



UNIVERSITAT DE
BARCELONA

Facultat de Matemàtiques
i Informàtica

GRAU DE MATEMÀTIQUES

Treball final de grau

Homología Persistente de las Redes Complejas

Autor: Guillermo Ródenas Alesina

Director: Dr. Javier J. Gutiérrez

Realitzat a: Departament de Matemàtiques

Barcelona, 21 de junio de 2020

Abstract

La homología persistente constituye a día de hoy una de las herramientas esenciales del TDA (*Topological Data Analysis*). Desarrollada principalmente en la última década, nace de la topología algebraica y hace uso de uno de los más importantes invariantes algebraicos de un espacio, la sucesión de grupos de homología, para determinar las características topológicas del espacio subyacente de una cierta nube de puntos. Se construye un complejo simplicial a partir de los datos iniciales y se define una filtración (sucesión de subcomplejos encajados) que permite estudiar cómo varían los rangos de grupos de homología (números de Betti) según evoluciona el parámetro que controla la dicha filtración. Esto proporciona un estudio transversal de las propiedades topológicas a todas las escalas, permitiendo desvelar cuáles de ellas son verdaderamente características del conjunto de estudio y cuáles son menos significativas o simplemente producto de ruido estadístico.

La semejanza entre un grafo y un complejo simplicial es el motor de la incursión de la homología persistente en la teoría de redes. Aún sin tener dimensionalidad, es fácil pensar que un grafo completo de tres vértices es un triángulo, y que uno completo con cuatro vértices constituye algo topológicamente equivalente a un tetraedro. De esta manera se convierte cualquier grafo en complejo simplicial, permitiendo la aplicación de las herramientas de la homología algebraica. Pero, ¿qué se gana estudiando un grafo de esta manera?

El estudio clásico de los grafos y redes complejas (que es una denominación más común en la ciencia aplicada, por denotar que caracterizan sistemas complejos) se ha centrado en medidas locales (grado de los nodos, longitud característica de los caminos, clustering,...) que suelen ser ciegas a las propiedades geométricas y topológicas del espacio que genera la propia existencia de los nodos y aristas. Aplicando la homología persistente a los grafos se obtiene una nueva y complementaria perspectiva que pretende extraer justamente estas características globales y libres de las restricciones que supone la elección de una escala de estudio. Esto supone un cambio radical de punto de vista que, como tal, descubre un mar de posibilidades de investigación y da rienda suelta a la creatividad y el optimismo.

El objetivo de este trabajo es presentar toda la información publicada hasta la fecha sobre la fusión de estos dos elementos, la homología persistente y los grafos, por tal de dar a conocer ideas y procedimientos que parecen tener –y han empezado a demostrar– mucho potencial en ámbitos como la neurociencia. La estructura del trabajo consta de una introducción breve a la topología algebraica, centrada únicamente en la homología simplicial y que desarrolla los conceptos hasta formalizar la homología persistente y las herramientas que la acompañan; y de una introducción a la teoría de redes complejas junto con sus modelos más típicos y medidas características. Por último, y como núcleo del estudio, se expone cómo interaccionan estas dos áreas. Se da métodos para la construcción de complejos simpliciales y filtraciones a partir de redes y se expone las conclusiones extraídas de su uso en toda una variedad de artículos matemáticos y aplicados, así como nuevas construcciones que surgen de la necesidad de extender el análisis a nuevos casos.

Se anima en todo momento al lector a dejarse ser creativo en el uso de la junción de estas dos materias: por la novedad de estas técnicas, es tan sólo cuestión de tiempo que la bibliografía que la utiliza vaya engrosándose.

Índice

1. Topología Algebraica	1
1.1. Homología simplicial	1
1.2. Homología persistente	4
1.2.1. Módulos de persistencia y funtorialidad	6
1.2.2. Distancias	9
1.2.3. Persistence landscapes	12
1.2.4. Alternativas al cálculo de distancias sobre barcodes	13
2. Redes Complejas	16
2.1. Conceptos básicos de la teoría de grafos	16
2.2. Propiedades de una red	17
2.3. Modelos de redes	19
3. Homología persistente de redes complejas	22
3.1. Construcción de complejos simpliciales y filtraciones en una red	22
3.1.1. Complejos simpliciales	23
3.1.2. Filtraciones	26
3.2. Resultados y análisis	31
3.2.1. Estudio de redes conocidas	31
3.2.2. Capacidad discriminativa de la homología persistente	33
3.2.3. Aplicaciones en la neurociencia aplicada	35
3.3. Otras construcciones sobre la persistencia homológica	37
3.3.1. Evolución temporal de redes	37
3.3.2. Persistence vineyard	39
3.3.3. Homological scaffold	40
4. Conclusiones	42

1. Topología Algebraica

La topología algebraica es la rama de las matemáticas que estudia la estructura de los espacios topológicos a través de la asignación de invariantes algebraicos. Aunque es difícil marcar un punto de inicio para esta área de conocimiento, fue a principios del siglo XX cuando empezó a desarrollarse por completo abriendo un mundo de herramientas de caracterización de los espacios topológicos, como los grupos de homotopía y de homología, el concepto central de este trabajo. Aún habiéndose generalizado las teorías homológicas de la mano de la teoría de categorías, se vuelve aquí al origen geométrico del concepto, la homología simplicial, para llegar a definir la homología persistente, herramienta que ha descubierto un nuevo modo de analizar las nubes de puntos y, como aquí se verá, también otros objetos matemáticos como son los grafos.

A continuación se hace un rápido paseo por las definiciones básicas necesarias para definir la homología simplicial y todo aquello que se necesitará para entender las sucesiones de homología persistente.

1.1. Homología simplicial

Definición 1.1. Sea \mathbb{R}^n el espacio afín euclidiano n -dimensional y v_0, \dots, v_k $k+1$ puntos afínmente independientes de \mathbb{R}^n , $k \leq n$. Llamamos *simplex k -dimensional* a la envoltura convexa de estos puntos, es decir:

$$\Delta(v_0, \dots, v_k) = \{x \in \mathbb{R}^n \mid x = \sum_i t_i v_i, 0 \leq t_i \leq 1, \sum_i t_i = 1\} \quad (1.1)$$

Si $\sigma = \Delta(v_0, \dots, v_p)$ es un *simplex p -dimensional*, una cara de σ de dimensión k es $\tau = \Delta(v_{i_0}, \dots, v_{i_k}) \subseteq \sigma$, y se denota $\tau < \sigma$.

Proposición 1.2. Los *simplex Δ^n* son compactos, arcoconexos, localmente arcoconexos, metrizables y contráctiles.

Definición 1.3. Un *complejo simplicial* K es un conjunto finito de *simplex* de \mathbb{R}^n , $K = \{\sigma_1, \dots, \sigma_r\}$ tal que:

1. Si σ_i es un *simplex* de K , entonces todas las caras de σ_i son de K
2. Si σ_i y σ_j son *simplex* de K , entonces o bien $\sigma_i \cap \sigma_j = \emptyset$ o bien $\sigma_i \cap \sigma_j$ es una cara de σ_i y de σ_j .

La *dimensión* de un *complejo simplicial* es la mayor dimensión de entre las de sus caras. Decimos que un *complejo simplicial* es *ordenado* si existe una *ordenación total* del conjunto de sus vértices $V_K = \{v_1, \dots, v_r\}$, de manera que $v_1 < \dots < v_r$. Dado un *simplex* ordenado $\Delta(v_{i_0}, \dots, v_{i_p})$ de K , lo denotaremos $[v_{i_0}, \dots, v_{i_p}]$.

Definición 1.4. Un *subcomplejo simplicial* de $K = \{\sigma_1, \dots, \sigma_p\}$, $\sigma_i \subseteq \mathbb{R}^n$ es un *complejo simplicial* $L = \{\tau_1, \dots, \tau_p\} \subseteq K$

A continuación se introduce el concepto de filtración, que tomará especial relevancia posteriormente en este trabajo.

Definición 1.5. Dado un complejo simplicial K , una filtración de K es una sucesión de subcomplejos

$$\emptyset = K_0 \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K_{n-1} \subseteq K_n = K. \quad (1.2)$$

Definición 1.6. Sea K un complejo simplicial ordenado. El grupo de cadenas p -dimensionales de K con $p \geq 0$ es el grupo abeliano libre generado por las caras p -dimensionales de K y es denotado $C_p(K)$,

$$C_p(K) = \bigoplus_{\substack{\sigma \subseteq K \\ \dim \sigma = p}} \mathbb{Z}[\sigma]. \quad (1.3)$$

Nótese que para todo $p > \dim K$, se tiene $C_p(K) = 0$.

Observación 1.7. $C_p(\Delta^n) \cong \mathbb{Z}^{\binom{n+1}{p+1}}$.

Definición 1.8. Sea K un complejo simplicial ordenado. Para todo $p \geq 1$, el operador diferencial $\partial_p : C_p(K) \rightarrow C_{p-1}(K)$ es un morfismo de grupos abelianos definido por:

$$\partial[v_{i_0}, \dots, v_{i_p}] = \sum_{k=0}^p (-1)^k [v_{i_0}, \dots, \widehat{v_{i_k}}, \dots, v_{i_p}], \quad (1.4)$$

donde $[v_{i_0}, \dots, \widehat{v_{i_k}}, \dots, v_{i_p}]$ es el $(p-1)$ -símplex ordenado obtenido de eliminar el vértice v_{i_k} . Se define $\partial_0 \equiv 0$.

Lema 1.9. Lema de Poincaré. $\partial^2 = 0$ (i.e., $\partial_k \partial_{k+1} = 0$ para todo $k \geq 0$).

Demostración. Sea $[v_0, \dots, v_k]$ un k -símplex.

$$\begin{aligned} \partial(\partial[v_0, \dots, v_k]) &= \partial \left(\sum_{l=0}^{k+1} (-1)^l [v_0, \dots, \widehat{v_l}, \dots, v_k] \right) = \sum_{l=0}^{k+1} (-1)^l \partial[v_0, \dots, \widehat{v_l}, \dots, v_k] = \\ &= \sum_{l=0}^{k+1} (-1)^l \left(\sum_{l' < l} (-1)^{l'} [v_0, \dots, \widehat{v_{l'}}, \dots, \widehat{v_l}, \dots, v_k] + \sum_{l' > l} (-1)^{l'-1} [v_0, \dots, \widehat{v_l}, \dots, \widehat{v_{l'}}, \dots, v_k] \right) = \\ &= \sum_{0 \leq l' < l \leq k} [(-1)^{l+l'} + (-1)^{l+l'-1}] [v_0, \dots, \widehat{v_{l'}}, \dots, \widehat{v_l}, \dots, v_k] = 0 \end{aligned}$$

ya que la primera parte de esta última expresión se anula. \square

Definición 1.10. Sea K un complejo simplicial ordenado, $C_p(K)$ sus grupos de cadenas p -dimensionales y ∂_p los correspondientes morfismos diferenciales. Llamamos complejo de cadenas a la sucesión de grupos abelianos finitamente generados

$$\cdots \xrightarrow{\partial_{n+1}} C_n(K) \xrightarrow{\partial_n} C_{n-1}(K) \xrightarrow{\partial_{n-1}} \cdots \xrightarrow{\partial_2} C_1(K) \xrightarrow{\partial_1} C_0(K) \quad (1.5)$$

Denotamos tal complejo como $C_*(K)$.

Definición 1.11. Para todo $p \geq 0$, definimos el grupo de los ciclos p -dimensionales de K como el núcleo de ∂_p , y el grupo de los bordes p -dimensionales de K como la imagen de ∂_{p+1} :

$$\begin{aligned} Z_p(K) &:= \ker \partial_p, \\ B_p(K) &:= \operatorname{Im} \partial_{p+1}. \end{aligned} \quad (1.6)$$

Observación 1.12. $Z_p(K)$ y $B_p(K)$ son subgrupos de $C_p(K)$ por ser el operador diferencial un morfismo de grupos.

Definición 1.13. Sea K un complejo simplicial. Para todo $p \geq 0$ el grupo de homología p -dimensional de K es el grupo cociente

$$H_p(K) := Z_p(K)/B_p(K). \quad (1.7)$$

Observación 1.14. Como $Z_p(K)$ es subgrupo de $C_p(K)$, que es finitamente generado, entonces $Z_p(K)$ es finitamente generado. En consecuencia, $H_p(K)$ también lo es y, por el teorema de estructura de los grupos abelianos finitamente generados, se tiene que

$$H_p(K) \cong \mathbb{Z}^r \oplus T_1 \oplus \dots \oplus T_m, \quad (1.8)$$

donde r es el rango del grupo de homología y $T_i \cong \mathbb{Z}/d_i\mathbb{Z}$ son grupos cíclicos finitos que constituyen la parte de torsión y cumplen $d_0|d_1|\dots|d_m$.

Definición 1.15. Llamamos número de Betti p -dimensional del complejo simplicial K , $\beta_p(K)$, al rango del grupo de homología p -dimensional de K .

$$\beta_p(K) := \text{rank } H_p(K) \quad (1.9)$$

Se acaba de presentar una de las herramientas más potentes con que cuenta la topología algebraica y en la que profundizará en este trabajo. De manera intuitiva, si se trabaja con coeficientes en \mathbb{Z} , los números de Betti p -dimensionales nos hablan sobre el número de agujeros p -dimensionales que tiene nuestro complejo. Ganemos un poco de intuición sobre el significado e importancia de los grupos de homología y números de Betti con un par de resultados sencillos que jugarán un papel importante en nuestro análisis de redes.

Teorema 1.16. Sea K un complejo simplicial no vacío. Entonces K tiene r componentes conexas si y solo si $H_0(K) \cong \mathbb{Z}^r$.

Demostración. Veremos que si K es conexo y no vacío, entonces $H_0(K) \cong \mathbb{Z}$. Sea $v_0 \in K$ un vértice. Dado cualquier otro vértice v_i , al ser K simplicialmente conexo existe una sucesión de vértices $w_{i0}, w_{i1}, \dots, w_{ir}$ con $w_{i0} = v_0$ y $w_{ir} = v_i$ tales que $[w_{ij}, w_{ij+1}]$ son aristas de K . La 0-cadena $v_i - v_0$ es un borde, ya que se puede considerar la siguiente 1-cadena:

$$c = [w_{i0}, w_{i1}] + \dots + [w_{ir-1}, w_{ir}] \in C_1(K)$$

que tiene borde:

$$\partial c = (w_{i1} - w_{i0}) + \dots + (w_{ir} - w_{ir-1}) = w_{ir} - w_{i0} = v_i - v_0.$$

Es necesario darse cuenta también de que el núcleo de ∂_0 es todo $C_0(K)$ por estar en la cola de la sucesión en el complejo de cadenas, así que $Z_0(K) = C_0(K)$. Ahora se construye un morfismo que permita llegar a la conclusión deseada a través del teorema de la isomorfía. Así se crea el llamado morfismo de aumentación.

$$\epsilon : C_0(K) \longrightarrow \mathbb{Z}$$

$$\sum \lambda_i v_i \mapsto \sum \lambda_i.$$

Será suficiente probar que es exhaustivo y que tiene núcleo $B_0(K)$ para aplicar el primer teorema de la isomorfía y concluir $H_0(K) \cong \mathbb{Z}$. Sea $\lambda \in \mathbb{Z}$, es claro que $\epsilon(\lambda v_0) = \lambda$, por lo que ϵ es exhaustiva. Veamos ahora las dos inclusiones:

$B_0(K) \subseteq \ker \epsilon$. Dado que las aristas de K generan $C_1(K)$, sus bordes generarán $B_0(K)$. Sea $\sigma = [v_1, v_2]$, entonces

$$\epsilon(\partial\sigma) = \epsilon(v_2 - v_1) = 1 - 1 = 0,$$

por lo que dado un elemento cualquiera $\partial\sigma$ de $B_0(K)$, este pertenece a $\ker \epsilon$.

$\ker \epsilon \subseteq B_0(K)$. Sea una 0-cadena que se anule por ϵ . Entonces

$$\sum_i \lambda_i v_i = \sum_i \lambda_i v_i - \sum_i \lambda_i v_0 = \sum_i \lambda_i (v_i - v_0) = \sum_i \lambda_i c_i = \partial(\sum_i \lambda_i c_i) \in B_0(K),$$

donde las últimas igualdades vienen del detalle con el que empezaba la demostración y de la linealidad del operador diferencial. Se tiene así la igualdad y por tanto $H_0(K) \cong \mathbb{Z}$. Si un complejo simplicial está constituido por diversas componentes conexas, $K = K_1 \cup \dots \cup K_n$, entonces $C_*(K) = C_*(K_1) \oplus \dots \oplus C_*(K_n)$, y su homología cumple $H_*(C_*(K_1)) \oplus \dots \oplus H_*(C_*(K_n)) \cong H_*(C_*(K_1)) \oplus \dots \oplus H_*(C_*(K_n))$ (se puede comprobar construyendo un isomorfismo a través de los morfismos de inclusión o de proyección evidentes, cualquiera de las dos opciones es válida por ser \oplus producto y coproducto). Por consiguiente, si K tiene r componentes conexas, entonces $H_0(K) \cong \mathbb{Z}^r$. \square

El siguiente resultado está enunciado para espacios topológicos y hace uso de la *homología singular*. Aunque no será usada explícitamente en el trabajo, puede aportar luz sobre la información que esconde, en particular, el primer grupo de homología y su relación con los *agujeros*. Así como el teorema que lo sigue, no será demostrado por no haber introducido toda la teoría de la homología singular ni las herramientas básicas y necesarias del álgebra homológica, como su teorema fundamental, las sucesiones de Mayer–Vietoris o el lema de los cinco. Se puede encontrar demostraciones en [2].

Teorema 1.17. *Teorema de Hurewicz. Sea X un espacio topológico arcoconexo y $n > 0$. Existe un homomorfismo canónico*

$$h_* : \pi_n(X) \longrightarrow H_n(X) \tag{1.10}$$

donde $\pi_n(X)$ es el grupo de homotopía n -dimensional. En particular, para $n = 1$ existe un isomorfismo con el abelianizado del grupo fundamental:

$$\overline{h_*} : \pi(X)_{ab} \longrightarrow H_1(X) \tag{1.11}$$

Por último, y confirmando una probable intuición del lector:

Teorema 1.18. $H_p(\Delta^n) \cong 0$ para todo $p \geq 0$ y para toda $n > 0$. (Todo simplex n -dimensional es contráctil)

1.2. Homología persistente

La homología persistente constituye un pilar del análisis de datos topológico (TDA, *Topological Data Analysis*) y fue desarrollada para extraer propiedades topológicas de espacios o funciones: qué características son intrínsecas del espacio a estudiar y cuáles son tan sólo ruido del conjunto de datos. La homología persistente hace un estudio a través de todas las escalas para determinar precisamente esto.

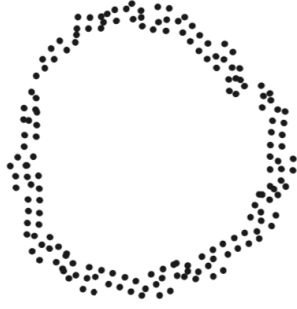


Figura 1: Nube de puntos en círculo estadístico

Vayamos al origen e imaginemos una muestra de puntos el espacio subyacente de los cuales queremos conocer por su topología. El siguiente ejemplo, extraído de [3], será de ayuda.

