Computational Fuzzy Topology for Data Analysis

Javier Perera-Lago[1],
Rocio Gonzalez-Diaz[1], and
Alvaro Torras-Casas[1]

Depto. de Matemática Aplicada I, Universidad de Sevilla, España
{jperera,rogodi,atorras}@us.es

Abstract. En este proyecto de tesis doctoral se propone una extensión de la homología persistente, una herramienta central del Análisis Topológico de Datos, al contexto de los conjuntos L-difusos. El objetivo es desarrollar un marco teórico y computacional que permita extraer información topológica de conjuntos de datos cuyos elementos incorporan información adicional, como etiquetas o grados de relevancia. Para ello se estudiarán la homología simplicial L-difusa, su extensión al ámbito persistente y posibles resultados de estabilidad frente a pequeñas perturbaciones en los datos. Así, se pretende establecer un vínculo entre la topología computacional y la teoría de conjuntos L-difusos, ampliando el alcance de las técnicas actuales de análisis topológico de datos.

Keywords: Análisis Topológico de Datos · Homología persistente · Conjuntos L-difusos.

1 Datos del proponente

El proponente de esta tesis es Javier Perera Lago, quien está matriculado a tiempo completo en el Programa de Doctorado de la Universidad de Sevilla desde octubre de 2023, bajo la dirección de la catedrática Rocío González Díaz y del doctor Álvaro Torras Casas. Los tres son miembros del grupo de investigación *Topology and Applications in Data & Science (TAPAS)*, liderado por Rocío. El proponente disfruta desde junio de 2025 de una beca predoctoral financiada por la Junta de Andalucía/CIIU y por la FSE+, con referencia DGP_PRED_2024_02465, y espera defender su tesis en octubre de 2028.

2 Resumen de la propuesta

El Análisis Topológico de Datos (Topological Data Analysis, TDA) es un área de investigación que emplea herramientas de la topología computacional para el estudio de conjuntos de datos. Su objetivo principal es extraer información de los datos mediante invariantes topológicos que puedan ser interpretados, analizados estadísticamente o utilizados por algoritmos de aprendizaje automático.

Dentro del TDA pueden distinguirse dos grandes líneas de investigación. La primera se centra en la construcción de descriptores topológicos capaces de

capturar propiedades relevantes de objetos complejos o de conjuntos de datos, habitualmente representados como nubes de puntos en espacios euclídeos. Entre las herramientas más destacadas de esta línea se encuentra la homología persistente [1], una extensión de la homología simplicial que ha dado lugar a numerosas generalizaciones, como la homología zigzag [2] o la homología multipersistente [3]. La segunda línea busca mejorar la visualización e interpretación de los datos mediante técnicas de reducción de dimensionalidad y agrupamiento que preserven información topológica relevante. Entre las herramientas más utilizadas en este contexto destacan UMAP [4] e IsUMap [5], que modelan los datos mediante conjuntos simpliciales difusos antes de proyectarlos a espacios de menor dimensión.

En numerosas aplicaciones, los elementos de un conjunto de datos contienen información adicional más allá de sus coordenadas o de las distancias que los relacionan. Por ejemplo, cada punto puede estar asociado a una categoría, una etiqueta, un grado de relevancia o cualquier otra forma de información jerárquica. Este tipo de estructuras puede modelarse de manera natural mediante conjuntos L-difusos. Surge entonces la siguiente pregunta: ¿es posible desarrollar una teoría de homología persistente para conjuntos L-difusos? Y, en caso afirmativo, ¿qué información adicional podría proporcionar para el análisis de datos?

La literatura existente ofrece únicamente respuestas parciales a esta cuestión. Los conjuntos simpliciales difusos se han utilizado principalmente en problemas de reducción de dimensionalidad [6,7]. Otros trabajos estudian la homología de espacios topológicos difusos [8]. Sin embargo, dicho enfoque difiere del que se propone en esta tesis, ya que se centra en espacios equipados con abiertos difusos y no en conjuntos de datos L-difusos. Otras propuestas estudian complejos simpliciales ponderados [9,10], aunque en ellas los invariantes topológicos obtenidos siguen siendo objetos clásicos y no incorporan una estructura difusa en el resultado final.

