Computation on Wasserstein Distance between 0-Dimensional Rips Persistence Diagrams

Bouchaib Azamir, Driss Bennis*

Abstract—**The use of topological summaries, specifically persistence diagrams, has seen significant progress in recent times. However, the computation becomes increasingly intricate and costly as the size of the data grows larger. The objective of this article is to determine mathematicaly the effects of stopping two filtrations at premature death times on the Wasserstein distance between the corresponding persistence diagrams. The established results provide a partial answer to a question raised in a preceding paper.**

Keywords: ***Persistent homology; Dynamic network; Wasserstein distance; time series.***

1 Introduction

Topological data analysis (TDA) combines algebraic topology and statistics to study data. It examines the geometric and topological structure of the underlying space, such as connected components, loops, and cavities, based on sampled data. TDA includes diverse techniques such as clustering, the Mapper algorithm, manifold estimation, and persistent homology (PH) see [7, 8, 10, 24, 30]. Being a key tool in TDA, persistent homology is an emerging method (algorithm) for detecting geometric and topological properties of a space with a topological structure. PH summarizes the topological features of the filtered dataset in a descriptor known as a persistence diagram (PD). The PD is a multiset in $\mathbb{R}^2$ where each point represents a feature like a connected component, hole, or cavity. The x and y coordinates of each point in PD indicate the corresponding inception "birth" and extinction "death" times during a topological feature's filtration. One of the powerful aspects of persistence diagrams is that they form a metric space through distances like the Wasserstein distance. This metric provides a means of assessing the similarity between two persistence diagrams. The notion of persistence was initially presented by Edelsbrunner, Letscher, and Zomorodian in [8], then further developed by Carlsson and Zomorodian in [30]. Subsequently, it has proven highly beneficial and has found applications in diverse scientific domains such as biology, image processing, sensor networks, etc. [24]. However, there are still issues in this field that require improvements, specifically regarding the complexity of algorithms used. As the data size increases, computing this distance becomes more complicated and resource-intensive. Extensive work has gone into developing methods to streamline data visualization, reducing the complexity of calculations involved. This has resulted in the exploration of various approaches for simplification (see for instance the papers and the references within [4, 5, 12, 15, 18, 19, 20, 21, 22, 25, 26, 27, 29, 31, 32].

In this study, we adopt a different yet complementary approach by examining the computational efficiency of PD simplification through modifications in filtration. Specifically, we explore how changes in the scaling parameter k affect execution time while maintaining key topological properties. Our findings reveal that certain values of k can significantly lower the computational cost without compromising stability. While [4] emphasizes optimizing the representation of persistence diagrams for subsequent tasks, our research offers a systematic assessment of their computational cost during the construction phase. These two viewpoints contribute to the overarching goal of enhancing the efficiency of persistence-based methods in TDA.

In the article [1], an algorithm was developed to decrease the execution time of the persistence homology algorithm while maintaining the expected outcome. This entails stopping filtration upon reaching a threshold selected using the closeness centrality concept. However, although this approach has shown effectiveness in decreasing execution time, demonstrating its ability to retain a substantial portion of the network's information has only been done visually, lacking a mathematical validation (see subsection 2.4 for the mathematical problem formulation). The article mainly focusses on some specific thresholds. In particular, we will study the interrelations between the Wasserstein distances of global persistence diagrams (obtained from complete filtrations) and those arising from such persistence diagrams stemming from filtrations sttoped early at a given scale. The rest of the paper is organized as follows:

- Section 2 presents the preliminaries and notations used in this paper. Moreover, it also sets up the problem in a mathematical way.
- Section 3 presents the two main theorems of this paper. Theorems 3.1 and 3.2 reveal a mathematical connection between the Wasserstein distances of Rips persistence diagrams that come from two distinct types of filtrations. One of these filtrations is complete, taking into account all the topological links, while the other is a filtration that stops prematurely at a certain threshold. This threshold is defined by real parameters α and k, as explained in Theorem 3.1, and it depends on the maximum and the second largest values from one of the

*Manuscript received December 19, 2024; revised May 3, 2025.

Bouchaib Azamir is a PhD candidate of the Faculty of Sciences, Mohammed V University in Rabat, Morocco (e-mail: bouchaib_azamir@um5.ac.ma).

Driss Bennis is a professor of the Faculty of Sciences, Mohammed V University in Rabat, Morocco (e-mail: d.bennis@um5r.ac.ma).

persistence diagrams, as described in Theorem 3.2.

- Section 4 goes over the key findings, which are illustrated and clarified through two examples of real dynamic weighted networks.
- Section 5 dives into how our approach stacks up in terms of computational efficiency, particularly looking at how different scaling parameters, used in the first main theorem, influence running time. It also compares our method to the auction-based Wasserstein distance algorithm put forth by Kerber et al. [16], showcasing the advantages of our approach in cutting down computation time.
- Finally, Section 6 provides conclusions and 7 outlines possible future research directions.

2 Preliminaries and mathematical problem formulation

In this section, we provide a brief overview of fundamental concepts related to specific simplicial complexes and their associated persistent homology. For a broader understanding of persistent homology, refer to [7, 10, 24, 30]. Using the notations given in this section, we end with a mathematical formulation of the problem.
A vector spaces considered in this paper are over $\mathbb{F}_2$, the field with two elements $\{0, 1\}$.

2.1 Simplicial complex and simplicial homology groups

In this subsection, we recall some fundamental aspects of simplicial complexes and their related homology groups which are very significant in the field of TDA.

Given a finite set V, a simplicial complex with the vertex set V is a set K of finite subsets of V such that the elements of V belong to K and for any $\sigma \in K$, any subset of σ belongs to K. The elements of K are called the simplices of K. The dimension of a simplex of K is just its cardinality minus 1. A p-simplex is a simplex σ with dimension p. The dimension of the simplicial complex K is the largest dimension of its simplices [24]. To illustrate this definition, consider the following example which describes a simplicial complex commonly used in TDA.

Example 2.1 (Vietoris-Rips complex)**.** *[24] Given a finite set of points $\mathbb{X}$ in a metric space (M, d) and a real number $\alpha \geq 0$. The Vietoris-Rips complex $\mathbb{VR}^{\alpha}(\mathbb{X})$, (Rips in short), is a simplicial complex whose simplices are sets $\{x_0, \ldots, x_k\}$ such that $d(x_i, x_j) \leq \alpha$ for all $0 \leq i, j \leq k$.*

For a given simplicial complex K, we denote the vector space generated by the p-simplices of K as $C_p(K)$. This space comprises all finite formal sums of p-simplices, referred to as p-chains. In other words, an element c belongs to $C_p(K)$ if it can be expressed as $c = \sum_j \gamma_j \sigma_j$ for scalars $\gamma_j \in \mathbb{F}_2$ and a family $(\sigma_j)_j$ of p-simplices.
For a positive integer p, we consider the linear map $\partial_p : C_p(K) \to C_{p-1}(K)$, known as the boundary map and defined on p-simplices as follows: for every p-simplex σ, $\partial_p(\sigma)$ is the formal sum of the $(p-1)$-dimensional faces, i.e., the subsets of σ with cardinality p. An element in the image of ∂_p is called a boundary. The boundary $\partial_p(c)$ of a chain $c = \sum_j \gamma_j \sigma_j$ is calculated by linearly extending ∂_p, as follows:

$$\partial_p(c) = \sum_j \gamma_j \partial_p(\sigma_j).$$

The p-chains with a boundary of 0 are referred to as p-cycles and collectively form a subspace $Z_p(K)$ of $C_p(K)$. On the other hand, p-chains that are the boundaries of $(p+1)$-chains are termed p-boundaries, constituting a subspace $B_p(K)$ of $C_p(K)$. It is noteworthy, though not challenging to demonstrate, that $\partial_p \circ \partial_{p+1} = 0$, which is equivalent to $B_p(K) \subseteq Z_p(K)$. Consider the quotient vector space $Z_p(K)/B_p(K)$, where p-boundaries that are not p-cycles have been annihilated. It is proved that the dimension of the vector space $B_1(K)/Z_1(K)$ represents the number of "holes" in the simplicial complex K, and the dimension of the vector space $B_0(K)/Z_0(K)$ corresponds to the number of connected components of K. The vector space $B_p(K)/Z_p(K)$, denoted as $H_p(K)$, is a key concept in algebraic topology, often called the p-th simplicial homology group of K. Its elements are recognized as homology classes.

2.2 Persistence diagrams

When a simplicial complex K can be expressed as the union of nested sequence subcomplexes, it often reveals deeper structural information about K. By examining how the homology groups evolve across this sequence, we gain insight into the topological features of K that persist across different scales. This idea forms the foundation of persistent homology, a key tool in TDA.
Consider the following chain of simplicial subcomplexes of a complex K, referred to as a filtration of K:

$$\mathcal{F}: \quad \emptyset \subseteq K_0 \subseteq K_1 \subseteq \cdots \subseteq K_p = K$$

The filtration provides insights into K by inducing a homomorphism on n^{th} simplicial homology groups for each dimension, through the inclusion map $K_i \to K_j$. The n-th persistent Betti number $\beta^n_{i,j}$ represents the rank of the vector space:

$$\beta^n_{i,j} = \operatorname{rank} \operatorname{Im}(f^n_{i,j}).$$

Persistent Betti numbers quantify the number of homology classes of dimension n that persist through the transition from K_i to K_j. A homology class $\alpha \in H_n(K_i)$ is said to be born upon entering K_i if α does not come from a previous subcomplex, i.e., $\alpha \notin Im(f^{i-1,i}_n)$. Similarly, if α is born in K_i, it dies upon entering K_j if the image of the map induced by $K_{i-1} \subseteq K_{j-1}$ does not contain the image of α, but the image of the map induced by $K_{i-1} \subseteq K_j$ does. In this case, the persistence of α is $j - i$. The birth-death pair of a homology class can be represented as coordinates of a point in the plane. Since several homology classes may share the same birth-death pair, the collected data of birth-death pairs or all homology classes can be presented as a multiset of points, where the coordinates are birth-death pairs. In fact, the coordinates of a point A can represent the birth-death pair of more than one homology class. The number of homology classes represented by A is the multiplicity of A and denoted by $m(A)$. For computational reasons (see [17]), the multiset is considered together with the diagonal Δ of points (x, x)

with infinite multiplicity. This extended structure is called the persistence diagram of the filtration $\mathcal{F}$.

In this paper, we focus on filtrations of the form

$$\emptyset \subseteq K_0 \subseteq K_1 \subseteq \cdots \subseteq K_p = K$$

, where each K_i corresponds to the Rips complex $\mathbb{VR}^{\alpha_i}(\mathbb{X})$ constructed from a finite set of points $\mathbb{X}$ in a metric space (M, d), and a sequence of scales

$$0 = \alpha_0 < \alpha_1 < \alpha_2 < \cdots < \alpha_n$$

. This type of filtration, known as the Vietoris-Rips filtration (or simply Rips filtration), is a fundamental tool in persistent homology. It builds simplicial complexes by connecting points in $\mathbb{X}$ based on their pairwise distances, with the parameter α_i controlling the scale at which these connections are made. We denote this filtration by $\mathcal{F}(\mathbb{X})$ (see, for example, Figure 1).