Dado que lo que conocemos del espacio es un conjunto finito de puntos, calcular su sucesión de grupos de homología es trivial y, más importantemente, inútil. Nace aquí la idea de intentar extraer un cierto espacio subyacente: se puede generar un complejo simplicial que guarde información métrica de la distribución inicial y que permitirá inferir sobre el conjunto. Una manera de abordar el problema, que será formalizada cuando se presente el complejo de Vietoris–Rips, sería considerar un simplex cuando la distancia entre cierta cantidad de nodos sea inferior a un cierto umbral. Pero esa cota estaría reduciendo el estudio a una escala de trabajo dada. Si se considera, sin embargo, una sucesión de complejos simpliciales creciente –que constituirá una filtración– para valores cada vez mayores de esa cota, podremos estudiar la homología del conjunto a todos los niveles. De manera intuitiva, aquellas características que encontremos en un intervalo más largo de valores de nuestro umbral –lo que se denominará tener mayor *persistencia*– serán las fundamentales del conjunto de puntos, mientras que aquellas que duren menos serán menos significativas o simplemente ruido estadístico. En el caso de la imagen anterior, se esperaría concluir que la topología del espacio subyacente es la de S^1 .

La homología persistente parte, pues, de una filtración del complejo simplicial. Formalicemos los conceptos.

Definición 1.19. *Sea K un complejo simplicial y sea una filtración:*

$$\emptyset = K_0 \subseteq K_1 \subseteq K_2 \subseteq \cdots \subseteq K_{n-1} \subseteq K_n = K \quad (1.12)$$

Las inclusiones inducen un morfismo en homología $i_p^{i,j} : H_p(K_i) \rightarrow H_p(K_j)$ para todo $p > 0$. El p -ésimo grupo de homología (i, j) -persistente es la imagen del morfismo inducido:

$$H_p^{i,j} := \text{Im } i_p^{i,j}, \quad 0 \leq i \leq j \leq n. \quad (1.13)$$

Definimos también el p -ésimo número de Betti (i, j) -persistente como $\beta_p^{i,j} = \text{rank } H_p^{i,j}$.

Por consiguiente, las clases de homología persistente son aquellas que se mantienen en los sucesivos pasos de la filtración definida. Más intuitivo por similitud con la definición de homología simplicial es el siguiente resultado que también se usa como definición en algunas referencias [6].

Proposición 1.20. $H_p^{i,j} \cong Z_p(K_i) / (Z_p(K_i) \cap B_p(K_j))$.

Demostración. $Im(i_p^{i,j}) \cong H_p(K_i)/ker(i_p^{i,j}) \cong H_p(K_i)/(Z_p(K_i) \cap B_p(K_j)) \cong Z_p(K_i)/(Z_p(K_i) \cap B_p(K_j))$. □

Para definir la persistencia de un ciclo dado es necesaria una noción de *nacimiento* y *muerte* del ciclo. Decimos que una clase de homología $[z]$ nace en K_i cuando $[z] \notin H_p^{i-1,1}$, y decimos que muere entrando en K_j si se une a una clase anterior en K_j , es decir, $i_p^{i,j-1}([z]) \notin H_p^{i-1,j-1}$ pero $i_p^{i,j}([z]) \in H_p^{i-1,j}$. En tal caso, el ciclo tiene una persistencia de $j - i$. Si no muere, se dice que su persistencia es infinita.

Podemos introducir así el llamado *diagrama de persistencia*, una representación gráfica de las clases de homología persistente del complejo a estudiar, donde cada clase es representada por un punto de coordenadas (nacimiento, muerte) en el plano. Por tal de definir bien estas herramientas gráficas es necesario dar estructura formal a la homología persistente.

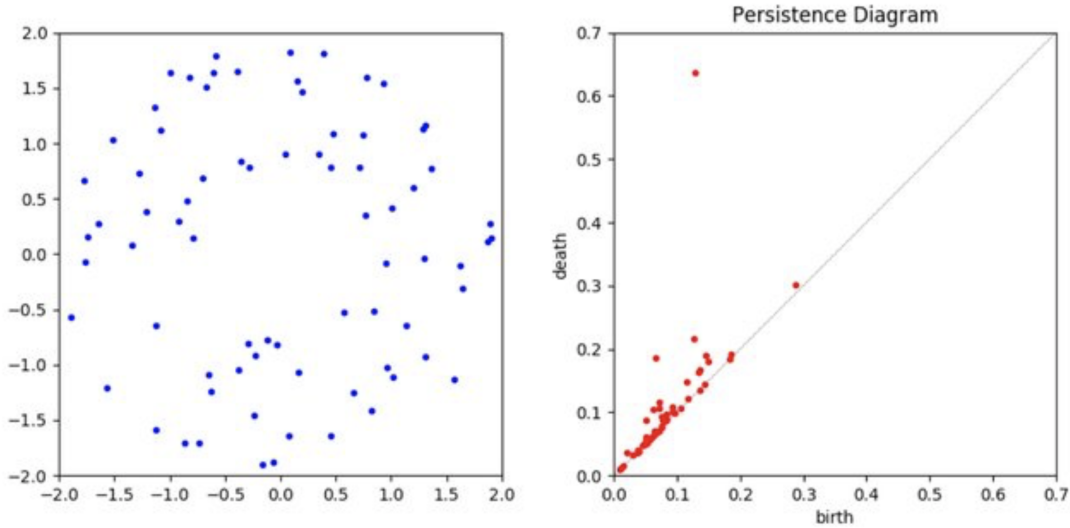


Figura 2: Nube de puntos y correspondiente diagrama de persistencia en dimensión 1. Se aprecia como el punto más distanciado de la diagonal detecta la estructura anular del espacio inicial, mientras que los demás ciclos mueren poco después de su aparición. Extraído de [37].

1.2.1. Módulos de persistencia y funtorialidad

A continuación se va a dar estructura formal a la homología persistente. Debido a la variedad de planteamientos que se encuentran en la bibliografía, se ha decidido aquí empezar con uno sencillo y breve para describir los resultados principales que se van a obtener, extraída de Polterovich [5]. Se utilizará como intuición del camino a seguir, sin detallar las cosas, para después formalizarlo siguiendo la línea de trabajo de Carlsson [6], que resulta más elegante. Vemos pues que podemos resumir la construcción en un par de definiciones y un resultado principal.

Primer acercamiento

Definición 1.21. *Un módulo de persistencia es un par (V, π) donde $V = \{V_t\}, t \in \mathbb{R}$ es una colección de espacios vectoriales sobre un cuerpo K y $\pi = \{\pi_{s,t}\}$ una colección de aplicaciones lineales $\pi_{s,t} : V_s \rightarrow V_t$ para todo $s, t \in \mathbb{R}, s \leq t$, que cumplen las cuatro condiciones siguientes:*

1. *Para cualesquiera $s \leq t \leq r$ se tiene $\pi_{s,r} = \pi_{t,r} \circ \pi_{s,t}$.*
2. *Para todo $t \in \mathbb{R}$ excepto un número finito de puntos existe un entorno U de t tal que $\pi_{s,r}$ es un isomorfismo para todo $s < r$ en U .*
3. *Para todo $t \in \mathbb{R}$ y para toda $s \leq t$ suficientemente cercana a t , la aplicación $\pi_{s,t}$ es un isomorfismo.*
4. *Existe un valor $s_- \in \mathbb{R}$ tal que $V_s = 0$ para todo $s \leq s_-$.*

El conjunto de espacios $P_{s,t} = \text{Im}(\pi_{s,t})$ se denomina homología persistente de V .

La primera condición de la definición es el núcleo del concepto de persistencia, mientras que las otras son añadidos que facilitan la posterior derivación de resultados. Las condiciones 2 y 4 aseguran que es suficiente asociar un número finito de espacios $P_{s,t}$ al módulo de persistencia, ya que tan sólo hay un número finito de puntos en los que $\pi_{s,t} \neq \text{Id}$, mientras que la tercera asegura la unicidad de descomposición de los módulos de persistencia en el *Normal Form Theorem* sin necesidad de definiciones o resultados intermedios. Veremos más adelante que la definición puede recortarse para ganar en elegancia y fundamentar mejor el concepto.

Definición 1.22. *Un barcode \mathcal{B} es un multiconjunto finito de intervalos, i.e. un conjunto de intervalos con una cierta multiplicidad, (I_i, m_i) , donde $I_i = (a, b]$, con $a \in \mathbb{R}, b \in \mathbb{R} \cup \{\infty\}$ y $m_i \in \mathbb{N}$.*

Teorema 1.23. *Normal Form Theorem. Sea (V, π) un módulo de persistencia sobre K . Entonces existe un conjunto $\{(I_i, m_i)\}_{i=1, \dots, N}$ de intervalos I_i con multiplicidades $m_i \in \mathbb{N}$, donde $I_i = (a_i, b_i]$ o $I_i = (a_i, \infty)$, $I_i \neq I_j$ para toda $i \neq j$, tales que*

$$V \cong \bigoplus_{i=1}^N K(I_i)^{m_i}. \quad (1.14)$$

donde $K(a, b)$ es un módulo de persistencia definido de la siguiente manera:

$$K(a, b)_t = \begin{cases} K & \text{si } t \in (a, b] \\ 0 & \text{si } t \notin (a, b] \end{cases} \quad \pi_{s,t} = \begin{cases} \text{Id} & \text{si } s, t \in (a, b] \\ 0 & \text{en otro caso} \end{cases} \quad (1.15)$$

Esta descomposición es además única. En consecuencia, a cada módulo de persistencia le corresponde un único barcode $\mathcal{B}(V)$, que consta de intervalos I_i con multiplicidad m_i .

Demostración. Ver el Capítulo 2 de [5]. □

Con esto queda construida una correspondencia para cada módulo de persistencia con un diagrama. Formalicemos estos resultados ahora partiendo de una definición más básica de módulo de persistencia. En consecuencia se deberá concretar las condiciones de finitud que permiten la construcción de la estructura.

Estructura de los módulos de persistencia

Definición 1.24. Un complejo de persistencia \mathcal{C} es una familia de complejos de cadenas $\{C_*^i\}_{i \geq 0}$ sobre un anillo R junto con una familia de morfismos de complejos de cadenas $\{f^i\}_i$, $f^i : C_*^i \rightarrow C_*^{i+1}$, de manera que se tiene el siguiente diagrama:

$$C_*^0 \xrightarrow{f^0} C_*^1 \xrightarrow{f^1} C_*^2 \xrightarrow{f^2} \dots \quad (1.16)$$

Observación 1.25. Un complejo simplicial filtrado con los morfismos inducidos por la inclusión es un complejo de persistencia.

Por claridad, se podría desplegar un complejo de persistencia de la manera siguiente:

$$\begin{array}{ccccccc} \partial_3 \downarrow & & \partial_3 \downarrow & & \partial_3 \downarrow & & \\ C_2^0 & \xrightarrow{f^0} & C_2^1 & \xrightarrow{f^1} & C_2^2 & \xrightarrow{f^2} & \dots \\ \partial_2 \downarrow & & \partial_2 \downarrow & & \partial_2 \downarrow & & \\ C_1^0 & \xrightarrow{f^0} & C_1^1 & \xrightarrow{f^1} & C_1^2 & \xrightarrow{f^2} & \dots \\ \partial_2 \downarrow & & \partial_2 \downarrow & & \partial_2 \downarrow & & \\ C_0^0 & \xrightarrow{f^0} & C_0^1 & \xrightarrow{f^1} & C_0^2 & \xrightarrow{f^2} & \dots \end{array} \quad (1.17)$$

De la misma manera podemos cambiar complejo de cadenas con coeficientes en un anillo R por R -módulo para obtener siguiente definición análoga.

Definición 1.26. Un módulo de persistencia es una familia de R -módulos M^i junto con homomorfismos $\phi^i : M^i \rightarrow M^{i+1}$.

Observación 1.27. Se puede definir fácilmente morfismos entre módulos de persistencia y comprobar que constituyen una categoría. La denotaremos por **R-PersMod**.

La definición es así directa, pero hará falta especificar la finitud del complejo para extraer ciertas conclusiones.

Definición 1.28. Un módulo de persistencia $\{(M_*^i, \phi^i)\}_{i \geq 0}$ es finito si cada uno de los módulos que lo componen es un R -módulo finitamente generado y si los morfismos ϕ^i son isomorfismos para $i \geq m$ para algún $m \in \mathbb{Z}$.

Observación 1.29. Dado un complejo simplicial filtrado K , como es finito, su complejo de persistencia es también finito y la homología del complejo de persistencia con los morfismos inducidos en homología son un módulo de persistencia finito.

Sea el anillo de polinomios $R[t]$ con la gradación estándar. Se busca ahora demostrar una equivalencia de categorías entre **R-PersMod** y la categoría de $R[t]$ -módulos finitamente generados no-negativamente graduados. Sea un módulo de persistencia $\mathcal{M} = \{(M^i, \phi^i)\}_{i \geq 0}$ sobre R . Se define un módulo graduado sobre $R[t]$ como

$$\alpha(\mathcal{M}) = \bigoplus_{i=0}^{\infty} M^i \quad (1.18)$$

La acción de t viene dada por:

$$t * (m^0, m^1, \dots) = (0, \phi^0(m^0), \phi^1(m^1), \dots) \quad (1.19)$$

Teorema 1.30. α define una equivalencia de categorías entre $\mathbf{R}\text{-PersMod}$ y la de $R[t]$ -módulos finitamente generados no-negativamente graduados.

Demostración. Requiere de la teoría de Artin-Rees de álgebra conmutativa. Puede hallarse en el Capítulo 5 de [9].

Este resultado nos da directamente la estructura de los módulos de persistencia. Sabemos que si el anillo R es un cuerpo, entonces $R[t]$ es un dominio de ideales principales y la estructura del $R[t]$ -módulo viene dada por

$$\left(\bigoplus_{i=1}^n F[t]\right) \oplus \left(\bigoplus_{j=1}^m F[t]/(t^{n_j})\right), \quad (1.20)$$

que expresan la parte libre y de torsión correspondientemente.

Observación 1.31. Si R no es un cuerpo no hay clasificaciones simples de los módulos de persistencia. Incluso la clasificación de módulos sobre $\mathbb{Z}[t]$ es extremadamente complicada.

Ahora tenemos una fácil descomposición de los módulos de persistencia. Facilitemos aún más su interpretación buscando una relación con los *barcodes*, que como hemos visto anteriormente, son conjuntos de intervalos con una multiplicidad dada.

Definición 1.32. Un \mathcal{P} -intervalo es un par ordenado (i, j) con $0 \leq i < j$ y donde $i \in \mathbb{Z}$, $j \in \mathbb{Z} \cup \{+\infty\}$.

Proposición 1.33. Existe una biyección entre el conjunto finito de \mathcal{P} -intervalos, \mathcal{S} y los $R[t]$ -módulos graduados finitamente generados.

Demostración. Sea $Q(i, j) = \sum_i R[t]/(t^{j-i})$ y $Q(i, +\infty) = \sum_i R[t]$. Nótese que hay una correspondencia de uno a uno entre los módulos definidos y los \mathcal{P} -intervalos. Para un conjunto de \mathcal{P} -intervalos $\mathcal{S} = \{(i_1, j_1), \dots, (i_n, j_n)\}$, $i_k \in \mathbb{Z}$, $j_k \in \bar{\mathbb{Z}}$, para todo $1 \leq k \leq n$, se define

$$Q(\mathcal{S}) = \bigoplus_{l=1}^n Q(i_l, j_l).$$

La correspondencia dada define la biyección. \square

Con esto acabamos de asegurar que cada \mathcal{P} -intervalo (i, j) representa un elemento de la base de cada espacio vectorial dado por el grupo de homología (i, j) -persistente, es decir, un ciclo que nace en el paso i de la filtración y muere recombinándose en el paso $j - 1$, es elemento de la base de un grupo cíclico finito. Los ciclos que no llegan a morir definen un elemento de la base de la parte libre de torsión del módulo de persistencia.

Ahora, recuperando la definición de *barcode* dada anteriormente (Definición 1.22), vemos que la representación contiene toda la información necesaria para describir el módulo de persistencia al que está asociado y que, de hecho, es equivalente.

1.2.2. Distancias

Es muy útil poder determinar cuán parecidos o cercanos son dos módulos de persistencia entre sí. Se verá en particular en este trabajo cómo se usa esto como herramienta para discernir entre objetos de estudio e incluso para clasificarlos. Definamos ahora una distancia entre módulos de persistencia.

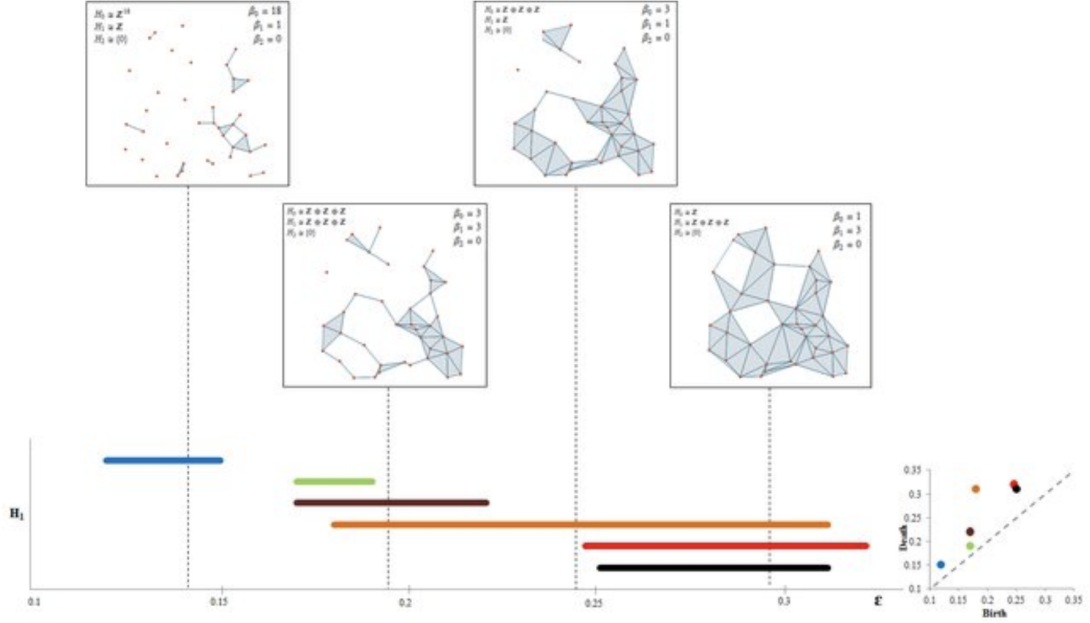


Figura 3: Comparación entre el barcode y el diagrama de persistencia de dimensión 1 del complejo simplicial filtrado resultante de una nube de puntos, que se muestra en diversas etapas de la filtración. Las mayores cavidades se ven representadas en el barcode por las barras de mayor longitud, y en el diagrama por los puntos más alejados de la diagonal. Extraída de [60].