La propuesta de esta tesis consiste en desarrollar una teoría de homología persistente L-difusa. Más concretamente, dado un conjunto de datos L-difuso $\mu : X \to L$, donde L es un retículo completamente distributivo, se pretende construir un módulo de homología persistente L-difuso en el que la información topológica también tome valores en L. El objetivo es establecer un nuevo puente entre la topología computacional y la teoría de conjuntos L-difusos, proporcionando un marco unificado para el estudio de distintos tipos de datos enriquecidos que actualmente se analizan mediante metodologías independientes.

La estructura de esta memoria es la siguiente. En la Sección 3 se presentan los conceptos teóricos necesarios para contextualizar la propuesta. En la Sección 4 se describen los objetivos científicos de la tesis. La Sección 5 expone la metodología seguida para alcanzar los objetivos ya cumplidos y la estrategia prevista para los restantes. Finalmente, en la Sección 6 se discute el impacto potencial de la investigación desde los puntos de vista científico, económico y social.

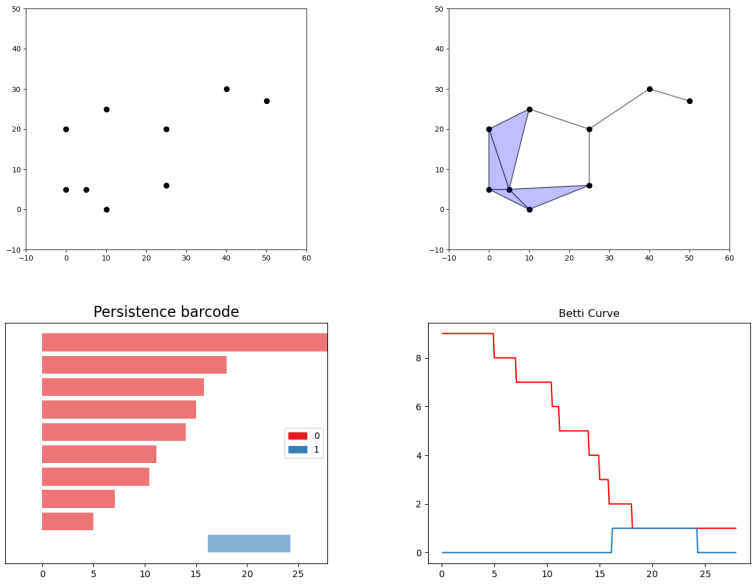


Fig. 1. Arriba a la izquierda, un conjunto de puntos $X \subset \mathbb{R}^2$. Arriba a la derecha, un complejo de Čech para un cierto parámetro r. Abajo, códigos de barras y curvas de Betti de la filtración completa para $d = 0, 1$.

3 Antecedentes

El contenido de la propuesta se basa en conceptos propios de la topología computacional y la teoría de conjuntos L-difusos. En esta sección, describiremos los conceptos mínimos necesarios para su comprensión.

3.1 Topología computacional

La topología computacional es una rama de las matemáticas que trata de aplicar la informática para abordar problemas topológicos. Recientemente, algunas herramientas de topología computacional como la homología persistente están siendo aplicadas para calcular propiedades topológicas de objetos complejos y conjuntos de datos e integrar esos resultados en modelos de aprendizaje automático, dando lugar al Análisis Topológico de Datos (TDA por sus siglas en inglés).

Supongamos que tenemos un conjunto de datos $X = \{x_1, \cdots, x_k\} \subset \mathbb{R}^n$ y un parámetro de proximidad $r \geq 0$. Podemos aproximar la forma de X mediante el complejo de Čech $\check{C}_r(X)$ [11], que es un complejo simplicial cuyo conjunto de vértices es X y que contiene al símplice $\sigma \subseteq X$ si y solo si existe una bola de radio r que contenga a todos los vértices de σ. A partir de una sucesión de parámetros $r_1 < r_2 < \cdots < r_m$ podemos crear una filtración de complejos simpliciales

$$\check{C}_{r_1}(X) \hookrightarrow \check{C}_{r_2}(X) \hookrightarrow \cdots \hookrightarrow \check{C}_{r_m}(X),$$

que será el objeto de estudio principal. Otras formas de construir filtraciones son el complejo de Vietoris-Rips [12] o el complejo Alpha [13].

