, \\
&r''_{\min} := \sqrt{\tau^2 - \epsilon(2\tau-\epsilon) - \frac{(2\tau^2 - r_{\min}^2 - \epsilon(2\tau-\epsilon))^2}{4\tau^2}}, \\
&\tilde{r}_b^2 := \frac{2\tau\left((r''_{\min})^2 + \epsilon(2\tau-\epsilon)\right)}{\tau + \sqrt{\tau^2 - ((r''_{\min})^2 + \epsilon(2\tau-\epsilon))}}, \qquad \tilde{r}_{\delta,b}^2 := \min\left\{\delta^2 + \epsilon(2\beta-\epsilon), \frac{1}{2}\tilde{r}_b^2\right\}.
\end{aligned}$$

Then $\mathbb{X}$ is homotopy equivalent to the ambient Čech complex $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}(\mathcal{X},r)$.

A similar approach also gives the homotopy equivalence between the target space and the Vietoris-Rips complex when the target space has positive reach.

▶ **Theorem 20.** *Let $\mathbb{X} \subset \mathbb{R}^d$ be a subset with reach $\tau > 0$ and let $\mathcal{X} \subset \mathbb{R}^d$ be a closed discrete set of points. Let $\{r_x > 0 : x \in \mathcal{X}\}$ be a set of radii indexed by $x \in \mathcal{X}$ with $r_{\min} := \min_{x\in\mathcal{X}}\{r_x\}$ and $r_{\max} := \max_{x\in\mathcal{X}}\{r_x\}$, and let $\epsilon := \max\{d_{\mathbb{X}}(x) : x \in \mathcal{X}\}$. Suppose $\mathbb{X}$ is covered by the union of balls centered at $x \in \mathcal{X}$ and radius δ as*

$$\mathbb{X} \subset \bigcup_{x\in\mathcal{X}} \mathbb{B}_{\mathbb{R}}(x,\delta). \tag{16}$$

Suppose that the maximum radius $r_{\max}$ is bounded as

$$r_{\max} \leq \sqrt{\frac{d+1}{2d}}\left(\tau - \epsilon\right). \tag{17}$$

Also, suppose δ satisfies the following condition:

$$\begin{gathered}\sqrt{\tilde{r}_b^2(r_{\max}) - (2\tau^2 - \tilde{r}_b^2(r_{\max}))\left(\frac{\tau}{\sqrt{\tau^2 - \tilde{r}_{\delta,b}^2(r_{\max})}} - 1\right)} + 2\delta \leq 2r_{\min},\\ \sqrt{\frac{d}{2(d+1)}}\left(\sqrt{\tilde{r}_b^2(r''_{\min}) - (2\tau^2 - \tilde{r}_b^2(r''_{\min}))\left(\frac{\tau}{\sqrt{\tau^2 - \tilde{r}_{\delta,b}^2(r''_{\min})}} - 1\right)} + 2\delta\right) \leq r''_{\min},\end{gathered} \tag{18}$$

where

$$r''_{\min} := \sqrt{\tau^2 - \epsilon(2\tau - \epsilon) - \frac{(2\tau^2 - r_{\min}^2 - \epsilon(2\tau-\epsilon))^2}{4\tau^2}},$$

$$\tilde{r}_b^2(t) := \frac{2\tau\left(t^2 + \epsilon(2\tau - \epsilon)\right)}{\tau + \sqrt{\tau^2 - (t^2 + \epsilon(2\tau-\epsilon))}}, \qquad \tilde{r}_{\delta,b}^2(t) := \min\left\{\delta^2 + \epsilon(2\tau-\epsilon), \frac{1}{2}\tilde{r}_b^2(t)\right\}.$$

Then $\mathbb{X}$ is homotopy equivalent to the Vietoris-Rips complex Rips$(\mathcal{X}, r)$.

▶ Remark 21. Compared to the restricted Čech complex (Corollary 10), the covering condition in (13) or (16) is more critical for the ambient Čech complex (Theorem 19) or the Vietoris-Rips complex (Theorem 20). Although the restricted Čech complex $\text{Čech}_{\mathbb{X}}(\mathcal{X}, r)$ is still homotopy equivalent to the union of restricted balls $\bigcup_{x\in\mathcal{X}} \mathbb{B}_{\mathbb{X}}(x, r_x)$ without the covering condition in (11), such homotopy equivalence does not hold for the ambient Čech complex or the Vietoris-Rips complex. This is since the upper triangle of the diagram in (12) only holds under the covering condition in (13) or (16). Furthermore, the covering condition in (13) or (16) is denser in that $\delta < r_x$ for all $x \in \mathcal{X}$, to construct an additional homotopy equivalence on the lower triangle of the diagram in (12).

The homotopy equivalences in Theorem 19 and 20 for the positive reach case is extended to the positive μ-reach case by applying Corollary 15 with the double offset of the target space. Corollary 22 shows that when the double offset of the target space of positive μ-reach is densely covered by the data points and if they are not too far apart, either the ambient Čech complex or the Vietoris-Rips complex can be used to recover the homotopy type of $\mathbb{X}$.