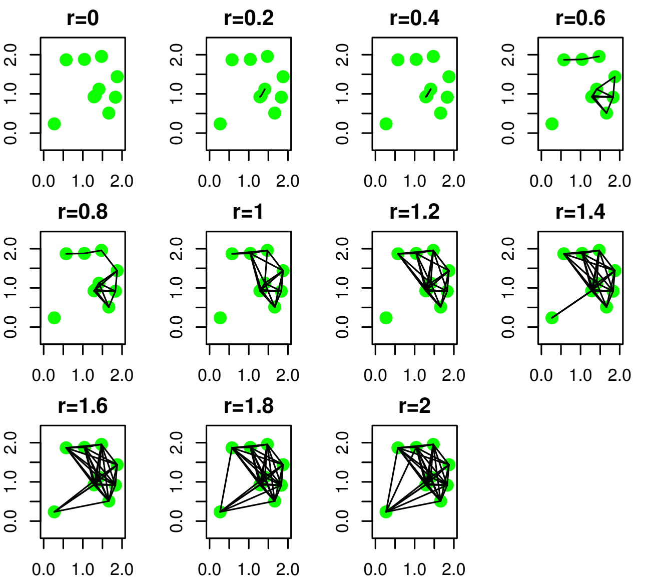

Figure 1: A Rips filtration composed of a total of eleven simplicial complexes. All connected components merge into a single entity for the first time at a scale of $r = 1.4$.

The persistence diagram $Dgm_p(\mathcal{F}(\mathbb{X}))$ associated with $\mathcal{F}(\mathbb{X})$ is known as a Vietoris-Rips (Rips in short) persistence diagram, where p is the dimension of the homology classes represented in $Dgm_p(\mathcal{F}(\mathbb{X}))$. To simplify notation, the Rips persistence diagram $Dgm_p(\mathcal{F}(\mathbb{X}))$ will be denoted by $\mathcal{D}_p(\mathbb{X})$. In our study, we deal only with $\mathcal{D}_0(\mathbb{X})$, the persistence diagram of the birth and death dates of connected components. Then, in this case, $\mathcal{D}_0(\mathbb{X})$ is the union of the two multisets $\widetilde{\mathcal{D}_0}(\mathbb{X})$ and Δ, where $\widetilde{\mathcal{D}_0}(\mathbb{X})$ is the multiset of off-diagonal points with finite multiplicity; that is, the points $(0, a_{i_{\mathbb{X}}})$ with $a_{i_{\mathbb{X}}} \in \mathbb{R}^+ \cup \{\infty\}$ and $m(0, a_{i_{\mathbb{X}}}) < \infty$, and Δ is the diagonal of point (x, x) with infinite multiplicity. Denote by $\underline{\mathcal{D}}_0(\mathbb{X})$ the underlying set of $\widetilde{\mathcal{D}_0}(\mathbb{X})$.

When a threshold value α_i is reached, the filtration process halts, resulting in a sub-filtration, which we denote as $\mathcal{F}_{\alpha_i}(\mathbb{X})$. The persistence diagram that represents the connected components related to this sub-filtration is referred to as $\mathcal{D}_{0,\alpha_i}(\mathbb{X})$.

The persistence diagrams play an important role in assessing the level of similarity among complexes since the space of persistence diagrams is provided with a range of metrics. In this paper, we consider the so-called p-Wasserstein distance where p is a positive integer. The p-Wasserstein distance $W_p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y}))$ is given by:

$$W_p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y})) = \inf_{\phi} \left(\sum_{x \in \mathcal{D}_0(\mathbb{X})} \|x - \phi(x)\|_\infty^p \right)^{1/p}$$

where ϕ runs over all bijections between $\mathcal{D}_0(\mathbb{X})$ and $\mathcal{D}_0(\mathbb{Y})$ and $\|(a, b)\|_\infty = \max\{|a|, |b|\}$.

2.3 The Wasserstein distance computation

In this subsection, we describe an algorithm wich allows to compute the p-Wasserstein distance, see [16] for more details. The calculation of the p-Wasserstein distance is based on concepts from graph theory, especially bipartite weighted graphs.
Let us recall what a bipartite weighted graph is. It is a graph, denoted as $G = (V, E, \omega)$, where the vertex set V can be split into two nonempty groups, A and B. This means that when you combine A and B, you get all of V, and they don't overlap at all. Each edge in this graph connects a point from A to a point in B. The way we divide V into A and B is known as a bipartition of G, and $\omega : E \to \mathbb{R}^+$ represents its weight function. Now, if two edges in G don't share any endpoints, we call them independent. A collection of these independent edges is referred to as a matching. If this matching M includes every vertex in V exactly once, we call it a perfect matching. Essentially, a perfect matching in our bipartite graph G creates a one-to-one correspondence between two separate subsets of V.
The $p-$Wasserstein cost of a perfect matching M is defined as $(\sum_{e \in E} \omega(e)^p)^{\frac{1}{p}}$. An optimal matching of G is a perfect matching whose cost is minimal among all perfect matchings of G.
Given two persistence diagrams $\mathcal{D}_0(\mathbb{X})$ and $\mathcal{D}_0(\mathbb{Y})$ with off-diagonal point sets $\underline{\mathcal{D}}_0(\mathbb{X})$ and $\underline{\mathcal{D}}_0(\mathbb{Y})$ respectively. Let $\underline{\mathcal{D}'}_0(\mathbb{X})$ and $\underline{\mathcal{D}'}_0(\mathbb{Y})$ be the sets of orthogonal projections, on the diagonal Δ, of $\underline{\mathcal{D}}_0(\mathbb{X})$ and $\underline{\mathcal{D}}_0(\mathbb{Y})$, respectively. Then we can define a bipartite weighted graph $G = (A \cup B, A \times B, c)$ where $A = \underline{\mathcal{D}}_0(\mathbb{X}) \cup \underline{\mathcal{D}'}_0(\mathbb{Y})$ and $B = \underline{\mathcal{D}}_0(\mathbb{Y}) \cup \underline{\mathcal{D}'}_0(\mathbb{X})$ (see [16]).

Persistence diagrams include points on the diagonal Δ with infinite multiplicity, which makes it possible to define bijections. The weights of the graph G are defined by the function:

$$c(a, b) = \begin{cases} \|a - b\|_\infty & \text{if } a \in \underline{\mathcal{D}}_0(\mathbb{X}) \text{ or } b \in \underline{\mathcal{D}}_0(\mathbb{Y}) \\ 0 & \text{otherwise} \end{cases}$$

These weights come purely from the persistence diagrams. The Wasserstein distance $W_p(X, Y)$ is defined as the cost of the graph G according to the Reduction Lemma in [7]. In [16], it was shown that no skew edge ever altered the minimal cost for optimal matching. An edge $(a, b) \in A \times B$ is said to be skew if one of the following conditions holds:

1. a is in $\underline{\mathcal{D}}_0(\mathbb{X})$, b is in $\underline{\mathcal{D}'}_0(\mathbb{X})$, and b is not equal to a', where a' is the orthogonal projection of a on Δ.

2. b is in $\underline{\mathcal{D}}_0(\mathbb{Y})$, a is in $\underline{\mathcal{D}'}_0(\mathbb{Y})$, and a is not equal to b', the orthogonal projection of b on Δ.

More precisely, they defined the bipartite graph

$$\widetilde{G} = (A \cup B, A \times B, \tilde{c})$$

which does not contain skew edges and where the cost function $\tilde{c}$ is defined by:

$$\tilde{c}(a,b) = \begin{cases} \|a-b\|_\infty & \text{if } a \in \underline{\mathcal{D}}_0(\mathbb{X}) \text{ or } b \in \underline{\mathcal{D}}_0(\mathbb{Y}) \\ \|a-a'\|_\infty & \text{if } a \in \underline{\mathcal{D}}_0(\mathbb{X}) \text{ or } b \in \underline{\mathcal{D}'}_0(\mathbb{X}) \\ \|b-b'\|_\infty & \text{if } a \in \underline{\mathcal{D}'}_0(\mathbb{Y}) \text{ or } b \in \underline{\mathcal{D}}_0(\mathbb{Y}) \\ 0 & \text{otherwise} \end{cases}$$

They proved that G and $\widetilde{G}$ have the same Wasserstein cost.

2.4 Formulation of the problem

Now, using the notations introduced in Subsection 2.3, we reformulate mathematically the problem mentioned in the introduction. The question explored in this paper, as introduced before, was implied in the paper [1]. To articulate it clearly, we suggest revisiting the original context.

In [11], Gidea used TDA to study a dynamic weighted financial network where the weights are determined by correlations between nodes representing stocks. Explicitly, for each stock i and a day t, the daily return $x_i(t)$ was calculated based on the adjusted closing prices $S_i(t)$ using the formula $x_i(t) = \frac{S_i(t+1)-S_i(t)}{S_i(t)}$. The correlation coefficient $C_{i,j}(t)$ between nodes $x_i(t)$ and $x_j(t)$ over the interval $[t-T,t]$ (where $T>0$) was used to define a distance function $d(i,j)(t) = \sqrt{2(1-C_{i,j}(t))}$ between nodes i and j. This distance metric allows the network to be seen as a time-evolving weighted network, or a (dynamic) graph $G_t(V,E_t,\omega_t)$, where V represents stocks and the weight function ω_t at time t assigns to each edge $e=(i,j)$ the distance $d(i,j)(t)$. In this scenario, we can think of the set of nodes V as a point cloud, and we denote $\mathcal{D}_0(V)$ by $\mathcal{D}_0(G_t)$. By fixing a persistence diagram $\mathcal{D}_0(G_{t_0})$ at an initial time t_0, the time series $(X_t)_t := (W_2^2(\mathcal{D}_0(G_t), \mathcal{D}_0(G_{t_0}))_t$ reflects the topological changes in the financial network over time. This can potentially detect significant changes in the network's topological structure before critical transitions, such as the peak of a financial crisis, indicating major shifts in stock correlations. This approach applies to other contexts, but for large networks, its execution time becomes very high. In the article [1], the study focused on the impact of early algorithm termination and only considering sub-filtrations upon reaching a threshold. It is shown that beyond a specific threshold determined by certain considerations, the new time series resembles the original and exhibits similar behavior (see Figures in [1]). The strength of the proposed methods lies in both suggesting an algorithmic approach for selecting this threshold and offering improved time efficiency compared to some existing simplification methods (see [1, Subsection 5.4]). Still, we haven't fully cracked the mathematical reasoning behind this result, and that's something we need to tackle. This paper aims to offer some positive initial insights in specific situations. Using notations mentioned in Subsection 2.2, the main question is stated as follows:

Question 1. Consider two persistence diagrams $\mathcal{D}_0(\mathbb{X})$ and $\mathcal{D}_0(\mathbb{Y})$. Let r and s be two thresholds of the corresponding filtrations respectively.
What relationship exists between $W_p^p(\mathcal{D}_{0,r}(\mathbb{X}), \mathcal{D}_{0,s}(\mathbb{Y}))$ and $W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y}))$?

