



**UNIVERSIDAD NACIONAL DE ROSARIO  
FACULTAD DE CIENCIAS ECONÓMICAS Y  
ESTADÍSTICA**

**DOCTORADO EN ESTADÍSTICA**

**DESARROLLO DE ÍNDICES ESTADÍSTICOS BASADOS  
EN REDES MULTICAPAS PARA LA  
SUSTENTABILIDAD DE ECOSISTEMAS  
PATAGÓNICOS**

**Claudia Alejandra Huaylla  
Doctoranda**

**Director:** Dr. Marcelo Kuperman

**Co-Director:** Dr. Lucas Garibaldi

Marzo 2025



# Agradecimientos

A mis dos directores, por guiarme y enseñarme durante todos estos años. Gracias Marcelo por tu dedicación y paciencia. Gracias por el tiempo y esfuerzo que le dedicaste a mi formación, sin tus palabras y correcciones precisas no hubiese podido lograr llegar a esta instancia tan anhelada. Gracias por confiar en mí y por brindarme tus conocimientos. Gracias Lucas por aceptarme en el IRNAD y por hacer posible que pueda asistir a cada congreso.

A Diego, mi compañero de vida, por todas las veces que estuvo para escucharme, por alentarme día a día, por leer cada trabajo y cada párrafo. Gracias por ser incondicional.

A mi padre, que si bien ya no está entre nosotros, se aseguró de que yo pueda terminar la carrera de grado y gracias a eso pude realizar un doctorado. A mi madre y a mi hermano por escucharme cada vez que los necesité y que, sin entender mucho de qué se trataba lo que estaba haciendo, supieron aguantar mis distancias y me apoyaron sin cuestionarlo.

A mi pequeño Aukan, quien logra sacarme una sonrisa día a día, y quien me mostró las prioridades de la vida.

A mis queridas amigas Julia y Gilda, por su constante apoyo y motivación para embarcarme en este posgrado. Gracias por creer en mí y estar siempre a mi lado.

A mis amigas que siempre estuvieron para escucharme cuando las necesitaba, Carla, Fernanda, Cecilia, Tati, Mariela.

A Marquitos y a Caro, que fueron los primeros en recibirme en el IRNAD y siempre estuvieron para responder mis dudas sobre la parte ecológica.



# Resumen

El análisis de redes complejas es fundamental para comprender cómo la estructura de sistemas complejos afecta la dinámica de diversos procesos. A pesar de su importancia, la cuantificación de la complejidad de estas redes representa un desafío, especialmente cuando las interacciones involucran múltiples entidades.

Para abordar este problema, se propuso una metodología de aleatorización controlada para detectar si la estructura de la red original codifica algún tipo de información sobre las interacciones asociadas a las especies del ecosistema estudiado y se exploraron índices para estudiar la complejidad de las redes ecológicas. Además, se utilizaron hipergrafos para modelar interacciones que implican más de dos entidades.

Los hallazgos más relevantes indican que la medida de modularidad es eficaz para caracterizar redes y que es fundamental emplear múltiples índices para evaluar su complejidad. El análisis de hipergrafos complementó estos resultados, ofreciendo una visión más profunda de las interacciones ecológicas.

En conclusión, los desarrollos presentados resaltan la relevancia de la teoría de redes, los modelos nulos y herramientas estadísticas avanzadas, como la entropía y K-complejidad, para la caracterización de sistemas ecológicos complejos. Estos enfoques no solo mejoran nuestra comprensión de las redes, sino que también tienen implicaciones significativas para la ecología y otros campos relacionados.

**Palabras clave:** COMPLEJIDAD, ENTROPÍA, MODULARIDAD, REDES ECOLÓGICAS



# Índice de contenidos

Agradecimientos	iii
Resumen	v
Índice de contenidos	vii
Índice de figuras	xi
Índice de tablas	xv
Índice de algoritmos	xvii
<b>1. Introducción</b>	<b>1</b>
1.1. Redes complejas . . . . .	1
1.2. Bosques en transformación: por qué es clave el estudio de sus interacciones ecológicas . . . . .	2
1.2.1. El papel de los bosques en la mitigación del cambio climático . . . . .	2
1.2.2. Transformaciones y fragmentación de los bosques . . . . .	3
1.2.3. Disminución de los bosques nativos y servicios ecosistémicos . . . . .	4
1.2.4. Los bosques andino-patagónicos . . . . .	4
1.2.5. Prácticas de manejo forestal sostenible . . . . .	4
1.2.6. Interacciones ecológicas en los bosques . . . . .	5
1.3. Especies clave . . . . .	6
1.4. Representación de las redes . . . . .	7
1.4.1. Caracterización de Redes . . . . .	8
1.5. Hipergrafos . . . . .	13
1.5.1. Características de los hipergrafos . . . . .	13
1.6. Objetivos de esta tesis . . . . .	16
<b>2. Decodificando información en redes ecológicas multicapas</b>	<b>19</b>
2.1. Introducción . . . . .	19
2.2. Metodología . . . . .	21

2.2.1.	Índices para caracterización de la red . . . . .	21
2.2.2.	Aleatorización controlada de la red . . . . .	25
2.2.3.	Caracterizando la topología de la red . . . . .	28
2.2.4.	Caso de estudio. Interacciones entre Polinizadores- Plantas- Herbívoros . . . . .	32
2.3.	Resultados . . . . .	34
2.3.1.	Implementación del algoritmo . . . . .	35
2.3.2.	Caracterización de nuestra muestra . . . . .	36
2.4.	Discusión . . . . .	38
2.4.1.	Aplicaciones del algoritmo . . . . .	40
2.4.2.	Identificación de especies clave . . . . .	40
<b>3.</b>	<b>Medidas de Complejidad de redes ecológicas bipartitas</b>	<b>43</b>
3.1.	Introducción . . . . .	43
3.2.	Metodología . . . . .	48
3.2.1.	Datos analizados . . . . .	48
3.2.2.	Medidas estadísticas frecuentemente usadas en la caracterización de las redes . . . . .	48
3.2.3.	Herramientas para estimar la complejidad de una red: <i>Singular</i> <i>value decomposition</i> y K-complejidad . . . . .	49
3.2.4.	Modelos nulos . . . . .	52
3.2.5.	Comparando medidas de complejidad . . . . .	53
3.2.6.	Medidas de complejidad aplicadas en 3 redes aleatorias conocidas	56
3.3.	Resultados . . . . .	60
3.3.1.	Modelos nulos . . . . .	60
3.4.	Discusión . . . . .	65
<b>4.</b>	<b>Hipergrafos: Un enfoque complementario</b>	<b>67</b>
4.1.	Introducción . . . . .	67
4.2.	Metodología . . . . .	68
4.2.1.	Construcción del hipergrafo a estudiar . . . . .	68
4.2.2.	Comparando comunidades versus hiperenlaces . . . . .	71
4.2.3.	La entropía como herramienta para determinar nodos influyentes	73
4.3.	Resultados . . . . .	74
4.3.1.	Hipergrafo del caso de estudio . . . . .	74
4.3.2.	Comparando comunidades versus hiperenlaces . . . . .	77
4.3.3.	Determinando los nodos influyentes usando la entropía . . . . .	77
4.4.	Discusión . . . . .	78

---

4.4.1. Implicancias de la comparación entre comunidades versus hiper-enlaces . . . . .	79
4.4.2. Identificación de nodos influyentes . . . . .	80
<b>5. Discusión</b>	<b>83</b>
5.1. Contribuciones . . . . .	84
5.1.1. Capítulo 2 . . . . .	84
5.1.2. Capítulo 3 . . . . .	86
5.1.3. Capítulo 4 . . . . .	87
5.2. Alcances de este trabajo . . . . .	89
5.3. Limitaciones y trabajos futuros . . . . .	89
<b>A. Material complementario A</b>	<b>93</b>
<b>B. Material complementario B</b>	<b>99</b>
<b>C. Pseudocódigo de los algoritmos principales</b>	<b>103</b>
<b>Bibliografía</b>	<b>109</b>
<b>Publicaciones asociadas</b>	<b>125</b>

\*



# Índice de figuras

1.1. Representación de una red no dirigida con 4 nodos . . . . .	7
1.2. Hipergrafo con 8 nodos y 4 hiperenlaces. Su matriz de incidencia es $E(H)$	15
2.1. Visualización de los pasos del algoritmo de Louvain. Figura adaptada de Blondel et al. (2008) . . . . .	24
2.2. Representación simplificada de un cambio en una red. (A) Grafo original. $A_{12} = 1, A_{45} = 1, A_{24} = 0, A_{15} = 0$ . (B) Grafo con un cambio. $A_{12} = 0, A_{45} = 0, A_{24} = 1, A_{15} = 1$ . . . . .	27
2.3. Ubicación y fotografías del sitio de estudio. Créditos fotográficos: Marina Alma y Marcos Nacif. . . . .	33
2.4. Metodología para determinar si una red tiene información codificada y para identificar las especies clave. . . . .	35
2.5. Modularidad vs. el número de cambios ( <i>swappings</i> ) en la red original. Cada valor es un promedio de 10 redes excepto por la red original. . . . .	36
2.6. Distribución de grado de los nodos de la red original. . . . .	37
2.7. Red proyectada de Polinizador-Planta- Gremios de Herbívoros. Las especies de plantas son representados en violeta y entre corchetes, los gremios de herbívoros en verde y entre paréntesis y los polinizadores en naranja. . . . .	37
2.8. Especies de plantas, polinizadores y gremios de herbívoros vs. <i>betweenness</i> normalizada. . . . .	39
2.9. (A) Grado vs. <i>betweenness</i> y (B) abundancia (Porcentaje de cobertura) vs. <i>betweenness</i> de 11 especies de plantas. . . . .	39
3.1. Grafo regular de grado 3 con 10 vértices . . . . .	45
3.2. Representación simplificada de un cambio, donde $A_{ij}$ es el elemento de la fila $i$ y la columna $j$ de la matriz de adyacencia y $D_i$ es el grado del nodo $i$ . (A) Grafo original. $A_{16} = 1, A_{59} = 1, A_{56} = 0, A_{19} = 0, D_4 = 2, D_6 = 2$ . (B) Grafo con un cambio. $A_{16} = 0, A_{59} = 0, A_{56} = 1, A_{19} = 1$ , donde la distribución de grado es la misma de la red original. (C) Grafo con un cambio, donde la distribución de grado ya no es la misma. $D_4 = 3, D_6 = 1$ . . . . .	53

3.3.	Estas figuras representan tres posibles escenarios para el comportamiento de los diferentes índices que consideramos para el estudio de la complejidad. Consideraremos tres escenarios posibles para el comportamiento del índice de K-complejidad versus los cambios/total de enlaces y luego ajustamos esos puntos utilizando un modelo lineal para identificar el signo del coeficiente principal. Tendremos tres escenarios posibles para ese coeficiente principal, replicamos esto para todos los índices calculados. A) $\beta_k > 0$ . K-complejidad versus cambios/total de enlaces. Cada valor es un promedio de 10 redes excepto la original. B) $\beta_k < 0$ . K-complejidad versus cambios/total de enlaces en la red original. Cada valor es un promedio de 10 redes excepto la original. C) $\beta_k \approx 0$ . K-complejidad versus cambios/total de enlaces en la red original. Cada valor es un promedio de 10 redes, excepto la original. . . . .	55
3.4.	A) Relación entre $\beta_k$ y la densidad para cada red, vale la pena resaltar que el grado de cada nodo se conserva en estas redes. B) Relación entre $\beta_{svd}$ y densidad para cada red. C) Relación entre $\beta_{modularidad}$ y densidad para cada grafo. D) Relación entre $\beta_{entropía}$ y densidad para cada red. . . . .	62
3.5.	A) Relación entre $\beta_k$ y la densidad de cada red; el número de enlaces se mantiene en estas redes pero su distribución de grado no. B) Relación entre $\beta_{svd}$ y la densidad para cada red. C) Relación entre $\beta_{modularidad}$ y la densidad para cada red. D) Relación entre $\beta_{entropía}$ y la densidad para cada red. . . . .	64
4.1.	Red bipartita de 10 nodos. 3 plantas (pl1, pl2 y pl3) y 7 polinizadores (P1, P2, P3, P4, P5, P6 y P7). . . . .	70
4.2.	Grafo de 10 nodos construido a partir de la red bipartita. . . . .	70
4.3.	Hipergrafo con 11 hiperenlaces, donde sus nodos están coloreados dependiendo de la comunidad a la que pertenecen. . . . .	76
4.4.	Valores de oARI ajustado vs. el número de cambios ( <i>swappings</i> ) en la red original. Cada valor es un promedio de 10 redes excepto por la red original. . . . .	77
A.1.	Red de interacciones entre planta-polinizador ordenada por comunidades. Cuatro comunidades fueron identificadas. Sus nodos son presentados como sigue: Las plantas en violeta y los polinizadores en color naranja. . . . .	94
A.2.	Red de interacción entre herbívoros-plantas ordenadas por comunidades. Se identificaron tres comunidades. Sus nodos se representan de la siguiente manera: plantas en color violeta y gremios de herbívoros en verde. . . . .	95

---

A.3. Red de las especies de plantas ordenada por comunidades. Dos comunidades fueron identificadas. . . . .	97
B.1. Relación entre anidamiento y densidad en la red hospedador-parásito (izquierda) y en la red planta-polinizador (derecha). . . . .	99
B.2. Relación entre SVD-entropía y densidad en la red hospedador-parásito (izquierda) y en la red planta-polinizador (derecha). . . . .	100



# Índice de tablas

2.1. Cambios en la estructura de la red después de eliminar una especie de planta. Solo fueron consideradas las especies de plantas con una <i>betweenness</i> normalizada más grande que $\alpha = 0.20$ . . . . .	38
4.1. Matriz de incidencia y grados . . . . .	73
4.2. Especies de plantas con su valor de entropía y con su grado . . . . .	78
A.1. Cambios en la estructura de la red después de eliminar una especie de planta. Las especies clave: <i>E.coccineum</i> y <i>R.magellanicum</i> . . . . .	93
A.2. Cambios en la estructura de la red después de eliminar una especie de planta. Especies clave: <i>R.cucullatum</i> and <i>N.antarctica</i> . . . . .	95
A.3. Cambios en la estructura de la red después de eliminar una especie de planta. . . . .	96
A.4. Cambios en la estructura de la red después de eliminar una especie de planta. . . . .	96
A.5. Lista de nodos . . . . .	98
B.1. Lista de redes de hospedador-parásito que se usaron en el capítulo 3 . . .	100
B.2. Lista de redes de planta-polinizador que se usaron en el capítulo 3 . . .	101
B.3. Tabla resumen . . . . .	102



# Índice de algoritmos

1.	Generación de Matriz de Adyacencia . . . . .	103
2.	Determinación de la Región No Nula en una Matriz . . . . .	104
3.	Algoritmo para desordenar una red conservando el grado de los nodos . . . . .	105
4.	Algoritmo para modificar la distribución de grado . . . . .	106
5.	Cálculo del Overlap Adjusted Rand Index . . . . .	108



# Capítulo 1

## Introducción

### 1.1. Redes complejas

El estudio de redes es una herramienta fundamental para describir y analizar cómo la estructura de un sistema complejo afecta la dinámica de diferentes procesos que puedan ocurrir en él [1–3]. En los últimos años, el formalismo de redes complejas ha despertado un enorme interés en múltiples disciplinas, ya que permite modelar y comprender sistemas que abarcan un amplio espectro de áreas de la ciencia, incluyendo fenómenos sociales, físicos, biológicos, tecnológicos y ecológicos[4–6].

Esto se debe a que las redes permiten modelar y comprender sistemas complejos compuestos por unidades interactuantes, describiendo la topología sobre la que se desarrollan las interacciones y proporcionando una representación de la estructura sobre la que ocurren estas interacciones.