▶ **Corollary 22.** *Let $\mathbb{X} \subset \mathbb{R}^d$ be a subset with positive μ-reach $\tau^\mu > 0$ and let $\mathcal{X} \subset \mathbb{R}^d$ be a closed discrete set of points. Let $\{r_x > 0 : x \in \mathcal{X}\}$ be a set of radii indexed by $x \in \mathcal{X}$ with $r_{\min} := \min_{x\in\mathcal{X}}\{r_x\}$ and $r_{\max} := \max_{x\in\mathcal{X}}\{r_x\}$. Let $s, t, \epsilon \geq 0$ with $\frac{t}{\mu} < s < \tau^\mu$, and let $\mathbb{Y} := (((\mathbb{X}^s)^\complement)^t)^\complement$ be the double offset, with $d_{\mathbb{Y}}(x) \leq \epsilon$ for all $x \in \mathcal{X}$. Suppose $\mathbb{Y}$ is covered by the union of balls centered at $x \in \mathcal{X}$ and radius δ as*

$$\mathbb{Y} \subset \bigcup_{x\in\mathcal{X}} \mathbb{B}_{\mathbb{R}}(x, \delta).$$

(i) *Suppose $r_{\max} \leq t - \epsilon$, and δ satisfies the following condition:*

$$\delta + \sqrt{r_{\max}^2 - \tilde{l}^2 + \epsilon(2t-\epsilon) - ((t-\epsilon)^2 - r_{\max}^2 + \tilde{l}^2 + (t-\epsilon_{\tilde{l}})^2)\left(\frac{t}{\sqrt{t^2 - \tilde{r}_{\delta,c}}} - 1\right)}$$
$$\leq r_{\min},$$
$$\sqrt{\frac{d}{2(d+1)}}\frac{r_{\max}}{r_{\min}}\left(\sqrt{\tilde{r}_b^2 - (2t^2 - \tilde{r}_b^2)\left(\frac{t}{\sqrt{t^2 - \tilde{r}_{\delta,b}^2}} - 1\right)} + 2\delta\right) \leq r''_{\min},$$

where

$$\tilde{l} := \frac{1}{2}\left(r_{\min} - t + \sqrt{(t-\epsilon)^2 - r_{\max}^2} - \delta\right), \qquad \epsilon_{\tilde{l}} := t - \sqrt{(t-\epsilon)^2 - r_{\max}^2} + \tilde{l},$$
$$\tilde{r}_{\delta,c}^2 := \min\left\{\delta^2 + \epsilon(2t-\epsilon), \frac{1}{2}(r_{\max}^2 - \tilde{l}^2 + \epsilon(2t-\epsilon) + \epsilon_{\tilde{l}}(2t - \epsilon_{\tilde{l}}))\right\},$$
$$r''_{\min} := \sqrt{t^2 - \epsilon(2t-\epsilon) - \frac{(2t^2 - r_{\min}^2 - \epsilon(2t-\epsilon))^2}{4t^2}},$$
$$\tilde{r}_b^2 := \frac{2t\left((r''_{\min})^2 + \epsilon(2t-\epsilon)\right)}{t + \sqrt{t^2 - ((r''_{\min})^2 + \epsilon(2t-\epsilon))}}, \qquad \tilde{r}_{\delta,b}^2 := \min\left\{\delta^2 + \epsilon(2t-\epsilon), \frac{1}{2}\tilde{r}_b^2\right\}.$$

Then $\mathbb{X}$ is homotopy equivalent to the ambient Čech complex Čech$_{\mathbb{R}^d}(\mathcal{X}, r)$.

(ii) *Suppose $r_{\max} \leq \sqrt{\frac{d+1}{2d}}(t-\epsilon)$, and δ satisfies the following condition:*

$$\sqrt{\tilde{r}_b^2(r_{\max}) - (2t^2 - \tilde{r}_b^2(r_{\max}))\left(\frac{t}{\sqrt{t^2 - \tilde{r}_{\delta,b}^2(r_{\max})}} - 1\right)} + 2\delta \leq 2r_{\min},$$
$$\sqrt{\frac{d}{2(d+1)}}\left(\sqrt{\tilde{r}_b^2(r''_{\min}) - (2t^2 - \tilde{r}_b^2(r''_{\min}))\left(\frac{t}{\sqrt{t^2 - \tilde{r}_{\delta,b}^2(r''_{\min})}} - 1\right)} + 2\delta\right) \leq r''_{\min},$$

where

$$r''_{\min} := \sqrt{t^2 - \epsilon(2t-\epsilon) - \frac{(2t^2 - r_{\min}^2 - \epsilon(2t-\epsilon))^2}{4t^2}},$$
$$\tilde{r}_b^2(t) := \frac{2t\left(t^2 + \epsilon(2t-\epsilon)\right)}{t + \sqrt{t^2 - (t^2 + \epsilon(2t-\epsilon))}}, \qquad \tilde{r}_{\delta,b}^2(t) := \min\left\{\delta^2 + \epsilon(2t-\epsilon), \frac{1}{2}\tilde{r}_b^2(t)\right\}.$$

Then $\mathbb{X}$ is homotopy equivalent to the Vietoris-Rips complex Rips$(\mathcal{X}, r)$.