Distancia de intercalación

Como para cualquier otra categoría, decimos que dos módulos de persistencia \mathcal{M} y \mathcal{N} son isomorfos si existen homomorfismos $\phi : \mathcal{M} \rightarrow \mathcal{N}$ y $\psi : \mathcal{N} \rightarrow \mathcal{M}$ tales que $\phi\psi = Id_{\mathcal{N}}$ y $\psi\phi = Id_{\mathcal{M}}$. Esta relación es muy fuerte sobre todo si pensamos que la motivación para estudiar el parecido entre módulos de persistencia proviene del estudio de conjuntos de datos diferentes y que contarán con una cierta cantidad de ruido. Por ese motivo se desarrolla una noción que pueda cuantificar la incertidumbre, el δ -intercalado (δ -interleaving). Definamos los conceptos que nos llevarán a una distancia entre módulos de persistencia.

Definición 1.34. *Homomorfismos desplazados.* Son aquellos que cambian el valor del índice de persistencia. Sean $\mathcal{M} = (M_i, m_{ij})$ y $\mathcal{N} = (N_i, n_{ij})$ R -módulos de persistencia y sea $\delta \in \mathbb{R}$. Un homomorfismo de grado δ es un conjunto ϕ de aplicaciones lineales $\phi_t : M_t \rightarrow N_{t+\delta}$ para cualquier t de manera que el siguiente diagrama commuta para todo $s \leq t$.

$$\begin{array}{ccc}
 M_s & \xrightarrow{m_s^t} & M_t \\
 \phi_s \downarrow & & \downarrow \phi_t \\
 N_{s+\delta} & \xrightarrow{n_{s+\delta}^{t+\delta}} & N_{t+\delta}
 \end{array} \tag{1.21}$$

Denotamos $Hom^\delta(\mathcal{M}, \mathcal{N}) := \{\text{homomorfismos } \mathcal{M} \rightarrow \mathcal{N} \text{ de grado } \delta\}$.

Observación 1.35. Está bien definido el endomorfismo que cambia índices de persistencia sobre el mismo módulo, $Id_{\mathcal{M}}^{\delta}$.

Definición 1.36. Decimos que dos módulos \mathcal{M} y \mathcal{N} están δ -intercalados si existen aplicaciones

$$\phi \in Hom^{\delta}(\mathcal{M}, \mathcal{N}), \quad \psi \in Hom^{\delta}(\mathcal{N}, \mathcal{M}) \quad (1.22)$$

de manera que $\psi\phi = Id_{\mathcal{M}}^{\delta}$ y $\phi\psi = Id_{\mathcal{N}}^{\delta}$.

Observación 1.37. De esta manera parece que podría anularse la distancia entre dos módulos no isomorfos. En la presentación dada en [5] la condición de semi-continuidad viene dada en la definición de módulo de persistencia, asegurando que eso no pase y convirtiendo así el δ -intercalado es una métrica.

Definición 1.38. Dados dos módulos de persistencia \mathcal{M} y \mathcal{N} , definimos la distancia de intercalación entre \mathcal{M} y \mathcal{N} como

$$d_{int}(\mathcal{M}, \mathcal{N}) = \inf\{\delta > 0 \mid \mathcal{M} \text{ y } \mathcal{N} \text{ están } \delta\text{-intercalados}\} \quad (1.23)$$

Distancias sobre barcodes

Definamos ahora una distancia sobre los diagramas de persistencia o barcodes que será más fácil de computar. La idea para la construcción es considerar el conjunto de todos los *matchings* entre intervalos de los barcodes de unos módulos de persistencia dados, asignar una cierta penalización a cada uno de los matchings y finalmente buscar la mínima penalización total en el conjunto de todos los matchings posibles.

El peso que se da a cada asignación viene dado por la norma del supremo L_{∞} , $\pi([x_1, x_2], [y_1, y_2]) = \|(y_1 - x_1, y_2 - x_2)\|_{\infty}$. Dados dos barcodes B_1 y B_2 , denominamos $\mathcal{D}(B_1, B_2)$ al conjunto de biyecciones $\theta : B_1 \rightarrow B_2$ para las que $\pi(I, \theta(I)) \neq 0$ para un número finito de puntos $I \in B_1$.

Definición 1.39. Sean B_1 y B_2 barcodes de dos módulos de persistencia distintos y sea $p > 0$. Definimos la distancia p -Wasserstein:

$$W_p(B_1, B_2) = \inf_{\theta \in \mathcal{D}(B_1, B_2)} \left(\sum_{I \in B_1} \pi(I, \theta(I))^p \right)^{\frac{1}{p}}. \quad (1.24)$$

Cuando $p = \infty$ obtenemos la denominada *Bottleneck distance* y la denotaremos d_{bot} .

Ahora que tenemos definidas distancias sobre los diagramas y sobre los módulos, encontrar una relación entre ellas nos permitirá inferir en las diferencias entre módulos de persistencia a partir de las distancias entre sus diagramas. Esa relación viene dada por el siguiente teorema de la isometría. Por la complejidad de su demostración (dividida en la demostración de las dos desigualdades como teoremas de estabilidad), no se entrará a desarrollar el teorema.

Teorema 1.40. *Teorema de la isometría.* Sean \mathcal{M} y \mathcal{N} dos módulos de persistencia y $\mathcal{B}(\mathcal{M})$, $\mathcal{B}(\mathcal{N})$ sus respectivos barcodes. Entonces $\mathcal{M} \mapsto \mathcal{B}(\mathcal{M})$ es una isometría, de manera que:

$$d_{int}(\mathcal{M}, \mathcal{N}) = d_{bot}(\mathcal{B}(\mathcal{M}), \mathcal{B}(\mathcal{N})). \quad (1.25)$$

Demostración. Consultar el capítulo 3 de [5] íntegramente dedicado a su demostración o el capítulo 4 de [10] para un acercamiento distinto a este mismo resultado. \square

1.2.3. Persistence landscapes

Los "paisajes de persistencia" son una manera de expresar toda la información contenida en un diagrama de persistencia de manera más fácilmente legible e interpretable, y fueron introducidos por Bubenik en 2015 [13]. Existen otras formas de resumir la información contenida en los diagramas de persistencia [16], pero se expone aquí esta como simple muestra de la ayuda que puede ser una reinención en la representación de la persistencia, por permitir ver los resultados más fácilmente y de manera más compacta.

Definición 1.41. Sea un módulo de persistencia $\mathcal{M} = (M_t, \pi_{st})$ y sean $\beta^{i,j}$ los números de Betti i, j -persistentes, es decir, la dimensión de $\text{Im}(\pi_{ij})$, con $i \leq j$. Un paisaje de persistencia es una secuencia de funciones $f_k : \mathbb{R} \rightarrow \mathbb{R} \cup \{\infty\}$ donde

$$f_k(t) := \sup \{m \geq 0 \mid \beta^{t-m, t+m} \geq k\}. \quad (1.26)$$

Dado un diagrama de persistencia $\mathcal{B} = \{(b_i, d_i)\}$, las funciones vienen dadas por:

$$f_k(t) = k\text{-ésimo valor más grande de } \min(t - b_i, d_i - t)_+ \quad (1.27)$$

donde el subíndice $+$ indica que se escoge el máximo entre el valor y 0.

Si nos fijamos, lo que se consigue es una rotación de 45 grados del diagrama de persistencia, haciendo de la diagonal el nuevo eje de abscisas. Se dibujan triángulos isósceles con base en el eje que tienen como vértice cada uno de los puntos del diagrama. El valor de k hace ir disminuyendo la importancia de los agujeros representados en el diagrama en la estructura topológica del conjunto, y el avance de t nos mueve a lo largo de la filtración.

Observación 1.42. Los paisajes de persistencia cumplen:

1. $f_k(t) \geq 0$ para toda k .
2. $f_k(t) \geq f_{k+1}(t)$ para toda k .
3. f_k es 1-Lipschitz para toda k .

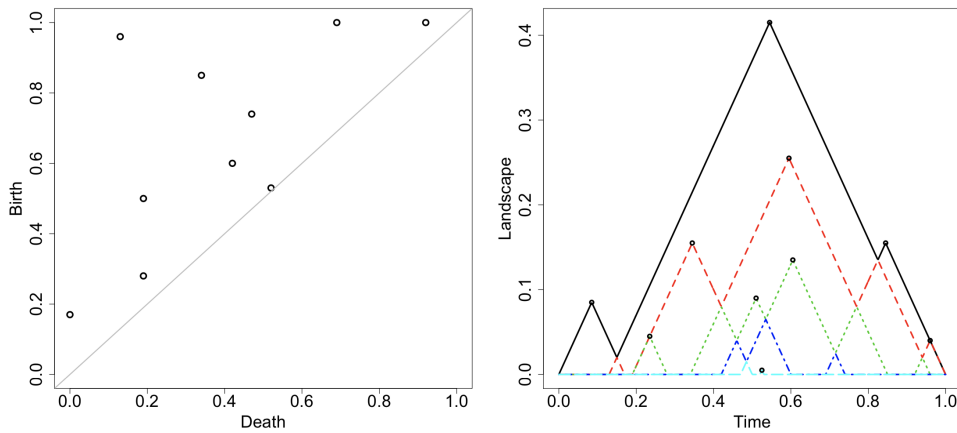


Figura 4: Diagrama de persistencia y su correspondiente paisaje de persistencia. Cada línea corresponde a una f_k : la negra para $k = 1$, roja para $k = 2$, etc. Extraído de [16].

Se demuestra en [13] su compatibilidad con las distancias definidas sobre barcodes y una variedad de ventajas estadísticas que ofrece. En la Figura 4 puede apreciarse la aportación de esta construcción.

Se tiene hasta ahora las herramientas necesarias para conocer la información homológica del espacio subyacente a un conjunto inicial de puntos o complejo simplicial, así como aquellas que permitirán compararlo con otros espacios a partir de la representación en barcodes y la inferencia sobre la estructura del módulo de persistencia que los caracterice.

1.2.4. Alternativas al cálculo de distancias sobre barcodes

Queda así asentado el valor que ofrece la bottleneck distance sobre los diagramas de persistencia, pero computacionalmente los cálculos son costosos y surgen alternativas más económicas.

Funciones indicatrices de persistencia

Rieck, Fugacci et al. proponen en [17] el uso de las funciones indicatrices de persistencia de un diagrama.

Definición 1.43. *Sea \mathcal{D} un diagrama de persistencia. Se define la función indicatriz de persistencia:*

$$\begin{aligned} \mathbb{1}_{\mathcal{D}} : \mathbb{R} &\longrightarrow \mathbb{N} \\ \epsilon &\mapsto \#\{(c, d) \in \mathcal{D} \mid \epsilon \in (c, d)\} \end{aligned} \tag{1.28}$$

que da el número de componentes activas en el diagrama para cada valor del parámetro ϵ .

La función indicatriz es de tipo escalón, y su integral es por consiguiente lineal. Se puede entonces usar la distancia L_p para calcular distancias entre indicatrices de persistencia:

$$d(\mathbb{1}_{\mathcal{D}_1}, \mathbb{1}_{\mathcal{D}_2}) = \left(\int_{\mathbb{R}} |\mathbb{1}_{\mathcal{D}_1}(x) - \mathbb{1}_{\mathcal{D}_2}(x)|^p dx \right)^{\frac{1}{p}} \tag{1.29}$$

que es mucho más sencilla de calcular que la p -Wasserstein. Aún siendo funciones claramente no inyectivas, las funciones indicatrices pueden ser útiles para hacer rápidos cálculos de disimilitud entre redes. La aplicación a redes reales encuentra además que los resultados obtenidos a través de la bottleneck distance y los obtenidos a través de este método están relacionados con un coeficiente de Pearson $R^2 \approx 0,96$, demostrando que las dos distancias están conectadas por una transformación lineal. Se concluye que el uso de funciones indicatrices con la distancia L_2 es un método suficientemente bueno para capturar diferencias entre redes.

Single Linkage Dendrograms

Una útil alternativa a la comparación de diagramas de persistencia mediante la Bottleneck distance se encuentra en la construcción de un *Single Linkage Dendrogram* y el uso de la distancia de Gromov-Hausdorff. Se expone primero un par de conceptos sobre el clustering aglomerativo para ganar visión sobre lo que se intenta.

Las técnicas de clustering jerárquico pretenden organizar una nube de puntos en módulos o clústers a todas sus escalas para ayudar a entender su estructura interna. En el clustering aglomerativo, se parte de un módulo formado por cada uno de los elementos a estudiar, y el sistema se va recombinando en módulos progresivamente mayores hasta que se engloba todo en una componente. Para hacer esto es necesario definir una distancia entre clusters para poder determinar en cada paso cuales son los más cercanos entre sí y considerarlos un único en el paso siguiente. El denominado *Single Linkage Clustering* usa la distancia dada por el mínimo de distancia entre puntos de clústers separados, i.e., dados dos clústers X e Y

$$d(X, Y) = \min_{x \in X, y \in Y} d(x, y). \quad (1.30)$$

Los resultados de estas técnicas pueden representarse visualmente en un *dendrograma*. Fijémonos en lo siguiente. Sea un complejo simplicial con la filtración de Vietoris-Rips (ver 3.1.2), y sean S_m^k, S_n^k dos componentes disconexas del complejo en el paso k de la filtración. Si hay dos nodos $x_i \in S_m^k$ y $x_j \in S_n^k$ tales que $d(x_i, x_j) < \epsilon_{k+1}$, entonces S_m^k y S_n^k desarrollan una arista en el siguiente paso y pasarán a estar conectados en el $k+1$. En otras palabras, S_m^k y S_n^k estarán conectados en el paso $k+1$ si

$$d(S_m^k, S_n^k) = \min_{x_i \in S_m^k, x_j \in S_n^k} d(x_i, x_j) < \epsilon_{k+1}, \quad (1.31)$$

que es la misma condición usada para unir clústers durante el Single Linkage Clustering. Por consiguiente, la secuencia de cómo las componentes se unen durante la filtración de Rips es idéntica a la secuencia del SLC. Los barcodes y dendrogramas son equivalentes, bajo reordenación de las componentes. La Figura 5 aclara esta equivalencia. En el estudio llevado a cabo en [18] se usan estas ideas reduciéndose a la homología 0-dimensional por su estrecha relación con la conexión, pero los autores ven posible extenderlo a una dimensión arbitraria usando una filtración adecuada en la red.

Se gana de esta manera también una nueva manera de calcular distancias entre diagramas. La distancia de Gromov-Hausdorff la propone Gromov al considerar el conjunto de todos los espacios métricos como espacio métrico en sí, y resulta una herramienta muy útil para calcular diferencias entre espacios métricos [14] [15]. Aunque la Single Linkage Distance no satisface la desigualdad triangular, sí satisface

$$\max(d(x_1, x_2), d(x_2, x_3)) \geq d(x_1, x_3) \quad (1.32)$$

lo que hace posible ver los dendrogramas como espacios ultramétricos. Sobre ellos, y tomando que los nodos tienen una localización fija, la distancia de Gromov-Hausdorff entre dos dendrogramas X e Y se discretiza y queda [18]:

$$d_{GH}(X, Y) = \frac{1}{2} \max_{i, j} |d_X(x_i, x_j) - d_Y(y_i, y_j)| \quad (1.33)$$

lo que hace sencillo su cálculo. En [19] o [18] se comprueba muy exitosamente el poder discriminativo de este método en redes complejas es mayor que el que ofrece la bottleneck distance o los típicos de la teoría de redes, y en [20] y [21] se usan estos conceptos para discriminar características anormales en redes cerebrales o encontrar propiedades morfológicas de las redes neuronales de adultos con sordera. Su eficacia y eficiencia, si nos podemos reducir a estudiar la homología 0-dimensional, está demostrada.

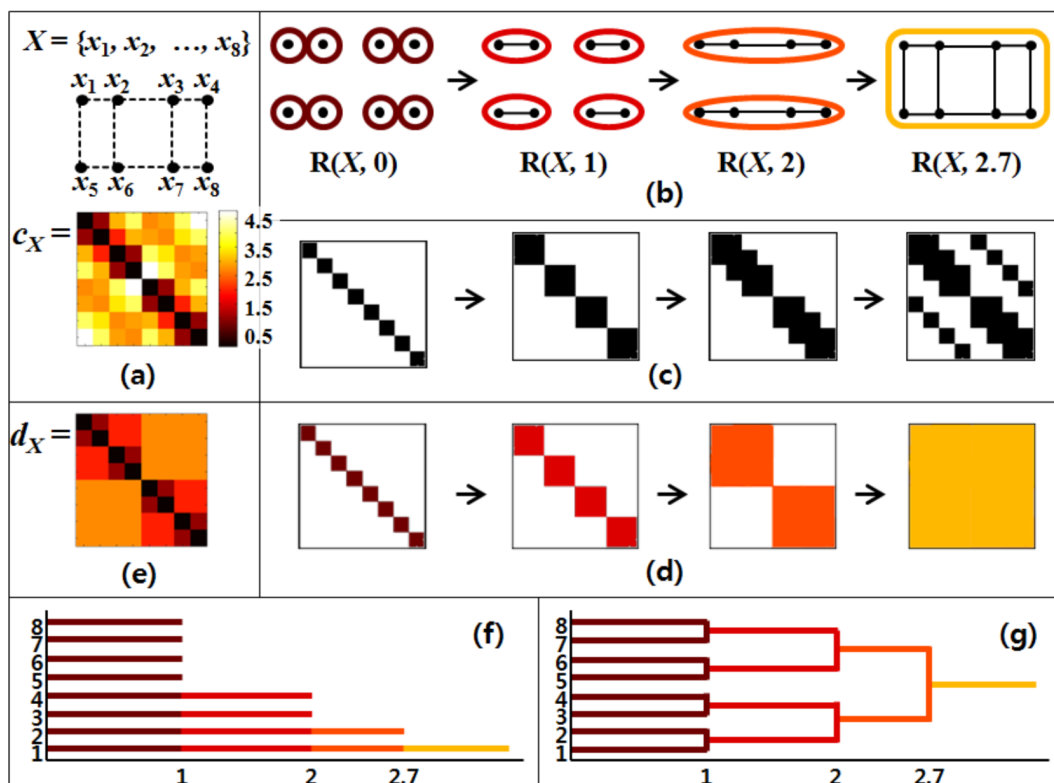


Figura 5: Comparación entre el barcode y el single linkage dendrogram de la red X representada en la esquina superior izquierda con su matriz de la métrica, en (a). En (b) los pasos de la filtración de Vietoris-Rips y en (c) la matriz de adyacencia del grafo en esos pasos. En (d) las matrices de componentes conexas y en (e) la matriz de la Single Linkage Distance. En las dos últimas figuras (f) y (g) se visualizan los cambios en la red a través del barcode de dimensión 0 y el Single Linkage Dendrogram, evidenciando su equivalencia. Extraído de [18].

2. Redes Complejas

No es poco frecuente que pensemos en un amasijo de neuronas o en un mapa de carreteras cuando tenemos que dar forma al concepto de Red Compleja. Tan cierta como la primera intuición es que la cantidad de sistemas que pueden ser modelados como interacción de distintos nodos o vértices entre sí es innombrable, desde modelos de contagio hasta colaboraciones entre investigadores. Las redes complejas no dejan de ser grafos, conjuntos de nodos que establecen relaciones binarias, haciendo hincapié en que modelan sistemas complejos, con un número de vértices suficientemente grande y una conectividad suficientemente rica como para empezar a revelar propiedades que no serían predictibles a pequeña escala.