Este enfoque no solo permite capturar la organización del sistema, sino que también ofrece herramientas para entender cómo su estructura influye en su dinámica y funcionamiento a distintas escalas espacio-temporales. Analizar estos sistemas implica estudiar sistemas complejos donde hay interacciones entre múltiples individuos. Los datos de redes ecológicas suelen ser recolectados en campo y suelen proveer información sobre la interacción entre dos individuos o especies. El número de individuos de cada sistema dependerá de los datos observados. Dado que los datos son recolectados por personas con un grado de error propio del ojo humano, resulta importante determinar si el sistema a estudiar es confiable y si se puede obtener algo relevante del mismo antes de extraer conclusiones sobre su estructura y dinámica.

El crecimiento actual del uso de redes complejas se vio reforzado por el desarrollo matemático y estadístico realizado en estos últimos años, enfocado especialmente en redes multicapas y bipartitas. Este avance también fue impulsado por el acceso a grandes volúmenes de datos. Las redes bipartitas, que contienen dos tipos de nodos y no permiten interacciones entre nodos del mismo grupo, son comunes en contextos como

redes ecológicas (por ejemplo, interacciones planta-polinizador), redes de colaboración (como autores y publicaciones) y sistemas de recomendación (usuarios y productos).

La versatilidad del enfoque basado en redes dio lugar al desarrollo de nuevos métodos para el análisis de datos y la identificación de patrones emergentes en sistemas complejos. A medida que el campo continúa evolucionando, el estudio de redes no solo ofrece una forma de representar sistemas, sino que también proporciona herramientas esenciales para comprender sus principios organizativos y predecir su comportamiento.

## **1.2. Bosques en transformación: por qué es clave el estudio de sus interacciones ecológicas**

### **1.2.1. El papel de los bosques en la mitigación del cambio climático**

Los bosques desempeñan un papel fundamental en la mitigación del cambio climático, actuando como importantes sumideros de carbono atmosférico. A nivel global, constituyen uno de los principales depósitos de carbono, almacenando aproximadamente el 77% de la biomasa viva terrestre en forma de biomasa leñosa viva [7]. Esta energía capturada en la biomasa forestal es aprovechada por diversos organismos, lo que genera múltiples interacciones ecológicas clave para la dinámica, redistribución y acumulación de nutrientes, carbono y energía en los ecosistemas. Estas interacciones pueden ser mutualistas, como la polinización y las micorrizas; antagonistas, como la folivoría y los hongos de madera; y comensales, como la descomposición y la humidificación.

Históricamente, los humanos hemos aprovechado los beneficios que surgen de estas interacciones, utilizando los bosques como fuente de combustible, material para construcción y alimentos. Sin embargo, debido a la expansión de la agricultura y ganadería extensiva, el sector forestal se ha enfocado en el establecimiento y manejo de plantaciones mono-específicas orientadas a la producción maderera. Actualmente, el cambio climático y la creciente preocupación ambiental han generado una revalorización de los bosques naturales multi-específicos, los cuales proveen una amplia gama de servicios ecosistémicos cruciales para la humanidad, como la regulación hídrica, control biológico, filtración de contaminantes y la provisión de servicios de polinización [8–11].

En este contexto, preservar y restaurar los bosques naturales se ha convertido en una estrategia clave para mitigar el cambio climático, ya que estos ecosistemas tienen la capacidad de absorber y almacenar grandes cantidades de dióxido de carbono de la atmósfera. Se estima que entre 2001 y 2019, los bosques a nivel global actuaron como sumideros netos, absorbiendo el doble de carbono que emitieron, equivalente a 7.600 millones de toneladas métricas de CO<sub>2</sub> por año. Por lo tanto, el manejo forestal soste-

nible y la expansión de la cobertura boscosa son estrategias de mitigación climática de gran potencial y bajo costo. La gestión sostenible de los bosques no sólo contribuye a la reducción de emisiones de gases de efecto invernadero, sino que también mejora la resiliencia de los ecosistemas y las comunidades ante los efectos del cambio climático. Los bosques bien gestionados pueden aumentar la productividad, conservar la biodiversidad, mantener la calidad del suelo y del agua, y actuar como amortiguadores contra fenómenos climáticos extremos, como inundaciones y sequías.

### **1.2.2. Transformaciones y fragmentación de los bosques**

A pesar de su importancia, los bosques sufren transformaciones negativas y se estima que 13 millones de hectáreas de bosques se destruyen en todo el mundo cada año. Entre 2000 y 2005 se perdieron 4.7 millones de hectáreas de superficie forestal, esto representa el 65 % de la deforestación mundial. Se sabe que la deforestación es responsable de una cuarta parte de las emisiones mundiales de gases de efecto invernadero, lo que causa que nuestro planeta mantenga una temperatura elevada, otra de las causas es la expansión agrícola y el uso del suelo. Estas transformaciones que sufren los bosques pueden ser caracterizadas por el nivel de destrucción y modificación del hábitat, la tasa de deforestación y los patrones de fragmentación [12, 13]. Conocer el estado de transformación de un bosque es un antecedente decisivo para una correcta planificación, manejo y restauración del paisaje, ya que esto nos indicará los principales atributos que se deben preservar o mejorar para la conservación de la biodiversidad y el bienestar humano. Es posible identificar paisajes poco modificados con una tasa baja de deforestación; estos tienen áreas extensas y conectadas de bosques nativos. Por otro lado, existen paisajes con una alta fragmentación, donde la cobertura boscosa ocupa un área reducida del paisaje, y los fragmentos de hábitat se encuentran dispersos, aislados y sin conexión entre ellos [14, 15]. La pérdida de biodiversidad suele traer efectos negativos en los procesos ecosistémicos, reduciendo los bienes y servicios que ofrece [16]. A pesar de que los bosques pueden mostrar resiliencia, esto no los excluye de verse afectados por el cambio de uso de la tierra, la deforestación, la intensificación agrícola y la urbanización. Estas presiones pueden comprometer su capacidad de recuperación y adaptación natural. La fragmentación del hábitat no solo afecta la biodiversidad, sino que también interfiere en los procesos ecológicos fundamentales, alterando las dinámicas de los ecosistemas y reduciendo su funcionalidad.

### 1.2.3. Disminución de los bosques nativos y servicios ecosistémicos

Un problema significativo y preocupante a nivel mundial es la disminución de bosques nativos y los servicios ecosistémicos que estos proporcionan [7, 17] a costa de la introducción de especies exóticas. Esto se debe a la baja rentabilidad del manejo forestal en comparación con otros usos del suelo, como la forestación de pino y la ganadería [18], así como a la falta de tecnologías que permitan a los productores utilizar de manera sostenible el bosque nativo desde una perspectiva económica y ambiental. Aunque diversos productos forestales, como la madera y la leña, tienen valor de mercado y pueden compararse directamente con otros usos del suelo, existen servicios ecosistémicos de regulación, como el control de la erosión del suelo y la polinización, que no tienen un valor de mercado. La ausencia de valor de mercado presenta un reto para desarrollar tecnologías de manejo que ofrezcan un retorno económico a los productores y, al mismo tiempo, proporcionen servicios ecosistémicos a la sociedad [19]. Es decir, la falta de incentivos económicos adecuados para la conservación de los bosques nativos puede llevar a su degradación y pérdida, lo que a su vez afecta la biodiversidad y la capacidad de estos ecosistemas para proporcionar servicios vitales.

### 1.2.4. Los bosques andino-patagónicos

En Argentina, los bosques andino-patagónicos son considerados entre los ecosistemas de mayor prioridad para la conservación y están presentes en numerosas áreas naturales protegidas [20]. La investigación en la ecología de estos bosques es extensa y destaca su importancia como reserva de vida silvestre con un gran potencial productivo [21–24]. Sin embargo, los bosques multi-específicos han sido poco explorados y la falta de información es un problema estructural en los matorrales de la Patagonia Norte Argentina. A pesar de que los matorrales actualmente producen bienes comerciales como madera, leña y forraje para la ganadería en la zona andina, la provisión de servicios ecosistémicos (SE) asociados a su diversidad podría mejorarse mediante una gestión adecuada. Los SE pueden ser divididos en tres categorías: de provisión (como alimentos, materiales y leña), de regulación (como polinización, regulación climática y control de la erosión) y culturales (como valor estético, recreativo y espiritual) [25, 26]. En muchos países, incluido Argentina, aún no está claro cómo manejar los bosques multi-específicos para garantizar la provisión sostenible de bienes de mercado y biocombustibles.

### 1.2.5. Prácticas de manejo forestal sostenible

En Patagonia Norte, se encuentran entre los principales cambios realizados por el ser humano, el reemplazo de los bosques por sistemas silvopastoriles, es decir por sis-

temas que combinan: árboles, arbustos forrajeros y pastos con la producción ganadera, como así también las plantaciones mono-específicas de coníferas exóticas. Éstas son plantaciones no nativas de rápido crecimiento, que se utilizan para el aprovechamiento leñero y la urbanización y que además generan cambios de cobertura o superficies del bosque nativo [27]. En el caso de ecosistemas forestales de bajo valor maderero, como los bosques de tipo matorral, una práctica posible para lograr un manejo forestal sustentable sería combinar la extracción de leña con plantaciones de árboles de interés productivo. Este sería un procedimiento adecuado, optimizando los costos al implementar cosechas sistemáticas para comercializar productos como leña y postes, en lugar de realizar podas no controladas. Entre las opciones forestales, una práctica alternativa de mitigación y/o práctica forestal superadora para revalorizar los recursos propios y mantener la biodiversidad sería la plantación de especies no invasivas y nativas [28, 29]. Estas plantaciones podrían agregar valor de mediano y largo plazo si se plantan a densidades que permitan la persistencia del matorral y la reducción de su combustibilidad. La forestación con especies nativas de alto valor está en línea con la política productiva (LN 25.080) y ambiental (LN 26.331) nacional y provincial (LRN 4.552).

### 1.2.6. Interacciones ecológicas en los bosques

En la naturaleza, diversos procesos ecológicos se asocian a una mayor regeneración y crecimiento del bosque y/o fomentan mayor diversidad en los siguientes niveles tróficos. Es necesario realizar un análisis que integre las respuestas simultáneas de plantas y diferentes grupos de invertebrados y microorganismos al manejo de la vegetación. Si bien hace años se ha introducido el concepto de redes de interacción [30] y se han realizado trabajos a lo largo de los últimos 10 años, poco se conoce sobre que sucede en algunos sitios de la Patagonía Argentina [31–34]. Es fundamental abordar esta metodología proporcionando información relevante a nivel local, regional y global.

Se ha estudiado que la riqueza de especies en sistemas productivos, además de su valor intrínseco, fomenta los servicios ecosistémicos [35, 36]. A pesar de que se ha estudiado sobre las consecuencias de los cambios en las redes de interacciones planta-insecto en el funcionamiento de los sistemas productivos [37, 38]. En la Patagonia Norte aún hay áreas que requieren mayor investigación, especialmente en su relación con los sistemas productivos y la conservación de la biodiversidad.

La herbivoría es una de las interacciones ecológicas más importantes en los bosques y se define como el proceso donde los animales consumen tejidos vegetales vivos, por ejemplo, hojas, tallos, raíces o frutos. Esta interacción es asimétrica, ya que el herbívoro se ve beneficiado, mientras la planta es perjudicada. En la Patagonia Norte, el tipo de herbivoría más estudiada en los sistemas forestales es la de herbívoros mamíferos [23],

este es un tipo de interacción muy diferente a la de artrópodo-árbol.

Los artrópodos son animales invertebrados que componen un cuarto de los organismos existentes y consumen entre el 5 y el 18 % de la cantidad de hojas de una planta, ya sea hojas secas o frescas [39, 40]. La herbivoría de artrópodos no sólo contribuye a la biodiversidad del bosque [41], sino que también puede convertirse en plagas y afectar el rendimiento de las plantas debido al consumo de hojas o la dispersión de enfermedades [42].

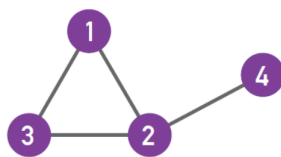
Otra interacción clave es la interacción entre planta-polinizador. Los polinizadores son animales que, al visitar las flores para alimentarse, transportan el polen de una flor a otra. Cumplen un rol fundamental en la reproducción de las plantas, al iniciarse así el desarrollo de las semillas y los frutos. Por este motivo son esenciales en los ecosistemas ya que están involucrados en la fertilización de alrededor del 90 % de especies de angiosperma [43, 44] y del 70 % de especies de cultivos en el mundo [45]. En las últimas décadas existe una creciente preocupación por la pérdida de polinizadores y sus servicios [43, 46], y una mayor focalización en las redes de interacciones para el diagnóstico y recomendaciones de manejo de los ecosistemas [35, 47, 48]. Perturbaciones causadas por fuego, sobrepastoreo o desmonte pueden causar distintos efectos sobre los polinizadores como reducir los sitios de anidamiento y disponibilidad de flores [44, 49] o, por el contrario, favorecer la floración de las especies del sotobosque [50], incrementando su abundancia.

### 1.3. Especies clave

En el ámbito de la ecología se utiliza el concepto de especie clave [51, 52]. Éste se refiere a una especie cuyo impacto en su ecosistema es considerablemente superior al esperado teniendo en cuenta su biomasa o abundancia. Estas especies desempeñan un rol fundamental en la estructura y funcionalidad de sus comunidades, y su presencia o ausencia puede tener efectos significativos en la biodiversidad y la salud del ecosistema. Ejemplos de especies clave son los depredadores que pueden regular las poblaciones de herbívoros, lo que a su vez permite el crecimiento de la vegetación y la conservación de la biodiversidad.

El tipo de interacción asociado a estas especies clave presenta un espectro amplio abarcando la depredación, la competencia y la simbiosis. Por ejemplo, las especies polinizadoras, como las abejas, son cruciales para la reproducción de muchas plantas, lo que afecta la producción de alimentos y la diversidad de especies vegetales. La desaparición de una especie clave puede llevar a cambios drásticos en la estructura del ecosistema, produciendo desequilibrios, sobrepoblación de ciertas especies y disminución de otras con la consiguiente pérdida de biodiversidad [45, 53, 54].

Teniendo en cuenta que las especies clave son esenciales para el mantenimiento de



**Figura 1.1:** Representación de una red no dirigida con 4 nodos

la estructura y la funcionalidad de los ecosistemas, su conservación es fundamental para preservar la biodiversidad y la resiliencia de los ecosistemas frente a cambios ambientales. La pérdida de estas especies puede conducir a la degradación del hábitat y a la disminución de la biodiversidad, lo que resalta la necesidad de estrategias de conservación efectivas que prioricen la protección de estas especies y sus hábitats. Por lo mismo, es importante poder identificarlas a fin de desarrollar estrategias óptimas para su preservación. Veremos más adelante que el rol de las especies clave es esencial al momento de restaurar ecosistemas degradados.

## 1.4. Representación de las redes

Las redes pueden ser representadas mediante un grafo. Sea  $G$  un grafo donde  $G = (V, E)$ , donde  $V$  es el conjunto de vértices, a los cuales llamaremos nodos y  $E$  es el conjunto de enlaces entre los nodos, es decir,  $E \subset V \times V$  [5, 55].

**Definición:** Se dice que un grafo es simple cuando no tiene múltiples enlaces y loops. Un loop se crea cuando un enlace involucra un único vértice. Decimos que un grafo tiene múltiples enlaces si dos vértices comparten más de un enlace.

**Observación:** El máximo número de enlaces en un grafo de  $N$  nodos es igual a  $\binom{N}{2}$ .

**Definición:** Sea  $V = \{v_1, v_2, \dots, v_n\}$  los vértices de un grafo. Un camino es una secuencia finita de números de la forma

$v_1v_2, v_2v_3, \dots, v_{n-1}v_n$ , esto también se puede escribir como  $v_1 \rightarrow v_2 \rightarrow v_3 \rightarrow \dots, v_n$ , donde  $v_1$  es el nodo inicial y  $v_n$  es el nodo final. Al número de enlaces se lo denomina longitud.

**Definición:** La distancia entre un par de nodos se define como el camino más corto que conecta a esos nodos. Si dos nodos no están conectados se dice que la distancia es infinita.