We end this section by introducing a sampling condition in which we can guarantee the covering conditions in Corollary 10 and Theorem 19, 20 are satisfied. Let P be the sampling distribution on $\mathbb{X}$. We assume that there exist positive constants a, b and ϵ_0 such that, for all $x \in \mathbb{X}$, the following inequality holds:

$$P(\mathbb{B}_{\mathbb{R}^d}(x, \epsilon)) \geq a\epsilon^b, \quad \text{for all } \epsilon \in (0, \epsilon_0). \tag{19}$$

This condition on P is also known as the (a, b)-condition or the standard condition [15, 14, 10]. It is satisfied, for example, if $\mathbb{X}$ is a smooth manifold of dimension b and P has a density with respect to the Hausdorff measure on it bounded from below by a.

Under this condition, we have the following covering lemma.

▶ **Lemma 23.** *Let $\{X_1, \ldots, X_n\}$ be an i.i.d. sample from the distribution P and let $\{r_n = (r_{n,1}, \ldots, r_{n,n})\}_{n\in\mathbb{N}}$ be a triangular array of positive numbers such that, for each n,*

$$2\left(\frac{\log n}{an}\right)^{1/b} \leq \min_i r_{n,i} \leq 2\epsilon_0. \tag{20}$$

Then, the probability that the sample is a r_n-covering of $\mathbb{X}$ is bounded as

$$P\left(\mathbb{X} \subset \bigcup_{i=1}^{n} \mathbb{B}_{\mathbb{R}^d}(X_i, r_{n,i})\right) \geq 1 - \frac{1}{2^b \log n}. \tag{21}$$

5.1 Conditions for homotopy reconstruction

In this subsection, we discuss the tightness of the conditions we have identified for guaranteeing the homotopy equivalence of the target space and the Čech complex and the Vietoris-Rips complex. We first argue that the maximum radius conditions in (14) and (17) are tight, as Example 24 shows that the Čech complex fails to be homotopy equivalent to $\mathbb{X}$ when $r_{\max} > \tau - d_{\mathbb{X}}(x)$ and the Vietoris-Rips complex fails to be homotopy equivalent to $\mathbb{X}$ when $r_{\max} > \sqrt{\frac{d+1}{2d}}\,(\tau - d_{\mathbb{X}}(x))$ and $d \leq 2$.

▶ **Example 24.** Let $\epsilon \in [0, 1)$ be fixed. Let $\mathbb{X} \subset \mathbb{R}^d$ be the unit sphere in $\mathbb{R}^d$, and let $\mathcal{X} = \{x_1, \ldots, x_n\} \subset (1-\epsilon)\mathbb{X}$ be a finite set of points on the sphere centered at 0 and of radius $1-\epsilon$. Suppose that for some $\delta > \epsilon$, $\mathbb{X}$ is covered by δ-balls centered at $\mathcal{X}$, that is, $\mathbb{X} \subset \bigcup_{x\in\mathcal{X}} \mathbb{B}_{\mathbb{R}}(x, \delta)$. The reach of $\mathbb{X}$ equals to its radius 1.

For the ambient Čech complex, if $r \in (0, 1-\epsilon]^n$ and condition (15) is satisfied, then $\mathbb{X}$ is homotopy equivalent to $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{X}}(\mathcal{X}, r)$ by Theorem 19. Now, suppose that $r_{\min} > 1-\epsilon$. Then $0 \in \mathbb{B}_{\mathbb{R}^d}(x_i, r_{x_i})$ for all i, hence for any $y \in \bigcup_{i=1}^n \mathbb{B}_{\mathbb{R}^d}(x_i, r_{x_i})$, a line segment connecting 0 and y is contained in $\bigcup_{i=1}^n \mathbb{B}_{\mathbb{R}^d}(x_i, r_{x_i})$ as well. Hence $\bigcup_{i=1}^n \mathbb{B}_{\mathbb{R}^d}(x_i, r_{x_i})$ is contractible, and then from the usual Nerve Theorem, $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}(\mathcal{X}, r) \simeq \bigcup_{i=1}^n \mathbb{B}_{\mathbb{R}^d}(x_i, r_{x_i}) \simeq 0$. On the other hand, the $d-1$-th homology group of $\mathbb{X}$ is $H_{d-1}(\mathbb{X}) = \mathbb{Z}$, so $\mathbb{X}$ and $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}(\mathcal{X}, r)$ are not homotopy equivalent.