A continuación se presentan una serie de conceptos básicos de teoría de grafos que se utilizarán en el apartado siguiente, así como algunas medidas usuales de teoría de redes clásica que nos darán un punto de comparación cuando hagamos el estudio con los nuevos métodos que nos ofrece la topología algebraica. Nótese la similitud entre el estudio típico de las redes complejas con la Física Estadística, debido a partes iguales a la descripción estadística que introdujeron Erdős y Renyi de los grafos aleatorios y a que han sido sistemas físicos complejos los que han motivado el desarrollo de esta materia. Por el carácter introductorio de este apartado, se exponen aquí tan sólo los conceptos que van a ser de utilidad en este trabajo.

2.1. Conceptos básicos de la teoría de grafos

Definición 2.1. *Un grafo es un par $G = (V, E)$ donde V y E son conjuntos y donde $E \subseteq V \times V$, i.e. los elementos de E son pares de elementos de V . Llamamos nodos o vértices a los elementos de V , mientras que los elementos de E son aristas o enlaces. $V(G)$ y $E(G)$ son el conjunto de nodos y aristas del grafo G , correspondientemente. El orden de G es el número de nodos con que cuenta y se denota $|G|$. Dos nodos son vecinos o adyacentes si existe una arista entre ellos. Llamamos grafo simple a aquél que no tiene aristas que conectan nodos con sí mismos.*

Definición 2.2. *Un subgrafo de G es un grafo cuyos vértices y aristas pertenecen a G .*

Observación 2.3. Vale la pena fijarse en el carácter binario de las relaciones en un grafo. Este será uno de los principales cambios al estudiar redes a través de la homología persistente.

Definición 2.4. *Decimos que un grafo es dirigido (también digrafo) si hay direccionalidad en las aristas, es decir, que tienen un nodo inicial y uno final.*

Definición 2.5. *El grado de un nodo es el total de vecinos que tiene. En grafos dirigidos, hay un grado de entrada y uno de salida. Si un grafo G de orden n cumple que todos los nodos son adyacentes entre ellos (todos tienen grado $n - 1$), decimos que G es un grafo completo que denotaremos K_n .*

Observación 2.6. Dado un grafo G , un subgrafo completo de G se denomina k -clique, donde k es el número de vértices del subgrafo.

Hasta ahora no hemos asociado a los enlaces ningún tipo de métrica o noción de cercanía, dos nodos están conectados o no lo están. Definimos los grafos ponderados para describir situaciones en que las conexiones no son necesariamente equivalentes y determinar una distancia.

Definición 2.7. Un grafo ponderado es un grafo simple $G = (V, E)$ con una aplicación $\omega : E(G) \rightarrow \mathbb{R}_+$ que se denomina peso

Definición 2.8. La matriz de adyacencia de un grafo no ponderado G de orden n es una matriz $n \times n$ donde a_{ij} es el número de aristas que unen v_i con v_j . Para grafos ponderados, a_{ij} toma el valor de la suma de pesos de las aristas entre v_i y v_j .

Distancia entre redes

Es importante darse cuenta de que el peso de una red ponderada no constituye a priori una distancia. No satisface necesariamente la desigualdad triangular, tampoco ninguna condición de simetría (es evidente en redes dirigidas) o siquiera que la distancia a un elemento sea nula únicamente cuando es a sí mismo, puesto que los pesos pueden ser también negativos. Por consiguiente no puede ser usado para construir ninguna correlación entre puntos de dos redes distintas. Para este propósito se construye una distancia de red.

Definición 2.9. Sean (X, ω_X) y (Y, ω_Y) dos redes ponderadas, y R una relación entre X y Y . La distorsión de la relación R es:

$$dis(R) := \max_{(x,y), (x',y') \in R} |\omega_X(x, x') - \omega_Y(y, y')| \quad (2.1)$$

Definición 2.10. Sea \mathcal{N} el conjunto de redes ponderadas y $X, Y \in \mathcal{N}$, se define la distancia de red, $d_{\mathcal{N}} : \mathcal{N} \times \mathcal{N} \rightarrow \mathbb{R}_+$:

$$d_{\mathcal{N}}(X, Y) := \frac{1}{2} \min_{R \in \mathcal{R}} dis(R) \quad (2.2)$$

donde \mathcal{R} son las correspondencias entre X e Y , es decir, el conjunto de relaciones tales que su imagen por la proyección a X o a Y son el espacio completo correspondiente.

Se puede comprobar que $d_{\mathcal{N}}$ es una seudométrica, y que las redes que se encuentran a distancia nula son totalmente caracterizables [26], por lo que esta distancia ofrece una herramienta de comparación entre redes distintas.

2.2. Propiedades de una red

En esta sección se describen las mediciones más usuales que se llevan a cabo en una red para caracterizarla, de carácter tanto local como global. De esta manera podremos tener una referencia en cuanto introduzcamos los métodos de la topología algebraica. Todas estas medidas son computables y existen herramientas de fácil acceso para hacerlo (la *Brain Connectivity Toolbox* está disponible para MATLAB y Python en [25]).

Characteristic path length

Esta longitud característica es una medida de la eficiencia en la comunicación de información en la red,

$$l_G = \frac{1}{n(n-1)} \sum_{i \neq j} d(v_i, v_j), \quad (2.3)$$

donde d es la distancia definida en la red.

Centralidad

Las medidas de centralidad son medidas locales que buscan determinar la importancia de un nodo dentro de la red. Una medida de centralidad básica es el grado del nodo. A continuación se exponen algunas alternativas más precisas.

Definición 2.11. *Closeness centrality.* Dado un grafo de orden N ,

$$C_i^{cl} = \frac{N}{\sum_{i \neq j} d(v_i, v_j)}. \quad (2.4)$$

Es el inverso de la suma normalizada de las distancias a todos los otros nodos. Cuanta menos distancia tenga hasta los otros nodos, más central será.

Definición 2.12. *Betweenness centrality.*

$$C_i^b = \sum_{j \neq i, k \neq i} \frac{\sigma_{jk}(i)}{\sigma_{jk}}, \quad (2.5)$$

donde σ_{jk} es la cantidad de caminos más cortos entre v_j y v_k y $\sigma_{jk}(i)$ son tan solo aquellos que pasan por v_i .

Se basa en la cantidad de veces que aparece el nodo en cuestión en los caminos más cortos entre cualesquiera otros dos nodos.

Definición 2.13. *Eigenvector centrality.* Dado un grafo G con matriz de adyacencia A , y sea $N(v)$ el conjunto de vecinos de v ,

$$C_i^{eig} = \frac{1}{\lambda} \sum_{v_j \in N(v_i)} C_j^{eig}, \quad (2.6)$$

donde λ es una constante. Equivalentemente, se puede reformular y considerar el vector

Queda claro que es el vector propio de la matriz de adyacencia con el valor propio λ . Se puede demostrar que existe un vector propio del valor propio máximo que tan sólo tiene componentes positivas (teorema de Perron–Frobenius), y se escoge ese mismo para dar valor a los nodos. Para computar la eigenvector centrality (este vector de valor propio maximal) se suele partir de un vector $C_0^{eig} = (1, \dots, 1)$ de mismo valor para todos los nodos y se hace iteraciones de $AC_k^{eig} = A_{k+1}$ para k creciente hasta que los valores se estabilizan. Hay otras maneras de encontrar este vector propio.

El valor que se asigna a un nodo depende, pues, del valor de otros nodos. Un nodo es importante en tanto que está conectado a otros nodos importantes. La *Katz centrality* o *PageRank* -algoritmo de Google para asignar la relevancia de los resultados de búsqueda son variantes de este tipo de centralidad.

Por último, un parámetro que parece guardar mucha información sobre la comunicabilidad en grafos no ponderados es el índice de Estrada. Como intuición, si $(A^n)_{ij}$ cuenta el número de caminos de longitud n entre los nodos i y j , y se valora como más importantes los caminos cortos llegamos a la siguiente definición.

Definición 2.14. *Subgraph centrality.* Dado un grafo G con matriz de adyacencia A ,

$$C_i^{sg} = \left(I + A + \frac{A^2}{2!} + \dots + \frac{A^k}{k!} + \dots \right)_{ii} = (exp(A))_{ii} \quad (2.7)$$

y también definimos la comunicabilidad entre los nodos i y j como $(exp(A))_{ij}$. Se define el índice de Estrada:

$$EE(G) = \sum_{i=1}^n (exp(A))_{ii} \quad (2.8)$$

Clustering

El coeficiente de *clustering* está relacionado con la tendencia de la red a crear triángulos, y tiene una versión global y una local. A continuación se dan las versiones para grafos no dirigidos y no ponderados. Se puede encontrar extensiones razonables de la definición en [27] y [28].

Definición 2.15. *El coeficiente de clustering global está definido como:*

$$C = \frac{3 \times \text{n}^\circ \text{ triángulos}}{\text{n}^\circ \text{ total de tripletes posibles}} \quad (2.9)$$

siendo un triplete una combinación cualquiera de tres vértices conectados todos entre sí.

Definición 2.16. *El coeficiente de clustering local está definido como:*

$$C_i = \frac{\text{n}^\circ \text{ triángulos con vértice en } i}{\text{tripletes posibles con vértice en } i} \quad (2.10)$$

A partir de los coeficientes locales se calcula también una medida de clustering global. Dado un grafo con n nodos:

$$\bar{C} = \frac{1}{n} \sum_{i=1}^n C_i \quad (2.11)$$

Modularidad

La modularidad Q intenta dar una intuición del grado de organización de una red en pequeñas comunidades. El valor Q que se obtiene es el resultado de optimizar la subdivisión de nuestra red en módulos. Los detalles de la expresión analítica y algoritmos para el cálculo no son relevantes para el trabajo y se pueden encontrar en [29]. Es suficiente con dar una idea intuitiva de lo que expresa: $Q \in [0, 1]$, $Q = 0$ para un grafo sin ningún tipo de preferencia de conexión y es mayor cuanto más internamente conectados estén pequeños grupos de nodos y cuanto más desconectados estén entre sí.

2.3. Modelos de redes

La perspectiva de estudio de los grafos dio un giro hacia 1960, cuando Paul Erdős y Alfred Rényi desarrollaron una descripción probabilística de las redes. Con ella se tratan colectividades estadísticas donde a cada grafo se le asigna una cierta probabilidad. Esto permite extraer resultados de existencia en los que no queda ninguna traza de aleatoriedad. En otros casos, se extraen resultados válidos para casi todo grafo (a.e), es decir, que el espacio de grafos que no lo cumplen es de medida nula. Aunque no se entrará a hablar sobre teoría de grafos aleatorios, es a partir de ella que se caracterizan los diversos tipos de redes que se presentan a continuación, conjuntos de grafos con una distribución probabilística asociada que determina el tipo de conectividad.

Modelo de Erdős-Rényi

Definición 2.17. Un grafo de Erdős-Rényi es un grafo $G(N, p)$ de N nodos en que cada posible enlace existe con una misma probabilidad p .

Observación 2.18. Es inmediato ver que el número de enlaces de un grafo $G(N, p)$ es una variable aleatoria de valor esperado $M = \binom{N}{2}p$.

Proposición 2.19. Sea $G(N, p)$ un grafo aleatorio. La probabilidad de que un nodo cualquiera tenga grado k viene dada por la distribución binomial, y se convierte por consiguiente en una distribución de Poisson en el límite.

$$P(k) = \binom{N-1}{k} p^k (1-p)^{N-1-k}. \quad (2.12)$$

Demostración. Fijado un vértice, existen $\binom{N-1}{k}$ posibilidades para escoger k vértices de los $N-1$ restantes. La probabilidad de que tengan enlaces es p^k y por tanto la probabilidad que el nodo conecte con k otros es $\binom{N-1}{k} p^k$. Además queremos que no conecte con ninguno de los $N-k-1$ restantes, evento que ocurrirá con probabilidad $(1-p)^{N-k-1}$. Haciendo el producto de probabilidades se obtiene el resultado. \square

Las redes de Erdős-Rényi tienen muy bajos coeficientes de clustering, pues dados tres nodos y dos aristas entre ellos, la tercera no tiene particular tendencia a trazarse, no formando así triángulos. Esto, junto con la distribución de grados, son las mayores diferencias que hay con modelos de redes reales.

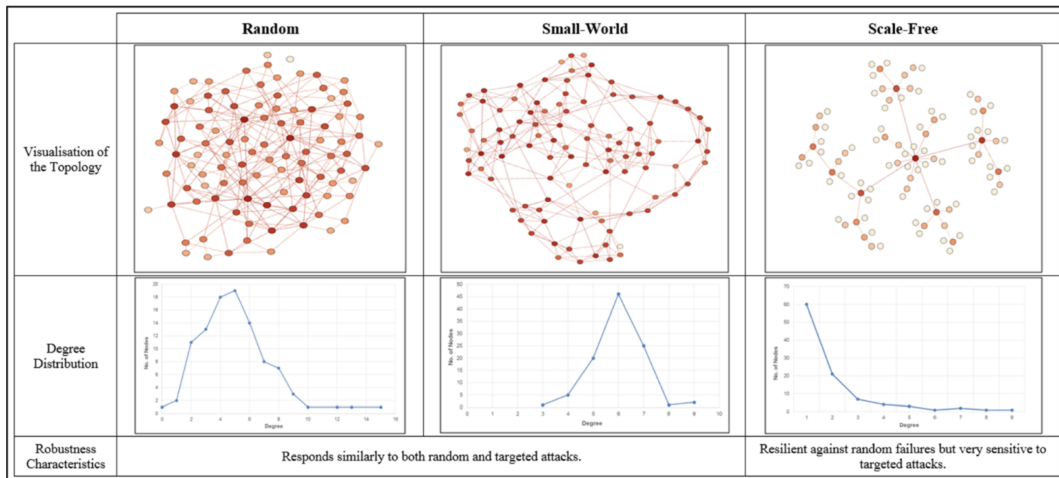


Figura 6: Tres redes de los tipos aquí descritos con su correspondiente distribución de grados. Extraído de [61].

Redes libres de escala

Una de las cosas más sencillas de detectar en una red real es una aproximación de la distribución de grados de sus nodos. A través de medidas directas sobre redes reales se ha visto que el modelo de Erdős-Rényi no es adecuado, y que estas exhiben más bien un comportamiento del tipo

$$P(k) = ck^{-\gamma} \quad k = m, \dots, K, \quad (2.13)$$

donde m y K son las cotas mínima y máxima del grado de los nodos y $c \approx (\gamma - 1)m^{\gamma-1}$ es una constante de normalización.

Existen diversos algoritmos para generar redes libres de escala y modelar su evolución. El primero fue el modelo de Barabási-Albert, basado en el *prefferential attachment*: en cada adición de un nodo a la red, se impone una probabilidad mayor de conexión con aquellos nodos que tengan grado más alto (*rich get richer*). Este modelo de red no consigue tener valores de clustering elevados, pero sí se aproxima más a la distribución de grados de una red real. Encontramos aquí un número muy reducido de neuronas que acumulan muchas conexiones, los *hubs*, mientras que la mayor parte de la red tiene conectividad baja.

Redes small-world o de Watts-Strogatz

Muchas redes complejas extraídas de interacciones reales cuentan con la propiedad de mundo pequeño o *small world*. Esto quiere decir que, aún habiendo un número enorme de nodos, con pocos pasos podemos trazar un camino entre cualesquiera de ellos. Ejemplos de esto es el popular resultado de que estamos conectados con cualquier ser humano del planeta con una media de 6 pasos intermedios.

Watts i Strogatz crearon un modelo de red que simulase este tipo de comportamiento y el primero que consigue simular unos valores de clustering más altos y parecidos a los encontrados en redes reales. La longitud característica es corta debido a la estructura *small world* y su distribución de grados es relativamente homogénea, perdiendo el realismo que ganaban las redes libres de escala en ese aspecto.

3. Homología persistente de redes complejas

Se ha introducido hasta ahora el objeto de estudio –las redes complejas– y la herramienta de estudio –la homología simplicial y persistente. Queda por consiguiente unir ambas para comprobar si el nuevo enfoque puede aportar luz al desarrollo de cualquiera de las dos ramas, ya sea desde la necesidad de distintas herramientas o desde el descubrimiento de nuevos patrones. Centrarse en los invariantes topológicos es una perspectiva fundamentalmente distinta que se complementa con la de la teoría de redes. La imagen siguiente consigue ejemplificar esto: mientras que los invariantes topológicos identificarán la fila de redes anulares como parecidas, las medidas más clásicas lo harán con las dos columnas, fijándose en propiedades más locales como el grado de los nodos o la longitud característica.

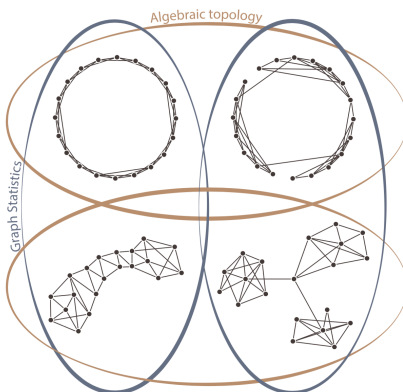


Figura 7: Mientras la teoría de grafos encontraría similitud entre las dos columnas de la imagen por tener misma secuencia de grados, la topología algebraica identificará las dos redes superiores como distintas de la fila inferior por su estructura anular. Extraído de [40].

Explicitemos qué separa el estudio de una red con el de una nube de puntos. En el caso de una nube de puntos, la construcción del complejo simplicial con el que se computa la homología persistente se basa únicamente en la distribución métrica de los puntos en el espacio en que queramos visualizarlos. Para una red es necesario tener en cuenta que hay ya una cierta métrica definida por las aristas y sus pesos, y que el complejo que se construya debe retener toda esa información en la medida que sea posible.

A continuación se divide esta sección en dos. Primero se estudia cómo es posible tal construcción de un complejo simplicial asociado a una red y, con la misma importancia, qué filtraciones se puede definir sobre el complejo que proporcionen información interesante. Un segundo punto se centrará en la interpretación de los resultados obtenidos al calcular las sucesiones de homología y en qué tipo de análisis puede ser útil. Se ha intentado exhaustivamente recoger y organizar las conclusiones que se han podido (o no) extraer del estudio de redes a través de la homología persistente.

3.1. Construcción de complejos simpliciales y filtraciones en una red

En teoría de redes el acercamiento más común para detectar características a una cierta escala es escoger un umbral de peso y estudiar el subgrafo que incluya únicamente las arista de peso menor. Algunas de las propiedades de la red inicial se pierden al fijar una

escala de esta manera, y nace la necesidad de poder hacer un estudio que pueda englobar todas las relaciones presentes en la red, pero que no pierda una visión a pequeña escala. Aquí entra en juego la homología persistente, que requerirá antes de una construcción simplicial sobre la red.

3.1.1. Complejos simpliciales

A continuación se presentan las construcciones de complejos simpliciales más comunes a partir de una nube de puntos. La información contenida en la red será entonces incorporada de diversas maneras. Posteriormente se proponen otras construcciones ya nacidas para ser aplicadas a redes.