**Definición:** Se dice que dos nodos son vecinos, si comparten un enlace. En el ejemplo de la figura 1.1, los vecinos del nodo 2 son los nodos 1, 3 y 4.

**Definición:** Se define el coeficiente de *clustering* de un grafo como  $C_i = \frac{2e_i}{k_i(k_i-1)}$  siendo  $k_i$  el número de vecinos del nodo  $i$  y  $e_i$  el número de enlaces entre los vecinos del nodo  $i$ .

**Definición:** Una componente conectada es un subgrafo donde todos los nodos están conectados y existe un camino entre todos los nodos del subgrafo, además ningún nodo del subgrafo está conectado con un nodo de otro subgrafo.

Una red está definida por un conjunto de nodos que representan diferentes grupos, por ejemplo, en una red de planta-polinizador tendremos 2 grupos de nodos, uno de polinizadores y otro de plantas y las relaciones entre ambos grupos van a estar representadas por medio de sus enlaces/vínculos. Cuando las redes son dirigidas, los enlaces tienen una dirección específica. En cambio, en las redes no dirigidas, los enlaces son bidireccionales, es decir, no tienen una dirección única. Los enlaces entre nodos pueden ser, por lo tanto, dirigidos o no dirigidos.

Cada red puede ser representada mediante una matriz, a la que llamaremos matriz de adyacencia ( $A$ ). Ésta resulta ser una matriz cuadrada, en el caso de tener  $N$  nodos la matriz será de  $N \times N$  y el elemento  $A_{ij}$  representa el enlace que une los nodos  $i$  y  $j$ . Además esos enlaces pueden tener asociado un peso, en nuestro ejemplo el grafo es sin peso, por lo tanto  $A_{ij} = 1$  si existe un enlace entre el nodo  $i$  y el nodo  $j$ , y  $A_{ij} = 0$  en caso contrario.

A modo de ejemplo a continuación mostramos la matriz de adyacencia del grafo de la figura 1.1.

$$\begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

De manera general, una red multicapa es un par  $M = (G, C)$  donde  $G = \{G_\alpha : 1 \leq M\}$  es una familia de capas y  $C$  es el conjunto de enlaces entre capas [56, 57]. A su vez, cada capa es un par  $G_\alpha = \{X_\alpha, E_\alpha\}$ , donde  $X$  es el conjunto de nodos de la capa y  $E$  el de los enlaces entre ellos. El conjunto de enlaces entre capas se define como  $C = \{E_{\alpha\beta} \subseteq X_\alpha \times X_\beta; \alpha, \beta \in 1, \dots, M, \alpha \neq \beta\}$ . Existe una gran variedad de redes multicapas entre las que se destacan dos grandes familias: redes multiplex y red de redes. Las primeras son redes donde cada capa tiene el mismo conjunto de nodos aunque la interacción entre ellos puede ser diferente en cada una. El ejemplo más trivial es el de dos nodos conectados en una capa, pero desconectados en otra. Por otro lado, una red de redes está formada por redes que están entrelazadas entre sí, pero que no contienen necesariamente los mismos nodos.

### 1.4.1. Caracterización de Redes

Para poder comprender la estructura de las redes se las ha caracterizado usando diferentes medidas de centralidad de los nodos, como el grado, la *betweenness* y

propiedades globales como el *clustering*, la asortatividad, la distribución de grado y la modularidad. Es esencial definir cada uno de estos conceptos para entender cómo contribuyen al análisis de las redes.

- Grado de un nodo: El grado ( $k$ ) de un nodo se define como el número de enlaces que ingresan o salen de un nodo. En el caso de una red dirigida existen dos tipos de grados:
- *In degree* ( $k_i$ ): Es el número de enlaces que ingresan al nodo.
- *Out degree* ( $k_o$ ): Es el número de enlaces que salen del nodo.

El grado total de un nodo es la suma entre  $k_i$  y  $k_o$ , es decir,  $k = k_i + k_o$ .

Si la red no es dirigida el grado de un nodo se define como la cantidad de vecinos que tiene ese nodo, es decir, la cantidad total de enlaces que tiene ese nodo.

- Densidad: La densidad ( $d$ ) es la razón entre la cantidad de enlaces que tiene la red ( $n$ ) y la cantidad de enlaces totales que puede admitir la red ( $\frac{N(N-1)}{2}$ ).

$$d = \frac{2n}{N(N-1)} \quad (1.1)$$

- Distribución de grado: Es la probabilidad que describe cómo se distribuyen los grados de los nodos en una red. La distribución de grado se representa por medio de una función de distribución que puede ser Binomial o de Poisson para una red aleatoria o en el caso de una red de escala libre sigue una ley de potencia.
- *Betweenness*: A veces llamada intermediación, es una medida de centralidad. Cuenta los caminos que pasan a través de un nodo y que conectan partes de la red.
- Modularidad: La modularidad cuantifica la intensidad de los enlaces entre comunidades vs los enlaces intra comunidades y también ayuda a caracterizar la estructura de la comunidad.
- Comunidades: Llamaremos comunidad a un subconjunto de nodos.
- Asortatividad: Se dice que una red es **asortativa** cuando los nodos con un alto grado de conexión están, en promedio, conectados a otros nodos con un alto grado de conexión, y los nodos con un bajo grado de conexión están, en promedio, conectados a otros nodos con un bajo grado de conexión. Por otro lado, se dice que una red es **disasortativa** cuando, en promedio, los nodos con un alto grado de conexión están conectados a nodos con un grado de conexión más bajo y, en promedio, los nodos con un bajo grado de conexión están conectados a nodos con un grado de conexión más alto.

- Anidamiento: El anidamiento describe un patrón de interacción en el cual las especies con menos interacciones tienden a relacionarse con un subconjunto específico que a su vez se relaciona con las especies más conectadas [58].

Con una comprensión clara de los índices anteriores, es posible aplicar distintos modelos y algoritmos que han sido desarrollados para generar redes con características específicas. Por ejemplo el modelo de Albert y Barabási para redes aleatorias de escala libre o redes de mundos pequeños. Particularmente interesante es el análisis de detección de comunidades de una red, que utiliza diferentes algoritmos basados en la maximización de la modularidad para una partición en comunidades. La mayor parte de los estudios se aplicaron inicialmente a redes de una sola capa, pero se han generalizado para redes multicapas en los últimos años. Dado que éste es un campo novedoso, en el que han aparecido en poco tiempo múltiples técnicas e índices, los avances matemáticos y estadísticos que se desarrollen en el futuro reciente van a ser de gran utilidad para poder seguir creando herramientas y modelos que permitan un avance cualitativo en el conocimiento de la estructura de las redes.

Si bien muchos aspectos de las redes fueron estudiados, hay un aspecto que no fue abordado exhaustivamente desde un punto de vista matemático, que es la complejidad de las redes. La palabra complejidad puede ser usada de diferentes formas dependiendo del área de estudio. En el área de ecología, muchas veces se suele usar para describir el estado de un ecosistema y puede ser asociada a un solo término como diversidad y resiliencia ecológica, pero no es suficiente considerar un índice para determinar la complejidad, será necesario comparar y contrastar diferentes medidas de complejidad. En el análisis de las redes, la complejidad puede describir la organización de los componentes del sistema y la relación entre ellos.

Existen diferentes índices para estudiar la complejidad de una red, entre ellos se encuentran, la modularidad, la entropía, K-complejidad y la Entropía SVD (*Singular Value Decomposition*). Muchos índices actualmente usados en el área de ecología fueron adaptados de otras áreas como física y teoría de la información. El enfoque usado en esta tesis tiene como base la teoría de la información y puede ser aplicado a diferentes tipos de redes. Diremos que una red es compleja si contiene información codificada en su sistema.

La entropía se define como

$$H(X) = - \sum_{i=1}^N P(x_i) \cdot \log(P(x_i)), \quad (1.2)$$

donde  $P(x_i) = \frac{\text{nodos en la comunidad } i}{\text{nodos totales}}$ ,  $N$  es el número de comunidades de la red, donde  $X = \{x_1, \dots, x_N\}$  es la variable aleatoria. Otro de los índices que puede ser usado para estudiar la complejidad de la red es la modularidad. Este índice cuantifica la intensidad

de los enlaces entre comunidades vs los enlaces intra comunidades y también ayuda a caracterizar la estructura de la comunidad.

El índice Entropía SVD (*Singular Value Decomposition*) se calcula factorizando la matriz de incidencia  $A$  donde  $A_{m,n} \in \mathbb{R}$ , una de las matrices resultantes de esta factorización es una matriz diagonal y usando los valores de esa matriz diagonal se obtiene el índice buscado, esto se desarrolla en detalle en el Capítulo 3 de la tesis. Otro de los índices es la complejidad de Kolmogorov (K-complejidad) y una forma de aproximar este valor es usando el algoritmo de probabilidad presentado por Solomonoff y Levin [59, 60].

Para detectar comunidades en redes complejas, se utilizan diversos algoritmos que se basan en diferentes enfoques y principios. Algunos de los más destacados son:

#### 1. Algoritmo de Louvain

Este algoritmo es uno de los más populares para la detección de comunidades. Se basa en la maximización de la modularidad, que mide la densidad de conexiones dentro de las comunidades en comparación con las conexiones entre ellas. El algoritmo funciona en dos fases: a) Fase de agrupamiento local: Cada nodo se asigna a la comunidad que maximiza la modularidad; b) Fase de agregación: Se construye una nueva red donde las comunidades detectadas se tratan como nodos, y se repite el proceso.

#### 2. Algoritmo de Girvan-Newman

Este método utiliza la eliminación de enlaces para identificar comunidades. Comienza calculando la “intermediación” de cada enlace, que mide cuántos caminos cortos pasan por él. Los enlaces con la mayor intermediación se eliminan iterativamente, y la red se divide en comunidades cuando ya no hay más enlaces que conecten diferentes grupos.

#### 3. Detección de comunidades basada en propagación

Estos algoritmos se basan en la idea de que los nodos se agrupan a través de un proceso de difusión. Un ejemplo es el algoritmo Label Propagation, donde cada nodo adopta la etiqueta más común de sus vecinos en iteraciones sucesivas, lo que permite que los nodos se agrupen en comunidades.

#### 4. Algoritmos de optimización de la modularidad

Además del algoritmo de Louvain, existen otros métodos que optimizan la modularidad, como el método de optimización de la modularidad de Blondel et al. [61], que también sigue un enfoque jerárquico y puede ser aplicado a redes de múltiples capas.

## 5. Métodos espectrales

Estos algoritmos utilizan propiedades espectrales de la matriz de adyacencia de la red. Por ejemplo, el uso del espectro de la matriz de Laplace permite identificar comunidades al observar los valores propios y vectores propios, que pueden indicar la estructura de la red.

## 6. Algoritmos de *clustering* jerárquico

Estos métodos construyen un árbol de decisiones (dendrograma) que permite agrupar nodos en diferentes niveles de granularidad. Se pueden utilizar medidas de similitud o distancia para decidir cómo agrupar los nodos. Cada uno de estos algoritmos tiene sus ventajas y desventajas, y la elección del método adecuado puede depender de las características específicas de la red que se esté analizando, así como de los objetivos del estudio.

## 7. Algoritmo de módulos bipartitos inducidos recursivamente (BRIM)

Este algoritmo fue propuesto por Barber en 2007 [62] y busca optimizar la modularidad asignando nodos a diferentes comunidades de manera sucesiva, con el objetivo de maximizar la contribución individual de cada nodo a la modularidad, teniendo en cuenta las asignaciones previas. En este proceso, cada conjunto de nodos induce recursivamente al otro conjunto de nodos. El algoritmo BRIM realiza estas asignaciones de nodos de cada tipo a comunidades hasta que se alcanza un valor máximo de la función.

## 8. Algoritmo Infomap

Este algoritmo identifica comunidades o módulos en redes complejas al optimizar una función. En redes ecológicas Infomap encuentra grupos de especies que están fuertemente conectadas entre sí, pero menos conectadas con otras especies. El algoritmo busca la partición de la red que minimiza la cantidad de nodos dentro de un modulo para describir las interacción dentro de la red, es decir, cómo las interacciones se distribuyen y se mueven entre las especies. Este se basa en una versión modificada del algoritmo de Louvain [63].

## 9. Algoritmo de Leiden

El algoritmo de Leiden busca optimizar la modularidad de una red. Asigna nodos a comunidades de manera inicial, luego realiza una optimización local de la modularidad, similar a Louvain, pero asegurándose de que las comunidades estén más conectadas entre sí. Finalmente refina las particiones mediante un proceso iterativo garantizando la estabilidad de las comunidades y su conectividad.

Nosotros haremos uso del algoritmo de Louvain.

## 1.5. Hipergrafos

Abordar el análisis con hipergrafos es fundamental cuando se necesita modelar y estudiar relaciones más complejas que las que permiten los grafos tradicionales [64, 65]. A diferencia de los grafos, que sólo permiten modelar relaciones entre dos nodos, los hipergrafos permiten representar interacciones entre más de dos entidades simultáneamente. Otra de las utilidades de los hipergrafos es representar interacciones indirectas, ya que las redes bipartitas no nos dan ese tipo de información. Esas interacciones en los hipergrafos se representan por hiperenlaces, los cuales conectan nodos que comparten una característica en común. Esto puede ser relevante en muchos contextos donde las relaciones entre más de dos elementos son comunes.

Gracias a los hipergrafos es posible expresar relaciones multidireccionales entre nodos. Por ejemplo, en áreas como la biología se pueden representar las redes de interacción proteína-proteína como un hipergrafo, mientras que en las ciencias sociales resultan útiles para representar redes de colaboración. Si la red de colaboración es estudiada como un grafo simple, sólo se puede identificar si los científicos han colaborado entre sí, pero no sabemos si un grupo de tres o más autores que aparecen vinculados fueron coautores del mismo artículo. Una forma de abordar este problema puede ser representando la red como un grafo bipartito, en este caso tenemos dos conjuntos de nodos, los autores y los nombres de los artículos. En este caso se pierde la homogeneidad porque cada conjunto de nodos tiene un significado diferente (autores versus nombre de artículos). Esto puede causar problemas en el análisis de propiedades topológicas como la densidad y el agrupamiento, ya que se asume que todos los nodos representan lo mismo. Los índices como la centralidad de grado o de intermediación pueden verse distorsionadas porque los nodos de un tipo (por ejemplo, artículos) no interactúan entre sí, lo que afectaría los cálculos. Al introducir diferencias entre dos clases de nodos podría distorsionar los resultados. Es por esto que surge la necesidad de modelar a la red de colaboración como un hipergrafo, donde los nodos representan a los autores y los hiperenlaces representan los grupos de autores que colaboraron en la publicación de un artículo.

El análisis de hipergrafos se vuelve esencial cuando las interacciones involucran a múltiples nodos simultáneamente. Esto permite una representación más rica y precisa de la estructura de la red, lo cual es clave para entender sistemas complejos en diversas disciplinas como biología, ecología, redes sociales y economía.