For the Vietoris-Rips complex, if $r \in \left(0, \sqrt{\frac{d+1}{2d}}(1-\epsilon)\right]^{d+1}$ and condition (18) is satisfied, then $\mathbb{X}$ is homotopy equivalent to $\mathrm{Rips}_{\mathbb{X}}(\mathcal{X}, r)$ by Theorem 20. Now, suppose each r_{x_i} is equal to some $r > \sqrt{\frac{d+1}{2d}}(1-\epsilon)$, and further suppose that $d \leq 2$ and $\delta < \frac{1}{2(1-\epsilon)}r_0 - \frac{\sqrt{3}}{4}$. When $d = 1$, then the Vietoris-Rips complex equals the ambient Čech complex, hence from the above argument, $\mathrm{Rips}(\mathcal{X}, r) = \check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}(\mathcal{X}, r) \simeq 0$. When $d = 2$, then $\mathrm{Rips}(\mathcal{X}, r) \cong \mathrm{Rips}\left(\frac{1}{1-\epsilon}\mathcal{X}, \frac{1}{1-\epsilon}r_0\right)$ and $\frac{1}{1-\epsilon}\mathcal{X} \subset \mathbb{X} \subset \bigcup_{i=1}^n \mathbb{B}_{\mathbb{R}^d}(\frac{1}{1-\epsilon}x_i, \delta)$ holds. Then $\frac{1}{1-\epsilon}r_0 - 2\delta > \frac{\sqrt{3}}{2}$, and hence from Proposition 3.8, Corollary 4.5, Proposition 5.2 of [2], either $\mathrm{Rips}\,(\mathcal{X},r)\simeq S^{2l+1}$ for some $l\geq 1$ or $\mathrm{Rips}\,(\mathcal{X},r)\simeq\vee^{c}S^{2l}$ for some $l\geq 1$ and $c\geq 0$. In either case, $H_1(\mathrm{Rips}\,(\mathcal{X},r))=0$. However, the $d-1$-th homology group of $\mathbb{X}$ is $H_{d-1}(\mathbb{X})=\mathbb{Z}$, so $\mathbb{X}$ and $\mathrm{Rips}\,(\mathcal{X},r)$ are not homotopy equivalent.

We then rephrase the conditions on $\epsilon:=\max\{d_{\mathbb{X}}(x):x\in\mathcal{X}\}$ and the covering radius δ in (15) and (18) in terms of the Hausdorff distance $d_H(\mathbb{X},\mathcal{X})$. For simplicity, we consider the case when all the radii r_x's are equal, and we denote that common value as r. In general, the Hausdorff distance $d_H(\mathbb{X},\mathcal{X})$ gives a bound for both ϵ and δ, that is, $\epsilon,\delta\leq d_H(\mathbb{X},\mathcal{X})$. Let $\rho:=\frac{d_H(\mathbb{X},\mathcal{X})}{\tau}$. For the Čech complex, a sufficient condition for (15) is that for some $\frac{r}{\tau}\in(0,1]$,

$$\rho+\sqrt{(\frac{r}{\tau})^2-\tilde{l}^2+\rho(2-\rho)-((1-\rho)^2-(\frac{r}{\tau})^2+\tilde{l}^2+(1-\rho_{\tilde{l}})^2)\left(\frac{1}{\sqrt{1-\tilde{r}^2_{\delta,c}}}-1\right)}\leq\frac{r}{\tau},$$

$$\sqrt{\frac{d}{2(d+1)}}\left(\sqrt{\tilde{r}_b^2-(2-\tilde{r}_b^2)\left(\frac{1}{\sqrt{1-\tilde{r}^2_{\delta,b}}}-1\right)}+2\rho\right)\leq r''_{\min},\tag{22}$$

where

$$\tilde{l}:=\frac{1}{2}\left(\frac{r}{\tau}-1+\sqrt{(1-\rho)^2-(\frac{r}{\tau})^2}-\rho\right),\qquad \rho_{\tilde{l}}:=1-\sqrt{(1-\rho)^2-(\frac{r}{\tau})^2}+\tilde{l},$$

$$\tilde{r}^2_{\delta,c}:=\min\left\{2\rho,\frac{1}{2}((\frac{r}{\tau})^2-\tilde{l}^2+\rho(2-\rho)+\rho_{\tilde{l}}(2-\rho_{\tilde{l}}))\right\},$$

$$r''_{\min}:=\sqrt{1-\rho(2-\rho)-\frac{(2-(\frac{r}{\tau})^2-\rho(2-\rho))^2}{4}},$$

$$\tilde{r}_b^2:=\frac{2\left((r''_{\min})^2+\rho(2-\rho)\right)}{1+\sqrt{1-((r''_{\min})^2+\rho(2-\rho))}},\qquad \tilde{r}^2_{\delta,b}:=\min\left\{2\rho,\frac{1}{2}\tilde{r}_b^2\right\}.$$

And for the Vietoris-Rips complex, the sufficient condition for (18) is

$$\sqrt{\tilde{r}_b^2(r_0)-(2-\tilde{r}_b^2(r_0))\left(\frac{1}{\sqrt{1-\tilde{r}^2_{\delta,b}(r_0)}}-1\right)}+2\rho\leq 2r_0,$$

$$\sqrt{\frac{d}{2(d+1)}}\left(\sqrt{\tilde{r}_b^2(r''_{\min})-(2-\tilde{r}_b^2(r''_{\min}))\left(\frac{1}{\sqrt{1-\tilde{r}^2_{\delta,b}(r''_{\min})}}-1\right)}+2\rho\right)\leq r''_{\min},\tag{23}$$

where

$$r_0:=\sqrt{\frac{d+1}{2d}}(1-\rho),\qquad r''_{\min}:=\sqrt{1-\rho(2-\rho)-\frac{(2-\frac{d+1}{2d}(1-\rho)^2-\rho(2-\rho))^2}{4}},$$

$$\tilde{r}_b^2(t):=\frac{2\left(t^2+\rho(2-\rho)\right)}{1+\sqrt{1-(t^2+\rho(2-\rho))}},\qquad \tilde{r}^2_{\delta,b}(t):=\min\left\{2\rho,\frac{1}{2}\tilde{r}_b^2(t)\right\}.$$

With the aid of a computer program, we can check that (22) is equivalent to $\rho\leq 0.01591\cdots$, and (23) is equivalent to $\rho\leq 0.07856\cdots$.