Dado un complejo simplicial Δ y un dominio de ideales principales $\mathbb{D}$, el módulo de d-cadenas $C_d(\Delta)$ es el $\mathbb{D}$-módulo generado por los d-símplices de Δ. El operador borde es un homomorfismo que asigna a cada generador $\sigma = [v_0, \ldots, v_d] \in C_d(\Delta)$ la $d-1$-cadena $\partial_d(\sigma) = \sum_{j=0}^{d} (-1)^j [v_0, \ldots, \hat{v}_j, \ldots, v_d]$, donde $\hat{v}_j$ significa que v_j se omite. Definimos el módulo de d-ciclos $Z_d(\Delta) \subseteq C_d(\Delta)$ como el núcleo de ∂_d, y el módulo de d-bordes $B_d(\Delta) \subseteq C_d(\Delta)$ como la imagen de ∂_{d+1}. Es fácil demostrar que $\partial_d \circ \partial_{d+1} = 0$ y, por tanto, $B_d(\Delta) \subseteq Z_d(\Delta)$. Entonces el módulo de d-homología de Δ se puede definir como el cociente $H_d(\Delta) = Z_d(\Delta)/B_d(\Delta)$ [14]. Cada elemento de $H_d(\Delta)$ se llama una clase de homología d-dimensional. Informalmente, las clases de dimensión 0 representan componentes conexas, las de dimensión 1 son agujeros, las de dimensión 2 son burbujas vacías y así sucesivamente.

Supongamos ahora que tenemos ahora una filtración:

$$\Delta_1 \hookrightarrow \Delta_2 \hookrightarrow \cdots \hookrightarrow \Delta_m$$

y que la homología de cada complejo se calcula con coeficientes en un cuerpo $\mathbb{F}$ (es decir, cada $H_d(\Delta_i)$ es un espacio vectorial). Cada inclusión $\Delta_i \hookrightarrow \Delta_j$ induce un homomorfismo $H_d(\Delta_i) \to H_d(\Delta_j)$. La homología persistente d-dimensional de la filtración sería la siguiente secuencia [15] (llamada módulo de persistencia):

$$H_d(\Delta_1) \to H_d(\Delta_2) \to \cdots \to H_d(\Delta_m).$$

Este módulo de persistencia nos indica cómo evoluciona la homología d-dimensional a lo largo de la filtración. Si una clase de homología $\alpha \neq 0$ pertenece a $H_d(\Delta_i)$ pero no pertenecía a $H_d(\Delta_{i-1})$, se dice que α *nace* en la etapa i de la filtración. Si $\alpha \neq 0$ pertenece a $H_d(\Delta_{j-1})$ pero su imagen en $H_d(\Delta_j)$ es nula, se dice que α *muere* en la etapa j de la filtración. Una clase de homología α que nace en la etapa i y muere en la etapa j tiene un intervalo de persistencia $[i, j)$. Una clase que nace en la etapa i pero no muere nunca tiene un intervalo de persistencia $[i, \infty)$. La unión de todos estos intervalos de persistencia es el código de barras de la filtración, que caracteriza completamente la homología persistente d-dimensional de Δ, y da una visión clara e intuitiva de su estructura algebraica.

Un código de barras es un multiconjunto de intervalos. Para facilitar su uso en estadística o para aprendizaje automático hace falta vectorizar el código de barras, es decir, construir un vector a partir del mismo que resuma su información. La vectorización más sencilla de todas sería la curva de Betti de dimensión d, que no es sino el vector $(\mathrm{Rank}(H_d(\Delta_1)), \mathrm{Rank}(H_d(\Delta_2)), \ldots, \mathrm{Rank}(H_d(\Delta_m))$ conteniendo el rango de cada espacio de homología en la filtración [16]. Otras vectorizaciones útiles son los landscapes [17] o las imágenes de persistencia [18].

En la Figura 1 se puede ver un ejemplo de un conjunto de datos, un complejo simplicial, un barcode y una curva de Betti.