### 1.5.1. Características de los hipergrafos

Un hipergrafo se representa como un par  $H = (V, E)$ , donde  $V = \{v_1, v_2, \dots, v_n\}$  es el conjunto de nodos y  $E = \{E_1, E_2, \dots, E_m\}$  es el conjunto de hiperenlaces, siendo

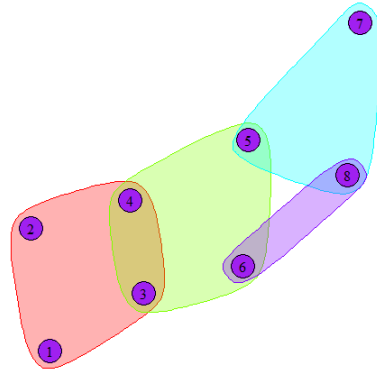
cada  $E_i \neq \emptyset$  y  $\cup_i E_i = V$  [5, 66]. Un hiperenlace  $E_i$  es un subconjunto de nodos. En los hipergrafos se dice que dos nodos son adyacentes en  $H = (V, E)$  si existe un hiperenlace  $E_i$  que contiene a ambos nodos. Decimos que un hipergrafo es simple si se cumple que  $E_i \subseteq E_j$  entonces  $i = j$ . Un grafo simple es un hipergrafo donde cada enlace tiene cardinalidad dos. Los hipergrafos pueden ser representados por una matriz de incidencia  $E(H) = E = [e_{ij}]$  donde cada  $e_{ij} \in \{0, 1\}$ , donde cada una de las  $N$  filas es asociada con un vértice y cada una de las  $m$  columnas es asociado con un hiperenlace, donde  $e_{ij} = 1$  si  $v_i \in E_j$  y  $e_{ij} = 0$  si  $v_i \notin E_j$ . El grado de cada nodo es la cantidad de hiperenlaces a los que pertenece. Si queremos obtener la matriz de adyacencia ( $A$ ) del grafo que da origen al hipergrafo  $H = (V, E)$  el resultado es una matriz cuadrada simétrica cuyas entradas  $a_{ij}$  es el número de hiperenlaces que contienen a ambos nodos  $v_i$  y  $v_j$ , donde la diagonal de la matriz es cero. Esta matriz puede ser obtenida usando la matriz de incidencia de la siguiente forma:

$$A = EE^T - D, \quad (1.3)$$

siendo  $E^T$  la matriz traspuesta de la matriz de incidencia y  $D$  la matriz diagonal cuyas entradas son los grados de cada nodo.

Sea  $E(H)$  la matriz de incidencia de un hipergrafo y en la Figura 1.2 podemos ver su representación gráfica.

$$E(H) = \begin{matrix} & E_1 & E_2 & E_3 & E_4 \\ \begin{matrix} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \\ V_6 \\ V_7 \\ V_8 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{matrix}$$



**Figura 1.2:** Hipergrafo con 8 nodos y 4 hiperenlaces. Su matriz de incidencia es  $E(H)$

Un camino de longitud  $l$  en  $H = (V, E)$  es definido por una secuencia de vértices  $(v_1, v_2, \dots, v_l, v_{l+1})$ , no necesariamente distintos, de manera que para cada  $i = 1, 2, \dots, l$  hay un hiperenlace que contiene a  $v_i$  y  $v_{i+1}$ . Decimos que un camino es cerrado (CW) si  $v_l = v_{l+1}$ .

Un camino simple es un camino en el cual todos los vértices y los hiperenlaces son distintos. Un ciclo es un (CW) en el cual todos los vértices y los hiperenlaces son distintos.

De la misma manera que a los grafos se los puede distinguir en dirigidos y no dirigidos con los hipergrafos podemos hacer lo mismo. En el caso de los hipergrafos no dirigidos, un hiperenlace es un subconjunto de nodos. No hay una dirección entre los nodos. En un hipergrafo dirigido, cada hiperenlace tiene una estructura más compleja, ya que puede dividirse en dos partes, por un lado el conjunto de nodos de origen y un conjunto de nodos de destino. Esto indica una dirección en la relación entre los nodos.

En los hipergrafos también es posible calcular el grado de cada nodo. Sea  $V$  un nodo, su grado será la suma de todas las filas en la que aparece ese nodo en la matriz de incidencia  $H$ , lo que es igual al número de hiperenlaces donde el nodo aparece [67].

De la misma forma que en los grafos podemos definir el coeficiente de *clustering* de un hipergrafo. El coeficiente de *clustering* mide el grado de cohesión que tiene una red. Watts y Strogatz definieron un coeficiente de agrupamiento local que describe la proporción de los conocidos de un vértice que se conocen entre sí [68]. El coeficiente de *clustering* de un hipergrafo se define como:

$$C_2(H) = \frac{6 \times \text{número de hiper-triángulos}}{\text{número de 2-caminos}}, \quad (1.4)$$

donde un hiper-triángulo es una secuencia de tres vértices/nodos diferentes y tres

hiperenlaces diferentes de la forma:  $v_i, E_p, v_j, E_q, v_k, E_r, v_i$ , en la cual los tres nodos son mutuamente adyacentes y un 2-camino es un camino de longitud 2, es decir, es una secuencia como  $v_i, E_p, v_j, E_q, v_k$ . Para calcular el coeficiente de *clustering* se suele usar el grafo asociado ( $G_H$ ) al hipergrafo.

## 1.6. Objetivos de esta tesis

Habiendo introducido el problema a estudiar y la metodología que se usará, describiremos los objetivos de este trabajo. Uno de los objetivos generales apunta a determinar y analizar el nivel de información que está codificada en una red ecológica. Para eso usaremos como base la teoría de la información. Por otro lado se busca diseñar metodologías de recolección de datos de campo, que se puedan aplicar en diferentes entornos ambientales.

Al mismo tiempo, buscamos desarrollar un formalismo basado en la teoría de redes multicapas con el objeto de analizar y caracterizar este tipo de redes asociadas a diferentes procesos ecológicos.

Se plantea estudiar las propiedades y características estadísticas de ciertos indicadores, los cuales esperamos aplicar en los campos de la Patagonia Norte. Entre estos indicadores encontramos los que caracterizan globalmente a la red como una estructura compleja y los que se centran en los nodos (centralidades) para estudiar las propiedades de los nodos individuales dentro de la red. El actual proyecto de doctorado, se encuentra dentro del marco del PICT-2016-0305 “Manejo productivo e interacciones biológicas en matorrales de Río Negro: valoración biofísica, social y económica de servicios ecosistémicos”

Ya a nivel de objetivos específicos, mencionamos

- i) Desarrollar una metodología innovadora para la evaluación de la información codificada en redes ecológicas. Aportar herramientas adecuadas para la caracterización de las redes multicapas que incluyan dos procesos ecológicos (herbivoría y polinización) con tres niveles (plantas, herbívoros y polinizadores), es decir, tres tipos de nodos, de forma simultánea.
- ii) Describir y desarrollar metodología para determinar la complejidad de una red, desarrollando modelos nulos para compararlos con la red original. Asimismo, comparar diferentes índices determinando cuál usar dependiendo de las características de la red a estudiar.
- iii) Construir hipergrafos a partir de grafos bipartitos determinando cuál es el mejor hipergrafo que se ajusta a cada grafo bipartito. Desarrollar una nueva metodología para comparar conjuntos no necesariamente disjuntos de nodos. Comparar

la información nueva obtenida usando la metodología de los hipergrafos con la información obtenida en los Capítulos 2 y 3.

El contenido de esta tesis se encuentra organizado en 5 capítulos, incluido este primer capítulo introductorio, un apartado de bibliografía y apéndices.

El Capítulo 2 corresponde al desarrollo del objetivo (i) de este trabajo: Desarrollar una metodología innovadora para la evaluación de la información codificada en redes ecológicas y caracterizar la red multicapa que incluye dos procesos ecológicos (herbivoría y polinización) con tres niveles (plantas, herbívoros y polinizadores) de forma simultánea. En el mismo se propone una metodología que permite revelar la información codificada de una red a través de un proceso de aleatorización controlada, manteniendo constante el grado de los nodos. Esta técnica se basa en la comparación de la modularidad de la red original con los valores obtenidos de sus versiones aleatorizadas. Al determinar que la red contiene información significativa, se procede a realizar un análisis de comunidades y una caracterización de propiedades topológicas adicionales, con el objetivo de identificar especies clave dentro del ecosistema.

En el Capítulo 3 desarrollamos el segundo objetivo específico: (ii) desarrollar metodología para determinar la complejidad de una red, desarrollando dos modelos nulos para compararlos con la red original. Asimismo, comparar diferentes índices determinando cuál usar dependiendo de las características de la red a estudiar. Aquí se introduce una herramienta novedosa para evaluar la complejidad de redes ecológicas. Se comparan dos índices relevantes: K-complejidad y la entropía mediante descomposición en valores singulares (SVD). Para llevar a cabo este análisis, se utilizaron datos reales junto con dos tipos de modelos nulos. El primer modelo conserva la distribución de grado, mientras que el segundo no lo hace. Ambos modelos se construyen a partir de redes aleatorizadas mediante el intercambio controlado de enlaces de las redes originales. Los resultados indican que K-complejidad es más sensible a las modificaciones en la red, lo que la posiciona como una herramienta más confiable para la determinación de la complejidad en redes ecológicas.

El Capítulo 4 se dedica a responder al último objetivo específico del trabajo: (iii) Estudiar la posibilidad de crear hipergrafos a partir de grafos bipartitos determinando cuál es el mejor hipergrafo que se ajusta a cada grafo bipartito. En este capítulo se crearon hipergrafos para modelar las interacciones ecológicas en la Patagonia Norte, incorporando nodos que representan especies de polinizadores, especies de plantas y gremios de herbívoros. Se creó una función en R para calcular la entropía de los hiperenlaces, lo que permitió evaluar la influencia de cada nodo en el hipergrafo. Este análisis reveló que dos especies de plantas, *R. magellanica* y *E. coccineum*, son las más influyentes, hallazgos que corroboran estudios previos realizados con metodologías diferentes. Además, se presenta un índice innovador que facilita la comparación entre

conjuntos de nodos que no son necesariamente disjuntos.

Para finalizar, en el Capítulo 5 discutimos ampliamente los resultados obtenidos, resumiendo las conclusiones y hallazgos más importantes de esta investigación. Planteamos también algunas posibles líneas de trabajos futuros y concluimos el trabajo final.

## Capítulo 2

# Decodificando información en redes ecológicas multicapas

### 2.1. Introducción

Las redes constituyen una forma eficaz de representación simplificada de sistemas complejos compuestos de muchas unidades interactivas. En esencia, una red captura la topología de las interacciones entre estas unidades a través de una abstracción que facilita el análisis matemático y la interpretación de patrones estructurales. La red básica se compone de nodos y conexiones; sin embargo, al integrar información adicional, como etiquetas en los nodos y pesos en las conexiones, es posible revelar detalles más profundos sobre el sistema subyacente. Aunque esta representación puede conllevar una pérdida parcial de información, las redes ofrecen ventajas significativas al permitir la representación y el análisis de patrones de conexión e interacciones entre las partes de un sistema.

A lo largo de las últimas décadas, la ciencia ha desarrollado un arsenal de herramientas matemáticas, computacionales y estadísticas para el análisis, el modelado y la comprensión de redes. Inicialmente, estas herramientas se enfocaron en representaciones simples, como el cálculo de los nodos más conectados y la distancia entre pares de nodos. Con el tiempo, se han desarrollado modelos de redes más sofisticados que permiten realizar predicciones matemáticas sobre los procesos que tienen lugar dentro de estas estructuras.

En el ámbito de la ecología, los modelos de redes más prevalentes en años recientes incluyen las redes bipartitas y las redes multicapa. En estas representaciones, los vértices pueden corresponder a especies dentro de un ecosistema, y los enlaces reflejan las relaciones entre estas especies. Las redes ofrecen una poderosa herramienta para entender y cuantificar fenómenos ecológicos complejos.

Las redes bipartitas se distinguen en dirigidas y no dirigidas. La principal diferencia

radica en la presencia de una dirección en las interacciones dentro de las redes dirigidas. Las redes bipartitas representan conexiones entre nodos de dos grupos distintos, sin enlaces entre nodos del mismo grupo [5, 69–71]. Esta estructura es especialmente útil para visualizar patrones de reciprocidad en interacciones esenciales para procesos coevolutivos [72] y para el análisis de relaciones tales como parásito-hospedador, planta-polinizador y planta-herbívoro [73–76].

En contraste, las redes multicapa ofrecen una representación más completa y realista de sistemas complejos al incorporar múltiples niveles de conectividad. Estas redes han permitido un entendimiento más profundo de mecanismos y procesos en diversos campos, incluyendo ciencias físicas, sociales, de ingeniería y ecológicas. Los sistemas complejos típicamente incluyen múltiples subsistemas y capas de conectividad que evolucionan, se adaptan y transforman mediante interacciones dinámicas tanto internas como externas, afectando a sus componentes a nivel local y global. Este dinamismo plantea desafíos significativos para observar, comprender, reconstruir y predecir sus dinámicas multiescala y multicomponente.

En los últimos años, ha habido un considerable avance en la comprensión de la estructura y dinámica de estos sistemas complejos. Se han introducido conceptos como redes de redes, redes multidimensionales, redes multiplex, redes interactivas y redes interdependientes, y se han propuesto enfoques matemáticos innovadores para abordar estos desafíos.

Este capítulo se enfoca en superar el paradigma tradicional basado en redes bipartitas para entender el ecosistema como una estructura más compleja, que puede ser representada a través de múltiples niveles interconectados. Específicamente, estudiaremos una red multicapa con dos niveles, es decir, con dos tipos de interacción y emplearemos la proyección de esta red para calcular diversos índices.

El propósito de esta investigación es aplicar el análisis de redes multicapa a problemas ecológicos relevantes, particularmente en el contexto de la restauración de ecosistemas, donde la identificación de especies clave es crucial. Nuestro objetivo principal es combinar la información derivada de la red multicapa con el análisis de sus proyecciones, calculando la modularidad y generando modelos nulos para comparar este índice. La proyección de la red permite calcular índices previamente estudiados y transformar el problema en una red tripartita.

Para alcanzar estos objetivos, caracterizaremos redes construidas a partir de datos de campo y examinaremos su *betweenness* y modularidad. Buscamos determinar si existe información codificada relevante en estas redes en comparación con versiones aleatorizadas. En particular, queremos entender cómo esta información afecta la modularidad de la red, ya que el análisis de la división en comunidades se basará en algoritmos que consideran la modularidad. Además, presentaremos un método sistemático para aumentar gradualmente el grado de aleatoriedad en una red y mostraremos cómo este

proceso afecta a la modularidad al eliminar correlaciones causales derivadas de interacciones ecológicas. Desarrollaremos un algoritmo útil para generar redes aleatorias, el cual se describirá en detalle, y propondremos un protocolo para identificar especies clave.

En concreto nuestros objetivos son (1) determinar si la red presenta una topología con características relevantes que la diferencien de una red aleatoria, (2) identificar las especies clave de la red.

La estructura del capítulo es la siguiente. En primer lugar describiremos la construcción de redes multicapa, incluyendo los detalles sobre los niveles y la proyección de la red. Luego aplicaremos el análisis a un caso específico relacionado con la restauración de ecosistemas para ilustrar la identificación de especies clave. Por último presentaremos los resultados obtenidos, su interpretación y las implicancias para la restauración de ecosistemas.

## 2.2. Metodología

### 2.2.1. Índices para caracterización de la red

A lo largo de este capítulo, examinaremos cómo las interacciones ecológicas significativas entre especies generan correlaciones que distinguen de manera clara una red ecológica real de una red aleatoria. Estas correlaciones se manifiestan como una cantidad de información codificada en el sistema, la cual puede ser evaluada mediante diversos índices y métricas de centralidad.

En una red ecológica real, las interacciones entre especies tienden a seguir patrones que reflejan procesos biológicos y ecológicos subyacentes. Estas interacciones no son aleatorias, sino que están guiadas por factores ecológicos específicos, como relaciones tróficas, competencia, mutualismo y parasitismo. Por tanto, la estructura de la red contiene información que se puede capturar mediante el cálculo de índices de centralidad y otros parámetros estructurales.

Cuando se compara la red ecológica real con una red aleatoria, se espera observar diferencias significativas en los valores de estos índices. En este capítulo, las redes aleatorias se construyen mediante un proceso que altera aleatoriamente las conexiones dentro de una red, sin conservar las correlaciones y patrones específicos presentes en la red original. Este proceso de aleatorización, llevado a cabo mediante modificaciones secuenciales y controladas en la topología de la red, tiende a desordenar las conexiones existentes y eliminar cualquier información codificada que estuviera almacenada en la red original.