Now, we consider two specific cases. First, we consider the noiseless case $\mathcal{X}\subset\mathbb{X}$, that is, the data points lie in the target space. For that case, $\epsilon=0$ and $\delta\leq d_H(\mathbb{X},\mathcal{X})$. For the Čech complex, the sufficient condition for (15) is that for some $\frac{r}{\tau}\in(0,1]$,

$$\rho + \sqrt{(\frac{r}{\tau})^2 - \tilde{l}^2 - (1 - (\frac{r}{\tau})^2 + \tilde{l}^2 + (1-\rho_{\tilde{l}})^2)\left(\frac{1}{\sqrt{1-\tilde{r}_{\delta,c}^2}} - 1\right)} \leq \frac{r}{\tau},$$
$$\sqrt{\frac{d}{2(d+1)}}\left(\sqrt{\tilde{r}_b^2 - (2-\tilde{r}_b^2)\left(\frac{1}{\sqrt{1-\tilde{r}_{\delta,b}^2}} - 1\right)} + 2\rho\right) \leq r''_{\min}, \tag{24}$$

where

$$\tilde{l} := \frac{1}{2}\left(\frac{r}{\tau} - 1 + \sqrt{1-(\frac{r}{\tau})^2} - \rho\right), \qquad \rho_{\tilde{l}} := 1 - \sqrt{1-(\frac{r}{\tau})^2} + \tilde{l},$$
$$\tilde{r}_{\delta,c}^2 := \min\left\{\rho^2, \frac{1}{2}((\frac{r}{\tau})^2 - \tilde{l}^2 + \rho_{\tilde{l}}(2-\rho_{\tilde{l}}))\right\},$$
$$r''_{\min} := \sqrt{1 - \frac{(2-(\frac{r}{\tau})^2)^2}{4}}, \qquad \tilde{r}_b^2 := \frac{2(r''_{\min})^2}{1+\sqrt{1-(r''_{\min})^2}}, \qquad \tilde{r}_{\delta,b}^2 := \min\left\{\rho^2, \frac{1}{2}\tilde{r}_b^2\right\}.$$

For the Vietoris-Rips complex, a sufficient condition for (18) is

$$\sqrt{\tilde{r}_b^2(r_0) - (2-\tilde{r}_b^2(r_0))\left(\frac{1}{\sqrt{1-\tilde{r}_{\delta,b}^2(r_0)}} - 1\right)} + 2\rho \leq 2r_0,$$
$$\sqrt{\frac{d}{2(d+1)}}\left(\sqrt{\tilde{r}_b^2(r''_{\min}) - (2-\tilde{r}_b^2(r''_{\min}))\left(\frac{1}{\sqrt{1-\tilde{r}_{\delta,b}^2(r''_{\min})}} - 1\right)} + 2\rho\right) \leq r''_{\min}, \tag{25}$$

where

$$r_0 := \sqrt{\frac{d+1}{2d}}(1-\rho), \qquad r''_{\min} := \sqrt{1 - \frac{(2-\frac{d+1}{2d}(1-\rho)^2)^2}{4}},$$
$$\tilde{r}_b^2(t) := \frac{2t^2}{1+\sqrt{1-t^2}}, \qquad \tilde{r}_{\delta,b}^2(t) := \min\left\{\rho^2, \frac{1}{2}\tilde{r}_b^2(t)\right\}.$$

With the aid of a computer program, we can check that (24) is equivalent to $\rho \leq 0.02994\cdots$, and (25) is equivalent to $\rho \leq 0.1117\cdots$.

Second, we consider the asymptotic case, where we sample more and more points and $\mathcal{X}$ forms a dense cover of $\mathbb{X}$, that is, $\sup_{y\in\mathbb{X}}\inf_{x\in\mathcal{X}}\|x-y\| \to 0$. Still, we have a noisy sample distribution, that is, $\sup_{x\in\mathcal{X}}\inf_{y\in\mathbb{X}}\|x-y\| \nrightarrow 0$, so the Hausdorff distance $d_H(\mathbb{X},\mathcal{X})$ need not go to 0. In this case, $\delta \to 0$ and $\epsilon \leq d_H(\mathbb{X},\mathcal{X})$. For the Čech complex, a sufficient condition for (15) is that for some $\frac{r}{\tau} \in (0,1]$,

$$\sqrt{(\frac{r}{\tau})^2 - \tilde{l}^2 + \rho(2-\rho) - ((1-\rho)^2 - (\frac{r}{\tau})^2 + \tilde{l}^2 + (1-\rho_{\tilde{l}})^2)\left(\frac{1}{\sqrt{1-\tilde{r}_{\delta,c}^2}} - 1\right)} \leq \frac{r}{\tau},$$
$$\sqrt{\frac{d}{2(d+1)}}\sqrt{\tilde{r}_b^2 - (2-\tilde{r}_b^2)\left(\frac{1}{\sqrt{1-\tilde{r}_{\delta,b}^2}} - 1\right)} \leq r''_{\min}, \tag{26}$$