3.2 Conjuntos L-difusos

La matemática difusa es aquella que se construye a partir del concepto de conjunto difuso, definido por primera vez en 1965 por Zadeh [19]. Dado un conjunto U al que llamamos *universo*, un subconjunto difuso se define como una aplicación $\mu : U \to [0,1]$ que muestra el grado de pertenencia de los elementos de U con respecto a μ. Así, un elemento $x \in U$ con $\mu(x) = 0$ no pertenece a μ, otro con $\mu(x) = 1$ pertenece totalmente a μ y otro con $0 < \mu(x) < 1$ pertenece parcialmente a μ. Los conjuntos difusos son una generalización de los conjuntos clásicos que sirven para modelar incertidumbre e imprecisión en ámbitos donde los límites de pertenencia no están claramente definidos, como en el razonamiento humano, en el lenguaje natural o las imágenes médicas. Se aplican en áreas como la inteligencia artificial, la toma de decisiones y el control difuso [20].

Las nociones de complemento, unión, e intersección se pueden definir para un conjunto difuso usando las funciones de pertenencia. El complemento de μ sería la función $\mu^c(u) = 1 - \mu(u)$. La unión de los subconjuntos difusos μ y ν se define como $(\mu \cup \nu)(u) = \max\{\mu(u), \mu(u)\}$, mientras que la intersección se define como $(\mu \cap \nu)(u) = \min\{\mu(u), \mu(u)\}$. También se pueden definir otros conceptos típicos de la teoría de conjuntos, como productos cartesianos, relaciones o aplicaciones, sentando las bases para entrar en contacto con otras ramas matemáticas.

Una extensión de los conjuntos difusos son los conjuntos L-difusos [21], en los que el codominio de la función de pertenencia no es $[0,1]$ sino un retículo completamente distributivo L. Aquí la relación de orden no es necesariamente total, pero existe un mínimo global $0 \in L$, un máximo global $1 \in L$, y cada par de elementos $\ell_1, \ell_2 \in L$ siempre tiene un ínfimo $\ell_1 \wedge \ell_2 \leq \ell_1, \ell_2$ y un supremo $\ell_1, \ell_2 \leq \ell_1 \vee \ell_2$. Cuando $L = \{0,1\}$, obtenemos los subconjuntos estándar de U.

También existen versiones L-difusas de estructuras más complejas que los conjuntos, como por ejemplo los complejos simpliciales o los $\mathbb{D}$-módulos.

Dado un complejo simplicial Δ, un subcomplejo L-difuso es una función $\mu : \Delta \to L$ tal que si τ es una cara de σ, entonces $\mu(\tau) \geq \mu(\sigma)$. Esta definición está inspirada en la de conjunto simplicial difuso de [6], y nos garantiza que para cada $\ell \in L$ el subconjunto $\mu^{\geq \ell} = \{\sigma \in \Delta \mid \mu(\sigma) \geq \ell\}$ (llamado corte de ℓ) siempre es un subcomplejo de Δ.

Dado un $\mathbb{D}$-módulo M, un submódulo L-difuso es una aplicación $\mu : M \to L$ tal que $\mu(0) = 1$, $\mu(a \cdot m) \geq \mu(m)$ para todo $a \in \mathbb{D}$ y todo $m \in M$, y tal que $\mu(m_1 + m_2) \geq \mu(m_1) \wedge \mu(m_2)$ para todo $m_1, m_2 \in M$. Esta definición nos garantiza que para cada $\ell \in L$ el subconjunto $\mu^{\geq \ell} = \{m \in M \mid \mu(m) \geq \ell\}$ (llamado corte de ℓ) siempre es un submódulo de M. A partir de esta sencilla definición se pueden definir otros conceptos como el submódulo L-difuso generado por una función $\mu : M \to L$ cualquiera o el cociente entre submódulos L-difusos. Más información sobre submódulos L-difusos se puede encontrar en [22].