In [1], it is observed through various examples that $W_p^p(\mathcal{D}_{0,r}(\mathbb{X}), \mathcal{D}_{0,s}(\mathbb{Y}))$ and $W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y}))$ show a strong linear relationship. This observation was also supported by the adjusted R-squared. Additionally, at specific thresholds, this linear relationship exhibits a slope near 1, resembling a translation.
Based on this observation, the question can be rephrased as follows:

Question 2. Is there a straight-line connection between $W_p^p(\mathcal{D}_{0,r}(\mathbb{X}), \mathcal{D}_{0,s}(\mathbb{Y}))$ and $W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y}))$ when we consider certain thresholds?

The purpose of this paper is to provide a positive answer to this question for two particular cases.

3 Main results

Before giving the first main result, we introduce some notations for the sake of simplification.
Since we deal with the connected components of the Rips persistence diagrams, all connected components appear at the time $t=0$. Then, the coordinates of the off-diagonal points in the persistence diagram are of the form $(0, a_{i_\mathbb{X}})$, where $a_{i_\mathbb{X}}$ is the death time of a connected component. After the death time $m_\mathbb{X} := \max(a_{i_\mathbb{X}}) < \infty$, all of these components merge into one that continues to infinity. Let us set $\alpha^+ := m_\mathbb{X} + \alpha$ for some real number $\alpha \geq 0$.
Now we are in position to set and prove the first main result. This theorem outlines a mathematical connection between the Wasserstein distances of persistence diagrams that come from two distinct types of filtrations. One is a complete filtration that takes into account all topological connections, while the other is a prematurely stopped filtration, which cuts off at a threshold set by α and a real number k.

Theorem 3.1. *Let $\mathcal{D}_0(\mathbb{X})$ and $\mathcal{D}_0(\mathbb{Y})$ be persistence Rips diagrams such that $m_\mathbb{X} \leq m_\mathbb{Y}$.*
For $\beta = k(m_Y - m_X) + \alpha$, where $\alpha \geq 0$ and $1 \leq k \leq 2$,

$$\begin{aligned} W_p^p\big(\mathcal{D}_{0,\beta^+}(\mathbb{X}), \mathcal{D}_{0,\alpha^+}(\mathbb{Y})\big) &= W_p^p\big(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y})\big) \\ &\quad + (k-1)^p\big(m_\mathbb{Y} - m_\mathbb{X}\big)^p \end{aligned}$$

Proof. Let ϕ_0 be a bijection from $\mathcal{D}_0(\mathbb{X})$ to $\mathcal{D}_0(\mathbb{Y})$ where the minimum cost is reached, i.e.,

$$W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y})) = \sum_{x \in \mathcal{D}_0(\mathbb{X})} \|x - \phi_0(x)\|_\infty^p$$

. The bijection ϕ_0 can be extended to $\mathcal{D}_{0,\beta^+}(\mathbb{X})$ by assigning the new points $(0, m_{\mathbb{X}}+\beta)$ and $(\frac{m_{\mathbb{Y}}+\alpha}{2}, \frac{m_{\mathbb{Y}}+\alpha}{2})$. Thus, there are two cases:

- Case 1: $\phi_0((0, m_{\mathbb{X}}+\beta)) = (\frac{m_{\mathbb{X}}+\beta}{2}, \frac{m_{\mathbb{X}}+\beta}{2})$ and $\phi_0((\frac{m_{\mathbb{Y}}+\alpha}{2}, \frac{m_{\mathbb{Y}}+\alpha}{2})) = (0, m_{\mathbb{Y}}+\alpha)$
- Case 2: $\phi_0((0, m_{\mathbb{X}}+\beta)) = (0, m_{\mathbb{Y}}+\alpha)$ and $\phi_0((\frac{m_{\mathbb{Y}}+\alpha}{2}, \frac{m_{\mathbb{Y}}+\alpha}{2})) = (\frac{m_{\mathbb{X}}+\beta}{2}, \frac{m_{\mathbb{X}}+\beta}{2})$

So, the additional cost associated to the new extension ϕ_0 for each one of the two cases will be, respectively, as follows:

$$\begin{aligned}
C_1 &= ||\phi_0((0, m_{\mathbb{X}}+\beta)) - (0, m_{\mathbb{X}}+\beta)||_\infty^p + \\
&\quad ||\phi_0\left(\left(\frac{m_{\mathbb{Y}}+\alpha}{2}, \frac{m_{\mathbb{Y}}+\alpha}{2}\right)\right) - \left(\frac{m_{\mathbb{Y}}+\alpha}{2}, \frac{m_{\mathbb{Y}}+\alpha}{2}\right)||_\infty^p \\
&= ||\left(\frac{m_{\mathbb{X}}+\beta}{2}, \frac{m_{\mathbb{X}}+\beta}{2}\right) - (0, m_{\mathbb{X}}+\beta)||_\infty^p + \\
&\quad ||(0, m_{\mathbb{Y}}+\alpha) - \left(\frac{m_{\mathbb{Y}}+\alpha}{2}, \frac{m_{\mathbb{Y}}+\alpha}{2}\right)||_\infty^p \\
&= \left(\frac{m_{\mathbb{X}}+\beta}{2}\right)^p + \left(\frac{m_{\mathbb{Y}}+\alpha}{2}\right)^p \\
&= \left(\frac{(1-k)m_{\mathbb{X}}+km_{\mathbb{Y}}+\alpha}{2}\right)^p + \left(\frac{1}{2}m_{\mathbb{Y}}+\frac{1}{2}\alpha\right)^p
\end{aligned}$$

And

$$\begin{aligned}
C_2 &= ||\phi_0((0, m_{\mathbb{X}}+\beta)) - (0, m_{\mathbb{X}}+\beta)||_\infty^p + \\
&\quad ||\phi_0\left(\left(\frac{m_{\mathbb{Y}}+\alpha}{2}, \frac{m_{\mathbb{Y}}+\alpha}{2}\right)\right) - \left(\frac{m_{\mathbb{Y}}+\alpha}{2}, \frac{m_{\mathbb{Y}}+\alpha}{2}\right)||_\infty^p \\
&= ||(0, m_{\mathbb{Y}}+\alpha) - (0, m_{\mathbb{X}}+\beta)||_\infty^p + \\
&\quad \underbrace{||\left(\frac{m_{\mathbb{X}}+\beta}{2}, \frac{m_{\mathbb{X}}+\beta}{2}\right) - \left(\frac{m_{\mathbb{Y}}+\alpha}{2}, \frac{m_{\mathbb{Y}}+\alpha}{2}\right)||_\infty^p}_{0} \\
&= ||(0, m_{\mathbb{Y}}+\alpha) - (0, (1-k)m_{\mathbb{X}}+km_{\mathbb{Y}}+\alpha)||_\infty^p \\
&= [(k-1)(m_{\mathbb{Y}}-m_{\mathbb{X}})]^p
\end{aligned}$$

Clearly $C_2 < C_1$, so

$$\begin{aligned}
&W_p^p\big(\mathcal{D}_{0,\beta^+}(\mathbb{X}), \mathcal{D}_{0,\alpha^+}(\mathbb{Y})\big) \\
&\qquad \le W_p^p\big(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y})\big) + \big[(k-1)(m_{\mathbb{Y}}-m_{\mathbb{X}})\big]^p \qquad (1)
\end{aligned}$$

It remains to prove the converse inequality. Consider a bijection $\psi : \mathcal{D}_{0,\beta^+}(\mathbb{X}) \longrightarrow \mathcal{D}_{0,\alpha^+}(\mathbb{Y})$.
If $\psi(\{(0, m_{\mathbb{X}}+\beta); (\frac{m_{\mathbb{Y}}+\alpha}{2}, \frac{m_{\mathbb{Y}}+\alpha}{2})\}) = \{(0, m_{\mathbb{Y}}+\alpha); (\frac{m_{\mathbb{X}}+\beta}{2}, \frac{m_{\mathbb{X}}+\beta}{2})\}$, then using the restriction η of ψ on $\mathcal{D}_0(\mathbb{X})$, we get:
$C_+ = C_0 + C$, where $C_+ = \sum_{x\in\mathcal{D}_{0,\beta^+}(\mathbb{X})} \|x-\psi(x)\|_\infty^p$, $C_0 = \sum_{x\in\mathcal{D}_0(\mathbb{X})} \|x-\eta(x)\|_\infty^p$ and C is either C_1 or C_2. Hence, the desired inequality holds.

Now, suppose that

$\psi((0, m_{\mathbb{X}}+\beta)) \neq (0, m_{\mathbb{Y}}+\alpha)$, i.e., $\psi((0, m_{\mathbb{X}}+\beta) = (0, a_{i_{\mathbb{Y}}}))$, where $a_{i_{\mathbb{Y}}} < m_{\mathbb{Y}}$. So there exists $(0, a_{i_{\mathbb{X}}}) \in \mathcal{D}_0(\mathbb{X})$ such that $\psi((0, a_{i_{\mathbb{X}}})) = (0, m_{\mathbb{Y}}+\alpha)$.
Consider the bijection η defined by:

$\eta(x) = \psi(x)$, if $x \in \mathcal{D}_0(\mathbb{X})$ and $\psi(x) \in \mathcal{D}_0(\mathbb{Y})$ and $\eta((0, a_{i_{\mathbb{X}}})) = (0, a_{i_{\mathbb{Y}}})$. We have

$$\begin{aligned}
C_+ &= \sum_{x\in\mathcal{D}_0(\mathbb{X})} \|x-\psi(x)\|_\infty^p + \|\psi((0, m_{\mathbb{X}}+\beta)) - (0, m_{\mathbb{X}}+\beta)\|_\infty^p \\
&\quad + \|\psi((\frac{m_{\mathbb{Y}}+\alpha}{2}, \frac{m_{\mathbb{Y}}+\alpha}{2})) - (\frac{m_{\mathbb{Y}}+\alpha}{2}, \frac{m_{\mathbb{Y}}+\alpha}{2})\|_\infty^p \\
&= \sum_{x\in\mathcal{D}_0(\mathbb{X})} \|x-\eta(x)\|_\infty^p + \|\psi((0, a_{i_{\mathbb{X}}})) - (0, a_{i_{\mathbb{X}}})\|_\infty^p + \\
&\quad \|\psi((0, m_{\mathbb{X}}+\beta)) - (0, m_{\mathbb{X}}+\beta)\|_\infty^p + \\
&\quad \|\psi((\frac{m_{\mathbb{Y}}+\alpha}{2}, \frac{m_{\mathbb{Y}}+\alpha}{2})) - (\frac{m_{\mathbb{Y}}+\alpha}{2}, \frac{m_{\mathbb{Y}}+\alpha}{2})\|_\infty^p - \\
&\quad \|\eta((0, a_{i_{\mathbb{X}}})) - (0, a_{i_{\mathbb{X}}})\|_\infty^p \\
&\ge \sum_{x\in\mathcal{D}_0(\mathbb{X})} \|x-\eta(x)\|_\infty^p + \|(0, m_{\mathbb{Y}}+\alpha) - (0, a_{i_{\mathbb{X}}})\|_\infty^p + \\
&\quad \|(0, a_{i_{\mathbb{Y}}}) - (0, m_{\mathbb{X}}+\beta)\|_\infty^p - \|(0, a_{i_{\mathbb{Y}}}) - (0, a_{i_{\mathbb{X}}})\|_\infty^p \\
&\ge \sum_{x\in\mathcal{D}_0(\mathbb{X})} \|x-\phi_0(x)\|_\infty^p + |m_{\mathbb{Y}}+\alpha-a_{i_{\mathbb{X}}}|^p + \\
&\quad |m_{\mathbb{X}}+\beta-a_{i_{\mathbb{Y}}}|^p - |a_{i_{\mathbb{Y}}}-a_{i_{\mathbb{X}}}|^p
\end{aligned}$$