Čech Complex

Para definir este complejo se introduce antes el concepto de *nervio* de un recubrimiento y un resultado esencial relacionado, que garantiza la similitud homotópica del complejo a construir con el espacio subyacente.

Definición 3.1. *Sea X un espacio topológico y $\mathcal{U} = \{U_i\}_{i \in I}$ un recubrimiento de X . El nervio del recubrimiento, denotado $N\mathcal{U}$, es el complejo simplicial que tiene I como conjunto de vértices y donde $\{i_0, \dots, i_p\}$ constituye una cara de $N\mathcal{U}$ si $U_{i_0} \cap \dots \cap U_{i_p} \neq \emptyset$.*

A continuación el teorema que asegura la similitud del complejo simplicial con el espacio subyacente. No se demostrará.

Teorema 3.2. *Nerve Theorem. Sea X un espacio topológico y U un recubrimiento numerable por abiertos de X . Si para todo $J \subseteq I$ no vacío se tiene que $\bigcap_{j \in J} U_j$ es contractible o vacío, entonces $N\mathcal{U}$ es topológicamente equivalente a X .*

Ahora, si el espacio a tratar es métrico, un recubrimiento sencillo viene dado por las bolas centradas en cada punto de un cierto radio. Generalizando este concepto llegamos a la siguiente definición.

Definición 3.3. *Sea X un espacio métrico. Sea $V \subseteq X$ un subconjunto para el cual $X = \bigcup_{v \in V} B_\epsilon(v)$, para $\epsilon > 0$, y $\mathcal{V} = \{B_\epsilon(v)\}_{v \in V}$ el consiguiente recubrimiento de X . Denominamos complejo de Čech $\check{C}(V, \epsilon)$ al nervio de tal recubrimiento, $N\mathcal{V}$.*

Para un grafo o red se puede adaptar el complejo de Čech para que mantenga el peso de las conexiones cambiando la métrica del espacio que tratábamos por el peso definido en la red, que es una distancia para grafos simples.

Vietoris-Rips Complex

Esta es tal vez la manera más intuitiva de generar un complejo simplicial a partir de una nube de puntos.

Definición 3.4. *Sea (X, d) un espacio métrico. El complejo de Vietoris-Rips de X con parámetro $\epsilon > 0$ es el complejo simplicial que tiene por vértices los puntos de X y donde $\{x_0, \dots, x_k\}$ constituye un k -simplex si, y sólo si $d(x_i, x_j) \leq \epsilon$, para todo $0 \leq i, j \leq k$. Lo denotamos $VR(X, \epsilon)$.*

Propone [26] extender el complejo a una red ponderada (X, ω_X) de la siguiente forma:

$$VR((X, \omega_X), \epsilon) = \{\sigma \in \mathcal{P}(\mathcal{V}(X)) \mid \omega_X(x_i, x_j) \leq \epsilon, \forall x_i, x_j \in \sigma\} \quad (3.1)$$

que extiende naturalmente el complejo a redes ponderadas. Para asegurar la similitud homotópica de este complejo respecto del espacio subyacente nos fijamos en su relación con el complejo de Čech.

Proposición 3.5. *Dado un espacio métrico X y $\epsilon > 0$, se tiene las inclusiones*

$$\check{C}(X, \epsilon) \subseteq VR(X, 2\epsilon) \subseteq \check{C}(X, 2\epsilon)$$

Demostración. Tan sólo hace falta fijarse que en el caso del complejo de Čech el parámetro ϵ corresponde al radio de las bolas, mientras que en el de Vietoris-Rips es la distancia entre los centros de las bolas, que quedará siempre entre las longitudes de uno y dos radios. \square

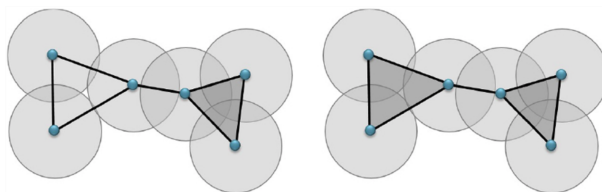


Figura 8: Comparativa entre los complejos simpliciales de Čech (izquierda) y de Vietoris-Rips (derecha). Vale la pena destacar que el parámetro para el de Čech es el radio de la bola, mientras que en Vietoris-Rips es la distancia entre los centros de dos bolas, facilitando la visualización de la Proposición 3.5. Extraído de [37].

Delaunay Complex

Este complejo aparece como simplificación computacional del complejo de Vietoris-Rips y se basa en el uso de un conjunto de puntos de referencia. Introducimos el recubrimiento por celdas de Voronoi.

Definición 3.6. *Sea X un espacio métrico y $\mathcal{L} \subseteq X$ un subconjunto, que llamamos de puntos de referencia. La celda de Voronoi asociada a un punto de referencia λ es*

$$V_\lambda = \{x \in X \mid d(x, \lambda) \leq d(x, \lambda'), \forall \lambda' \in \mathcal{L}\}. \quad (3.2)$$

Observación 3.7. El conjunto de celdas de Voronoi de un subconjunto $\mathcal{L} \subseteq X$ forma un recubrimiento de X

Definición 3.8. *Sea X un espacio métrico y $\mathcal{L} \subseteq X$ un subconjunto. Definimos el complejo de Delaunay como el nervio del recubrimiento de X por celdas de Voronoi con puntos de referencia en \mathcal{L} .*

La principal ventaja que ofrece este complejos es que genera símplex de dimensiones bajas, tantas como celdas de Voronoi colindantes tenga cada una de ellas. Esto facilita la computación y convierte el complejo de Delaunay en una herramienta habitual en la

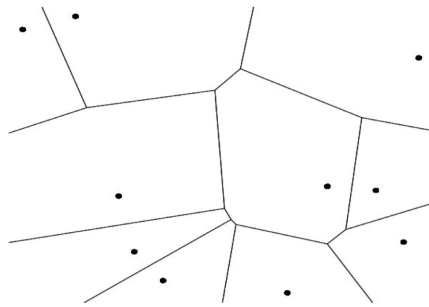


Figura 9: Celdas de Voronoi alrededor de unos puntos de referencia. Extraído de [62].

geometría computacional. Para espacios finitos se puede modificar la construcción para que sea más usual que dos puntos equidisten de puntos de referencia y obtener así más simplex de más dimensiones, introduciendo un margen ϵ para producir un cierto solapamiento de regiones. El complejo resultante es el *witness complex*. No existen referencias que adapten esta construcción para redes, y por consiguiente no se formaliza la definición del witness complex. Se puede encontrar en [3] junto con alguna otra variante similar, como el *alpha complex*.

A continuación los complejos simpliciales que parten de un grafo.

Clique Complex

La manera más natural de generar un complejo simplicial en una red es a través de los *cliques*. Es fácil ver la relación entre un k -clique y un $(k - 1)$ -simplex. Para una red dada podemos crear su complejo simplicial asociado de la siguiente manera.

Definición 3.9. Sea G un grafo. El *clique complex* o *flag complex* de G , $\mathcal{C}(G)$, es el complejo simplicial con los vértices de G resultante de asociar un k -simplex a cada $(k+1)$ -clique de G .

Codifica, por consiguiente, la misma información que el grafo en cuestión completando el esqueleto de la red con la estructura simplicial más llena posible.

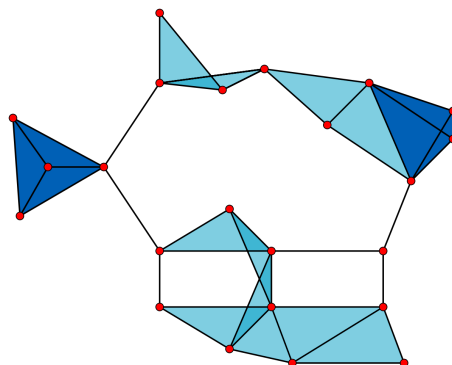


Figura 10: Clique complex de un grafo. Extraído de [63].

Neighborhood Complex

Definición 3.10. Dado un grafo G , el *neighborhood complex* de G , $\mathcal{N}(G)$, es el complejo simplicial que tiene como caras los subconjuntos de vértices de G con un vecino común.

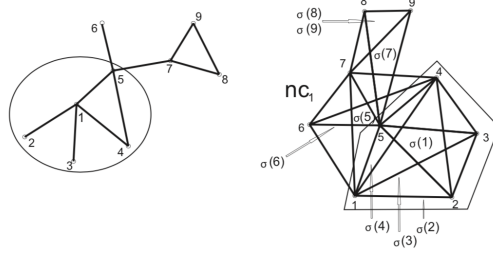


Figura 11: Neighborhood Complex de la red de la izquierda. Los simplex $\sigma(i)$ corresponden al vértice i del grafo. Extraído de [38].

Es conveniente recordar que la construcción de complejos simpliciales no se reduce a la adaptación de una red, sino que podría ser una alternativa al estudio de cualquier sistema. En [30] se exponen diversos complejos simpliciales que han permitido sacar conclusiones que quedaban fuera del alcance del modelado de relaciones binarias de una red, en el campo de la neurociencia.

3.1.2. Filtraciones

Contamos con maneras para construir complejos simpliciales, pero para computar la homología persistente requerimos de una filtración. La intención del estudio que se quiera llevar a cabo será esencial para definir la filtración y tal elección cambiará los resultados obtenidos y su interpretación, pues el orden en que van a irse incorporando los simplex generará unos o otros ciclos a lo largo del proceso. Algunas filtraciones no están pensadas para hacerse sobre grafos pesados o dirigidos; se incluye en el Anexo 1 una tabla extraída de [31] que resume la validez de las filtraciones. Las primeras que se presentan son consecuencia inmediata del proceso de construcción del complejo simplicial.

Filtración de Vietoris–Rips

Dado un espacio métrico X , esta filtración es consecuencia de considerar complejos de Vietoris–Rips para valores crecientes del parámetro.

$$\{VR(X, \epsilon) \hookrightarrow VR(X, \epsilon')\}_{\epsilon < \epsilon'}. \quad (3.3)$$

Esta filtración es estable a perturbaciones de la información inicial (de la nube de puntos o red). Veámoslo con el siguiente resultado:

Proposición 3.11. Sean (X, ω_X) , (Y, ω_Y) dos redes ponderadas y sean $Dgm_k^{\mathcal{R}}(X)$, $Dgm_k^{\mathcal{R}}(Y)$ para todo $k \geq 0$ los respectivos diagramas de persistencia correspondientes al complejo de Vietoris–Rips con su misma filtración. Entonces

$$d_B(Dgm_k^{\mathcal{R}}(X), Dgm_k^{\mathcal{R}}(Y)) \leq 2d_{\mathcal{N}}(X, Y). \quad (3.4)$$

Demostración. Se puede encontrar un análogo en la proposición 15 de [26], aplicada a la filtración sink and source de Dowker. Requiere de lemas previos que no se cubren en el trabajo y por consiguiente aquí se omite.

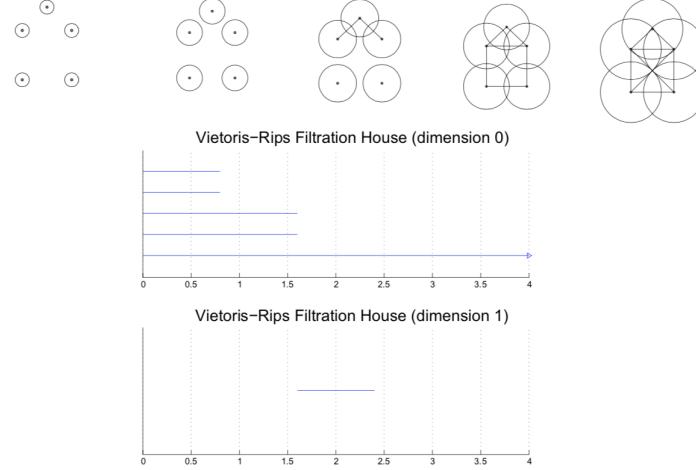


Figura 12: Filtración de Vietoris-Rips de una nube de puntos en forma de casa y barcodes resultantes. Extraído de [32].

Filtración de Čech

Análogamente al caso anterior, esta filtración es simplemente resultado de considerar complejos de Čech para valores crecientes del parámetro.

$$\{\check{C}(X, \epsilon) \hookrightarrow \check{C}(X, \epsilon')\}_{\epsilon < \epsilon'}. \quad (3.5)$$

Filtración por Clique Complex

Una manera de generar una filtración sencilla es partir del clique complex del grafo, $Cl(G)$, y introducir en cada paso de la filtración una dimensión mayor de cliques.

$$K_0 \subseteq K_1 \subseteq \dots \subseteq K_n = K$$

donde cada subcomplejo viene dado por $K_i = \sum_{j=1}^i S_j$, siendo S_j el esqueleto j -dimensional de $Cl(G)$ (el complejo simplicial formado por todos los j -simplex de G).

Via Teoría de Morse

Se introduce aquí toda un estudio sistemático que permite definir filtraciones en redes no dirigidas y no ponderadas en base a la teoría de Morse discreta, partiendo del clique complex del grafo inicial. Debe recordarse que el orden en que se decide incorporar los simplex en la filtración determinará las características topológicas del complejo durante su evolución, y que debe por consiguiente ser escogido intencionadamente para el propósito deseado.

Definición 3.12. Sea una función $f : K \rightarrow \mathbb{R}$. Para cada p -simplex $\alpha^p \in K$ se definen los conjuntos

$$U_\alpha^f = \{\beta^{p+1} \mid \alpha^p < \beta^{p+1}, f(\beta) \leq f(\alpha)\}, \quad (3.6)$$

$$V_\alpha^f = \{\gamma^{p-1} \mid \gamma^{p-1} < \alpha^p, f(\alpha) \leq f(\gamma)\}, \quad (3.7)$$

es decir, U_α^f contiene $(p+1)$ -simplex que contienen a α^p como cara y cuyo valor por f es más pequeño o igual al de $f(\alpha)$. Análogamente para $(p-1)$ -simplex contenidos en α^p que tengan un valor superior o igual a α cuando son evaluadas por f .

Definición 3.13. Dado un complejo simplicial K , una función $f : K \rightarrow \mathbb{R}$ es una función de Morse discreta si para cada $\alpha^p \in K$ se cumple:

$$\#U_\alpha^f \leq 1 \quad \text{y también} \quad \#U_\alpha^f \leq 1. \quad (3.8)$$

Dada una función de Morse discreta sobre K , decimos que un simplex $\alpha^p \in K$ es crítico si

$$\#U_\alpha^f = 0 \quad \text{y también} \quad \#U_\alpha^f = 0. \quad (3.9)$$

Observación 3.14. Un p -simplex α^p es crítico si para todo $(p+1)$ -simplex β que lo contenga como cara se cumple $f(\alpha) < f(\beta)$ y si para todas sus caras $(p-1)$ -dimensionales γ se cumple $f(\alpha) > f(\gamma)$.

Ejemplo 3.15. Una posible función de Morse es la función dimensión, que da la dimensión del simplex evaluado. Para esta función es fácil ver que todo simplex del complejo sería crítico.

Construyamos ahora una filtración a partir de los conceptos que se acaban de definir. Sea $r > 0$ un número real, definimos un subcomplejo de nivel $K(r)$:

$$K(r) = \bigcup_{f(\beta) \leq r} \bigcup_{\alpha \leq \beta} \alpha, \quad (3.10)$$

es decir, el complejo que junta los simplex que toman un valor por f menor o igual a r con todas sus caras. Dado que no todos los simplex tienen por qué ser críticos, en $K(r)$ puede haber algunas caras α con pesos mayores a r .

Como queremos un complejo simplicial finito, sea $\{f(\sigma)\}_{\sigma \in K}$ el conjunto de valores asignados a los simplex de K y sea $\{w_k\}_{k=0, \dots, n}$ la secuencia de valores únicos en el conjunto anterior, que es finita por ser K un complejo simplicial finito. Se tiene ahora la secuencia de inclusiones de subcomplejos de nivel siguiente:

$$\emptyset \subseteq K(w_0) \subseteq K(w_1) \subseteq \dots \subseteq K(w_{n-1}) \subseteq K(w_n) = K \quad (3.11)$$

Esta secuencia da una filtración del complejo simplicial K a partir de la que ya se puede estudiar la homología persistente. Aún así, la cantidad de valores distintos que toman los simplex de K suele ser inmensa, por lo que conviene reducirlos. El siguiente lema adaptado de Forman en [33], permite esto mismo.

Lema 3.16. Sea toda la construcción de la filtración expuesta hasta ahora. Si no existen simplex críticos α tales que $f(\alpha) \in (a, b]$, entonces $K(b)$ es homotópicamente equivalente a $K(a)$

Como la homología es invariante bajo equivalencia homotópica, podemos reducir el conjunto de valores en que dividir la filtración a los resultantes de evaluar tan sólo los simplex críticos, pues no se pierde información eliminando todos los otros. Esto resulta en una subsecuencia $\{w_{c_k}\}_{k=0,\dots,m}$, donde $m \leq n$ y correspondiente subcomplejo de nivel que simplifican la computación.

Es esencial, por consiguiente, escoger una función de Morse discreta que vaya a aportarnos información útil según el orden en que vaya a ir incorporando los simplex.

Filtraciones Sink and Source de Dowker

Las filtraciones *Sink* y *Source* de Dowker fueron propuestas en [26] y se construyen alrededor de una autoridad central en redes dirigidas y ponderadas, lo que las hace potencialmente útiles para el análisis de redes que cuenten con *hubs*. Sea un grafo $G = (V, E)$ con pesos dados por $w : V \times V \rightarrow \mathbb{R}$, se define el complejo simplicial δ -*sink* de Dowker:

$$\mathcal{D}_\delta^{si} := \{\sigma = [x_0, \dots, x_n] \mid \exists x' \in V \text{ tal que } w(x_i, x') \leq \delta \forall x_i \in \sigma\} \quad (3.12)$$

Es decir, que existe un nodo al que le llegan enlaces con pesos menores a δ de cada uno de los simplex del complejo. De manera dual se define el complejo simplicial δ -*source* de Dowker:

$$\mathcal{D}_\delta^{so} := \{\sigma = [x_0, \dots, x_n] \mid \exists x' \in V \text{ tal que } w(x', x_i) \leq \delta \forall x_i \in \sigma\} \quad (3.13)$$

Es decir, que en este caso existe un nodo referencia que lanza aristas de peso menor o igual a δ a cada uno de los demás nodos del simplex. A partir de estos complejos simpliciales la construcción de la filtración viene dada por valores crecientes del parámetro

$$\{\mathcal{D}_\delta^{si} \hookrightarrow \mathcal{D}_{\delta'}^{si}\}_{0 \leq \delta \leq \delta'} \quad \{\mathcal{D}_\delta^{so} \hookrightarrow \mathcal{D}_{\delta'}^{so}\}_{0 \leq \delta \leq \delta'}. \quad (3.14)$$

Nótese que la filtración empezará incluyendo los simplex que incluyen los nodos más potentes de la red para ir incorporando después aquellos de conectividades más bajas o con menor peso. Los simplex de mayores dimensiones serán incorporados al principio.