En detalle, cuando se transita de una red ecológica real a su versión aleatorizada, se efectúan cambios en la topología de la red que eliminan las correlaciones estructurales

originales. Estos cambios afectan la distribución de los enlaces y, en consecuencia, la estructura global y local de la red. La variación en los índices de centralidad y otros parámetros estructurales a medida que se realizan estos cambios proporciona una medida de la cantidad de información codificada en la red original. Específicamente, al calcular los índices de centralidad en cada etapa del proceso de aleatorización, se observa que los valores varían desde los de la red original hasta aquellos que se aproximan a los valores observados en redes aleatorias con topologías comparables, es decir, redes que comparten la misma distribución de grado.

Este análisis tiene implicaciones importantes para la comprensión de la estructura y dinámica de redes ecológicas. Si una red real mantiene características de centralidad y correlaciones que difieren significativamente de sus versiones aleatorizadas, esto indica que la red contiene información codificada relevante, que refleja interacciones ecológicas no aleatorias y procesos biológicos específicos. Este enfoque no solo permite la evaluación de la estructura de la red, sino que también proporciona una herramienta para identificar patrones de interacción que son fundamentales para la comprensión de los sistemas ecológicos. Es decir, al comparar una red ecológica real con una red aleatorizada, podemos evaluar la cantidad de información codificada en la red original y determinar cómo las interacciones ecológicas significativas influyen en la estructura y funcionalidad de la red. Este análisis permite una mejor comprensión de la red ecológica en su totalidad, revelando las complejidades y dinámicas que las simples redes aleatorias no pueden capturar.

Las redes pueden ser caracterizadas a través de diversos índices que reflejan diferentes aspectos de su estructura y dinámica. Entre estos índices, las medidas de centralidad son particularmente útiles para evaluar la importancia y el papel de los nodos dentro de la red. Existen múltiples medidas matemáticas de centralidad que abordan distintos conceptos, tales como la densidad, la *betweenness* (centralidad de intermediación), el grado de un nodo, y la modularidad, entre otros [56, 57]. Algunas de estas medidas proporcionan una perspectiva local, enfocándose en las características de nodos individuales, mientras que otras ofrecen una visión global de la red, capturando la estructura y las relaciones entre todos los nodos.

Para comparar la red original con sus versiones aleatorizadas en el presente estudio, hemos seleccionado la modularidad como medida clave. La modularidad es una métrica que evalúa la calidad de la división de la red en comunidades. Específicamente, cuantifica la intensidad de los enlaces dentro de las comunidades en comparación con los enlaces entre comunidades. Esta medida permite identificar conjuntos disjuntos de nodos que tienden a compartir propiedades similares o desempeñar funciones similares, basándose principalmente en la topología de la red.

La elección de la modularidad se justifica por su capacidad para ponderar la existencia de correlaciones en la red y extraer valores residuales que permiten evaluar

la estructura de la red en comparación con una red aleatoria. En otras palabras, la modularidad mide hasta qué punto los nodos dentro de una comunidad están más densamente conectados entre sí en comparación con los nodos en otras comunidades. Esta métrica es especialmente útil en el análisis de redes porque ofrece una forma efectiva de caracterizar la estructura de la comunidad dentro de la red.

La modularidad se calcula utilizando varios algoritmos, entre los cuales se destacan el algoritmo de Newman-Girvan [70, 77–79], el algoritmo Infomap [63] y el algoritmo de Louvain [61]. En el presente estudio, hemos optado por el algoritmo de Louvain debido a su eficiencia y capacidad para identificar comunidades con alta modularidad. El algoritmo de Louvain trabaja agrupando nodos en comunidades de manera que maximiza una función de modularidad definida como:

$$Q = \frac{1}{2m} \sum_{i,j} [A_{ij} - \frac{k_i k_j}{2m}] \delta(c_i, c_j) \quad (2.1)$$

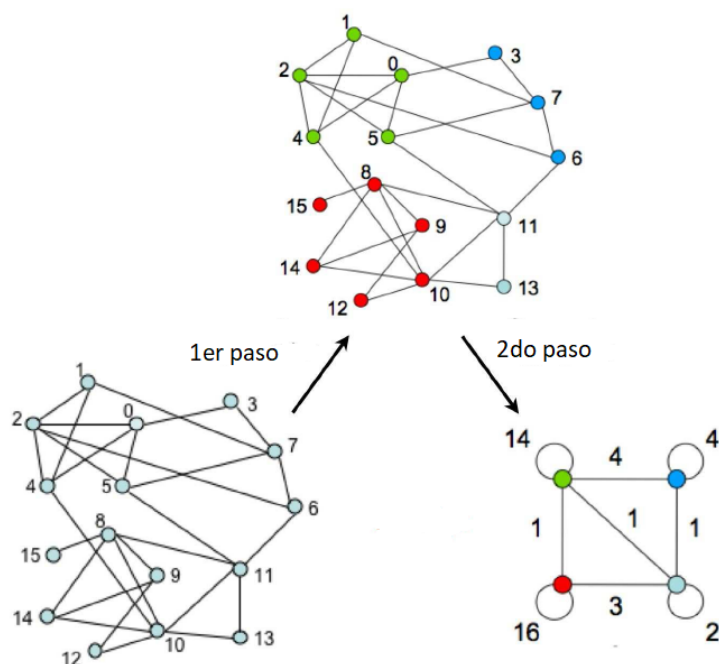
donde  $A_{ij}$  es el peso del enlace entre los nodos  $i$  y  $j$ ,  $k_i$  y  $k_j$  son los grados de los nodos  $i$  y  $j$ ,  $m$  es el número total de enlaces en la red, y  $\delta(c_i, c_j)$  es una función delta que toma el valor 1 si  $i$  y  $j$  están en la misma comunidad y 0 de lo contrario.

El algoritmo de Louvain realiza una agrupación jerárquica de nodos, evaluando iterativamente diferentes particiones de la red y seleccionando aquella que maximiza la función de modularidad. Esta metodología permite identificar comunidades en la red y evaluar la robustez de la estructura de la red a medida que se introduce aleatorización.

El paso 1 de este algoritmo considera inicialmente cada nodo de la red como una comunidad en sí misma. Por lo tanto, si hay  $n$  nodos, inicialmente hay  $n$  comunidades. El algoritmo se basa en la maximización de la función de modularidad, que evalúa la densidad de enlaces dentro de las comunidades en comparación con los enlaces entre comunidades. En la ecuación (2.1)  $A_{ij}$  = representa el peso de los enlaces entre  $i$  y  $j$ ,  $k_i = \sum_j A_{ij}$ , es decir que es el grado del nodo  $i$ ,  $c_i$  = es la comunidad a la que pertenece el vértice  $i$ ,  $\delta(c_i, c_j) = 1$  si  $c_i = c_j$ , 0 si no y  $m = \frac{1}{2} \sum_{ij} A_{ij}$ .

Cada nodo se mueve a la comunidad de uno de sus vecinos si hacerlo incrementa la modularidad. Este proceso se repite iterativamente para todos los nodos hasta que no se pueda mejorar más la modularidad moviendo nodos individuales.

En el paso 2, se arma una nueva red donde cada nodo es una comunidad encontrada en el paso 1. Se generan autoenlaces que corresponden a enlaces intra-comunidad.



**Figura 2.1:** Visualización de los pasos del algoritmo de Louvain. Figura adaptada de Blondel et al. (2008)

El proceso del paso 1 y 2 se repite en esta nueva red y continuamos hasta que ya no hay más cambios en la función de modularidad (Figura 2.1).

Esperamos que, al aumentar el grado de aleatoriedad en la red, la modularidad disminuya de forma monótona. Este comportamiento reflejará la pérdida progresiva de la estructura comunitaria original y la convergencia hacia una red aleatoria, en la que las conexiones entre nodos se distribuyen de manera más uniforme y menos estructurada. Este análisis nos proporcionará una comprensión más profunda de la información codificada en la red ecológica real en comparación con una red aleatorizada.

Además del cálculo de comunidades, es fundamental determinar cuál configuración comunitaria es más adecuada para un conjunto de datos específico. Este proceso de evaluación se facilita mediante el uso de la métrica conocida como Variación de la Información (VI) [80]. La Variación de la Información proporciona una perspectiva teórica robusta para comparar diferentes particiones de una red y evaluar la calidad del clustering en términos de información contenida en las comunidades.

La Variación de la Información es una métrica que se fundamenta en la teoría de la información [81] y mide la cantidad de información compartida entre dos particiones de una red. En esencia, esta medida evalúa cuánta información tiene un cluster sobre otro y la cantidad de información contenida dentro de cada cluster. La VI permite cuantificar las diferencias entre dos agrupamientos, proporcionando una medida objetiva para seleccionar la partición que mejor refleja la estructura subyacente de los datos.

Para formalizar esta métrica, consideremos una partición dada de una red en clus-

ters, denotada como  $\mathcal{C}$ . Sea  $P(k) = \frac{n_k}{n}$ , siendo  $n_k$  la cantidad de elementos que tiene el cluster  $k$  y  $n$  el número total de elementos.

La Variación de la Información se basa en el cálculo de la entropía asociada a la partición  $\mathcal{C}$ , que se define como:

$$H(\mathcal{C}) = - \sum_{k=1}^K P(k) \log(P(k)). \quad (2.2)$$

$VI(\cdot, \cdot)$  mide la incertidumbre o la dispersión de los elementos dentro de los clusters. Una partición que maximiza la entropía puede ser indicativa de una configuración de comunidades menos coherente, mientras que una entropía más baja sugiere una partición en la que los elementos están más concentrados en clusters bien definidos.

Consideremos que  $P(k, k')$  representa la probabilidad de que un punto pertenezca a  $C_k$  en el clustering  $\mathcal{C}$  y  $C'_{k'}$  en el clustering  $\mathcal{C}'$ , la definimos como

$$P(k, k') = \frac{|C_k \cap C'_{k'}|}{n}. \quad (2.3)$$

Además se define la información mutua ( $I(\mathcal{C}, \mathcal{C}')$ ) entre dos clusterings  $\mathcal{C}, \mathcal{C}'$ , es decir, la información que un clustering tiene sobre otro como:

$$I(\mathcal{C}, \mathcal{C}') = \sum_{k=1}^K \sum_{k'=1}^{K'} P(k, k') \log \left( \frac{P(k, k')}{P(k)P(k')} \right). \quad (2.4)$$

Se define la variación de la información como

$$VI(\mathcal{C}, \mathcal{C}') = H(\mathcal{C}) + H(\mathcal{C}') - 2I(\mathcal{C}, \mathcal{C}'), \quad (2.5)$$

$VI(\mathcal{C}, \mathcal{C}')$  es la suma entre la cantidad de información que perdemos sobre la clusterización  $\mathcal{C}$  y la cantidad de información que ganamos cuando pasamos de  $\mathcal{C}$  a  $\mathcal{C}'$ .

El cálculo de la Variación de la Información permite comparar diferentes configuraciones comunitarias y seleccionar la partición que optimiza la separación y la cohesión de las comunidades dentro de la red. En el contexto del análisis de redes ecológicas, esta métrica es particularmente útil para determinar la configuración más adecuada de comunidades que mejor representa la estructura ecológica y las interacciones entre las especies.

### 2.2.2. Aleatorización controlada de la red

El grado de un nodo se define como el número de vecinos que este tiene, y puede calcularse tanto en redes dirigidas como no dirigidas [82].

En el caso de las redes no dirigidas, el grado promedio de la red se obtiene calculando el promedio de los grados de todos los nodos. Este valor proporciona una medida del

nivel general de conectividad en la red. Por otro lado, en las redes dirigidas, existen dos tipos de grados que capturan diferentes aspectos de la conectividad de un nodo:

- *In-degree*: este grado cuenta el número de enlaces que ingresan a un nodo, es decir, el número de conexiones que recibe desde otros nodos.
- *Out-degree*: este grado contabiliza el número de enlaces que salen desde un nodo hacia otros nodos, reflejando la cantidad de conexiones que el nodo emite.

El algoritmo propuesto en este capítulo para desordenar una red conserva el grado de cada nodo. Esto implica que la distribución de grado en la red permanece inalterada, es decir, el número de conexiones o vecinos de cada nodo no se modifica durante el proceso de desordenamiento, manteniendo así la estructura de la red en términos de sus relaciones locales.

Toda red puede representarse matemáticamente a través de una matriz de adyacencia, una estructura que describe las conexiones entre los nodos de la red. A lo largo de este trabajo, nos referiremos a la red de estudio como la red original. Para evaluar las diferencias entre la red original y sus versiones aleatorizadas, es fundamental analizar cómo se comportan ciertas propiedades estructurales, como la modularidad, en función del número de cambios realizados en la red.

Como ya se dijo, la modularidad es una medida clave que refleja el grado en que la red puede dividirse en comunidades o módulos, es decir, grupos de nodos más densamente conectados entre sí que con el resto de la red. En el caso de que la red original difiera significativamente de una red aleatoria, se espera que la modularidad disminuya a medida que la red se aleatoriza. Este decrecimiento es un indicio de la ruptura de las estructuras comunitarias inherentes a la red original, debido a los cambios introducidos en la estructura global de la red.

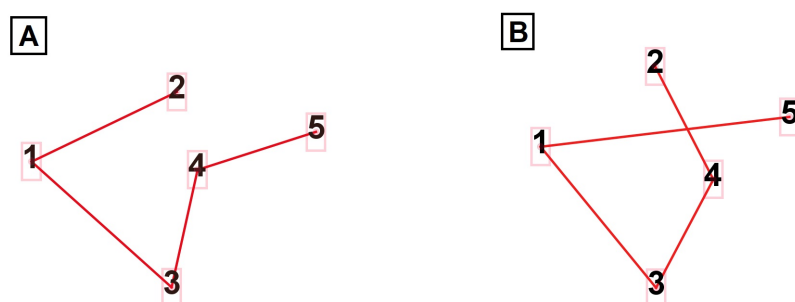
Este comportamiento es particularmente relevante al estudiar redes complejas, donde la preservación o alteración de características como el grado de los nodos y la modularidad puede ofrecer una visión profunda sobre la organización subyacente y la robustez estructural de la red.

Algoritmo para desordenar la red: (Descripción detallada)

Paso 1: Calcular el grado de cada nodo.

Paso 2: Establecer un número de cambios  $\Delta$  el cual nos dará una nueva matriz de adyacencia.

Definimos un cambio: Elegir dos nodos y dos vecinos de esos nodos ( $A_{ij} = 1, A_{i'j'} = 1$ ),  $j$  el cual no debe ser un vecino de  $i'$ , y  $j'$  no debe ser vecino de  $i$  ( $A_{ij'} = 0, A_{i'j} = 0$ ). Como nos interesa preservar el grado de cada nodo, intercambiamos ambos vecinos ( $A_{i'j} = 1, A_{ij'} = 1$ ) (Figura 2.2), de esta forma nos garantizamos que el grado de cada nodo no cambia.



**Figura 2.2:** Representación simplificada de un cambio en una red. (A) Grafo original.  $A_{12} = 1$ ,  $A_{45} = 1$ ,  $A_{24} = 0$ ,  $A_{15} = 0$ . (B) Grafo con un cambio.  $A_{12} = 0$ ,  $A_{45} = 0$ ,  $A_{24} = 1$ ,  $A_{15} = 1$

Paso 3: Calcular el número de enlaces distintos entre la red original y la nueva red. Para esto hemos comparado la nueva matriz de adyacencia con la original.

Paso 4: Calcular la modularidad de la nueva red.

Paso 5: Repetir el paso 2, 3 y 4, varias veces hasta obtener una nueva red incrementando el grado de desorden cada vez. Una vez que la modularidad ha alcanzado un valor estable, no realizamos más cambios.

Siguiendo el procedimiento mencionado anteriormente, las redes obtenidas al incrementar el nivel de aleatorización preservan la distribución de grado. Es importante resaltar que este método nos permite aleatorizar la red de una forma controlada y además nos permite conservar el grado de los nodos. En cada paso, se eligen un par de enlaces de forma aleatoria y se intercambian dos de los nodos adyacentes de cada enlace. Cuando el procedimiento conduce a una conexión duplicada, rechazamos el cambio propuesto y seguimos adelante con el algoritmo. El algoritmo correspondiente fue programado utilizando R [83].