One can show that

$$|m_{\mathbb{Y}}+\alpha-a_{i_{\mathbb{X}}}|^p + |m_{\mathbb{X}}+\beta-a_{i_{\mathbb{Y}}}|^p - |a_{i_{\mathbb{Y}}}-a_{i_{\mathbb{X}}}|^p \ge C_2$$

Therefore, the converse inequality of (1) holds. ■

Now we give the second main theorem. It involves the value proceeding the largest value of the ordinates of the off-diagonal points of the persistence diagrams. For this, additional notations must be introduced.
First, recall the notation of the underlying set $\underline{\mathcal{D}}_0(\mathbb{X})$=$\{(0, a_{i_{\mathbb{X}}}) | a_{i_{\mathbb{X}}} \in \mathbb{R}^+\}$ of a persistence Rips diagram $\mathcal{D}_0(\mathbb{X})$. We denote $n_{\mathbb{X}}$ the maximum of values $a_{i_{\mathbb{X}}}$ such that $(0, a_{i_{\mathbb{X}}}) \in \underline{\mathcal{D}}_0(\mathbb{X}) - \{(0, m_{\mathbb{X}})\}$.

This theorem allows us to compute the Wasserstein distance by taking into account the maximum and the second largest values from one of the persistence diagrams. The output provides a quantification of the error in two situations that happen when the filtrations are truncated too early.

Theorem 3.2. *Let $\mathcal{D}_0(\mathbb{X})$ and $\mathcal{D}_0(\mathbb{Y})$ be persistence Rips diagrams. Assume that $m_{\mathbb{X}} < m_{\mathbb{Y}}$ and that $|\underline{\mathcal{D}}_0(\mathbb{Y})| \ge 2$. Then, for every $\alpha \ge m_{\mathbb{Y}} - n_{\mathbb{Y}}$,*

$$W_p^p(\mathcal{D}_{0,\alpha^+}(\mathbb{X}), \mathcal{D}_{0,\alpha^+}(\mathbb{Y})) = W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y})) + (m_{\mathbb{Y}}-m_{\mathbb{X}})^p$$

when $m_{\mathbb{X}} > \frac{m_{\mathbb{Y}}+n_{\mathbb{Y}}}{2}$.
Otherwise,

$$W_p^p(\mathcal{D}_{0,\beta^+}(\mathbb{X}), \mathcal{D}_{0,\alpha^+}(\mathbb{Y})) = W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y})) + (\frac{m_{\mathbb{Y}}-n_{\mathbb{Y}}}{2})^p$$

where $\beta = \frac{m_{\mathbb{Y}}+n_{\mathbb{Y}}}{2} - m_{\mathbb{X}} + \alpha$.

Proof. 1) Assume that $m_{\mathbb{X}} > \frac{m_{\mathbb{Y}}+n_{\mathbb{Y}}}{2}$. Let ϕ_0 be the bijection from $\mathcal{D}_0(\mathbb{X})$ to $\mathcal{D}_0(\mathbb{Y})$ where the minimum cost is reached. We extend ϕ_0 to

$$\underline{\mathcal{D}}_0(\mathbb{X}) \cup \underline{\mathcal{D}'}_0(\mathbb{Y}) \cup \{(0, m_{\mathbb{X}}+\alpha), (\frac{m_{\mathbb{Y}}+\alpha}{2}, \frac{m_{\mathbb{Y}}+\alpha}{2})\}$$

so that the resulting cost remains minimal. The bijection ϕ_0 can be extended in two ways:

- Case 1.1:

$$\phi_0((0, m_{\mathbb{X}}+\alpha)) = (\frac{m_{\mathbb{X}}+\alpha}{2}, \frac{m_{\mathbb{X}}+\alpha}{2})$$

and

$$\phi_0((\frac{m_{\mathbb{Y}}+\alpha}{2},\frac{m_{\mathbb{Y}}+\alpha}{2}))=(0,m_{\mathbb{Y}}+\alpha)$$

- Case 2.1:

$$\phi_0((0,m_{\mathbb{X}}+\alpha))=(0,m_{\mathbb{Y}}+\alpha)$$

and

$$\phi_0((\frac{m_{\mathbb{Y}}+\alpha}{2},\frac{m_{\mathbb{Y}}+\alpha}{2}))=(\frac{m_{\mathbb{X}}+\alpha}{2},\frac{m_{\mathbb{X}}+\alpha}{2})$$

As done in the proof of Theorem 3.1, we will show that the additional cost

$$C=||\phi_0((0,m_{\mathbb{X}}+\alpha))-(0,m_{\mathbb{X}}+\alpha)||_\infty^p + ||\phi_0((\frac{m_{\mathbb{Y}}+\alpha}{2},\frac{m_{\mathbb{Y}}+\alpha}{2}))-(\frac{m_{\mathbb{Y}}+\alpha}{2},\frac{m_{\mathbb{Y}}+\alpha}{2})||_\infty^p$$

is minimum for the second case.

- Case 1.1: We have,

$$\begin{aligned}C_{1.1}&=\left\|\phi_0\big((0,m_{\mathbb{X}}+\alpha)\big)-(0,m_{\mathbb{X}}+\alpha)\right\|_\infty^p\\&\quad+\left\|\phi_0\left(\tfrac{m_{\mathbb{Y}}+\alpha}{2},\tfrac{m_{\mathbb{Y}}+\alpha}{2}\right)\right.\\&\quad\left.-\left(\tfrac{m_{\mathbb{Y}}+\alpha}{2},\tfrac{m_{\mathbb{Y}}+\alpha}{2}\right)\right\|_\infty^p\\&=\left\|\left(\tfrac{m_{\mathbb{X}}+\alpha}{2},\tfrac{m_{\mathbb{X}}+\alpha}{2}\right)-(0,m_{\mathbb{X}}+\alpha)\right\|_\infty^p\\&\quad+\left\|(0,m_{\mathbb{Y}}+\alpha)-\left(\tfrac{m_{\mathbb{Y}}+\alpha}{2},\tfrac{m_{\mathbb{Y}}+\alpha}{2}\right)\right\|_\infty^p\\&=\left(\tfrac{m_{\mathbb{X}}+\alpha}{2}\right)^p+\left(\tfrac{m_{\mathbb{Y}}+\alpha}{2}\right)^p\end{aligned}$$

- Case 2.1:

$$\begin{aligned}C_{2.1}&=\left\|\phi_0\big((0,m_{\mathbb{X}}+\alpha)\big)-(0,m_{\mathbb{X}}+\alpha)\right\|_\infty^p\\&\quad+\left\|\phi_0\left(\tfrac{m_{\mathbb{Y}}+\alpha}{2},\tfrac{m_{\mathbb{Y}}+\alpha}{2}\right)-\left(\tfrac{m_{\mathbb{Y}}+\alpha}{2},\tfrac{m_{\mathbb{Y}}+\alpha}{2}\right)\right\|_\infty^p\\&=\left\|(0,m_{\mathbb{Y}}+\alpha)-(0,m_{\mathbb{X}}+\alpha)\right\|_\infty^p\\&\quad+\underbrace{\left\|\left(\tfrac{m_{\mathbb{Y}}+\alpha}{2},\tfrac{m_{\mathbb{Y}}+\alpha}{2}\right)-\left(\tfrac{m_{\mathbb{X}}+\alpha}{2},\tfrac{m_{\mathbb{X}}+\alpha}{2}\right)\right\|_\infty^p}_{0}\\&=(m_{\mathbb{Y}}-m_{\mathbb{X}})^p\end{aligned}$$

Let us show that $C_{2.1}<C_{1.1}$.
Using the hypothesis $m_{\mathbb{X}}>\frac{m_{\mathbb{Y}}+n_{\mathbb{Y}}}{2}$, we get

$$(m_{\mathbb{Y}}-m_{\mathbb{X}})^p<(\frac{m_{\mathbb{Y}}+\alpha}{2})^p$$

and since $\frac{m_{\mathbb{Y}}+\alpha}{2}$ is a positive real, we conclude that $C_{2.1}<C_{1.1}$. Hence,

$$\begin{aligned}&W_p^p\big(\mathcal{D}_{0,\alpha^+}(\mathbb{X}),\mathcal{D}_{0,\alpha^+}(\mathbb{Y})\big)\\&\qquad\le W_p^p\big(\mathcal{D}_0(\mathbb{X}),\mathcal{D}_0(\mathbb{Y})\big)+\big(m_{\mathbb{Y}}-m_{\mathbb{X}}\big)^p\end{aligned}\tag{2}$$

Now, we proceed in the same way as in the proof of Theorem 3.1 and we keep the same notations of ψ and η. The goal is to prove that

$$\begin{aligned}\sum_{x\in\mathcal{D}_{0,\beta^+}(\mathbb{X})}\|x-\psi(x)\|_\infty^p&\ge\sum_{x\in\mathcal{D}_0(\mathbb{X})}\|x-\eta(x)\|_\infty^p+C_{2.1}\\&\ge\sum_{x\in\mathcal{D}_0(\mathbb{X})}\|x-\phi_0(x)\|_\infty^p+C_{2.1}.\end{aligned}$$

This requires to show that

$$|m_{\mathbb{Y}}+\alpha-a_{i_{\mathbb{X}}}|^p+|m_{\mathbb{X}}+\alpha-a_{i_{\mathbb{Y}}}|^p\ge|a_{i_{\mathbb{Y}}}-a_{i_{\mathbb{X}}}|^p+(m_{\mathbb{Y}}-m_{\mathbb{X}})^p$$

In fact this follows from the fact that

$$|m_{\mathbb{Y}}+\alpha-a_{i_{\mathbb{X}}}|^p\ge|a_{i_{\mathbb{Y}}}-a_{i_{\mathbb{X}}}|^p$$

And

$$|m_{\mathbb{X}}+\alpha-a_{i_{\mathbb{Y}}}|^p\ge(m_{\mathbb{Y}}-m_{\mathbb{X}})^p$$

Hence,

$$W_p^p(\mathcal{D}_{0,\alpha^+}(\mathbb{X}),\mathcal{D}_{0,\alpha^+}(\mathbb{Y}))\ge W_p^p(\mathcal{D}_0(\mathbb{X}),\mathcal{D}_0(\mathbb{Y}))+(m_{\mathbb{Y}}-m_{\mathbb{X}})^p\tag{3}$$

The inequalities (2) and (3) complete the proof of the first result of this theorem.