Chowdhury y Mémoli, autores de [26], usan el teorema de Dowker (originalmente en [34]) para probar que de estas dos filtraciones resultan módulos de persistencia isomorfos, y deducen la equivalencia de sus respectivos diagramas de persistencia. Este resultado lo definen como *dualidad de Dowker*, explicitado a continuación:

Proposición 3.17. *Sea (X, w_X) una red ponderada y $k \in \mathbb{Z}_+$. Entonces $Dgm_k^{si}(X) = Dgm_k^{so}(X)$ y puede entonces denotarse como diagrama de persistencia de Dowker k -dimensional, $Dgm_k^{\mathcal{D}}(X)$.*

Demostración. Ver Corolario 20 de [26]. □

Esta filtración es estable. Análogamente a la Proposición 3.11, se tiene:

Proposición 3.18. *Sean $(X, \omega_X), (Y, \omega_Y)$ dos redes ponderadas y sean $Dgm_k^{\mathcal{R}}(X), Dgm_k^{\mathcal{R}}(Y)$ $\forall k \geq 0$ los respectivos diagramas de persistencia correspondientes al complejo de Vietoris-Rips con su misma filtración. Entonces*

$$d_B(Dgm_k^{\mathcal{D}}(X), Dgm_k^{\mathcal{D}}(Y)) \leq 2d_{\mathcal{N}}(X, Y). \quad (3.15)$$

Demostración. Ver Proposición 15 de [26]. □

Estos autores demuestran también que los complejos de Dowker son sensibles a asimetrías en la información inicial, sugiriendo que constituyen un invariante más fuerte que los diagramas de persistencia de Vietoris–Rips o de Čech para redes dirigidas.

Weight Rank Clique Filtration

De manera muy intuitiva, se parte del grafo inicial G y para cada valor de un parámetro δ se genera el subgrafo G_δ que contiene todos los vértices del inicial y sólo las aristas que tengan peso menor o igual a δ . El clique complex de cada uno de estos subgrafos constituye una filtración el complejo simplicial de la red al completo.

$$\{Cl(G_\delta) \hookrightarrow Cl(G_{\delta'})\}_{\delta < \delta'}. \quad (3.16)$$

Si hay una métrica subyacente, esta filtración no es otra que la de Vietoris–Rips. Permite hacer el estudio sobre grafos que no necesariamente vivan en un espacio métrico. También puede plantearse de la manera inversa: incorporar antes las aristas que pasen de una cierta cota inferior y ir reduciéndola hasta cero. El propósito será el que guíe cómo aplicar el método.

Filtración por Cliques basada en vértices

Esta filtración es válida para redes que sean ponderadas sobre los vértices, es decir, que cuenten con una función peso $w : V \rightarrow \mathbb{R}$. Puede adaptarse de un grafo ponderado considerando que el peso de un vértice es la suma de pesos de todas sus aristas entrantes o salientes, según el estudio que se quiera llevar a cabo. En cualquier caso, para todo $\delta \in \mathbb{R}$ se considera el subgrafo $G_\delta = (V_\delta, E_\delta)$ tal que

$$V_\delta := \{v \in V \mid w(v) \leq \delta\}, \quad (3.17)$$

$$E_\delta := \{e = \{u, v\} \in E \mid \max(w(u), w(v)) \leq \delta\}. \quad (3.18)$$

Se define la filtración considerando, igual que en el caso anterior, el clique complex de cada uno de los subgrafos.

$$\{Cl(G_\delta) \hookrightarrow Cl(G_{\delta'})\}_{0 \leq \delta \leq \delta'}. \quad (3.19)$$

También análogamente puede decidirse proceder a la inversa para incorporar primero los vértices más pesados.

Filtración por k -Cliques

Esta filtración se centra en el estudio de los k -cliques en un grafo determinado. Fijada una $k \in \mathbb{Z}_+$, en un grafo ponderado sobre los vértices. Se asigna peso a un símplex determinado:

$$w(\sigma) := \{\max w(v) \mid v \text{ es vértice de } \sigma\}. \quad (3.20)$$

Creamos ahora un grafo de conectividad entre k -cliques $G^k = (V^k, E^k)$, donde hay un vértice para cada k -clique de la red y donde las aristas están definidas así:

$$E^k := \{(\sigma, \sigma') \in V^k \times V^k \mid \sigma \text{ y } \sigma' \text{ son adyacentes}\}. \quad (3.21)$$

Se extiende la función peso a estas aristas: $w((\sigma, \sigma')) := \max(w(\sigma), w(\sigma'))$. Se aplica sobre G^k ahora la Vertex-Based Clique Filtration

$$\{Cl(G_\delta^k) \hookrightarrow Cl(G_{\delta'}^k)\}_{0 \leq \delta \leq \delta'}. \quad (3.22)$$

Puede interesar hacer una asignación distinta de pesos a los simplex de la red original.

En [31] se describe brevemente todas las filtraciones sobre redes que se han usado en artículos de toda naturaleza, incluyendo aún algunas que aquí no se mencionan.

3.2. Resultados y análisis

Se comenta en esta sección toda conclusión a la que ha permitido llegar el estudio de la homología persistente de complejos simpliciales filtrados extraídos inicialmente de redes.

3.2.1. Estudio de redes conocidas

Es de utilidad estudiar cómo es la homología persistente de redes que se conocen bien para ganar capacidad de interpretación.

Redes de Erdős-Rényi

Las conclusiones extraídas en este apartado provienen del estudio del clique complex de grafos aleatorios $G(n, p)$, que recordemos que coincide con el de Vietoris-Rips si contamos con una métrica. En [35] puede encontrarse un extenso trabajo centrado en esto.

A grandes rangos, lo que se observa es que el k -ésimo número de Betti es distinto de cero en un rango delimitado de valores de p , y que se anula con alta probabilidad para cualquier otro. La presencia de ciclos k -dimensionales se encuentra más adelante en el espectro de valores de p para valores crecientes de k , de manera que para $p = 0$ se empiece con $|H_0(G(n, p))| = n$. La siguiente imagen lo ejemplifica.

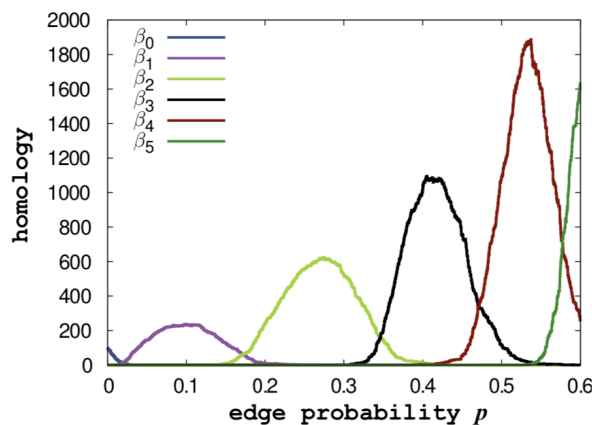


Figura 13: Números de Betti para un grafo aleatorio en función de p para una red con 100 vértices. Extraído de [35].

Se puede formalizar esto con tres resultados: el primero, corolario de un teorema de Erdős y Rényi, da un umbral brusco para el cual un grafo aleatorio es conexo; el segundo,

de Kahle [35] determina la desaparición de los demás grupos de homología para $k \geq 1$, y el tercero su aparición para una ventana de valores determinada. Demostraciones en [36].

Definición 3.19. *Se dice que un suceso A ocurre con alta probabilidad si*

$$\lim_{N \rightarrow \infty} P(G(N, p) \text{ cumple } A) = 1. \quad (3.23)$$

Proposición 3.20. *Sea $G(n, p)$ un grafo aleatorio y $w \rightarrow \infty$ cuando $n \rightarrow \infty$ a un ritmo arbitrario. Si*

$$p \geq \frac{\log n + w(n)}{n} \quad (3.24)$$

entonces G es conexo con alta probabilidad. Si por el contrario tenemos la desigualdad opuesta, G es no conexo con alta probabilidad.

Teorema 3.21. *Sea $G(n, p)$ un grafo aleatorio y sean $k \geq 1$, $\alpha > 0$ fijadas, $p = n^{-\alpha}$. Entonces:*

1. *Si $\alpha > \frac{1}{k}$, entonces $H_k(Cl(G)) = 0$ con alta probabilidad.*
2. *Si $\alpha < \frac{1}{2k+1}$ entonces $H_k(Cl(G)) = 0$ con alta probabilidad.*

Teorema 3.22. *Sea $G(n, p)$ un grafo aleatorio y sea $\alpha > 0$ fija, $p = n^{-\alpha}$. Si $\frac{1}{k+1} < \alpha < \frac{1}{k}$, entonces $H_k(Cl(G)) \neq 0$ con alta probabilidad.*

Horak et al. trabajan con simulaciones en [38] para llegar a resultados concordantes con lo que se acaba de presentar. El valor de p determina el aspecto que tendrán los diagramas en los últimos pasos de filtración, que deberían asemejarse al punto correspondiente de la Figura 14. La filtración se hace por Clique Complexes. Una única línea aparece para β_0 a partir del primer paso de la filtración, de acuerdo con la conexión del grafo, y aparecen entonces ciclos unidimensionales abundantes, sin ciclos de mayores dimensiones. Con un valor mayor de p se consigue tener grupos de homología no nulos de más dimensiones. Los dos ejemplos son ilustrativos.

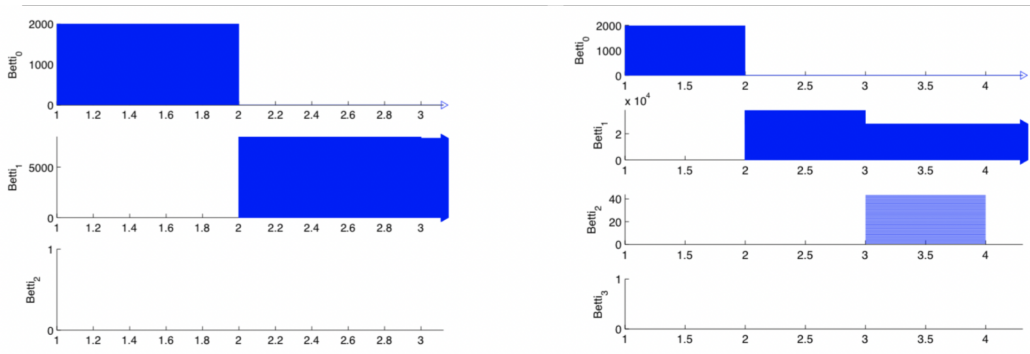


Figura 14: Barcodes de dos simulaciones de grafo aleatorio. A la izquierda, $n = 2000$ y $p = 0,005$. A la derecha, $n = 2000$ y $p = 0,02$. Los resultados coinciden con lo esperado por la Figura 13. Extraído de [38].

Redes libres de escala

Se usa aquí un modelo generalizado de preferential attachment que mejora el de Barabási-Albert para poder crear nuevos módulos a parte del inicial. Se controlan tres parámetros: la conectividad media, la probabilidad de emergencia de un nuevo módulo y la atractividad de un nodo. Jugando con estos valores se pueden crear redes libres de escalas con distintos grados de modularidad y de clustering. Siendo estas características que podrían cambiar los módulos de persistencia asociados, se estudian como casos separados.

En el caso de redes modulares y con altos valores de clustering se encuentran primer, segundo y tercer grupos de homología no nulos y que permanecen hasta los últimos pasos de filtración una vez aparecidos. La persistencia de estos ciclos es un claro signo de la robustez de la red ante la adición o eliminación de símplex. Si de ahí se pierde la modularidad aparece en el diagrama de persistencia el cuarto grupo de homología, con pocos generadores con poca persistencia. H_3 desaparece para los valores de estudio que se utilizaron. El caso modular pero con bajo clustering es el que cuenta con menos grupos de homología y menos generadores.

Observación 3.23. La persistencia del n -ésimo número de Betti implica que en el complejo hay una ausencia de un objeto n -dimensional. Esto se traduce a la red a través de la ausencia de relaciones n -ádicas: para $n = 2$ no habría relaciones diádicas; para $n = 3$ triádicas, etc.

Parece ser que la presencia de H_3 y H_4 sugiere que el clustering genera más robustez en la red que la modularidad, ya que la organización interna de la red en pequeñas comunidades es un prerequisite para la emergencia de subgrafos completos y, por consiguiente, de mayores k -agujeros.

3.2.2. Capacidad discriminativa de la homología persistente

Parece esencial ahora comprobar que la homología persistente constituye un buen método para caracterizar una red. En [39] se presentan cuatro modelos de red en forma de anillo y se calcula en ellas diversas medidas típicas de Teoría de Redes. Puede verse en la imagen inferior que ninguna de ellas capta un patrón común, debido a su carácter local. De manera evidente, la homología persistente descubriría su estructura global sin problemas al revelar un ciclo unidimensional muy duradero y que, por consiguiente, esconde una propiedad estructural importante de la red.

La incorporación de cliques a lo largo de la filtración ofrece también la posibilidad de no perder mucha información sobre el clustering, que parece ser una de las características que discriminan mejor las redes localmente. En conclusión, se tiene un método de identificación de redes de carácter más global.

Por primera vez planteado en [41], parece que el estudio de la persistencia homológica de redes muy distintas las divide en dos subgrupos: aquellas con estructura geométrica y aquellas que no la tienen (aleatorias). Se hizo un duplicado randomizado de cada red (i.e. redistribución de enlaces y pesos, conservando las secuencias de grados y pesos) y se llevó a cabo el estudio topológico sobre las redes y sus duplicados midiendo la Bottleneck distance entre los diagramas de persistencia obtenidos.

Redes de clase I: Las distribuciones de ciclos son marcadamente distintas que en sus

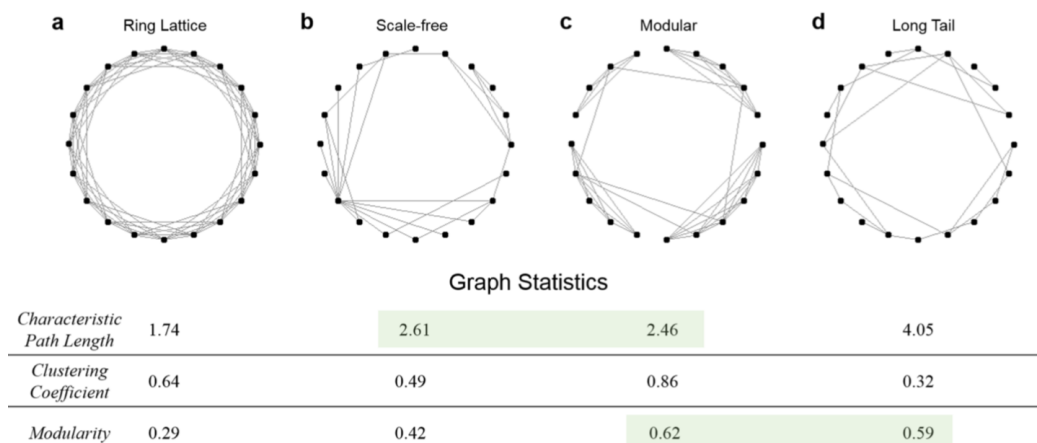


Figura 15: Resultados obtenidos de aplicar tres medidas clásicas de la Teoría de Redes. Ninguna de ellas capta la topología anular común de las cuatro redes. Extraído de [39].

versiones randomizadas: persistencia más corta y nacimientos más pronto y durante más tiempo, sugiriendo que se responde a la desigualdad triangular (ya que la probabilidad de generar el tercer enlace de un 2-símplex es mayor al existir los dos anteriores). La organización de la red consta de una estructura geométrica intrínseca.

Redes de clase II: Las distribuciones de ciclos son muy parecidas a las de sus versiones randomizadas. La organización de la red no parece responder a ninguna estructura geométrica subyacente

En [42] se aplica este concepto a la actividad neuronal registrada en el hipocampo de ratas y se comprueba su eficacia, siendo suficiente con las curvas de Betti para distinguir estructura aleatoria de geométrica. Además resulta ser invariante ante transformaciones no-lineales de la matriz de adyacencia y permite inferir sobre la organización geométrica de la red neuronal sin recurrir a estímulos externos.

Una posible manera de medir cuantitativamente estas diferencias es el parámetro de superficialidad de la red [41], *network hollowiness* y su derivado *chain-length normalized hollowiness*, definidos para cualquier $k \geq 0$:

$$h_k = \frac{1}{N_{g_k}} \sum_{g_k} \frac{p_{g_k}}{T}, \quad (3.25)$$

$$\bar{h}_k = \frac{1}{N_{g_k}} \sum_{g_k} \frac{\lambda_{g_k} p_{g_k}}{N T}, \quad (3.26)$$

donde $\{g_k\}$ es el conjunto de generadores de H_k y $N_{g_k} = \dim H_k$. El primero es una medida de la persistencia media, y el segundo tiene en cuenta los generadores según su persistencia y longitud. Las redes de clase I parecen presentar valores inferiores de hollowiness respecto sus copias aleatorizadas.

Otro parámetro que también parece ser distinto en las redes de estos dos tipos es el llamado *gap spectral*. Dado un grafo G y su matriz de adyacencia A , el gap espectral es la diferencia entre valores propios de A , que en caso que A sea simétrica será constante. Las redes de clase I parecen tener un gap mucho mayor que las de clase II [41]. El interés de esto viene por estar este parámetro relacionado con la conectividad de la red. Si definimos

la matriz laplaciana de G , $L(G) = D - A(G)$ donde $d_{ij} = \delta_{ij} \sum_k a_{ik}$, con valores propios $\lambda_1^L \leq \lambda_2^L \leq \dots \leq \lambda_N^L$, el ratio propio del Laplaciano es $R_L := \frac{\lambda_N^L}{\lambda_2^L}$. Dado un sistema dinámico definido en una red, se sabe que tiene un estado sincrónico linealmente estable si $R_L < \beta$, donde β es un parámetro puramente dinámico. Por tener un valor mayor de gap espectral, las redes de clase I parecen ser más difícilmente sincronizables. Esto abre una nueva puerta todavía por desarrollar: la conexión entre la estructura homológica de la red y sus propiedades dinámicas.

Sizemore [39] lleva algo más allá la caracterización de redes ampliando en dos el abanico. Distingue entre cuatro clases naturales a partir de su homología persistente: la primera redes muy estructuradas con un núcleo denso (e.g. Scale-Free), la segunda redes con clústers de tamaño moderado (e.g. Watts-Strogatz), la tercera grafos aleatorios sujetos a restricciones estructurales (e.g. redes aleatorias pero divididas en módulos) y por último los grafos puramente aleatorios (e.g. de Erdős-Rényi). En la Figura 16 hay un intuitivo mapa conceptual junto con los resultados del estudio persistente de las redes.

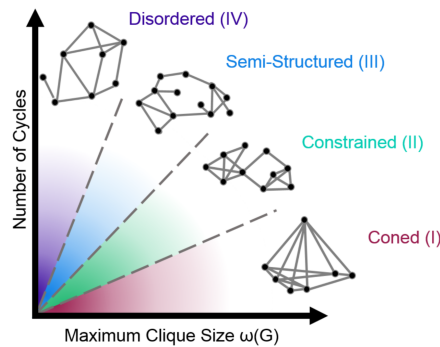


Figura 16: Los cuatro tipos de red diferenciados por Sizemore en función del número de ciclos y el tamaño de los cliques. Extraído de [39].