4 Objetivos

La homología simplicial presentada en la Sección 3 se calcula sobre un conjunto de datos en el que todos los elementos tienen el mismo estatus y no hay una

jerarquía predefinida entre ellos. Sin embargo, es común encontrar conjuntos de datos donde, además de las coordenadas, cada elemento tiene una etiqueta que proporciona información adicional sobre el mismo. Por ejemplo, cuando los datos tienen asignado un peso o se agrupan en categorías o clases disjuntas. Este tipo de conjuntos de datos etiquetados se pueden modelizar como conjuntos L-difusos. Por ello, el objetivo principal de este proyecto de tesis es estudiar cómo se pueden generalizar los resultados ya conocidos sobre homología simplicial y homología persistente para conjuntos de datos L-difusos. Este proyecto se puede dividir en los tres siguientes objetivos científicos:

4.1 Objetivo 1: Homología simplicial L-difusa

El primer objetivo de esta tesis es definir la homología L-difusa asociada a un subcomplejo simplicial L-difuso. En particular, queremos encontrar un método que a cada subcomplejo L-difuso $\mu : \Delta \to L$ le asocie un submódulo L-difuso $\eta_d : H_d(\Delta) \to L$. Esta construcción debe ser coherente con la definición estándar de la homología simplicial, de forma que coincida con ella cuando trabajemos con el retículo $L = \{0, 1\}$. Queremos estudiar las propiedades más importantes de esta construcción, y encontrar un método computacional para calcular el valor $\eta_d([h]) \in L$ asociado a cada clase de homología $[h] \in H_d(\Delta)$ así como la estructura de los cortes $\eta_d^{\geq \ell}$.

4.2 Objetivo 2: Homología persistente L-difusa

Supongamos ahora que tenemos una filtración de subcomplejos L-difusos:

$$\begin{array}{ccccccc} \Delta_1 & \hookrightarrow & \Delta_2 & \hookrightarrow & \cdots & \hookrightarrow & \Delta_n \\ \downarrow{\scriptstyle \mu_1} & & \downarrow{\scriptstyle \mu_2} & & & & \downarrow{\scriptstyle \mu_n} \\ L & & L & & \cdots & & L \end{array}$$

El segundo objetivo de esta tesis es definir la homología persistente L-difusa asociada a esta filtración, de forma que sea coherente con la homología persistente estándar y con la homología L-difusa de cada subcomplejo L-difuso $\mu_i : \Delta_i \to L$ en la filtración. Queremos definir invariantes útiles para esta construcción y encontrar un método computacional para extraer dichos invariantes.

4.3 Objetivo 3: Estabilidad de la homología persistente L-difusa

Supongamos que tenemos dos datasets X_1 y X_2. Consideremos las filtraciones de Čech construidas sobre ellos, denotadas por $\check{C}(X_1)$ y $\check{C}(X_2)$, definidas para todos los parámetros de escala en $\mathbb{R}_{\geq 0}$, y sean $\mathbf{B}_1$ y $\mathbf{B}_2$ los códigos de barras asociados a dichas filtraciones. La similitud entre los datasets X_1 y X_2 puede medirse con la distancia de Gromov-Hausdorff $d_{GH}(X_1, X_2)$ [23], mientras que la similitud entre los códigos de barras $\mathbf{B}_1$ y $\mathbf{B}_2$ puede medirse con la distancia bottleneck $d_B(\mathbf{B}_1, \mathbf{B}_2)$ [24].

Un resultado fundamental en la teoría de homología persistente es la estabilidad con respecto a perturbaciones pequeñas [25], concretada en la desigualdad:

$$d_B(\mathbf{B}_1, \mathbf{B}_2) \leq 2\, d_{GH}(X_1, X_2).$$

Esta cota garantiza que datasets similares producen códigos de barras similares. En consecuencia, la homología persistente varía de manera controlada ante pequeñas perturbaciones de los datos. Resultados análogos pueden demostrarse para otras filtraciones como las de Vietoris-Rips y Alpha [25].

El tercer objetivo de esta tesis consiste en obtener resultados de estabilidad análogos para la homología persistente L-difusa, proporcionando garantías de robustez similares a las que existen en la teoría clásica de homología persistente.

5 Metodología y Plan de trabajo

En esta sección se describe la estrategia que el proponente prevé seguir para alcanzar los objetivos previamente establecidos, así como, en el caso de aquellos ya completados, se ofrece un breve resumen de los resultados obtenidos.