En nuestro caso, para cada valor de  $\Delta$  consideramos 10 redes y su correspondiente modularidad fue promediada. Además, el algoritmo previamente mencionado fue aplicado a redes con más de un nivel de manera de determinar si su estructura fue significativamente diferente de sus versiones aleatorizadas. El número de cambios no está definido *a priori*. Se ha analizado cómo el índice calculado varía cuando aumentamos el desorden y se determinó detener el algoritmo cuando el valor se estabiliza y permanece casi constante a pesar de que el número de cambios aumente. Consideramos que la estabilidad ha sido alcanzada si la modularidad no cambió entre el número de cambios entre  $n$  y su doble  $2n$ . Para determinar ese cambio se realiza un test de hipótesis donde se comprueba si los valores de la red original y los valores entre  $n$  y  $2n$  son significativamente diferentes a nivel 0.05. La red original multicapa fue proyectada,

como lo propone Boccaletti et al. [56], para identificar especies claves usando herramientas conocidas. La cantidad de enlaces de la red proyectada será el resultado de sumar todos los enlaces de todas las capas y la cantidad de nodos será la suma de los nodos de cada capa. La matriz de adyacencia fue construida y en esta instancia incluye 3 tipos de nodos.

### 2.2.3. Caracterizando la topología de la red

En el presente análisis, se ha decidido mantener la distribución de grado de la red original, lo que elimina la necesidad de llevar a cabo un estudio sobre la distribución de grado de sus versiones aleatorizadas. Este enfoque se justifica en que tanto la red original como sus versiones aleatorizadas comparten la misma distribución de grado, lo que sugiere que el grado de cada nodo por sí solo no es suficiente para captar plenamente su papel o relevancia dentro de la red. Sin embargo, es importante destacar que la distribución de grado sigue siendo una herramienta valiosa para evaluar si una muestra puede estar sesgada o incompleta. La distribución de grado también nos permite identificar la presencia de hubs dentro de la red y comprender su influencia. Las redes ecológicas pueden seguir una distribución de grado que obedece a una ley de potencia pero esto no fue probado para todas las redes [84–86]. Existen evidencias que sugieren que las redes biológicas pueden ser tanto de escala libre como jerárquicas, como se ha documentado en estudios previos [87–89]. A pesar del número relativamente bajo de nodos en nuestra red ecológica, se asume que su distribución de grado no debería diferir significativamente de una típica distribución con forma de ley de potencia.

Es relevante mencionar que aunque una red aleatoria no necesariamente exhibe esta característica común a las redes biológicas, las redes resultantes del proceso de aleatorización propuesto sí mantendrán esta propiedad, dado que se conserva la distribución de grado original. Al conservar la distribución de grado, se preservan los patrones de conectividad de la red, lo que facilita una caracterización precisa de los nodos y sus interacciones. Este enfoque permite obtener una visión más clara de cómo funciona el ecosistema en su conjunto.

Además, el estudio detallado de estas características topológicas puede ofrecer comprensión sobre cómo las perturbaciones en la red afectan su estructura y funcionalidad. Por ejemplo, investigaciones recientes han demostrado que las redes tróficas presentan patrones similares a los observados en otras redes complejas, lo cual sugiere una cierta universalidad en los principios organizativos y funcionales dentro de sistemas complejos [90]. Esto abre nuevas avenidas para explorar cómo diferentes tipos de interacciones dentro de las redes ecológicas pueden influir en su resiliencia y adaptabilidad ante cambios ambientales.

Barabási y Albert [91] mostraron que la distribución de grado de muchas redes

reales se aproximan por una ley de potencia con la forma

$$P(\text{grado} = k) \approx k^{-\gamma}, \quad (2.6)$$

siendo  $P(k)$  la probabilidad de encontrar un nodo con grado  $k$  y  $\gamma$  denota el exponente de la ley de potencia.

La ecuación (2.6) indica que la probabilidad de encontrar un nodo con grado  $k$  creciente disminuye como la potencia de  $\gamma$ . Este decrecimiento más lento significa que la probabilidad de encontrar nodos con  $k$  grande es mayor en la ley de potencia en comparación con otras distribuciones. En la mayoría de los sistemas de ley de potencia,  $2 < \gamma < 3$  [92]. Los valores más altos de  $\gamma$  indican una pendiente más pronunciada en la distribución, un régimen de ley de potencia más restringido y una menor probabilidad de encontrar nodos más conectados. En la distribución de la ley de potencia, la mayor parte de los nodos tienen un grado bajo pero también tenemos nodos con altos valores de  $k$ .

Clauset [93] propuso un algoritmo para determinar si la distribución de grado de un conjunto de datos sigue una ley de potencia. Este procedimiento se puede desglosar en los siguientes pasos:

- **Estimación de parámetros:** En esta fase inicial, se deben estimar los parámetros  $k_{min}$  y  $\gamma$  de la ecuación (2.6). El parámetro  $k_{min}$  representa el valor mínimo del grado considerado, mientras que  $\gamma$  es el exponente que define la pendiente de la distribución.
- **Evaluación de la bondad de ajuste:** Una vez que se han estimado los parámetros, se procede a evaluar la bondad de ajuste entre los datos observados y la ley de potencia utilizando el estadístico de Kolmogorov-Smirnov (K-S). Este test no paramétrico permite medir la discrepancia entre la función de distribución empírica y la función de distribución acumulativa esperada bajo la ley de potencia. Si el p-valor resultante es superior a 0.1, se puede considerar que la ley de potencia es una hipótesis válida para describir los datos; en caso contrario, se rechaza esta hipótesis.
- **Comparación con hipótesis alternativas:** En este paso final, se compara la ley de potencia con otras hipótesis alternativas mediante una prueba de razón de verosimilitud (*likelihood ratio*). Para cada alternativa considerada, si el valor calculado del *likelihood ratio* es significativamente diferente de cero, su signo indicará si dicha alternativa es favorecida en comparación con el modelo de ley de potencia.

Este enfoque sistemático permite no sólo determinar si los datos se ajustan a una ley de potencia, sino también evaluar su robustez frente a diferentes modelos alternativos,

contribuyendo así a una comprensión más profunda del comportamiento estructural y funcional del sistema analizado.

Con respecto al paso 1, para estimar  $\gamma$  se usa el método de máxima verosimilitud donde se define la ecuación (2.7) y  $k_{min}$  debe cumplir que  $k_{min} \leq k, \forall k$ .

$$\hat{\gamma} \approx 1 + n \cdot \left[ \sum_{i=1}^n \ln \frac{k_i}{k_{min} - \frac{1}{2}} \right]^{-1}. \quad (2.7)$$

En el segundo paso del análisis, se propone evaluar la bondad de ajuste utilizando el estadístico de Kolmogorov-Smirnov (K-S), una herramienta estadística que cuantifica la credibilidad de la hipótesis planteada. Este método se basa en medir la “distancia” entre la distribución de los datos empíricos y el modelo que se está evaluando. La distancia calculada se compara con las distancias obtenidas de conjuntos de datos simulados, que son generados a partir del mismo modelo. El p-valor se define como la fracción de distancias en los datos simulados que superan la distancia empírica observada. Si el p-valor es alto (cercano a 1), esto sugiere que las diferencias entre los datos empíricos y el modelo pueden atribuirse únicamente a fluctuaciones estadísticas, lo que valida el modelo como un ajuste adecuado. Por otro lado, un p-valor bajo indica que el modelo no es un ajuste posible para los datos observados, lo que sugiere la necesidad de considerar otras distribuciones o modelos alternativos.

En el tercer paso, la idea fundamental detrás de la prueba de *likelihood ratio* (razón de verosimilitud) es calcular la verosimilitud de los datos bajo dos distribuciones en cuestión. La distribución que presenta la mayor verosimilitud se considera como el mejor ajuste para los datos. Alternativamente, se puede calcular el *likelihood ratio*, o su logaritmo, lo cual proporciona una medida más intuitiva sobre qué distribución es más adecuada. El resultado del logaritmo del *likelihood ratio* será positivo o negativo dependiendo de cuál distribución ofrezca un mejor ajuste; un valor cero indicaría un empate entre las distribuciones evaluadas. Este enfoque permite no sólo validar la hipótesis inicial sobre la ley de potencia, sino también explorar y comparar diversas alternativas, enriqueciendo así el análisis y comprensión del comportamiento estructural del sistema estudiado.

Si la distribución de grado de la red original se encuentra demasiado alejada de una distribución de ley de potencia, podría ser que la muestra no sea completa ni insesgada. Por otra parte, el hecho de que la distribución de grado se aproxime a una ley de potencia no garantiza la precisión de la muestra. Utilizar la distribución de grado en estas redes muy pequeñas, es decir, de pocos nodos, para analizar su completitud podría dar lugar a resultados poco fiables. Por consiguiente, se tendrá en cuenta esta limitación de la distribución de grado cuando extraemos información de la misma. En particular, a la hora de identificar especies clave, nos centraremos en la medida de *betweenness*, la cual es una medida de centralidad entre nodos y que ya fue explicada

en la sección anterior.

El problema de identificar nodos que desempeñan un rol central en su estructura es uno de los temas principales del análisis tradicional de redes complejas. Hay muchos parámetros que miden la relevancia estructural de cada nodo, entre ellos el grado, *closeness* y *betweenness* [56, 57]. Nos focalizaremos en dos de ellas: el grado y la *betweenness*. Como ya se ha mencionado, el grado de un nodo puede decir algo sobre su relevancia, pero no deja de ser una medida local y carece de información global sobre la red.

En este capítulo, estamos interesados en identificar los nodos específicos que son responsables de la cohesión de la red, los cuales a partir de ahora serán considerados especies clave. Recordemos que la *betweenness* puede proveer información global valorable. Este índice mide la proporción de caminos más cortos entre todos los pares de nodos que pasan a través de un nodo dado y esos caminos conectan partes de la red [94]. Si una cantidad de nodos con alta *betweenness* son borrados, la estructura de la red podría colapsar. No basta con constatar su papel clave, hay que verificar su importancia para preservar la arquitectura de la red. Finalmente, el número de comunidades es calculado con el objetivo de complementar esta estructura.

El uso del grado para identificar especies claves no necesariamente arroja los mismos resultados como el uso de la *betweenness*. Si bien parece intuitivo relacionar una especie clave con un nodo con grado alto, esto no debería ser así, ya que en este caso no se está considerando la estructura de la red. Nosotros estamos buscando especies que al ser eliminadas de la red el ecosistema colapse. Se recurre a la *betweenness* por ser un índice global. Generalmente la *betweenness* y el grado no están relacionados pero en redes ecológicas podrían estar correlacionadas [95]. Si se elimina un nodo con alto grado puede producir un cambio drástico en la cohesión de la red ecológica, su *betweenness* debe ser también alta. La inversa no es necesariamente cierta.

### Algoritmo para la selección de especies clave:

- Paso 1: Calcular el valor de la *betweenness* para cada nodo.
- Paso 2: Normalizar cada valor del paso 1 y dividirlo por el valor más alto de la *betweenness*.
- Paso 3: Ordenar los valores desde los más altos hasta los más bajos.
- Paso 4: Graficar la *betweenness*  $y_i$  en función del orden  $x_i$ , obteniendo una curva  $f(x_i) = y_i$ . La curva resultante presenta un punto de máxima pendiente en un valor específico  $x_c$ , que se define como  $\alpha_c = f(x_c)$ .
- Paso 5: Elegir un valor de  $\alpha \leq \alpha_c$ , de tal forma que la región de la máxima pendiente esté cubierta.

- Paso 6: Calcular el número de valores más grandes que  $\alpha$ .  $j =$  Número de valores más grandes que  $\alpha$ .  $\alpha \in \mathbb{R}$ .
- Paso 7: Borrar sólo el nodo  $i$  de la lista. Luego, calcular el número de comunidades y componentes conectadas.
- Paso 8: Repetir el paso 6 para  $i \in \{1, \dots, j\}$ .
- Paso 9: Armar una tabla con el número de comunidades y componentes conectadas antes y después de borrar cada especie.

Para llevar a cabo el análisis de la *betweenness*, se elabora una lista con los valores de la *betweenness* normalizada, ordenándolos desde el más alto hasta el más bajo. A continuación, se grafica la *betweenness*  $y_i$  en función del orden  $x_i$ , lo que permite obtener una curva  $f(x_i) = y_i$ .

Esta representación gráfica es fundamental para identificar patrones en la distribución de la *betweenness* entre los nodos de la red. La curva resultante presenta un punto de máxima pendiente en un valor específico  $x_c$ , que se define como  $\alpha_c = f(x_c)$ . Los puntos situados por encima de este valor son considerados relevantes y se agrupan en lo que denominamos la “región de interés”. Para que un nodo sea considerado significativo dentro de este análisis, debe cumplir con la condición  $\alpha \leq \alpha_c$ .

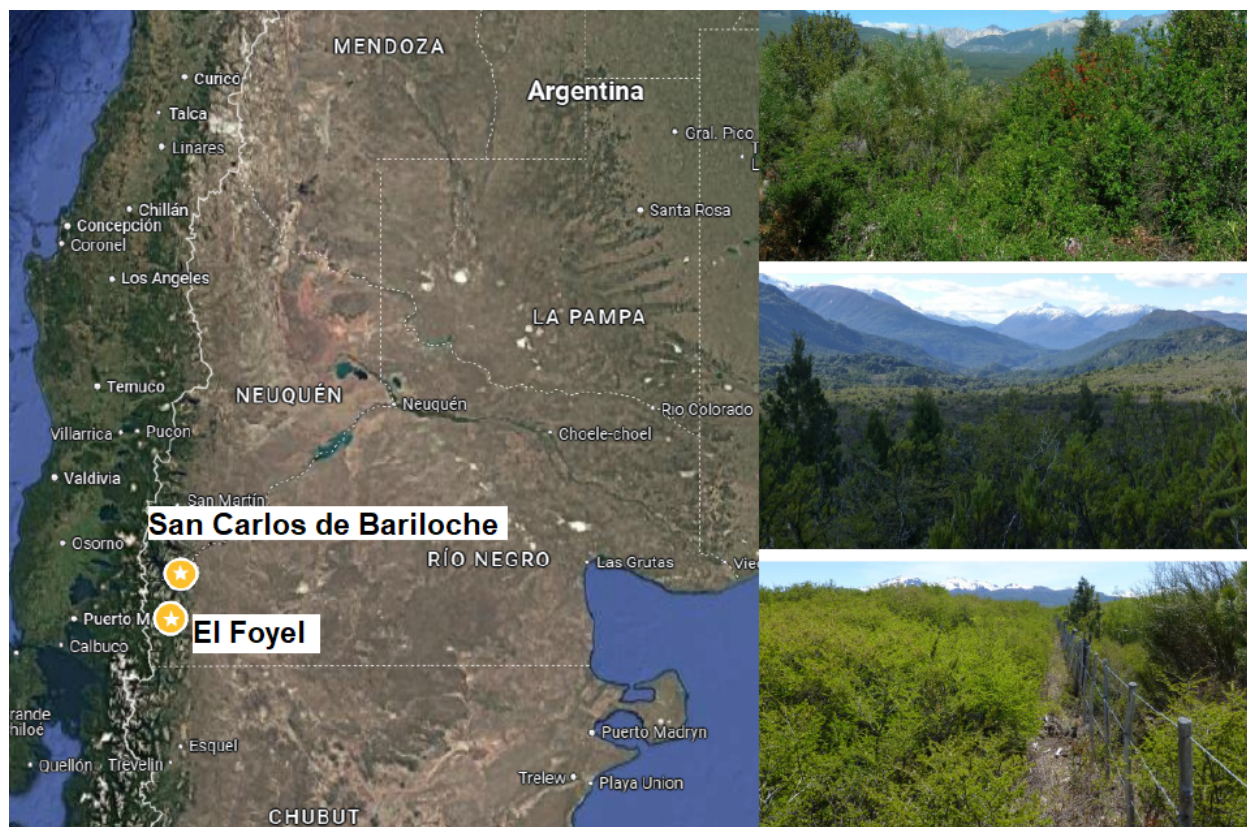
En esta etapa, nuestro enfoque se centra en identificar aquellos nodos cuya eliminación provoca una pérdida notable de cohesión en la red. A estos nodos los denominaremos “especies clave”. Se seleccionan específicamente aquellos nodos cuya eliminación resulta en un aumento en el número de comunidades o componentes conectados en los que puede dividirse la red. Es importante señalar que el concepto de “comunidad” utilizado aquí se refiere al término empleado en el contexto de redes complejas, y no al concepto ecológico tradicional.