2) Now we treat the second case where $m_{\mathbb{X}}\le\frac{m_{\mathbb{Y}}+n_{\mathbb{Y}}}{2}$.
There are two cases to discuss:

- Case 1.2: We have,

$$\begin{aligned}C_{1.2}&=||\phi_0((0,m_{\mathbb{X}}+\beta))-(0,m_{\mathbb{X}}+\beta)||_\infty^p+\\&\quad||\phi_0\left(\left(\frac{m_{\mathbb{Y}}+\alpha}{2},\frac{m_{\mathbb{Y}}+\alpha}{2}\right)\right)-\left(\frac{m_{\mathbb{Y}}+\alpha}{2},\frac{m_{\mathbb{Y}}+\alpha}{2}\right)||_\infty^p+\\&\quad||\left(\frac{m_{\mathbb{X}}+\beta}{2},\frac{m_{\mathbb{X}}+\beta}{2}\right)-(0,m_{\mathbb{X}}+\beta)||_\infty^p+\\&\quad||(0,m_{\mathbb{Y}}+\alpha)-\left(\frac{m_{\mathbb{Y}}+\alpha}{2},\frac{m_{\mathbb{Y}}+\alpha}{2}\right)||_\infty^p\\&=\left(\frac{m_{\mathbb{X}}+\beta}{2}\right)^p+\left(\frac{m_{\mathbb{Y}}+\alpha}{2}\right)^p+\\&\quad\left(\frac{m_{\mathbb{Y}}+n_{\mathbb{Y}}}{4}+\frac{\alpha}{2}\right)^p+\left(\frac{m_{\mathbb{Y}}+\alpha}{2}\right)^p\end{aligned}$$

- Case 2.2:

$$\begin{aligned}C_{2.2}&=||\phi_0((0,m_{\mathbb{X}}+\beta))-(0,m_{\mathbb{X}}+\beta)||_\infty^p+\\&\quad||\phi_0\left(\left(\frac{m_{\mathbb{Y}}+\alpha}{2},\frac{m_{\mathbb{Y}}+\alpha}{2}\right)\right)-\left(\frac{m_{\mathbb{Y}}+\alpha}{2},\frac{m_{\mathbb{Y}}+\alpha}{2}\right)||_\infty^p\\&=||(0,m_{\mathbb{Y}}+\alpha)-(0,m_{\mathbb{X}}+\beta)||_\infty^p+\\&\quad\underbrace{||\left(\frac{m_{\mathbb{X}}+\beta}{2},\frac{m_{\mathbb{X}}+\beta}{2}\right)-\left(\frac{m_{\mathbb{Y}}+\alpha}{2},\frac{m_{\mathbb{Y}}+\alpha}{2}\right)||_\infty^p}_{0}\\&=\left(\frac{m_{\mathbb{Y}}-n_{\mathbb{Y}}}{2}\right)^p\end{aligned}$$

Since

$$(\frac{m_{\mathbb{Y}}-n_{\mathbb{Y}}}{2})^p<(\frac{m_{\mathbb{Y}}+\alpha}{2})^p$$

And

$$(\frac{m_{\mathbb{Y}}+n_{\mathbb{Y}}}{4}+\frac{\alpha}{2})^p>0$$

We conclude that $C_{2.2}<C_{1.2}$.
Hence,

$$W_p^p(\mathcal{D}_{0,\beta^+}(\mathbb{X}),\mathcal{D}_{0,\alpha^+}(\mathbb{Y}))\le W_p^p(\mathcal{D}_0(\mathbb{X}),\mathcal{D}_0(\mathbb{Y}))+\left(\frac{m_{\mathbb{Y}}-n_{\mathbb{Y}}}{2}\right)^p\tag{4}$$

As before, it remains to show that

$$|m_{\mathbb{Y}}+\alpha-a_{i_{\mathbb{X}}}|^p+|m_{\mathbb{X}}+\alpha-a_{i_{\mathbb{Y}}}|^p\ge|a_{i_Y}-a_{i_{\mathbb{X}}}|^p+(\tfrac{m_{\mathbb{Y}}-n_{\mathbb{Y}}}{2})^p.$$

This is possible by distinguishing the two cases

$a_{i_{\mathbb{X}}}<a_{i_{\mathbb{Y}}}$ and $a_{i_{\mathbb{X}}}\ge a_{i_{\mathbb{Y}}}$ and by using the fact that $\alpha\ge m_{\mathbb{Y}}-n_{\mathbb{Y}}$ So,

$$W_p^p(\mathcal{D}_{0,\beta^+}(\mathbb{X}),\mathcal{D}_{0,\alpha^+}(\mathbb{Y}))\ge W_p^p(\mathcal{D}_0(\mathbb{X}),\mathcal{D}_0(\mathbb{Y}))+\left(\frac{m_{\mathbb{Y}}-n_{\mathbb{Y}}}{2}\right)^p\tag{5}$$

Inequalities (4) and (5) complete the proof of the second equality. ■

In the next section, we provide algorithms to implement and illustrate the theorems. Here, we present a simple example to better understand the applicability of Theorems 3.1 and 3.2. Specifically, we examine the geometric setting in which these results hold. Our goal is to describe what types of finite point clouds in a metric space can give rise to the persistence diagrams considered in the theorems.

Let $\mathbb{Y}$ be a finite point cloud in a metric space (X,d), and let $D_0(\mathbb{Y})$ be its associated persistence diagram capturing the connected components. For a given $\alpha \geq 0$, define the threshold $\alpha^+ := m_{\mathbb{Y}} + \alpha$, where $m_{\mathbb{Y}}$ is the largest death time in $D_0(m_{\mathbb{Y}})$. We define the extended set

$$\mathbb{Y}_{\alpha^+} := \mathbb{Y} \cup \{x \in X \mid d(x, \mathbb{Y}) = \alpha^+\}.$$

Any discrete subset of $\mathbb{Y}_{\alpha^+}$ that contains $\mathbb{Y}$ gives rise to a Rips persistence diagram in which all connected components merge precisely at time α^+, that is, a diagram of the form $D_{0,\alpha^+}(\mathbb{Y})$.

To illustrate this, consider a point cloud $\mathbb{Y}$ formed by three points in $\mathbb{R}^2$, arranged as a right triangle with sides 1, 2, and $\sqrt{5}$. For this cloud, we have $n_Y = 1$, $m_{\mathbb{Y}} = 2$, and $\mathbb{Y}_{3^+}$ is the union of Y with three circular arcs (see Figure 2).

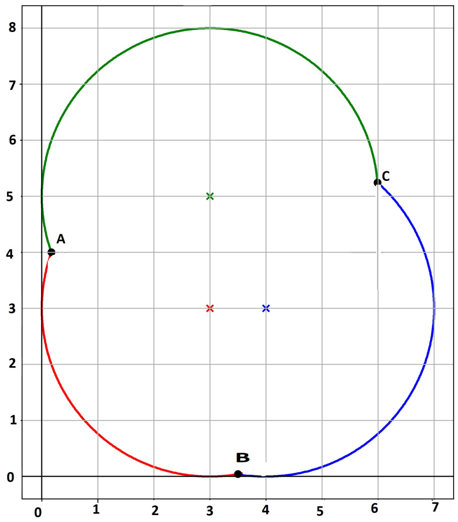

Figure 2: Point cloud $\mathbb{Y}$ of three points in the plane and its extended set $\mathbb{Y}_{3^+}$.

Now consider a second point cloud $\mathbb{X}$, consisting of four points forming a square of side 1, so that $m_{\mathbb{X}} = 1$. Let $\beta = 2.5$, and define the extended set $\mathbb{X}_{\beta^+}$ by adding four arcs at distance $\beta^+ = m_{\mathbb{X}} + \beta = 2.5$ (see Figure 3). These values of α and β follow the assumptions of Theorem 3.2: since $m_{\mathbb{Y}} = 2, n_{\mathbb{Y}} = 1$, and $m_{\mathbb{X}} = 1$, we choose $\alpha = m_{\mathbb{Y}} - n_{\mathbb{Y}} = 1$, and since $m_{\mathbb{X}} \leq \frac{m_{\mathbb{Y}}+n_{\mathbb{Y}}}{2}$, we compute

$$\beta = \frac{m_{\mathbb{Y}} + n_{\mathbb{Y}}}{2} - m_{\mathbb{X}} + \alpha = 1.5.$$

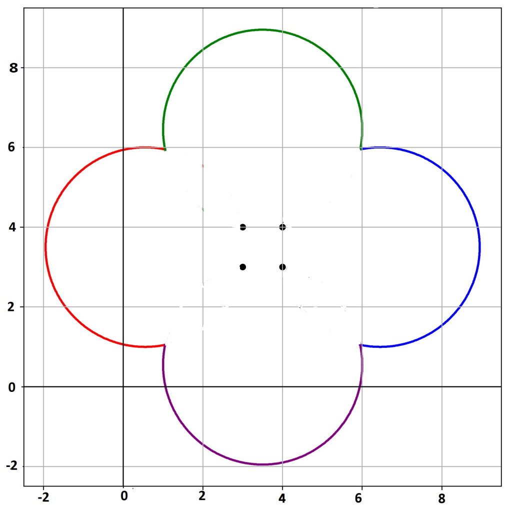

Figure 3: Point cloud $\mathbb{X}$ and its extended set $\mathbb{X}_{2.5^+}$.

Now, let A be a finite point cloud such that $\mathbb{X} \subset A \subset \mathbb{X}_{\beta^+}$, and let B be a finite point cloud such that $\mathbb{Y} \subset B \subset \mathbb{Y}_{\alpha^+}$. Let's take a look at the Rips persistence diagrams $D_0(A)$ and $D_0(B)$. Their maximum death times are β^+ and α^+, respectively. This means we can apply Theorem 3.2. Specifically, we have the equation:

$$W_p^p(D_0(A), D_0(B)) = W_p^p(D_0(\mathbb{X}), D_0(\mathbb{Y})) + \left(\frac{m_{\mathbb{Y}} - n_{\mathbb{Y}}}{2}\right)^p$$

It turns out that the quantity $W_p^p(D_0(A), D_0(B))$ stays the same for any pair of point clouds A and B selected in the way we've discussed.

4 Implementation and illustration of main theorems

In this section, we carry out the implementation of Theorems 3.1 and 3.2 via two different algorithms and demonstrate their applications using real-world instances represented as weighted networks.