5.1 Métodos para el Objetivo 1

Todos los hitos asociados al Objetivo 1 han sido completados. Para ello, hemos estudiado la bibliografía especializada en teoría de conjuntos L-difusos, revisando conceptos como los retículos completamente distributivos, los submódulos y subespacios vectoriales L-difusos, los subcomplejos L-difusos, el principio de extensión, así como el cociente entre submódulos L-difusos. Algunas de estas herramientas no existían como tal en la literatura previa, sino que han sido adaptadas para nuestro propio objetivo. Este es el caso de la noción de subcomplejo L-difuso, inspirada en trabajos previos sobre conjuntos simpliciales difusos.

A partir de estas nociones, se ha definido el submódulo de homología L-difuso $\eta_d : H_d(\Delta) \to L$ mediante un procedimiento análogo al empleado en la construcción clásica de la homología simplicial descrita en la Sección 3.1, sustituyendo cada uno de los objetos algebraicos involucrados por sus correspondientes versiones L-difusas. El desarrollo completo de esta teoría, junto con sus propiedades fundamentales y una caracterización que permite su cálculo mediante operaciones matriciales, ha sido recogido en [26]. Dicho trabajo ha sido enviado a la revista *Fuzzy Sets & Systems* y se encuentra actualmente en proceso de revisión.

5.2 Métodos para el Objetivo 2

El primer paso para abordar el Objetivo 2 consiste en formular rigurosamente el concepto de homología persistente L-difusa. En la teoría clásica, la homología persistente se modela mediante un módulo de persistencia, es decir, una sucesión de espacios vectoriales conectados por aplicaciones lineales

$$H_d(\Delta_1) \to H_d(\Delta_2) \to \cdots \to H_d(\Delta_n).$$

En nuestro contexto, a cada espacio de homología $H_d(\Delta_i)$ le asociamos un subespacio L-difuso $\eta_{d,i}$, obteniendo un diagrama de la forma

$$\begin{array}{ccccccc} H_d(\Delta_1) & \longrightarrow & H_d(\Delta_2) & \longrightarrow & \cdots & \longrightarrow & H_d(\Delta_n) \\ \downarrow{\scriptstyle \eta_{d,1}} & & \downarrow{\scriptstyle \eta_{d,2}} & & & & \downarrow{\scriptstyle \eta_{d,n}} \\ L & & L & & \cdots & & L. \end{array}$$

a la que podríamos llamar módulo de persistencia L-difuso.

Una vez definido este objeto, será necesario estudiar la noción de morfismo entre subespacios L-difusos y analizar la estructura algebraica de los módulos de persistencia L-difusos. Un posible enfoque para ello surge de un resultado clásico de [27], donde se demuestra una equivalencia de categorías entre los módulos de persistencia con coeficientes en un cuerpo $\mathbb{F}$ y los módulos graduados sobre el anillo de polinomios $\mathbb{F}[x]$.

Esta equivalencia sugiere que los módulos de persistencia L-difusos podrían estudiarse como un único submódulo L-difuso de un módulo graduado adecuado. De ser así, muchas de las herramientas desarrolladas en el Objetivo 1 podrían trasladarse directamente al contexto persistente. En particular, seríamos capaces de calcular el grado de pertenencia asociado a cada clase de homología, el cual puede variar a lo largo de la filtración. Además, para cada $\ell \in L$, podríamos considerar el corte correspondiente. Cada uno de estos cortes sería un módulo de persistencia clásico y, por tanto, tendría asociado su propio código de barras. Esto permitiría describir la homología persistente L-difusa mediante una familia de códigos de barras ordinarios indexados por los elementos de L.

Con respecto a las vectorizaciones, hasta el momento hemos desarrollado un método para calcular curvas de Betti L-difusas. Para ello, hemos estudiado en qué condiciones tiene sentido definir el rango de un submódulo L-difuso $\mathrm{Rank}(\eta_{d,i})$ y hemos demostrado que, en dichos casos, este puede calcularse aplicando homología zigzag. El resultado es una familia de curvas (una por cada $\ell \in L$) cuya suma recupera la curva de Betti clásica. Esta línea de investigación se encuentra actualmente en desarrollo y los resultados obtenidos serán previsiblemente enviados al congreso SoCG 2027.