Sin buscar una caracterización, al contar con una distancia entre barcodes se puede usar los diagramas de persistencia para ver la similitud entre redes. En [43] se construyen las filtraciones a partir de la teoría de Morse discreta de redes aleatorias, scale-free (con el modelo de Barabási-Albert), de Watts-Strogatz y aleatorias sobre espacios esféricos y hiperbólicos. Basta un vistazo a la Figura 17 para ver que en efecto la homología persistente es suficiente para distinguir entre estas redes en la mayoría de casos.

Cabe recordar que se puede recurrir a la transformación de los barcodes en dendrogramas y utilizar entre ellos la distancia de Gromov-Hausdorff, que parece suponer una mejora en el rendimiento [18] [19] o bien utilizar indicatrices de persistencia para simplificar la computación [17], como se explica en la Sección 1.2.3.

3.2.3. Aplicaciones en la neurociencia aplicada

Por la inmediatez de la relación entre las redes complejas y el cerebro, he considerado interesante dedicar un pequeño apartado a recoger algunos de los estudios sobre redes neuronales más interesantes que se han llevado a cabo a través de la homología persistente, pretendiendo que se incentive la creatividad y se abran nuevas perspectivas de estudio. Por la variedad de artículos que han usado estas ideas no se entrará en detalle en cada uno de ellos, pero se anima a querer ver cómo la homología persistente toma su vertiente

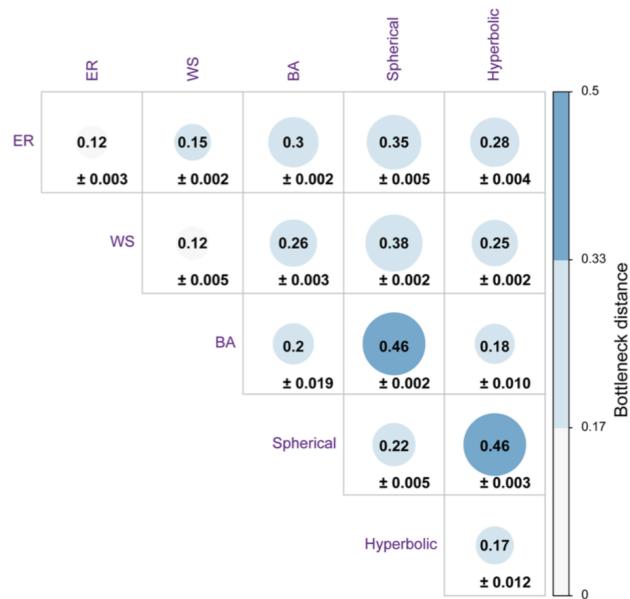


Figura 17: Valores de las bottleneck distances entre cinco tipos de red distintos. Son suficientes para poder discernir entre ellas. Extraído de [43].

más aplicada en la ciencia de las redes.

La primera incursión de estos conceptos en la neurociencia parece ser [44], donde se usa para discriminar las imágenes por TEP (*Tomografía por Emisión de Positrones*) de pacientes pediátricos control, con TDAH y con autismo. Nótese que son estos autores los que acaban aplicando la equivalencia de barcodes con dendrogramas que se ha expuesto en la Sección 1.2.4. En [42] se detecta la organización geométrica del hipocampo de ratas a partir de patrones de actividad, y en [45] se usa para determinar que en cerebros de ratón con depresión inducida la conectividad global decrece y es mucho más localizada. La misma idea pasando los barcodes a dendrogramas sirve en [46] para detectar actividad metabólica anormal en ratas con epilepsia. Ya sin partir de una red, pero para inferir sobre ellas, en [47] se usa la información dinámica de activación del hipocampo para reconstruir la estructura del campo visual de ratones a partir de un complejo simplicial (a través de un proceso desarrollado en [48]). En [49] se consigue detectar y caracterizar patrones de coactivación repetitivos y sincrónico en cultivos corticales, sugiriendo un posible núcleo que sostiene la actividad global de la red. Una reciente investigación de 2018 [50] usa la homología persistente para extraer nueva información del conectoma humano (estructura neuronal de la red cerebral) filtrado con la Rank Weight Clique Filtration. Se obtuvo una presencia de cliques y cavidades en más cantidad de la esperada por el modelo nulo creado, así como regiones que ponían en contacto áreas de origen temprano y tardío en largos loops, poniendo en relieve su importante papel en el control de las funciones cerebrales.

Vale la pena destacar y desarrollar un poco el trabajo de 2018 de Spremann y Dunn en [51], que proponen un método para extraer propiedades topológicas de variables desconocidas y encubiertas en la actividad neuronal. Aunque no parte de una red, se expone aquí por funcionar como alternativa a una caracterización de la red neuronal que normalmente sería hecha desde la perspectiva de la teoría de redes. Supongamos que se lleva a cabo un experimento en el que se capturan disparos de neuronas a lo largo del tiempo,

obteniendo para cada una de las N neuronas del sistema

$$s_i = (s_i(1), \dots, s_i(T)) \in \{-1, 1\}^T, \quad i = \{1, \dots, N\} \quad (3.27)$$

junto con muestras de la trayectoria en un espacio de configuraciones abstracto del espécimen en cuestión $\alpha(1), \dots, \alpha(T)$ (ver en este mismo artículo o en [48] la explicación de las *place cells*). Obviando los detalles técnicos, el procedimiento consiste en lo siguiente. Se computan las distancias entre eventos $D(s_i, s_j) \forall 1, j \leq N$, y se toman como pesos en un grafo completo G con los N vértices dados por cada neurona. A continuación se computan los grupos de homología persistente sobre el clique complex de este grafo, $PH_*(Cl(G))$. Si los resultados son consistentes con la homología persistente de una red de Erdős-Rényi, entonces no hay nada más que aprender. Si difiere, entonces debe pensarse si hay variables que se creen que influyen en este proceso dado. Si las hay, desarrollan un proceso para eliminar su contribución de manera estadística y se obtiene un nuevo set s'_1, \dots, s'_n . Si no hay más estímulos a considerar, $PH_*(Cl(G))$ guarda información sobre el espacio de estados de la variable encubierta.

A través de este método se puede encontrar factores influyentes inesperados en determinados experimentos y puede resultar un modelo muy útil. El siguiente esquema facilita la comprensión del proceso.

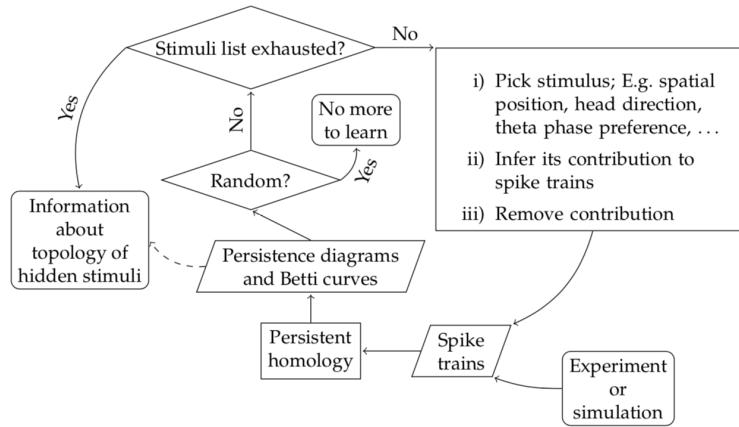


Figura 18: Modelo a seguir propuesto por Spreemann y Dunn para descubrir variables ocultas a partir de la homología persistente. Extraído de [51].

3.3. Otras construcciones sobre la persistencia homológica

A raíz de llevar a cabo estudios a través de la homología persistente surgen necesidades o ideas que puedan mejorar el entendimiento de estas herramientas o incluso extenderlo. El propósito de este apartado es comentar algunas de estas propuestas que han aparecido recientemente y que parecen acomodar la persistencia homológica a otros usos.

3.3.1. Evolución temporal de redes

Como bien se ha podido comprobar, la homología persistente realiza un análisis estático en el tiempo, sin considerar que pueda haber cambios en los pesos de las aristas o incluso en la existencia de nodos o conexiones. La idea de estudiar un grafo temporalmente dependiente a través de la homología persistente es tan sencilla como discretizar

la dimensión temporal y obtener diagramas de persistencia para un número finito de instancias, para después calcular bottleneck distances entre cada uno de los pasos. Este acercamiento al estudio de la variación temporal de redes se lleva a cabo por primera vez en [52] con filtraciones de Vietoris-Rips.

Formalizando un poco el método, dado un grafo $G = G(t)$ que evoluciona con el tiempo, se consideran n momentos o instancias y por consiguiente se tiene n grafos distintos $\{G_1, \dots, G_n\}$ donde $G_i = G(t_i)$ para todo $i \in \{1, \dots, n\}$. Se considera el complejo de Vietoris Rips para cada grafo y su correspondiente filtración, y se computan sus grupos de homología persistente $PH_*(VR(G_i))$ para toda i . Se calcula ahora las distancias

$$d_{ij} = d_{bot}(\mathcal{D}(VR(G_i)), \mathcal{D}(VR(G_j))) \quad \forall i, j \in \{1, \dots, n\} \quad (3.28)$$

y se colocan en una tabla para una fácil lectura. Es muy sencillo sobre esta tabla determinar entre qué momentos del tiempo parece haber habido mayores cambios en la estructura topológica del grafo. A continuación un ejemplo con tan sólo 3 instancias muestra cómo en el último salto temporal parece haber cambiado más significativamente la red.

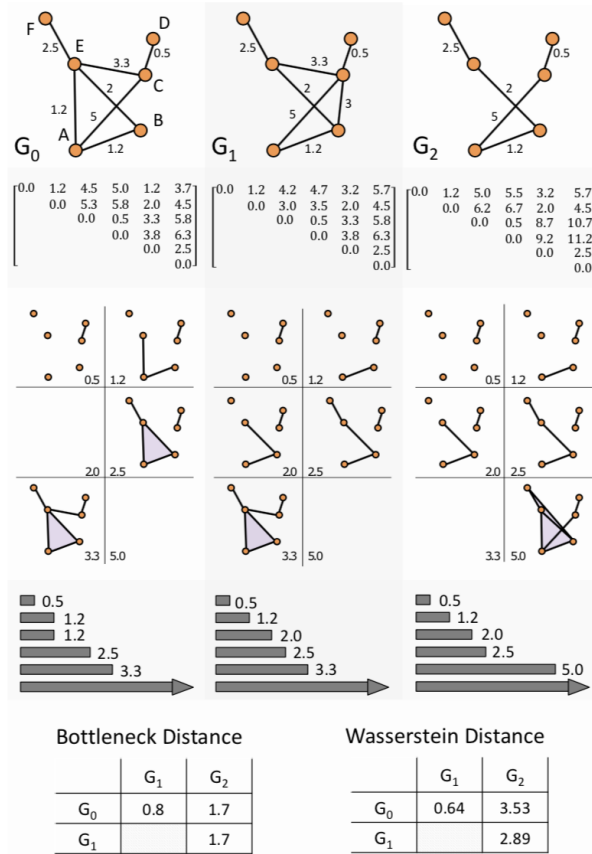


Figura 19: Tres instancias de una red variable en el tiempo. En cada columna, la correspondiente matriz de distancias de camino más corto, complejo simplicial en cada paso de la filtración y barcode 0-dimensional. Al final tabla de valores de distancias entre barcodes (donde la de la derecha es la distancia 2-Wasserstein). Extraído de [52].

3.3.2. Persistence vineyard

Esta herramienta fue desarrollada en 2006 por Cohen-Steiner et al. [53] como prueba de la estabilidad de los diagramas de persistencia planteada en el apartado 1.2.2., y es una representación de la persistencia homológica durante la perturbación de los datos iniciales. Pensado inicialmente sobre una nube de puntos, ha llegado a ser útil en el estudio de redes complejas en [54], donde fue capaz de identificar patrones robustos de conectividad funcional dinámica en cerebros de sujetos durante estados de juego y de descanso. Una adaptación completa a redes variables con el tiempo está por desarrollar.

Definición 3.24. *Sea una nube de puntos dinámica $\mathcal{X}(t) = \{x_1(t), \dots, x_N(t)\}$ que varía en un intervalo finito de tiempo, de 0 a 1 por simplicidad. El persistence vineyard es la familia uniparamétrica de diagramas de persistencia obtenidos de la nube de puntos para cada t .*

$$V(\mathcal{X}) = \{\mathcal{D}(\mathcal{X}(t)) \mid t \in [0, 1]\}. \quad (3.29)$$

Estas representaciones temporales tienen buenas propiedades métricas que se pueden encontrar demostradas en [55]. Un primer y importante ejemplo es el siguiente.

Proposición 3.25. *Si una nube de puntos dinámica $\mathcal{X}(t)$ es continua respecto la distancia de Hausdorff, su vineyard $V(\mathcal{X})$ es continuo respecto la bottleneck distance.*

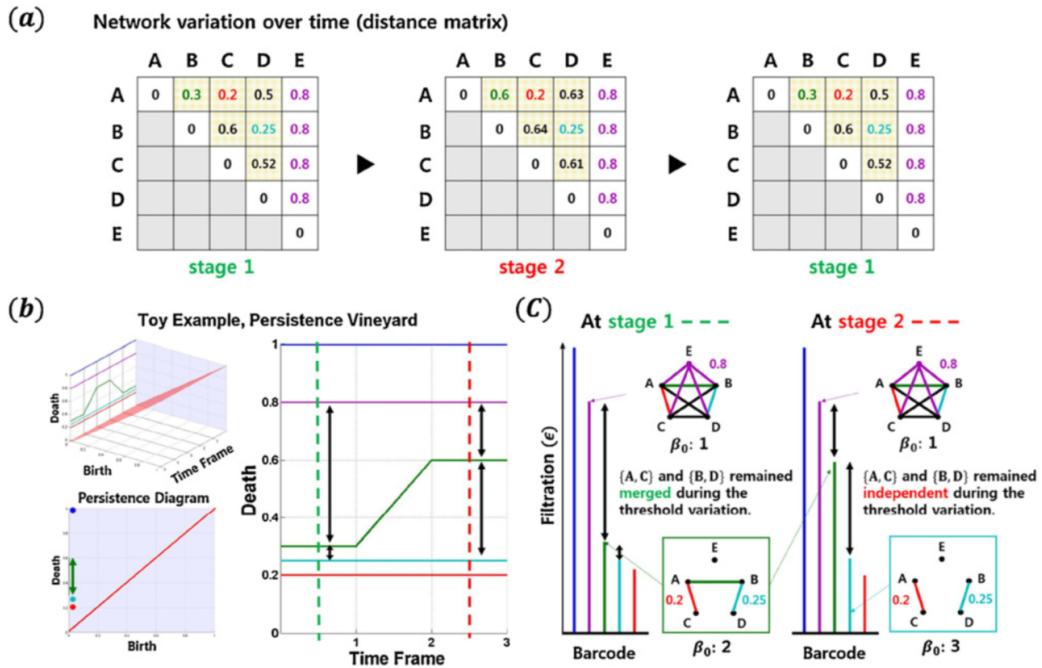


Figura 20: Pequeña variación en el peso de una arista. En (a) las matrices de distancia, (b) el correspondiente persistence vineyard y en (c) los barcodes para cada etapa, mostrando el cambio de longitud de las barras que queda plasmado en (b). Extraído de [54].

En general las variaciones iniciales no inducen cambios en las secuencias de grupos de homología, pero en determinados casos puede desaparecer o crearse un agujero durante el proceso $0 < t \leq 1$. En tal caso se introduce en la representación un switch, un diagrama del momento en que cambian los elementos del diagrama. Dos agujeros distintos se unen,

uno de los puntos se lleva a la diagonal y desaparece. Estudiando estos puntos de cambio se puede estudiar las fronteras de los cambios de estado del sistema. Así se consigue en [54] determinar propiedades dinámicas y temporalmente cambiantes de la conectividad funcional del cerebro sin caer en las restricciones de las medidas de la teoría de grafos. En la Figura 20 se encuentra un ejemplo muy sencillo de vineyard.

3.3.3. Homological scaffold

Esta herramienta definida en [56] ayuda a visualizar el papel de enlaces concretos que formen parte de ciclos con largas persistencias, y codifica la importancia de las aristas durante la filtración, dando el número de ciclos a los cuales pertenece una arista dada en cada paso. Un grave problema del método es su dependencia en la elección de un representante para cada clase de equivalencia, por lo que no es un invariante topológico. En un artículo de abril de 2020 [57] se intenta corregir estos problemas proponiendo una construcción más canónica que esté bien definida dadas ciertas condiciones (se acaba reduciendo a dimensión 1), el *minimal scaffold*. Se presentan cronológicamente las definiciones para entender mejor la motivación y las aportaciones.

La presencia de grupos de homología no triviales implican que la estructura es no contráctil, haciendo que la información de la red deba fluir por ciertos canales que toman el papel de puentes. Este es el papel que se busca subrayar con la construcción. Sea $G = (V, E)$ un grafo ponderado por $w : E \rightarrow \mathbb{R}_+$ y sea una filtración de su clique complex $Cl(G)$. Sea $\{b_i\}_i$ un conjunto de ciclos generadores de persistencia homológica sobre $\mathbb{Z}/2\mathbb{Z}$, que queda totalmente identificado por su soporte, un conjunto de aristas de G . Así, se puede representar $\{b_i\}_i$ por una matriz que tenga los ciclos b_i por columnas y los elementos de E como filas. La suma de cada fila constituye un nuevo peso para G , la cantidad de ciclos de persistencia a los que pertenece cada arista. Formalizado:

Definición 3.26. *Sea G un grafo ponderado y su correspondiente clique complex filtrado. Sea $\{b_i\}_i$ un conjunto de ciclos generadores de persistencia homológica sobre $\mathbb{Z}/2\mathbb{Z}$. Sea $h_w : E \rightarrow \mathbb{R}_+$ definida por:*

$$h_W(e) := \sum_i \mathbb{1}_{b_i}(e), \quad (3.30)$$

donde $\mathbb{1}_{b_i}$ es una función que es 1 si e aparece en el ciclo b_i y 0 en caso contrario. El homological scaffold de G es el grafo ponderado $H(W)$ con los vértices de V y las aristas de G no nulas por h_W con el peso dado por h_W .

En la Figura 21 se muestra un complejo simplicial que nace de una nube de puntos marcando sus ciclos generadores de PH_1 junto con su homological scaffold. A continuación se intenta resolver la arbitrariedad de esta definición escogiendo de todos los posibles representantes aquellos que minimicen la distancia total. En dimensiones mayores a 1 está probado [58] que el problema es de tipo NP, es decir, computacionalmente intratable. Por este motivo los scaffolds son una herramienta útil –o más concretamente, bien definida– únicamente para dimensión 1. Para este caso encontrar una base minimal de homología se puede hacer mediante un algoritmo desarrollado por Dey et al. [59], garantizando su existencia para complejos simpliciales finitos. Así se redefine el *minimal scaffold*:

Definición 3.27. *El minimal scaffold es el homological scaffold dado por una base minimal de homología y su respectivo peso $h_{W,min}$.*

La unicidad del minimal scaffold viene dada por la unicidad de la base minimal, se estudian varias situaciones en [57]. En este mismo artículo se aplica con éxito a diversas redes, demostrando identificar las aristas esenciales y las prescindibles de un grafo y dando una representación compacta de las características topológicas de la red.