4.1 Implementation of Theorem 3.1

To put Theorem 3.1 to good use, we recommend trying out the following Algorithm 1. This approach will allow us to effectively apply Theorem 3.1 to the two real dynamic networks we talked about at the beginning of Subsection 2.4.

Algorithm 1 Computation of $W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y}))$ with Theorem 3.1

Require: Rips filtrations $\mathcal{F}(\mathbb{X})$ and $\mathcal{F}(\mathbb{Y})$.
Ensure: The Wasserstein distance $W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y}))$.
1. Extract $m_{\mathbb{X}}$ and $m_{\mathbb{Y}}$ from the filtrations $\mathcal{F}(\mathbb{X})$ and $\mathcal{F}(\mathbb{Y})$.
2. Compare $m_{\mathbb{X}}$ and $m_{\mathbb{Y}}$.
3. Pick $\alpha > 0$, $1 \leq k \leq 2$ and compute β.
4. Add the point $(0, m_{\mathbb{X}} + \alpha)$ to the persistence diagram $\mathcal{D}_0(\mathbb{X})$ and $(0, m_{\mathbb{Y}} + \beta)$ to $\mathcal{D}_0(\mathbb{Y})$.
5. Compute the distance $W_p^p(\mathcal{D}_{0,\beta^+}(\mathbb{X}), \mathcal{D}_{0,\alpha^+}(\mathbb{Y}))$.
6. $W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y})) = W_p^p(\mathcal{D}_{0,\beta^+}(\mathbb{X}), \mathcal{D}_{0,\alpha^+}(\mathbb{Y})) - (1 - \frac{1}{k})^p(\beta - \alpha)^p$

We use $(X_t)_t$ to denote the time series of Wasserstein distances between persistence diagrams that come from full filtrations in the dynamic network. Similarly, $(\widetilde{X}_t)_t$ will represent the time series of the Wasserstein distance between persistence diagrams derived from subfiltrations, when the thresholds used are unambiguous.

Example 4.1. *This example is constructed from the multivariate time series of the closing prices of the four cryptocurrencies Bitcoin, Ethereum, Litecoin, and Ripple, from August 24, 2016, to February 19, 2020. This data is available on the website* `www.investing.com`

Figure 4 represents three plots of the time series of 2-Wasserstein distances between the persistence diagrams for 1224 instances and the reference diagram for $t_0 = 1005$. Black, blue, green and red graphs refer to the values of the parameter k for Thoerem 3.1 of 2, 1.7, 1.5 and 1 respectively. It should be noted that the red graph represents the Wasserstein distance between diagrams of complete filtration. Since the factor $(k-1)^p(m_{\mathbb{Y}} - m_{\mathbb{X}})^p$ is strictly increasing with respect to the real $1 \leq k \leq 2$, it follows that the graph for any value of $1 \leq k \leq 2$ will lie between the (red) graph for k = 1 and the (black) one for k = 2.

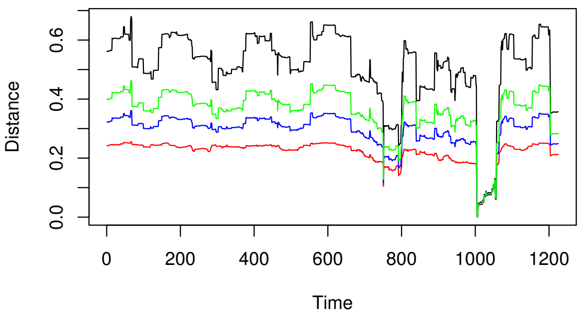

Figure 4: The graphs displaying $(\widetilde{X}_t)_t$ values for $k = 2$ (in black), 1.7 (in blue), $k = 1.5$ (in green) and $k = 1$ (in red) .

Time series of the difference $(\widetilde{X}_t - X_t)_t$ in the Figure 4 illustrates nothing but the factor $(k-1)^p(m_{\mathbb{Y}} - m_{\mathbb{X}})^p$, for different chosen values of k: the graph corresponding to $k = 2$ is drawn in black, for $k = 1.7$ in blue, and for $k = 1.5$ in green. According to Theorem 3.1, we have the following inequalities:

$$0 \leq W_p^p(\mathcal{D}_{0,\beta^+}(\mathbb{X}), \mathcal{D}_{0,\alpha^+}(\mathbb{Y})) - W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y})) \leq (m_{\mathbb{Y}} - m_{\mathbb{X}})^p$$

The black graph represents the time series of the right-hand side of this inequality, which is why it appears higher than the others. Additionally, as k approaches 1, the graph of the time series $(\widetilde{X}_t - X_t)_t$ gets closer to 0.

Now, we explore another kind of weighted networks.

Example 4.2. *The Dow Jones Industrial Average (DJIA) is a stock market index representing 30 major publicly traded companies in the United States. This second example of dynamic network is derived from the DJIA stocks listed as of February 19, 2008, and the data considered corresponds to the period between January 2004 and September 2008. This data was studied in [11] and also in Subsection* 6.1 *of [1].*

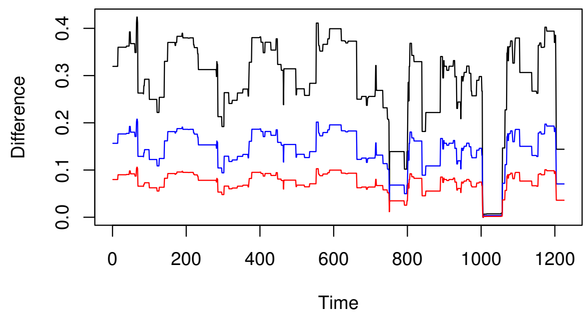

Figure 5: Plots of the time series $((\widetilde{X}_t - X_t)t)$ for the values $k = 2$ (in black), 1.7 (in blue) and 1.5 (in red).

Figure 6 is the superposition of the graphs of the time series, of the Wasserstein distances between the persistence diagrams and a reference diagram, $(X_t)_t$.
The graphs depicted correspond to the values: $k = 1$, $k = 1.7$ and $k = 2$ of theorem 3.1.
We notice that the graphs look nearly identical to each other. This similarity arises because the dead times, $m_{\mathbb{X}}$ and $m_{\mathbb{Y}}$, are almost the same for each pair of persistence diagrams.

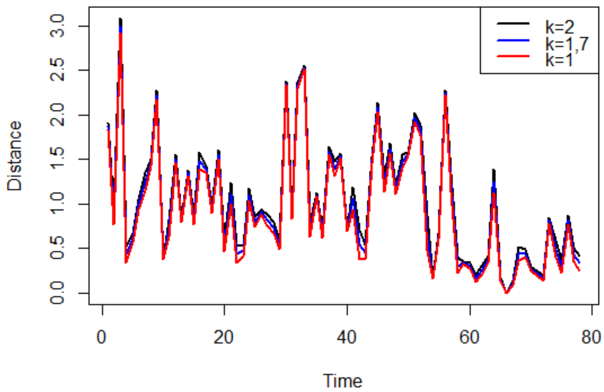

Figure 6: 2-Wasserstein distances between persistence diagrams with Rips filtration with $k = 2$ (in black), $k = 1.7$ (in blue) and $k = 1$ (in red).

4.2 Implementation of Theorem 3.2

We implemented Theorem 3.2 in this subsection by way of the subsequent Algorithm2.

Algorithm 2 Computation of $W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y}))$ using Theorem 3.2

Require: Rips filtrations $\mathcal{F}(\mathbb{X})$ and $\mathcal{F}(\mathbb{Y})$.
Ensure: The Wasserstein distance $W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y}))$.
1. Extract $m_{\mathbb{X}}$, $m_{\mathbb{Y}}$, $n_{\mathbb{X}}$ and $n_{\mathbb{Y}}$ from the filtrations.
2. $\alpha \leftarrow m_{\mathbb{X}} - n_{\mathbb{X}}$
if $m_{\mathbb{X}} > \frac{m_{\mathbb{Y}} + n_{\mathbb{Y}}}{2}$ **then**
$W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y})) = W_p^p(\mathcal{D}_{0,\alpha^+}(\mathbb{X}), \mathcal{D}_{0,\alpha^+}(\mathbb{Y})) - (m_{\mathbb{Y}} - m_{\mathbb{X}})^p$
else
$\beta \leftarrow \frac{m_{\mathbb{Y}} + n_{\mathbb{Y}}}{2} - m_{\mathbb{X}} + \alpha$
$W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y})) = W_p^p(\mathcal{D}_{0,\beta^+}(\mathbb{X}), \mathcal{D}_{0,\alpha^+}(\mathbb{Y})) - (\frac{m_{\mathbb{Y}} - n_{\mathbb{Y}}}{2})^p$
end if
2. Return the Wasserstein distance $W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y}))$.

We illustrate Theorem 3.2 using the same examples of dynamic networks as in Subsection 4.1.

Example 4.3. *In this example, Theorem 3.2 was applied to the data in Example 4.1.*

Figure 7 depicts the time series of 2-Wasserstein distances related to cryptocurrency data, represented by (X_t) (in blue) and $(\widetilde{X}_t)$ (in black), following the assumptions made in theorem 3.2. The visual analysis shows a higher similarity between the graphs and a smaller discrepancy between them. Moreover, the representative graph for (X_t) lies below that for $(\widetilde{X}_t)$.

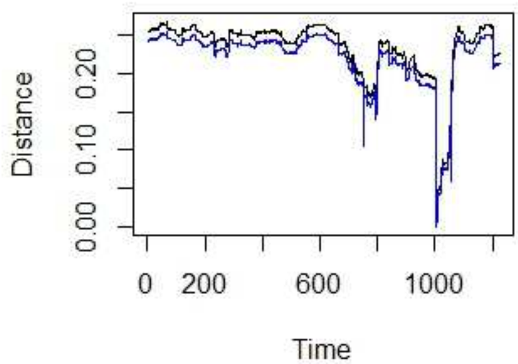

Figure 7: 2-Wasserstein distances: complete Rips (blue) vs.early-stopped filtration (black) (Theorem 3.2).

Example 4.4. *In this example, we apply Theorem 3.2 to the data used in Example 4.2.*

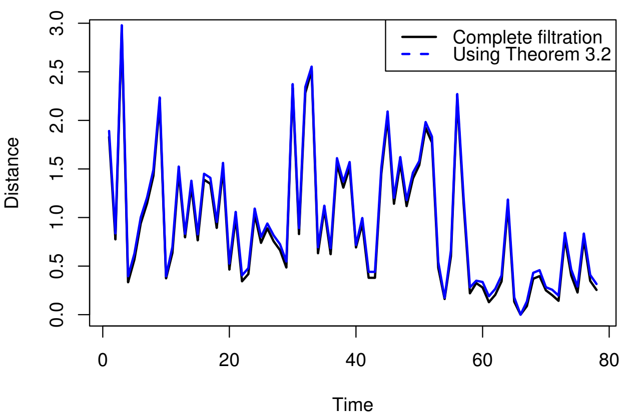

Figure 8: 2-Wasserstein distances between persistence diagrams. Complete Rips filtration (black) vs. early sttoped filtration (blue) computed using Theorem 3.2.