5.3 Métodos para el Objetivo 3

La distancia de Gromov-Hausdorff, introducida en la Sección 4.3, permite comparar la similitud entre espacios métricos. Existe una caracterización equivalente de esta distancia en términos de correspondencias entre los puntos de ambos espacios, en la cual se mide la distorsión inducida al comparar distancias y se selecciona aquella correspondencia que minimiza dicha distorsión [28].

En la literatura se han propuesto diversas adaptaciones de esta noción para contextos en los que los puntos del espacio están equipados con información adicional como una jerarquía. En estos casos, el conjunto de correspondencias admisibles se ve restringido, lo que da lugar a variantes de la distancia de Gromov-Hausdorff adaptadas a dicha estructura. Un ejemplo de este tipo de construcción aparece en las demostraciones de estabilidad del *six-pack* en TDA cromático [29].

Por nuestra parte, creemos que es posible definir una adaptación adecuada de la distancia de Gromov-Hausdorff para conjuntos de datos L-difusos, que sirva para demostrar resultados de estabilidad. Para ello será necesario revisar la bibliografía existente en este tipo de métricas, para determinar si las técnicas ya disponibles pueden aplicarse directamente o si es necesario desarrollar nuevas herramientas inspiradas en ellas.

6 Relevancia

En esta sección describimos los beneficios que el proponente espera aportar a la sociedad mediante el desarrollo de la presente tesis doctoral.

6.1 Impacto científico

Al combinar herramientas de topología computacional y teoría de conjuntos L-difusos, esta investigación permitiría extender las técnicas habituales del análisis topológico de datos a tipos de datos que incorporan información adicional en sus puntos, como conjuntos cromáticos (en los que cada punto posee una etiqueta categórica) o conjuntos ponderados (en los que cada punto lleva asociado un grado de relevancia).

La literatura actual ha abordado estas situaciones mediante enfoques distintos. Por una parte, los conjuntos cromáticos han dado lugar al área conocida como Chromatic TDA [30]; por otra, los conjuntos ponderados pueden estudiarse mediante técnicas como la homología multipersistente [3]. La propuesta desarrollada en esta tesis permitiría tratar ambos escenarios dentro de un mismo marco matemático, facilitando la comparación sistemática entre los distintos enfoques y ampliando el conjunto de herramientas disponibles para la comunidad científica.

6.2 Impacto económico

A medio y largo plazo, el desarrollo de nuevas herramientas para el análisis topológico de datos podría facilitar el diseño de métodos más eficientes para la exploración y el procesamiento de grandes volúmenes de datos. Asimismo, las herramientas computacionales desarrollados durante la tesis podrían servir de base para futuras colaboraciones con grupos de investigación y empresas tecnológicas interesadas en inteligencia artificial y ciencia de datos.

6.3 Impacto social

Aunque esta investigación tiene un carácter fundamentalmente teórico, sus posibles aplicaciones se enmarcan en áreas de gran relevancia social. Por ejemplo, algunos datasets genéticos recopilan el grado de expresión de ciertos genes, lo cual podría ser modelado en términos de conjuntos L-difusos si se encuentra el retículo L adecuado. Una mejora en el análisis de estos datos podría contribuir indirectamente al avance de campos como la medicina o la biología.