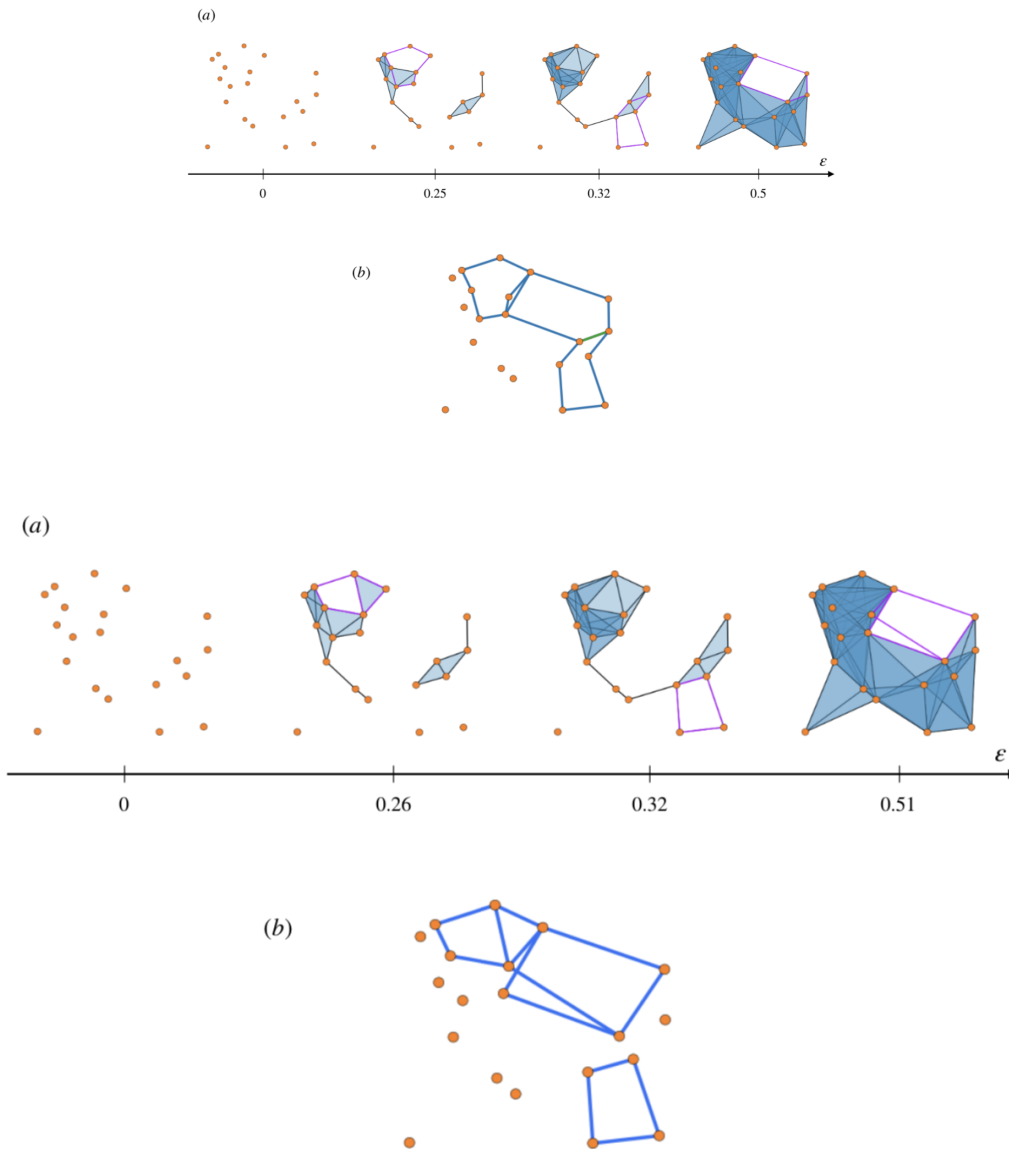


Figura 21: Scaffold estándar y scaffold minimal para el mismo complejo simplicial. En rosa se marca los generadores escogidos, y la arista verde del estándar indica que forma parte de dos generadores distintos. Extraído de [57].

4. Conclusiones

La aplicación de la homología persistente sobre redes complejas resulta ser especialmente natural y descubre una nueva faceta de estudio que ya ha empezado a dar fruto y complementar la información extraída del análisis clásico sobre los grafos. Las propiedades estructurales desveladas por la persistencia hacen referencia a aspectos raramente estudiados, sobre todo en ciencias esencialmente aplicadas como la neurociencia, y pueden así dar pie fácilmente a nuevas conclusiones y descubrimientos.

La utilidad de su uso reside, en cualquier caso, en una elección sensata en la construcción de un complejo simplicial sobre la red y más importantemente incluso, de una filtración adecuada. El orden en que se generen los subcomplejos que constituyen la filtración determinará en gran parte el orden en que las características topológicas aparezcan, así como su persistencia. Precisamente por esto, también el tipo de resultados que se obtengan serán de naturalezas muy diversas, y está en las manos del investigador decidir aplicar estas ideas con creatividad para llegar a conclusiones novedosas. Se espera que la presentación aquí hecha de una amplia variedad de métodos sea suficiente para poder tener un criterio a la hora de escoger lo que más convenga en un estudio dado.

La fiabilidad de los estudios de persistencia homológica sobre redes está comprobada por haber sido estudiada en los modelos típicos de red como lo son los grafos de Erdős-Rényi o las redes libres de escala, y haber demostrado que es capaz de caracterizarlas. La aleatoriedad de una red se ve claramente diferenciada en los diagramas de persistencia frente a las redes que cuentan con una cierta organización geométrica, y incluso pueden distinguirse así diferentes grados de agregación o de dependencia geométrica. Paralelamente, la comparación de barcodes a través de la bottleneck distance también es una herramienta discriminativa que, usada con grafos de referencia, puede usarse para catalogar las redes de estudio según su semejanza con las tipologías ya conocidas.

Así pues, viendo que van apareciendo mecanismos para extender el uso de la homología persistente en redes a grafos dependientes del tiempo o a la extracción de variables ocultas de un cierto proceso, se puede asegurar que la junción de estas dos áreas de trabajo se encuentra en pleno desarrollo y parece que tiene aún mucho que decir. Hasta el momento ha sido motivo de ilusión y promesa de nuevos descubrimientos; esperemos que no tarde el momento en que sean los resultados los que den para que se hable de estos métodos.

Referencias

- [1] J. R. Munkres: Elements of Algebraic Topology, *Addison-Wesley Publishing Company*, 1984.
- [2] V. Navarro: Topologia Algebraica, *notas del curso impartido en la Universitat de Barcelona*, 2017.
- [3] G. Carlsson: Persistent Homology and Applied Homotopy Theory, *arXiv: 2004.00738v1*, 2020.
- [4] G. Carlsson: Topology and Data, *Bulletin of the American Mathematical Society Vol.46*, (2): 255-308, 2009.
- [5] L. Polterovich, D. Rosen, K. Samvelyan, J. Zhang: Topological Persistence in Geometry and Analysis, *arXiv: 1904.04044v1*, 2019.
- [6] A. Zomorodian, G. Carlsson: Computing Persistent Homology, *Discrete Computational Geometry*, 33: 249-274, 2005.
- [7] H. Edelsbrunner, J. Harer: "Computational Topology: an introduction", *American Mathematical Society*, 2010.
- [8] H. Edelsbrunner, J. Harer: Persistent Homology - a Survey, en *Surveys on Discrete and Computational Geometry: Twenty Years Later*, *American Mathematical Society*, 2008.
- [9] D. Eisenbud: Commutative Algebra with a View Toward Algebraic Theory, *Springer-Verlag*, New York, 1995.
- [10] F. Chazal, V. De Silva, M. Glisse, s. Oudot: The Structure and Stability of Persistence Modules, *arXiv: 1207.3674v3*, 2013.
- [11] D. Cohen-Steiner, H. Edelsbrunner, J. Harer: Stability of Persistence Diagrams, *Discrete Computational Geometry* (37):103–120, 2007.
- [12] C. Casacuberta: Geometric Uses of Persistent Homology, *Notes from UB Topology Seminar*, 2020.
- [13] P. Bubenik: Statistical Topological Data Analysis using Persistence Landscapes, *Journal of Machine Learning Research* 16: 77-102, 2015.
- [14] M. Gromov: Metric Structures for Riemannian and Non-Riemannian Spaces, *Birkhäuser*, Basel, 2007.
- [15] G. Carlsson, F. Mémoli: Characterization, Stability and Convergence of Hierarchical Clustering Methods, *Journal of Machine Learning Research* 11 1425-1470, 2010.
- [16] E. Berry, Y. Chen, J. Cisewski-Kehe, B. T. Fasy: Functional Summaries of Persistence Diagrams, *arXiv: 1804.01618v1*, 2018.
- [17] B. Rieck, U. Fugacci, J. Lukasczyk, H. Leitte: "Clique Community Persistence: A Topological Visual Analysis Approach for Complex Networks", *IEEE Transactions on Visualization and Computer Graphics*, TVCG.2017.2744321, 2018.

- [18] H. Lee, H. Kang, M. K. Chung, B-N. Kim, D. S. Lee: Persistent Brain Homology from the Perspective of Dendrogram, *IEEE Transactions to Medical Imaging* 31(12): 2267-2277, 2012.
- [19] H. Lee, H. Kang, M. K. Chung, B-N. Kim, D. S. Lee: Computing the Shape of Brain Networks Using Graph Filtration and Gromov-Hausdorff Metric. In: Fichtinger G., Martel A., Peters T. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI. *Lecture Notes in Computer Science*, vol 6892. Springer, Berlin, Heidelberg, 2011.
- [20] H. Lee, H. Kang, M. K. Chung, B-N. Kim, D. S. Lee: Weighted Functional Brain Network Modeling via Network Filtration. Presentado en *NIPS 2012 Workshop on Algebraic Topology and Machine Learning*, 2012.
- [21] E. Kim, H. Kang, H. Lee, H-J. Lee, M-W. Suh, J-J. Song, S-H. Oh, D.S. Lee: Morphological brain network assessed using graph theory and network filtration in deaf adults. *Hearing Research* 315, 88-98, 2014.
- [22] E. Ben-Naim, H. Frauenfelder, Z. Toroczkai (Eds.), *Complex Networks*, Springer, Berlin, 2004.
- [23] S. Boccaletti, V. Latora, Y. Moreno et al., *Complex Networks: Structure and Dynamics*, Physics Reports (424) 175-308, 2006.
- [24] R. Pastor-Satorras, M. Rubi, A. Diaz-Guilera (Eds.), *Statistical Mechanics of Complex Networks*, Springer, Berlin, 2003.
- [25] Brain Connectivity Toolbox disponible en <https://sites.google.com/site/bctnet/> a fecha de Junio de 2020.
- [26] S. Chowdhury, F. Mémoli: A Functorial Dowker Theorem and Persistent Homology of Asymmetric Networks, *arXiv*: 1608.05432v3, 2018.
- [27] G. Kalna, D. J. Higham: Clustering Coefficients for Weighted Networks, *University of Strathclyde Mathematics, Research Report 3*, 2006.
- [28] G. Fagiolo: Clustering in Complex Directed Networks, *arXiv*: physics/0612169v3, 2007.
- [29] M. E. J. Newman: Modularity and community structure in networks, *PNAS*, vol.103, (23):8577-8582, 2006.
- [30] C. Giusti, R. Ghrist, D. S. Bassett: "Two'a company, three (or more) is a simplex. Algebraic-topological tools for understanding higher-order structure in neural data.", *Journal of Computational Neuroscience* 41: 1-14, 2016.
- [31] M. E. Aktas, E. Akbas, A. El Fatmaoui: Persistence homology of networks: methods and applications. *Applied Network Science* 4:61, 2019.
- [32] B. Stolz: *Computational Topology in Neuroscience*, 2014.
- [33] R. Forman: A User's Guide to Discrete Morse Theory, *Séminaire Lotharingien de Combinatoire* 48, Article B48c, 2002.

- [34] C. H. Dowker: Homology Groups of Relations. *The Annals of Mathematics* 2nd Ser, Vol 56, n.1, 84-95, 1952.
- [35] M. Kahle: Topology of random simplicial complexes: a survey, *arXiv*: 1301.7165v2, 2013.
- [36] M. Kahle: Sharp Vanishing Thresholds for Cohomology of Random Flag Complexes, *arXiv*: 1207.0149v3, 2013.
- [37] O. Bobrowski, M. Kahle: Topology of random geometric complexes: a survey, *Journal of Applied and Computational Topology* vol.1, p.331–364, 2018.
- [38] D. Horak, S. Maletić, M. Rajković: Persistent Homology of Complex Networks, *arXiv*: 0811.2203v2, 2008.
- [39] A. E. Sizemore, C. Giusti, D. Bassett: Classification of weighted networks through mesoscale homological features, *arXiv*: 1512.06457v2, 2016.
- [40] A. E. Sizemore, J. E. Phillips-Cremins, R. Ghrist, D. D. Bassett: "The importance of the whole: Topological data analysis for the network neuroscientist", *Network Neuroscience* 3(3), 656–673, 2019.
- [41] G. Petri, M. Sciamiero, I. Donato, F. Vaccarino: Topological Strata of Weighted Complex Networks, *PLoS ONE* 8(6): e66506, 2013.
- [42] C. Giusti, E. Pastalkova, C. Curto, V. Itsvok: Clique topology reveals intrinsic geometric structure in neural correlations, *PNAS* vol.112, (44): 13455-13460.
- [43] H. Kannan, E. Saucan, I. Roy, A. Samal: persistent homology of unweighted complex networks via discrete Morse theory, *Scientific Reports* (9): 13817, 2019.
- [44] H. Lee, M. K. Chung, H. Kang, B. Kim and D. S. Lee: Discriminative persistent homology of brain networks, *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, Chicago, pp. 841-844, 2011.
- [45] A. Khalid, B. S. Kim, M. K. Chung, J. C. Ye, D. Jeon: Tracing the evolution of multi-scale functional networks in a mouse model of depression using persistent brain network homology, *Neuroimage* (101): 351-363, 2014.
- [46] H. Choi, Y. K. Kim, H. Kang, H. Lee, H-J. Im, D. W. Hwang, E. E. Kim, J-K. Chung, D. S. Lee: "Abnormal metabolic connectivity in the pilocarpine-induced epilepsy rat model: A multiscale network analysis based on persistent homology", *Neuroimage* (99): 226-236, 2014.
- [47] C. Curto, V. Itskov: Cell Groups Reveal Structure of Stimulus Space, *PLoS Comput Biol* 4(10): e1000205, 2008.
- [48] C. Curto: What can topology tell us about the neural code?, *arXiv*: 1605.01905v1, 2016.
- [49] V. Pirino, E. Riccomagno, S. Martinoia, P. Massobrio: A topological study of repetitive co-activation networks in in vitro cortical assemblies, *Phys. Biol.* 12 016007, 2015.

- [50] A. E. Sizemore, C. Giusti, A. Kahn, J. M. Vetter, R. F. Betzel, D. S. Bassett: Cliques and cavities in the human connectome, *Journal of Computational Neuroscience* (44): 115-145, 2018.
- [51] G. Spreemann, B. Dunn, M. B. Botnan, N. A. Baas: Using persistent homology to reveal hidden information in neural data, *arXiv: 1510.06629v1*, 2015.
- [52] M. Hajij, B. Wang, C. Scheidegger, P. Rosen: Visual Detection of Structural Changes in Time-Varying Graphs Using Persistent Homology, *arXiv: 1707.06683v2*, 2017
- [53] D. Cohen-Steiner, H. Edelsbrunner, D. Morozov: Vines and vineyards by updating persistence in linear time, *Proc. 22nd Annual Symposium in Computational Geometry*, pp. 119–126, 2006.
- [54] J. Yoo, E. Y. Kim, Y. M. Ahn, J. C. Ye: Topological persistence vineyard for dynamic functional brain connectivity during resting and gaming stages, *Journal of Neuroscience Methods* 267 1–13, 2016.
- [55] E. Munch: Applications of Persistent Homology to Time Varying Systems, 2013.
- [56] G. Petri, P. Expert, F. Turkheimer, R. Carhart-Harris, D. Nutt, P. J. Hellyer and F. Vaccarino: Homological scaffolds of brain functional networks, *J. R. Soc. Interface* 11: 20140873, 2014.
- [57] M. Guerra, A. De Gregorio, U. Fugacci, G. Petri, F. Vaccarino: Homological Scaffold via Minimal Homology Bases, *arXiv: 2004.11606v1*, 2020.
- [58] Chen, C., Freedman, D.: Hardness results for homology localization, *Discrete Computational Geometry* 45(3), 425–448, 2011.
- [59] T. K. Dey, A. N. Hirani, B. Krishnamoorthy: Optimal Homologous Cycles, Total Unimodularity, and Linear Programming, *arXiv: 1001.0338v3*, 2011.
- [60] Imagen de J.F. Cuenca, A. Iske: Persistent Homology for Defect Detection in Non-Destructive Evaluation of Materials, 2016.
- [61] Imagen de S. Perera, M.G.H. Bell, M.C.J. Bliemer: Network science approach to modelling the topology and robustness of supply chain networks: a review and perspective, *Applied Network Science*, (2):33, 2017.
- [62] Imagen de J. L. Bravo Trinidad: Apuntes sobre Diagramas de Voronoi, *Universidad de Extremadura*.
- [63] Imagen de https://en.wikipedia.org/wiki/Clique_complex (visitado en junio de 2020).

Anexo 1: Tabla de filtraciones posibles sobre una red

Tabla extraída de [31].

	Undirected network	Directed network	Weighted network	Unweighted network	Metric network	Dynamic network	Simplicial Complex	Ordered-tuple complex	Stable
Vietoris-Rips filtration (VR) section	✓		✓				✓		✓
Dowker sink and source filtration (DSS) section		✓	✓				✓		✓
Clique complex filtration (CCL) section	✓			✓			✓		
Vertex-based clique filtration (VBCL) section	✓		✓				✓		
k-clique filtration (kCL) section	✓		✓				✓		
Weighted simplex filtration (WS) section	✓		✓				✓		
Vertex function based filtration (VFB) section	✓		✓				✓		✓
Intrinsic Čech filtration (IC) section	✓			✓	✓		✓		
Functional metric graph filtration (FMG) section	✓		✓		✓		✓		✓
Power filtration (POW) section	✓			✓			✓		
Temporal filtration (TMP) section						✓	✓		
Zigzag simplicial filtration (ZSF) section	✓			✓		✓	✓		✓
Digraph filtration using persistent path homology (PPH) section		✓					✓		✓
Vietoris-Rips filtration under sym_α section	✓		✓				✓		✓
Directed Vietoris-Rips filtration section		✓	✓					✓	✓
Associated filtration of directed graphs section		✓	✓				✓		✓
Preorder filtration section		✓	✓					✓	✓