where

$$\tilde{l} := \frac{1}{2}\left(\frac{r}{\tau} - 1 + \sqrt{(1-\rho)^2 - (\frac{r}{\tau})^2}\right), \qquad \rho_{\tilde{l}} := 1 - \sqrt{(1-\rho)^2 - (\frac{r}{\tau})^2} + \tilde{l},$$
$$\tilde{r}_{\delta,c}^2 := \min\left\{\rho(2-\rho), \frac{1}{2}((\frac{r}{\tau})^2 - \tilde{l}^2 + \rho(2-\rho) + \rho_{\tilde{l}}(2-\rho_{\tilde{l}}))\right\},$$
$$r''_{\min} := \sqrt{1 - \rho(2-\rho) - \frac{(2-(\frac{r}{\tau})^2 - \rho(2-\rho))^2}{4}},$$
$$\tilde{r}_b^2 := \frac{2\left((r''_{\min})^2 + \rho(2-\rho)\right)}{1 + \sqrt{1 - ((r''_{\min})^2 + \rho(2-\rho))}}, \qquad \tilde{r}_{\delta,b}^2 := \min\left\{\rho(2-\rho), \frac{1}{2}\tilde{r}_b^2\right\}.$$

And for the Vietoris-Rips complex, a sufficient condition for (18) is

$$\sqrt{\tilde{r}_b^2(r_0) - (2 - \tilde{r}_b^2(r_0))\left(\frac{1}{\sqrt{1-\tilde{r}_{\delta,b}^2(r_0)}} - 1\right)} \le 2r_0,$$
$$\sqrt{\frac{d}{2(d+1)}}\sqrt{\tilde{r}_b^2(r''_{\min}) - (2 - \tilde{r}_b^2(r''_{\min}))\left(\frac{1}{\sqrt{1-\tilde{r}_{\delta,b}^2(r''_{\min})}} - 1\right)} \le r''_{\min}, \tag{27}$$

where

$$r_0 := \sqrt{\frac{d+1}{2d}}(1-\rho), \qquad r''_{\min} := \sqrt{1 - \rho(2-\rho) - \frac{(2 - \frac{d+1}{2d}(1-\rho)^2 - \rho(2-\rho))^2}{4}},$$
$$\tilde{r}_b^2(t) := \frac{2\left(t^2 + \rho(2-\rho)\right)}{1 + \sqrt{1 - (t^2 + \rho(2-\rho))}}, \qquad \tilde{r}_{\delta,b}^2(t) := \min\left\{\rho(2-\rho), \frac{1}{2}\tilde{r}_b^2(t)\right\}.$$

With the aid of a computer program, we can check that (26) is equivalent to $\rho \le 0.03440\cdots$, and (27) is equivalent to $\rho \le 0.07712\cdots$.

6 Discussion and Conclusion

Above we have provided conditions under which the ambient Čech complex $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}(\mathcal{X}, r)$ and the Rips complex $\mathrm{Rips}(\mathcal{X}, r)$ are homotopy equivalent to the target space $\mathbb{X}$ when the target space $\mathbb{X}$ has positive μ-reach τ^μ and the data points $\mathcal{X}$ being contained in the ϵ-offset $\mathbb{X}^\epsilon$ of $\mathbb{X}$. In this section, we further discuss our results and compare them with existing ones. For the comparison purpose, we consider the case when all the radii r_x's are equal, and we denote the common value as r. In these settings, an analogous homotopy equivalence between the ambient Čech complex $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}(\mathcal{X}, r)$ and the target space $\mathbb{X}$ is presented in [6] and [27].

First, we compare the upper bound for the maximum parameter value r in $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}(\mathcal{X}, r)$ or $\mathrm{Rips}(\mathcal{X}, r)$. When $\mu = 1$ so that $\tau^\mu = \tau$, our result suggests that the homotopy equivalences hold when $r \le \tau - \epsilon$ for $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}(\mathcal{X}, r)$ and $r \le \sqrt{\frac{d+1}{2d}}(\tau - \epsilon)$ for $\mathrm{Rips}(\mathcal{X}, r)$. As we have seen in Example 24, these bounds are optimal bounds. In [27], such a bound for $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}(\mathcal{X}, r)$ is $\frac{(\tau+\epsilon)+\sqrt{\tau^2+\epsilon^2-6\tau\epsilon}}{2}$ (see Proposition 7.1). Then our bound is strictly sharper than this when $\epsilon > 0$ since

$$\frac{(\tau+\epsilon) + \sqrt{\tau^2+\epsilon^2-6\tau\epsilon}}{2} < \frac{(\tau+\epsilon) + \sqrt{\tau^2+9\epsilon^2-6\tau\epsilon}}{2} = \tau - \epsilon.$$

In [6], a necessary condition for $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}(\mathcal{X}, r)$ in Section 5.3 is $r \le \tau - 3\epsilon$, so our upper bound is strictly better when $\epsilon > 0$.