The black graph in Figure 8 corresponds to an interrupted filtration according to the assumptions of Theorem 3.2.

According to the previous figures, it is clear that the difference between $W_p^p(\mathcal{D}_{0,\beta^+}(\mathbb{X}), \mathcal{D}_{0,\alpha^+}(\mathbb{Y}))$ and $W_p^p(\mathcal{D}_0(\mathbb{X}), \mathcal{D}_0(\mathbb{Y}))$ is smaller under the assumptions of Theorem 3.2.

The Theorems mentioned above use persistence diagrams as descriptors of point clouds without mentioning the clouds themselves. We will now focus on the impact of the previous theorems on these point clouds.

Following these instances of the Theorems 3.1 and 3.2. The following section is devoted to establishing the benefit of using the method described in Theorem 3.1. Indeed, there is significant savings in running time.

5 Computational Efficiency Analysis

This section deals with the computational gains of the reduction principle stemming from the two central results, Theorems 3.1 and 3.2. To investigate the saving brought by Theorem 3.1, we consider running time for increasing values of the parameter k, from $k = 1$ to $k = 2$. We apply our approach to the above-mentioned cryptocurrency network.

5.1 Cost reduction with Theorem 3.1

Table 1 summarizes the results, showing the mean execution time and its standard deviation across multiple repetitions. The computational gain is expressed as a ratio compared to the original filtration ($k = 1$).

Table 1: Mean execution time and standard deviation for different scaling parameters k. The computational gain is expressed as a ratio compared to $k = 1$.

Parameter k	Mean time (s)	Sd time (s)	Ratio
1.0	10.547	4.409	1.00
1.1	6.669	2.180	0.63
1.25	6.619	1.939	0.63
1.3	6.150	2.010	0.58
1.5	6.819	2.540	0.65
1.6	6.590	2.080	0.63
1.7	7.700	2.920	0.73
1.8	7.081	2.470	0.67
1.9	5.990	1.828	0.57
2.0	6.940	2.960	0.66

The results show that increasing the value of k can lead to a significant reduction in computation time. Specifically, $k = 1.3$ and $k = 1.9$ provide the best balance, cutting the mean execution time by 42% (Ratio = 0.58) and 43% (Ratio = 0.57), respectively.

Among these options, $k = 1.9$ results in the lowest mean execution time (5.99 s), making it the most efficient choice. On the other hand, $k = 1.3$ has a slightly higher mean execution time (6.15 s) but a lower standard deviation (Sd = 2.01), which indicates better stability.

Other selected parameter values, such as $k = 1.25$ and $k = 1.6$, also provide remarkable computational improvements (reducing execution time approximately by 37%), but with less stability compared to $k = 1.3$ due to somewhat higher variability.

Therefore, $k = 1.3$ is the optimal choice for achieving a balance between efficiency and stability, while $k = 1.9$ delivers the fastest computation, albeit with a bit more variance.

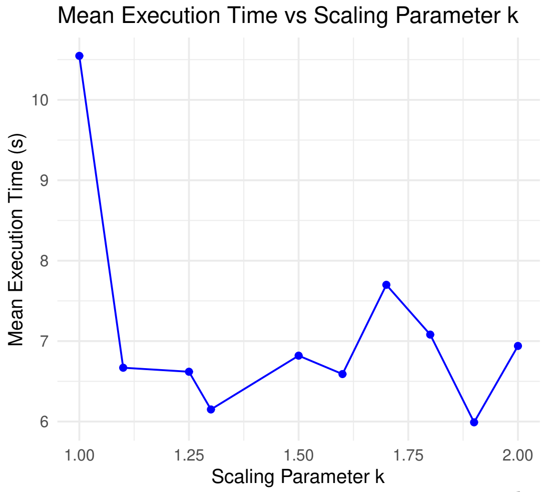

Figure 9: Execution time vs scaling parameter k.

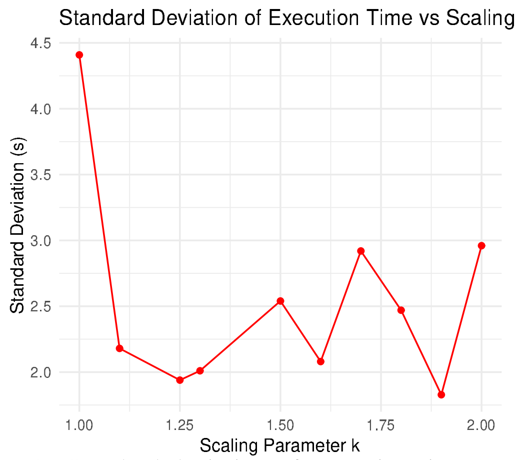

Figure 10: Standard deviation of execution time vs scaling parameter k.

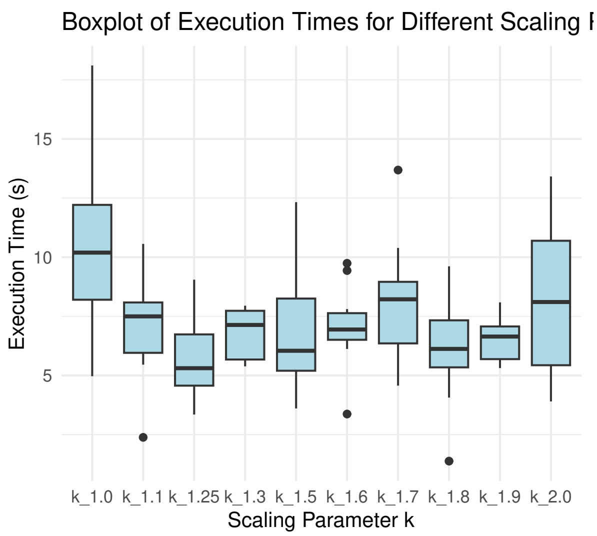

Figure 11: Boxplot of execution times for different scaling parameters k.

To analyze how different k values behave, we applied the k-means clustering algorithm [23]. This clustering focused on two key features: the average execution time and the standard deviation for each k.

The Elbow criterion was used to identify the optimal number of clusters, which relies on a plot of the sum of squared errors against k (for further discussions, see [14]). The "elbow" point at $k = 3$ supports our selection (refer to Figure 13).

The outcomes are illustrated in Figure 12, with a detailed description of each cluster provided below.

- *Cluster 0 (Blue):*
  - *Characteristics:* This cluster is characterized by $k = 1$ which represents the original persistence diagrams without any modifications.
  - *Values of k:* $k = 1$
  - *Interpretation:* The execution times for $k = 1.0$ are considerably higher compared to other configurations. This indicates that using the original persistence diagrams is computationally intensive, highlighting the effectiveness of simplification in lowering computational costs.
- *Cluster 1 (Green):*
  - *Characteristics:* This cluster shows low execution times with minimal variability, suggesting high stability.
  - *Values of k:*. $k = 1.3, 1.9, 2.0$
  - *Interpretation:* These values indicate the most efficient configurations, achieving a great balance between speed and stability. Their consistently low execution times make them the top choices for computational efficiency.
- *Cluster 2 (Orange):*
  - *Characteristics:* This cluster features moderate execution times and variability.
  - *Values of k:* $k = 1.1, 1.25, 1.5, 1.6, 1.7, 1.8.$
  - *Interpretation:* These configurations provide intermediate performance, balancing speed and computational stability. While they are not as optimal as those in Cluster 1, they still serve as viable alternatives when needed.

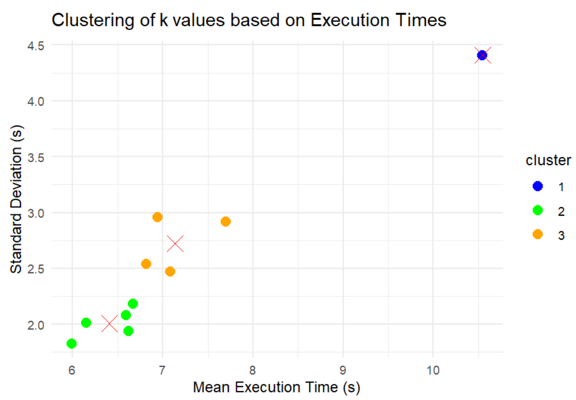

Figure 12: Clustering of k values based on execution times. Each cluster is represented by a unique color, and centroids are shown as red crosses.

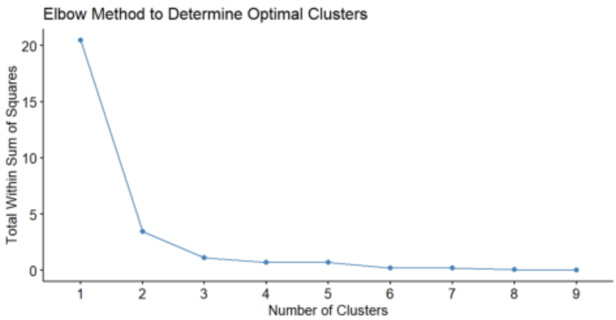

Figure 13: Elbow method to determine the optimal number of clusters.

5.2 Cost reduction with Theorem3.2

In contrast to Theorem 3.1, Theorem 3.2 does not include scale parameters. In this subsection, we compare the execution times for calculating Wasserstein distances between persistence diagrams derived from complete filtrations and those from truncated filtrations, as outlined in Theorem 3.2.

To quantify the impact of the simplification method on computational cost, we measured the execution time, for the original diagrams, across multiple runs. The results, before and after applying the simplification, are summarized in Table 2.

Table 2: Comparison of execution times, for original diagrams and after applying the Theorem 3.2 .

	Mean time (s)	Sd time (s)
Original Diagrams	10.547	4.409
After Simplification	5.99	1.399

The results show a 43.22% reduction in execution time, connecting the simplification process to a significant decrease in computational load. Additionally, the variability in execution times was greatly reduced, with values changing from 4.41 s to 1.4 s, indicating a much tighter computation. These testing outcomes align with the theoretical expectations outlined in Theorem 3.2.

5.3 Comparison with the Auction Algorithm for Wasserstein Distances Method

The *Auction Algorithm* is an iterative optimal assignment problem solving algorithm. It frames the assignment problem as a competitive auction where agents bid on objects based on how utilitarian the objects are to them. At each round, an agent selects the object that provides the maximum value, increases the bid, and assigns the rewarded object to the current highest bidder. This step is repeated iteratively until all assignments are found [3].

The method introduced by Kerber et al.[16] is an auction-based algorithm aimed at optimizing the computation of Wasserstein distances between persistence diagrams. Traditional methods typically involve solving an optimal transport problem, which can be quite resource-intensive. The auction algorithm redefines this problem, enabling quicker convergence while still providing theoretical accuracy guarantees.