Un alto número de comunidades indica una red con menor cohesión, lo que puede tener implicaciones significativas para la estabilidad y funcionalidad del ecosistema. La identificación de estas especies clave es crucial, ya que su preservación puede ser fundamental para mantener la integridad estructural y funcional del sistema ecológico. Este análisis no sólo proporciona información sobre las interacciones dentro de la red, sino que también ofrece herramientas prácticas para la gestión y conservación de las especies que desempeñan roles críticos en sus respectivos ecosistemas.

#### 2.2.4. Caso de estudio. Interacciones entre Polinizadores-Plantas- Herbívoros

Para evaluar este protocolo, se utilizaron datos de campo con el objetivo de estudiar la relación entre las interacciones planta-polinizador y planta-herbívoro. Los datos

fueron recolectados en “El Foyel”, una región situada en la provincia de Río Negro, en la Patagonia Argentina (Figura 2.3). Las recolecciones de datos se llevaron a cabo simultáneamente durante la primavera de 2014 y el verano de 2015. Estos datos han sido publicados en dos estudios independientes: uno enfocado en las interacciones planta-polinizador [96] y otro en las interacciones planta-herbívoro [97].



**Figura 2.3:** Ubicación y fotografías del sitio de estudio. Créditos fotográficos: Marina Alma y Marcos Nacif.

En el caso de las interacciones planta-polinizador, se realizó un censo periódico del número de visitas de potenciales polinizadores, donde se registraron las visitas durante un periodo de 10 minutos para cada especie de flor (número de visitas por flor cada 10 minutos) a lo largo de toda la etapa de floración. Es importante destacar que el número de observaciones varió debido a diferencias en la duración de la floración, así como en la abundancia y distribución espacial de las especies.

Para medir la herbivoría causada por artrópodos, se seleccionaron al azar veinticuatro muestras por especie (para más detalles, ver Nacif et al. [97]). En cada individuo, se muestrearon pequeñas ramas marcadas a alturas intermedias y se registró la herbivoría en 18 hojas por rama. La herbivoría se cuantificó calculando la frecuencia de daños en las hojas (número de hojas dañadas sobre el total), siguiendo las metodologías propuestas por Garibaldi et al. [10]. Los artrópodos herbívoros y el daño foliar fueron

clasificados en once gremios diferentes.

En cuanto a las especies vegetales analizadas, se registraron las frecuencias de polinizadores y las tasas de herbivoría para varias especies nativas de la Patagonia, incluyendo:

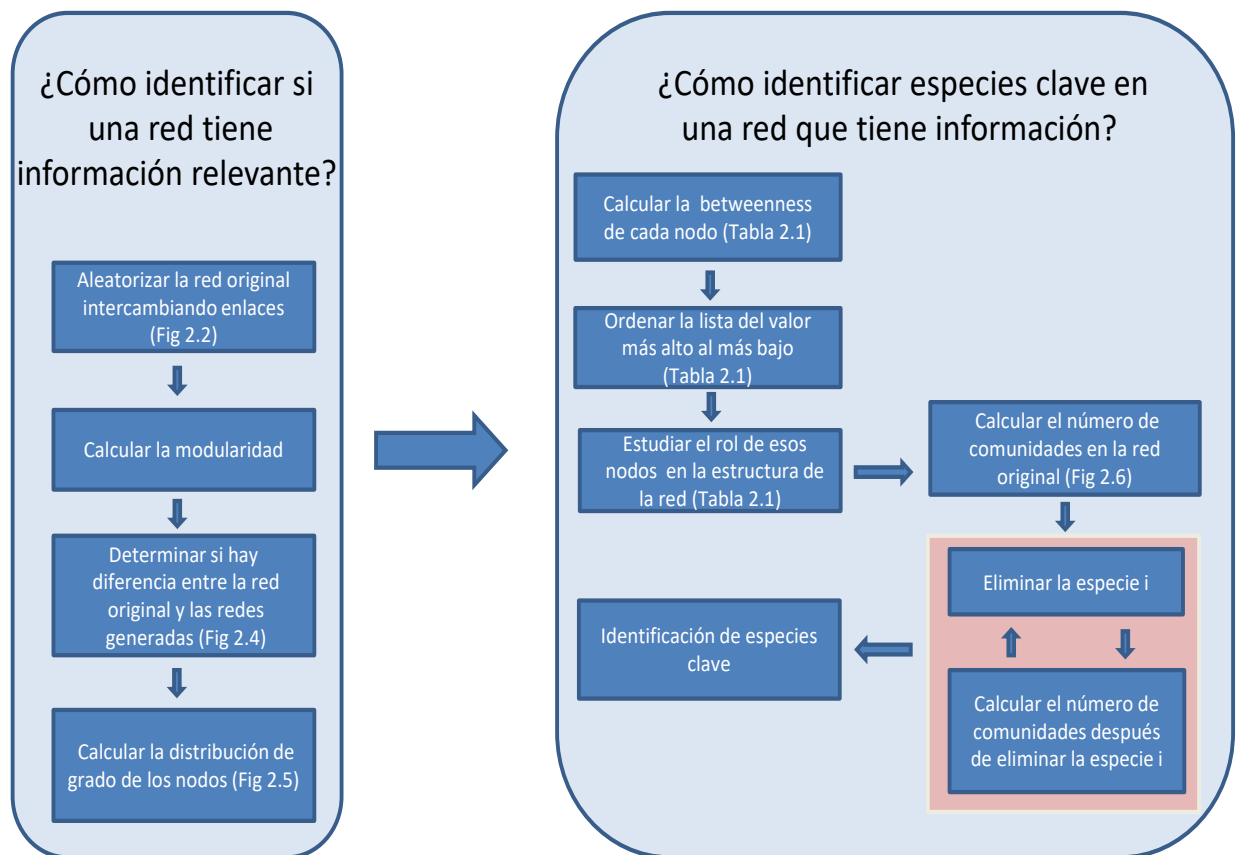
- *Azara microphylla* (Flacourtiaceae)
- *Berberis darwinii* (Berberidaceae)
- *Embothrium coccineum* (Proteaceae)
- *Gaultheria mucronata* (Ericaceae)
- *Lomatia hirsuta* (Proteaceae)
- *Maytenus chubutensis* (Celastraceae)
- *Maytenus boaria* (Celastraceae)
- *Nothofagus antarctica* (Nothofagaceae)
- *Ribes cucullatum* (Grossulariaceae)
- *Ribes magellanicum* (Grossulariaceae)
- *Schinus patagonicus* (Anacardiaceae)

Dentro de nuestra red, se representaron dos tipos de interacción: una capa compuesta por polinizadores y especies vegetales, y otra que incluye especies vegetales y gremios de herbívoros. Se identificaron ocho plantas comunes en ambas redes bipartitas, además de tres plantas que presentan interacciones exclusivas con herbívoros. La proyección de esta red resultó en una matriz de adyacencia [56], que fue utilizada para nuestro estudio. Esta matriz incluye 11 especies representativas de plantas leñosas típicas en un bosque mixto del norte de la Patagonia, 11 gremios de herbívoros y 25 polinizadores agrupados por morfotipos [96] (Ver Tabla A.5). La abundancia de cada planta se cuantificó como porcentaje de cobertura [98].

A continuación presentamos un resumen detallado de la metodología propuesta 2.4.

## 2.3. Resultados

El algoritmo descripto anteriormente ayudó a evaluar la importancia de la estructura topológica de una red añadiendo un grado de desorden mientras se preservaba el grado de cada nodo. Posteriormente, se aplicó un protocolo paso a paso para identificar las especies clave en nuestro caso de estudio.

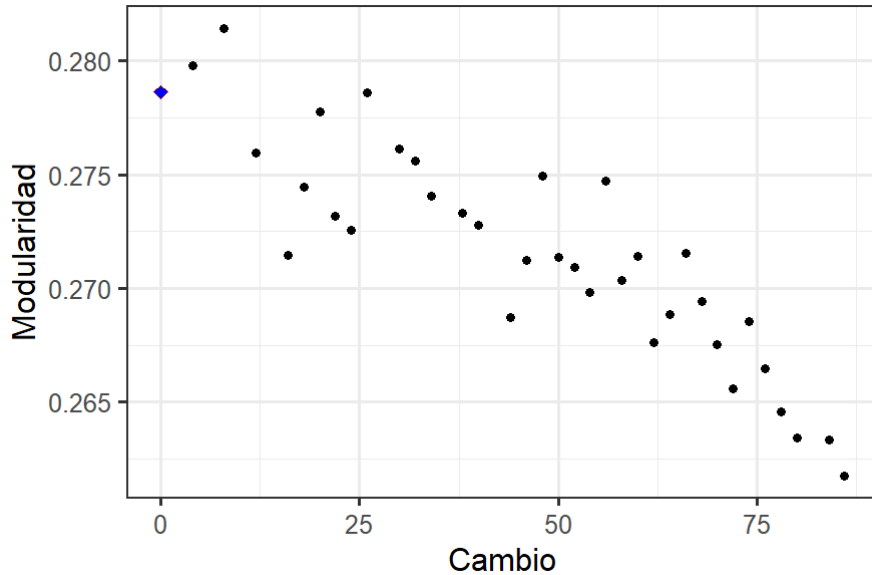


**Figura 2.4:** Metodología para determinar si una red tiene información codificada y para identificar las especies clave.

### 2.3.1. Implementación del algoritmo

La modularidad fue calculada en función del grado de aleatoriedad incorporado a la red original intercambiando los nodos adyacentes de dos enlaces elegidos al azar. Para cada número de cambios, generamos diez redes diferentes y calculamos el promedio de la modularidad en función de ese valor. Como se muestra a continuación, la modularidad disminuye a medida que aumenta el número de cambios (Figura 2.5). Además, el desorden de la red crece a medida que los valores de modularidad pasan de los originales a los aleatorios correspondientes. Este cambio monótono es un indicador de que la red original tenía cierta información codificada que se destruyó al incorporar la aleatoriedad. Se realizó una prueba de hipótesis, en la que se utilizó el estadístico *t*. Se obtuvo un p-valor de 0.031 y al nivel 0.05 se rechazó la hipótesis nula. La conclusión fue

que, tras realizar una serie de cambios, la modularidad obtenida es significativamente diferente de la original.



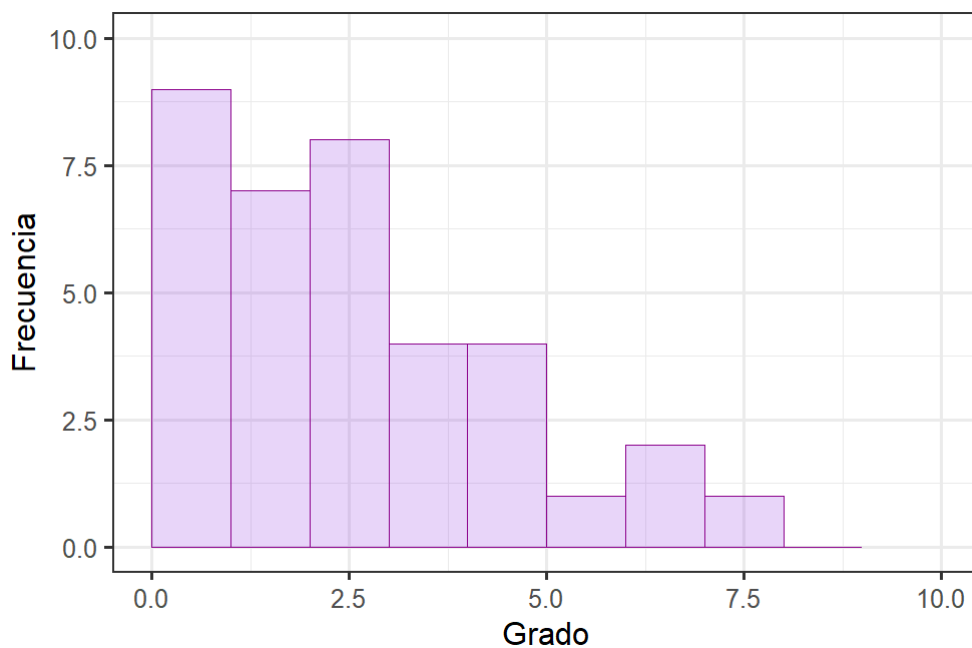
**Figura 2.5:** Modularidad vs. el número de cambios (*swappings*) en la red original. Cada valor es un promedio de 10 redes excepto por la red original.

### 2.3.2. Caracterización de nuestra muestra

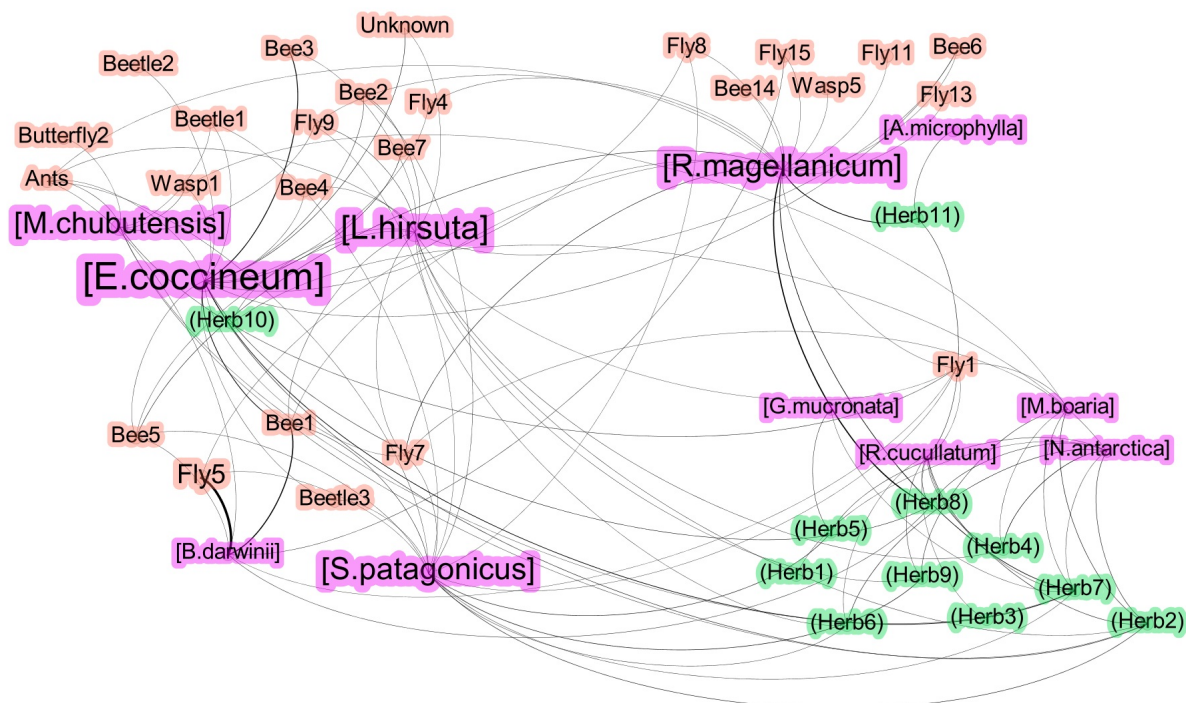
En nuestra muestra, el grado de cada nodo fue calculado y la distribución de grado también. Como se muestra en la Figura 2.6, la frecuencia de los nodos muestra atributos típicos de una red ecológica. La característica interesante es que la frecuencia de cada nodo decrece con el grado [91]. Sin embargo, esto no es un indicador de que la red sea diferente a una aleatoria, y por la cantidad de nodos que tienen las redes no es posible distinguir una ley de potencia de una exponencial en una red aleatoria [5].