Second, we compare the condition for the maximum possible ratio of the Hausdorff distance $d_H(\mathbb{X}, \mathcal{X})$ and the μ-reach τ^μ. For this case, as we have seen in Section 5.1, we can check that $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}(\mathcal{X}, r)$ is homotopy equivalent to $\mathbb{X}$ when $\frac{d_H(\mathbb{X},\mathcal{X})}{\tau} \leq 0.01591\cdots$. This result is worse than $3-\sqrt{8} \approx 0.1716\cdots$ in Proposition 7.1 in [27] or $\frac{-3+\sqrt{22}}{13} \approx 0.1300\cdots$ in Section 5.3 in [6]. Again from Section 5.1, we can check that $\mathrm{Rips}(\mathcal{X}, r)$ is homotopy equivalent to $\mathbb{X}$ when $\frac{d_H(\mathbb{X},\mathcal{X})}{\tau} \leq 0.07856\cdots$. This result is better than $\frac{2\sqrt{2-\sqrt{2}}-\sqrt{2}}{2+\sqrt{2}} \approx 0.03412\cdots$ in Section 5.3 in [6].

Then we consider two specific cases. In the noiseless case $\mathcal{X} \subset \mathbb{X}$, the data points lie in the target space. In this case, as we have seen in Section 5.1, we can verify that $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}(\mathcal{X}, r)$ is homotopy equivalent to $\mathbb{X}$ when $\frac{d_H(\mathbb{X},\mathcal{X})}{\tau} \leq 0.02994\cdots$, and $\mathrm{Rips}(\mathcal{X}, r)$ is homotopy equivalent to $\mathbb{X}$ when $\frac{d_H(\mathbb{X},\mathcal{X})}{\tau} \leq 0.1117\cdots$.

In the asymptotics case, as we sample more and more points from the target space, $\mathcal{X}$ forms a dense cover on $\mathbb{X}$, that is, $\sup_{y\in\mathbb{X}} \inf_{x\in\mathcal{X}} \|x-y\| \to 0$. For this case, as we have seen in Section 5.1, we can check that $\check{\mathrm{C}}\mathrm{ech}_{\mathbb{R}^d}(\mathcal{X}, r)$ is homotopy equivalent to $\mathbb{X}$ when $\frac{d_H(\mathbb{X},\mathcal{X})}{\tau} \leq 0.03440\cdots$, and $\mathrm{Rips}(\mathcal{X}, r)$ is homotopy equivalent to $\mathbb{X}$ when $\frac{d_H(\mathbb{X},\mathcal{X})}{\tau} \leq 0.07712\cdots$.

Finally, we emphasize that our result also allows the radii $\{r_x\}_{x\in\mathcal{X}}$ to vary across the points $x \in \mathcal{X}$. Considering different radii is of practical interest if each data point has different importance. For example, one might want to use large radii on the flat and sparse region, while to use small radii on the spiky and dense region. However, there remain significant technical difficulties to allow for a different radius per each data point. As it can be seen in Figure 2, an uneven distribution of radii might lead to nonhomotopic between the Čech complex (or the Vietoris-Rips complex) and the target space. This situation has been studied in [12] for the union of balls under the reach condition, but not the Vietoris-Rips complex or under the μ-reach case. Theorem 20 or Corollary 22 first tackles this homotopy reconstruction problem with different radii for the Vietoris-Rips complex or under the μ-reach condition.

References

1 Eddie Aamari, Jisu Kim, Frédéric Chazal, Bertrand Michel, Alessandro Rinaldo, and Larry Wasserman. Estimating the reach of a manifold. *Electronic Journal of Statistics*, 13(1):1359–1399, 2019. doi:10.1214/19-EJS1551.

2 Michał Adamaszek and Henry Adams. The Vietoris-Rips complexes of a circle. *Pacific Journal of Mathematics*, 290(1):1–40, 2017. doi:10.2140/pjm.2017.290.1.

3 Michał Adamaszek, Henry Adams, and Florian Frick. Metric reconstruction via optimal transport. *SIAM Journal on Applied Algebra and Geometry*, 2(4):597–619, 2018. doi:10.1137/17M1148025.

4 Henry Adams and Joshua Mirth. Metric thickenings of euclidean submanifolds. *Topology and its Applications*, 254:69–84, 2019. doi:10.1016/j.topol.2018.12.014.

5 Paul Alexandroff. Über den allgemeinen Dimensionsbegriff und seine Beziehungen zur elementaren geometrischen Anschauung. *Mathematische Annalen*, 98(1):617–635, 1928. doi:10.1007/BF01451612.

6 Dominique Attali, André Lieutier, and David Salinas. Vietoris–rips complexes also provide topologically correct reconstructions of sampled shapes. *Computational Geometry*, 46(4):448–465, 2013. 27th Annual Symposium on Computational Geometry (SoCG 2011). doi:10.1016/j.comgeo.2012.02.009.

7 Anders Björner. *Topological Methods*, page 1819–1872. MIT Press, Cambridge, MA, USA, 1996.

8 Frédéric Chazal, David Cohen-Steiner, and André Lieutier. A sampling theory for compact sets in euclidean space. *Discrete & Computational Geometry*, 41(3):461–479, 2009. doi:10.1007/s00454-009-9144-8.

9 Frédéric Chazal, David Cohen-Steiner, André Lieutier, and Boris Thibert. Shape smoothing using double offsets. In *Proceedings of the 2007 ACM Symposium on Solid and Physical Modeling*, SPM '07, page 183–192, New York, NY, USA, 2007. Association for Computing Machinery. doi:10.1145/1236246.1236273.