This approach has gained popularity in topological data analysis TDA because of its efficiency in managing large-scale datasets. However, it is important to note that while it streamlines the transport computation, it does not simplify the complexity of the persistence diagrams themselves.

To evaluate the computational advantages of simplifying the diagrams, we compare the runtime of the Kerber et al. method when used on the original persistence diagrams ($k = 1$) with the standard $k = 1$ computation. The results show that:

- The average execution time for Kerber et al.'s method is 9.967 s, with a standard deviation of 4.665 s.
- The average time taken to process persistence diagrams from complete filtrations (k=1) is 10.547 s, with a standard deviation of 4.409 s.

This demonstrates that using auction-based optimization results in a computational improvement of approximately 5.5% when applied to persistence diagrams obtained from complete filtrations ($k = 1$).

When $k = 1$ is exceeded, our method effectively simplifies the persistence diagrams, leading to lower computational costs while maintaining important topological features. The execution times and their standard deviations for various k values are shown in Table 1.

When k exceeds 1, our method alone greatly reduces execution time.
- For example, at $k = 1.3$, the execution time drops to 6.150 s, which is 38.3% faster than Kerber's method at $k = 1$.
- At $k = 1.9$, the time is further reduced to 5.990 s, indicating a 40% improvement.

Table 3 shows the percentage increase in execution speed for each k parameter when compared to the auction algorithm method.

Table 3: Speed improvement compared to the Auction Algorithm (Kerber et al.[16]).

Parameter k	Mean Time (s)	Improvement (%)
1.1	6.669	33.06%
1.25	6.619	33.56%
1.3	6.150	38.29%
1.5	6.819	31.54%
1.6	6.590	33.85%
1.7	7.700	22.73%
1.8	7.081	28.94%
1.9	5.990	39.87%
2.0	6.940	30.35%

Thus, for large-scale computations, simplifying the persistence diagrams (using Theorm 3.1) is a more effective strategy than relying on auction-based Wasserstein optimization alone.

6 Conclusions

This paper examines the relationship between Wasserstein distances of Rips persistence diagrams from complete filtrations and those from early-stopped filtrations. Theorems 3.1 and 3.2 establish the relationships between these two distances. Theorem 3.1 suggests a whole range of thresholds corresponding to real numbers k between 1 and 2. The difference between the two Wasssetin distances depends on the instants at which all the connected components merge into one in each of the two filtrations, and the two Wasserstein distances coincide when $k = 1$. Theorem 3.2 involves the first and second largest finite values of one of the two persistence diagrams studied. Both theorems have been illustrated using time series from real data.

In addition to our theoretical findings, we assessed the computational efficiency of simplifying persistence diagrams with various values of k. Our experimental results indicate that:

- Increasing k leads to a significant reduction in the computational cost of Wasserstein distance calculations, achieving up to a *39.87% speed improvement* compared to the original persistence diagrams ($k = 1$).
- The most favorable trade-offs are found at $k = 1.3$ and $k = 1.9$, where execution times decrease by *38.29% and 39.87%*, respectively.
- Theorem 3.2 resulted in a savings of *43.22%* in execution time.

Moreover, we compared our method to the auction-based Wasserstein distance algorithm proposed by Kerber et al. [16], which enhances Wasserstein computation without altering the diagrams. Our results show that:

- The auction algorithm reduces computation time by 5.5% for $k = 1$, but it does not achieve the efficiency improvements seen with diagram simplification.
- For values of $k \geq 1.3$, the computational benefits from simplification alone exceed those provided by the auction algorithm at $k = 1$.
- The variability in execution time is generally lower for certain values of k, suggesting both stability and efficiency.

7 Future Work

This paper primarily examines Rips persistence diagrams, but a promising direction for future research is to explore whether similar computational improvements can be achieved with other types of persistence diagrams, like alpha complexes or witness complexes. Additionally, integrating diagram simplification with auction-based optimization could yield further computational benefits, especially when dealing with very large datasets. Another key area to investigate is identifying the optimal thresholds that maximize the benefits outlined in Theorem 3.1. Applying the findings from this work to other fields may also uncover broader implications and insights. Lastly, conducting a more in-depth statistical analysis to assess the trade-offs between computation time and accuracy loss across various filtration techniques could be valuable.

References

[1] B. Azamir, D. Bennis, and B. Michel, “A Simplified Algorithm for Identifying Abnormal Changes in Dynamic Networks,” *Physica A: Statistical Mechanics and Its Applications*, vol. 607, 2022, 128224.

[2] Z.-Z. Bai, G. H. Golub, and M. K. Ng, “Hermitian and Skew-Hermitian Splitting Methods for Non-Hermitian Positive Definite Linear Systems,” *SIAM Journal on Matrix Analysis and Applications*, vol. 24, 2006, pp. 603–626.

[3] D. P. Bertsekas, “Auction Algorithms for Network Flow Problems: A Tutorial Introduction,” *Computational Optimization and Applications*, vol. 1, no. 1, 1992, pp. 7–66.

[4] M. Carrière, F. Chazal, Y. Ike, T. Lacombe, M. Royer, and Y. Umeda, “A Computationally Efficient Framework for Vector Representation of Persistence Diagrams,” *Journal of Machine Learning Research*, vol. 23, no. 1129, 2022, pp. 1–43.

[5] L. Cheng-Te and L. Shou-De, “Egocentric Information Abstraction for Heterogeneous Social Networks,” *Proceedings of the International Conference on Advances in Social Network Analysis and Mining (ASONAM ’09)*, IEEE, 2009, pp. 255–260.

[6] Y.-M. Chung and A. Lawson, “Persistence Curves: A Canonical Framework for Summarizing Persistence Diagrams,” arXiv:1904.07768, 2019.

[7] H. Edelsbrunner and J. Harer, *Computational Topology: An Introduction*, American Mathematical Society, 2010.

[8] H. Edelsbrunner, D. Letscher, and A. Zomorodian, “Topological Persistence and Simplification,” *Discrete & Computational Geometry*, vol. 28, 2002, pp. 511–533.

[9] B. T. Fasy and B. Wang, “Exploring Persistent Local Homology in Topological Data Analysis,” *Proceedings of the 2016 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2016, pp. 6430–6434.

[10] R. Ghrist, “Barcodes: The Persistent Topology of Data,” *Bulletin of the American Mathematical Society*, vol. 45, 2008, pp. 61–75.

[11] M. Gidea, “Topological Data Analysis of Critical Transitions in Financial Networks,” *Proceedings of the Third International Winter School and Conference on Network Science*, Springer, 2017.

[12] M. Heesch, J. H. H. Nielsen, and B. B. Kragelund, "Ranking Beta Sheet Topologies of Proteins," *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering and Computer Science 2010*, WCECS 2010, 20-22 October, 2010, San Francisco, USA, pp. 624–628.

[13] J. Hein, “Discrete Mathematics,” Discrete Mathematics and Logic Series, Jones and Bartlett Publishers, 2003.

[14] A. K. Jain and R. C. Dubes, “Algorithms for Clustering Data,” Prentice Hall, 1988.

[15] T. D. Jorgensen, B. P. Haynes, and C. C. F. Norlund, "Evolving Artificial Neural Networks through Topological Complexification," *Engineering Letters*, vol. 17, no. 1, pp. 25–31, 2009.

[16] M. Kerber, D. Morozov, and A. Nigmetov, “Geometry Helps to Compare Persistence Diagrams,” *ACM Journal of Experimental Algorithmics*, vol. 22, 2017. doi:10.1145/3064175.

[17] T. Krishna Dey and Y. Wang, *Computational Topology for Data Analysis*, Cambridge University Press, 2022.

[18] H.-J. Li, L. Wang, Z. Bu, J. Cao, and Y. Shi, “Measuring the Network Vulnerability Based on Markov Criticality,” *ACM Transactions on Knowledge Discovery from Data*, vol. 16, 2021, pp. 1–24.

[19] H.-J. Li, Z. Wang, J. Cao, J. Pei, and Y. Shi, “Optimal Estimation of Low-Rank Factors via Feature Level Data Fusion of Multiplex Signal Systems,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 34, 2022, pp. 2860–2871.

[20] H.-J. Li, L. Wang, M. Perc, and Y. Zhang, “Optimization of Identifiability for Efficient Community Detection,” *New Journal of Physics*, vol. 22, 2020, 063035.

[21] Y. Liu, T. Safavi, A. Dighe, and D. Koutra, “Graph Summarization Methods and Applications: A Survey,” *ACM Computing Surveys*, vol. 51, 2018, article 62.

[22] H.-J. Li, W. Xu, S. Song, W.-X. Wang, and M. Perc, “The Dynamics of Epidemic Spreading on Signed Networks,” *Chaos, Solitons & Fractals*, vol. 151, 2021, 111294.

[23] J. B. MacQueen, “Some Methods for Classification and Analysis of Multivariate Observations,” *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1*, University of California Press, 1967, pp. 281–297.

[24] B. Michel and F. Chazal, “Fundamental and Practical Aspects for Data Scientists,” *Frontiers in Artificial Intelligence*, vol. 4, 2021, pp. 1–28.

[25] Z. M. Zin, M. Khalid, E. Mesbahi, and R. Yusof, “Data Clustering and Topology Preservation Using 3D Visualization of Self Organizing Maps,” *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering 2012*, vol. II, pp. 696–701, London, U.K., 4–6 July 2012.

[26] O. P. Oladimeji, O. A. Ajayi, and M. A. Adewuyi, "A Suitable Scheme for Influential Load Nodes Detection in Power Systems Based on Network Structural Characteristics Theory," *Engineering Letters*, vol. 32, no. 3, pp. 630–649, 2024.

[27] C. Qiu, H.-J. Li, J. Pei, and W. Xu, “Fast Markov Clustering Algorithm Based on Belief Dynamics,” *IEEE Transactions on Cybernetics*, 2022, pp. 1–10.

[28] B. Summa, B. Fasy, C. Wenk, and Y. Qin, “Comparing Distance Metrics on Vectorized Persistence Summaries,” *NeurIPS 2020 Workshop on Topological Data Analysis and Beyond*, 2020.

[29] Y. Wang, Y. Liu, and S. Xing, "Structural Topology Optimization Using Genetic Algorithms," *Lecture Notes in Engineering and Computer Science: Proceedings of The World Congress on Engineering 2013*, WCE 2013, 3-5 July, 2013, London, U.K., pp. 1933–1937.

[30] A. Zomorodian and G. Carlsson, “Computing Persistent Homology,” *Discrete & Computational Geometry*, vol. 33, 2005, pp. 249–274.

[31] Y. Zhang, Y. Zhang, and X. Du, "The Research on WSNs Scale-free Topology for Prolonging Network Lifetime," *Engineering Letters*, vol. 29, no. 1, pp. 238-243, 2021.

[32] F. Zhou, H. Toivonen, and S. Malher, “Network Simplification with Minimal Loss of Connectivity,” *Proceedings of the 10th IEEE International Conference on Data Mining (ICDM 2010)*, IEEE, 2010, pp. 659–668.