References

1. H. Edelsbrunner, D. Letscher, and A. Zomorodian. Topological persistence and simplification. *Discrete & Computational Geometry*, 28(4):511–533, 2002.
2. G. Carlsson, V. de Silva, and D. Morozov. Zigzag persistent homology and real-valued functions. In *Proceedings of the Twenty-Fifth Annual Symposium on Computational Geometry*, pages 247–256, 2009.
3. G. Carlsson and A. Zomorodian. The theory of multidimensional persistence. In *Proceedings of the Twenty-Third Annual Symposium on Computational Geometry*, pages 184–193, 2007.
4. L. McInnes, J. Healy, and J. Melville. UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv preprint arXiv:1802.03426*, 2018.
5. P. Joharinad, H. Fahimi, L. S. Barth, J. Keck, and J. Jost. IsUMap: Manifold learning and data visualization leveraging Vietoris–Rips filtrations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 39(17):17699–17706, 2025.
6. D. I. Spivak. Metric realization of fuzzy simplicial sets. *Preprint*, 4, 2009.
7. L. S. Barth, H. Fahimi, P. Joharinad, J. Jost, J. Keck, and T. J. Mikhail. Fuzzy simplicial sets and their application to geometric data analysis. *Applied Categorical Structures*, 33(5):31, 2025.
8. W.-j. Liu and C.-y. Zheng. Singular homology groups of fuzzy topological spaces. *Fuzzy Sets and Systems*, 15(2):199–207, 1985.
9. R. J. MacG. Dawson: Homology of weighted simplicial complexes. Cahiers de Topologie et Géométrie Différentielle Catégoriques **31**(3), 229–243 (1990)
10. S. Ren, C. Wu, and J. Wu: Weighted persistent homology. The Rocky Mountain Journal of Mathematics **48**(8), 2661–2687 (2018)
11. G. Carlsson: Topology and data. Bulletin of the American Mathematical Society **46**(2), 255–308 (2009)
12. A. Zomorodian: Fast construction of the Vietoris–Rips complex. Computers & Graphics **34**(3), 263–271 (2010)
13. J. Giesen, F. Cazals, M. Pauly, and A. Zomorodian: The conformal alpha shape filtration. The Visual Computer **22**(8), 531–540 (2006)
14. A. Hatcher: Algebraic Topology. Cambridge University Press, Cambridge (2002). Available at: `https://pi.math.cornell.edu/~hatcher/AT/ATpage.html`
15. A. Zomorodian and G. Carlsson. Computing persistent homology. In *Proceedings of the Twentieth Annual Symposium on Computational Geometry*, pages 347–356, 2004.
16. C. Giusti, R. Ghrist, and D. S. Bassett: Two's company, three (or more) is a simplex. Journal of Computational Neuroscience **41**(1), 1–14 (2016)
17. P. Bubenik: Statistical topological data analysis using persistence landscapes. Journal of Machine Learning Research **16**, 77–102 (2015)
18. H. Adams, T. Emerson, M. Kirby, R. Neville, C. Peterson, P. Shipman, S. Chepushtanova, E. Hanson, F. Motta, and L. Ziegelmeier: Persistence images: A stable vector representation of persistent homology. Journal of Machine Learning Research **18**(8), 1–35 (2017)
19. L. A. Zadeh: Fuzzy sets. Information and Control **8**(3), 338–353 (1965)
20. M. Eftekhari, A. Mehrpooya, F. Saberi-Movahed, and V. Torra: How fuzzy concepts contribute to machine learning. Springer (2022)
21. J. A. Goguen: L-fuzzy sets (1966)
22. M. Mashinchi and M. M. Zahedi: On L-fuzzy primary submodules. Fuzzy Sets and Systems **49**(2), 231–236 (1992)

23. P. Petersen: Gromov–Hausdorff convergence of metric spaces. In: Proceedings of Symposia in Pure Mathematics, pp. 489–504 (1993)
24. D. Cohen-Steiner, H. Edelsbrunner, and J. Harer: Stability of persistence diagrams. Discrete & Computational Geometry **37**(1), 103–120 (2007)
25. F. Chazal, V. de Silva, and S. Oudot: Persistence stability for geometric complexes. *Geometriae Dedicata* **173**(1), 193–214 (2014). Springer.
26. J. Perera-Lago, A. Torras-Casas, and R. Gonzalez-Diaz: *L*-fuzzy simplicial homology. arXiv preprint arXiv:2604.08166 (2026).
27. D. Eisenbud: *Commutative Algebra with a View Toward Algebraic Geometry*. Springer, New York (1995), Chap. 5.
28. N. J. Kalton and M. I. Ostrovskii: Distances between Banach spaces. Forum Mathematicum **11**(1), 17–48 (1999)
29. O. Draganov, S. Rosenmeier, and N. Zava: Gromov–Hausdorff distance between chromatic metric pairs and stability of the six-pack. arXiv preprint arXiv:2507.17994 (2025).
30. S. Cultrera di Montesano, O. Draganov, H. Edelsbrunner, and M. Saghafian: Chromatic topological data analysis. arXiv preprint arXiv:2406.04102 (2024).