10 Frédéric Chazal, Brittany Fasy, Fabrizio Lecci, Bertr, Michel, Aless, ro Rinaldo, and Larry Wasserman. Robust topological inference: Distance to a measure and kernel distance. *Journal of Machine Learning Research*, 18(159):1–40, 2018. URL: http://jmlr.org/papers/v18/15-484.html.

11 Frédéric Chazal and André Lieutier. Weak feature size and persistent homology: Computing homology of solids in $\mathbb{R}^n$ from noisy data samples. In *Proceedings of the Twenty-First Annual Symposium on Computational Geometry*, SCG '05, page 255–262, New York, NY, USA, 2005. Association for Computing Machinery. doi:10.1145/1064092.1064132.

12 Frédéric Chazal and André Lieutier. Smooth manifold reconstruction from noisy and non-uniform approximation with guarantees. *Computational Geometry*, 40(2):156–170, 2008. doi:10.1016/j.comgeo.2007.07.001.

13 Frédéric Chazal and Steve Yann Oudot. Towards persistence-based reconstruction in euclidean spaces. In *Proceedings of the Twenty-Fourth Annual Symposium on Computational Geometry*, SCG '08, page 232–241, New York, NY, USA, 2008. Association for Computing Machinery. doi:10.1145/1377676.1377719.

14 Antonio Cuevas. Set estimation: another bridge between statistics and geometry. *Boletín de Estadística e Investigación Operativa. BEIO*, 25(2):71–85, 2009.

15 Antonio Cuevas and Alberto Rodríguez-Casal. On boundary estimation. *Advances in Applied Probability*, 36(2):340–354, 2004. doi:10.1239/aap/1086957575.

16 Vin de Silva and Robert Ghrist. Coverage in sensor networks via persistent homology. *Algebraic & Geometric Topology*, 7:339–358, 2007. doi:10.2140/agt.2007.7.339.

17 Tamal K. Dey. *Curve and Surface Reconstruction: Algorithms with Mathematical Analysis*. Cambridge Monographs on Applied and Computational Mathematics. Cambridge University Press, 2006. doi:10.1017/CBO9780511546860.002.

18 Jürgen Eckhoff. Chapter 2.1 - Helly, Radon, and Carathéodory type theorems. In P.M. GRUBER and J.M. WILLS, editors, *Handbook of Convex Geometry*, pages 389–448. North-Holland, Amsterdam, 1993. doi:10.1016/B978-0-444-89596-7.50017-1.

19 Herbert Edelsbrunner and John Harer. *Computational Topology: An Introduction*. American Mathematical Society, 2010. URL: http://www.ams.org/bookstore-getitem/item=MBK-69.

20 Herbert Edelsbrunner and Ernst P. Mücke. Three-dimensional alpha shapes. *ACM Transactions on Graphics*, 13(1):43–72, January 1994. doi:10.1145/174462.156635.

21 Herbert Federer. Curvature measures. *Transactions of the American Mathematical Society*, 93:418–491, 1959. doi:10.2307/1993504.

22 Karsten Grove. Critical point theory for distance functions. In *Differential geometry: Riemannian geometry (Los Angeles, CA, 1990)*, volume 54 of *Proc. Sympos. Pure Math.*, pages 357–385. Amer. Math. Soc., Providence, RI, 1993.

23 Allen Hatcher. *Algebraic topology*. Cambridge University Press, Cambridge, 2002.

24 Jean-Claude Hausmann. On the Vietoris-Rips complexes and a cohomology theory for metric spaces. In *Prospects in topology (Princeton, NJ, 1994)*, volume 138 of *Annals of Mathematics Studies*, pages 175–188. Princeton Univ. Press, Princeton, NJ, 1995.

25 Janko Latschev. Vietoris-Rips complexes of metric spaces near a closed Riemannian manifold. *Archiv der Mathematik*, 77(6):522–528, 2001. doi:10.1007/PL00000526.

26 John M. Lee. *Introduction to smooth manifolds*, volume 218 of *Graduate Texts in Mathematics*. Springer, New York, second edition, 2013.

27 Partha Niyogi, Stephen Smale, and Shmuel Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete & Computational Geometry*, 39(1-3):419–441, 2008. doi:10.1007/s00454-008-9053-2.

28 Vanessa Robins, James D. Meiss, and Elizabeth Bradley. Computing connectedness: disconnectedness and discreteness. *Physica D: Nonlinear Phenomena*, 139(3):276–300, 2000. doi:10.1016/S0167-2789(99)00228-6.

29 Amit Singer and Hau-Tieng Wu. Vector diffusion maps and the connection laplacian. *Communications on Pure and Applied Mathematics*, 65(8):1067–1144, 2012. doi:10.1002/cpa.21395.

30 Jean-Luc Starck, Vicent Martínez, D. Donoho, Ofer Levi, Philippe Querre, and Enn Saar. Analysis of the spatial distribution of galaxies by multiscale methods. *EURASIP Journal on Applied Signal Processing*, 2005(15):2455–2469, 2005. doi:10.1155/ASP.2005.2455.