La forma de la distribución ayuda a detectar posibles sesgos en la muestra. El valor de la modularidad, número de comunidades, densidad y el grado de cada nodo fue calculado. Esta red tiene 47 nodos, 111 enlaces y su densidad es  $\frac{111}{2162}$  ( $\approx 0.051$ ). La longitud del promedio de los caminos es 2.585 y su modularidad es 0.278. Un total de 4 comunidades fueron detectadas: 0, 1, 2, 3 con 34.04 %, 21.28 %, 14.89 % y 29.79 % de los nodos, respectivamente (Figura 2.7).

Para determinar las especies clave, construimos una lista de especies, normalizamos su *betweenness* y las ordenamos desde sus valores más altos. Se ha elegido  $\alpha = 0.20$  porque este valor cubre la región de interés y parte de la región que no lo es (Tabla 2.1). Una vez que la especie con alta *betweenness* fue eliminada, (*R. magellanicum*), el número de comunidades fue calculado. Luego *E. coccineum* fue eliminada pero sin eliminar *R. magellanicum* y el número de comunidades fue calculado. Finalmente cuando



**Figura 2.6:** Distribución de grado de los nodos de la red original.



**Figura 2.7:** Red proyectada de Polinizador-Planta- Gremios de Herbívoros. Las especies de plantas son representados en violeta y entre corchetes, los gremios de herbívoros en verde y entre paréntesis y los polinizadores en naranja.

**Tabla 2.1:** Cambios en la estructura de la red después de eliminar una especie de planta. Solo fueron consideradas las especies de plantas con una *betweenness* normalizada más grande que  $\alpha = 0.20$ .

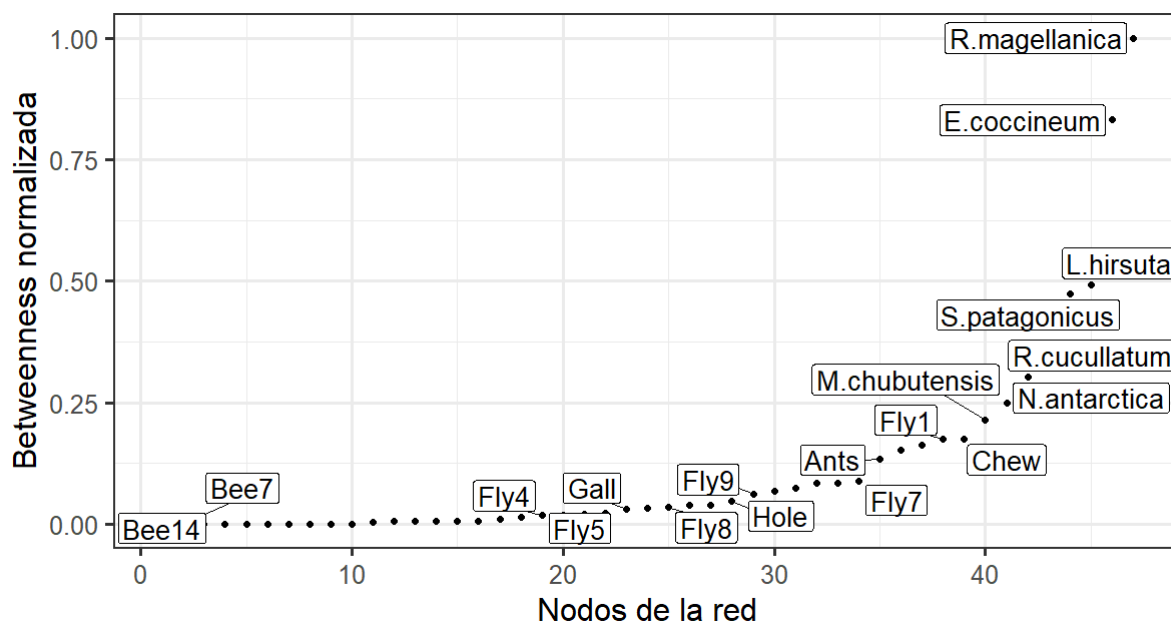
Planta eliminada	<i>Betweenness</i>	<i>Betweenness</i> normalizada	Comunidades (Antes)	Comunidades (Después)	Componentes conectadas (Antes)	Componentes conectadas (Después)
<i>R. magellanicum</i>	307.79	1	4	8	1	5
<i>E. coccineum</i>	256.16	0.83	4	6	1	3
<i>L. hirsuta</i>	151.88	0.49	4	5	1	1
<i>S. patagonicus</i>	145.82	0.47	4	4	1	1
<i>R. cucullatum</i>	93.36	0.30	4	5	1	2
<i>N. antarctica</i>	76.88	0.25	4	5	1	2
<i>M. chubutensis</i>	66.14	0.21	4	5	1	2

*R. magellanicum* y *E. coccineum* fueron eliminados de forma separada, el número de comunidades y componentes conectadas aumentó en ambos casos, indicando que estas eran especies clave (Tabla 2.1). Notar que al remover *R. magellanicum* la red comienza a ser más desconectada. La *betweenness* fue calculada para todos los nodos de la base de datos, sin embargo las que tenían más alta *betweenness* fueron las plantas. Hicimos principal foco en las plantas, ya que es más fácil introducirlas para restaurar el ecosistema, a diferencia de incorporar un polinizador o una especie de artrópodo, ya que éstos podrían ser más difíciles [99, 100].

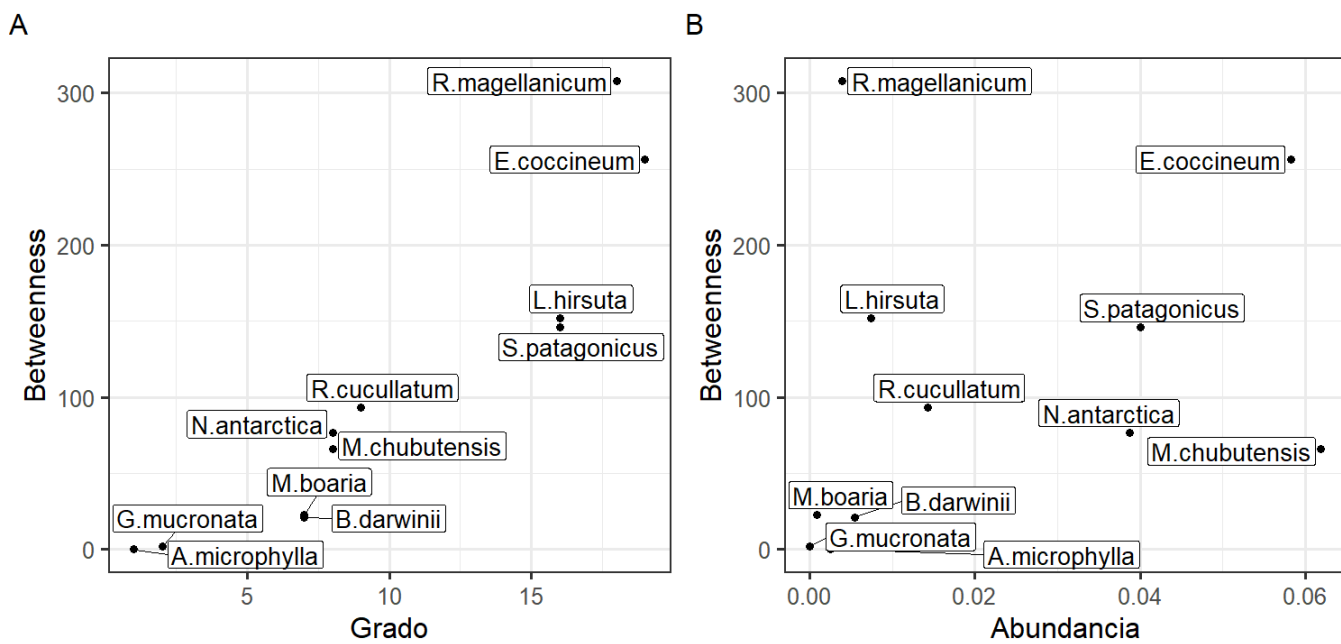
Además, el grado de cada nodo fue calculado y las 2 especies con más alto grado fueron: *R. magellanicum* y *E. coccineum*. En general no hay correlación entre el grado y la *betweenness*; sin embargo encontramos que en este caso en particular hay correlación entre ellos (Figura 2.9A) usando la correlación de Spearman y obteniendo un  $\rho = 0.98$ . Finalmente, estudiamos la correlación entre abundancia de cada planta y su *betweenness*, sin embargo, no hay correlación entre ellos (Spearman  $\rho$  factor = 0.49, Figura 2.9B). Observamos que los nodos con más alta *betweenness* no eran los que tenían una mayor abundancia. Además, las redes bipartitas de planta- polinizador, herbívoro-planta fueron analizados y de éstos se observó que las especies clave son diferentes que cuando analizamos la red de polinizador-planta-herbívoro (Tabla A.1, Tabla A.2). Si analizamos la red de plantas, la cual fue obtenida proyectando las plantas que pertenecen a la red de planta-polinizador, no podemos asegurar que esas especies de plantas sean las especies clave porque la información original se ha perdido (Tabla A.3). Algo similar sucede con las plantas de la red de planta-herbívoro (Tabla A.4).

## 2.4. Discusión

Se presentó un método para evaluar cualitativamente la división en comunidades de una red al incorporar una aleatorización gradual. Este enfoque permite eliminar



**Figura 2.8:** Especies de plantas, polinizadores y gremios de herbívoros vs. *betweenness* normalizada.



**Figura 2.9:** (A) Grado vs. *betweenness* y (B) abundancia (Porcentaje de cobertura) vs. *betweenness* de 11 especies de plantas.

información valiosa almacenada en la red, mientras se conservan algunas de sus características fundamentales. Este método fue utilizado para identificar especies clave dentro de un sistema ecológico.

Se evaluaron los cambios en la red original a medida que aumentaba la aleatoriedad, preservando la distribución de grado. Es relevante destacar que la modularidad de la red disminuyó de manera monótonica hasta alcanzar un valor correspondiente a una red completamente aleatoria, manteniendo la restricción de conservar la distribución de grado original. Finalmente, se caracterizó la distribución de grado de los nodos y se identificaron las especies clave. Nuestros resultados indicaron que la estructura de la red contenía información relevante sobre el sistema ecológico, lo que nos permitió detectar las especies clave *Ribes magellanicum* y *Embothrium coccineum*. Este trabajo se distingue de estudios previos que analizaban estructuras de redes bipartitas y medidas de centralidad, ya que se evaluó previamente si una red contenía información relevante antes de trabajar sobre ella [101–107]. Además, se desarrolló un método para validar los resultados obtenidos. Aunque existen estudios que realizan aleatorización conservando el grado de los nodos, nuestro trabajo se diferencia al aplicar una aleatorización progresiva, considerando además el número de cambios realizados [108].

### 2.4.1. Aplicaciones del algoritmo

Este algoritmo constituye una herramienta valiosa para determinar si una red contiene información relevante en su sistema, revelada a través de su estructura comunitaria. En este estudio, al incrementar los cambios en la red, el valor de la modularidad mostró una tendencia decreciente, indicando una pérdida de información en el sistema.

A medida que se incrementa el desorden en la red, esta pierde su estructura original y se asemeja a una red aleatoria, que carece de una estructura comunitaria definida. Por lo tanto, podemos concluir que la red estudiada contenía información valiosa que podría ser cuantificada en futuras investigaciones. La red aleatoria constituye nuestro modelo nulo; si no disponemos de un modelo nulo adecuado, debemos modificar nuestra metodología. El algoritmo mencionado puede aplicarse en redes con dos o más tipos de interacción; sin embargo, es necesario realizar una proyección previa.

### 2.4.2. Identificación de especies clave

El análisis de redes con dos o más tipos de interacción ofrece un enfoque más completo que aquellos desarrollados anteriormente centrados únicamente en interacciones herbívoro-planta o polinizador-planta. Según los datos recogidos, el efecto que cada nivel tiene sobre el otro puede ser considerado al analizar las interacciones conjuntamente. Debido a las características específicas de las especies y a los resultados obtenidos en estudios anteriores, se esperaba que ciertas plantas como *Nothofagus antarctica*,

*Maytenus boaria* y *Ribes cucullatum* tuvieran vínculos más fuertes con los gremios de herbívoros que con los polinizadores [97]. Además, *E. coccineum* y *R. magellanicum* poseen flores muy atractivas y sus hojas son altamente palatables; por lo tanto, estas especies pueden tener interacciones significativas tanto con herbívoros como con polinizadores. Cabe destacar que *E. coccineum* fue muy abundante entre las especies leñosas muestreadas. En este análisis, las especies clave *R. magellanicum* y *E. coccineum* desempeñan un papel estabilizador en el ecosistema. Entre las especies forestales leñosas, estas dos tienen funciones relevantes no sólo en términos de interacciones ecológicas sino también como recursos forestales. Las flores de ambas especies son altamente preferidas por polinizadores como insectos y aves [96], y sus hojas soportan una alta diversidad de gremios herbívoros [97]. Por ejemplo, *E. coccineum* tiene un alto valor potencial para leña y madera para muebles; su alta tasa de rebrote permite implementar estrategias alternativas para su manejo [109]. Además, las hojas, corteza, tallo y raíces de ambas especies tienen múltiples aplicaciones médicas [110, 111], y son apreciadas como especies ornamentales, especialmente *E. coccineum* [112]. Asimismo, el fruto de *R. magellanicum* proporciona un valioso doble servicio como alimento para fauna del bosque y como producto potencial para consumo humano (mermeladas y licores) [113]. Por lo tanto, estas dos especies son clave debido a su relación entre polinizadores, herbívoros y plantas leñosas con uso antrópico; el seguimiento y comprensión del papel que desempeñan en las redes ecológicas es esencial para la conservación y gestión sostenible de estas comunidades.

Además, la construcción de comunidades de las redes puede revelar resultados interesantes respecto a su composición. En nuestro caso particular, encontramos diferencias significativas entre las cuatro comunidades identificadas, variando en términos de polinizadores, especies vegetales y composición del gremio herbívoro. Notablemente, aquellas comunidades cuyas plantas eran muy visitadas por polinizadores mostraron ser menos preferidas por los herbívoros. Las especies vegetales con flores atractivas para insectos polinizadores pertenecen a tres de las comunidades mencionadas [23]. Es importante señalar que estas plantas también presentan hojas menos palatables; sin embargo, *R. magellanicum* es una excepción [114]. Por otro lado, la comunidad más consumida por diversos gremios de herbívoros estuvo compuesta por plantas que fueron polinizadas con menor frecuencia por insectos; especies como *Gaultheria mucronata*, *Maytenus boaria*, *Nothofagus antarctica* y *Ribes cucullatum* tienen hojas consideradas de alta calidad [115, 116]. Este enfoque matemático podría convertirse en una estrategia útil para determinar si existe un equilibrio en los estratos de la red y proporcionar información valiosa para análisis posteriores. En general, los resultados obtenidos tienen sentido biológico y sugieren replicar este enfoque en otros sistemas o ampliarlo.

Otro método para identificar especies clave consiste en evaluar la abundancia relativa de cada especie vegetal; sin embargo, esto depende del contexto específico del

estudio. En nuestro caso particular, no se observó correlación entre abundancia y *betweenness centrality*. La especie más abundante fue *Maytenus chubutensis*, aunque esta es menos preferida por los polinizadores y su eliminación no genera cambios significativos en la estructura general de la red. Curiosamente, *E.coccineum*, además de ser nuestra segunda especie clave, también es la segunda más abundante y presenta el grado más alto entre las 11 especies analizadas. Esto sugiere una correlación positiva entre el grado y *betweenness centrality*; sin embargo, generalmente no están correlacionados.

En años recientes, tanto biología como gestión ambiental se han centrado en el estudio de especies clave; sin embargo su definición y cuantificación sigue siendo un desafío significativo. Jordán y Scheuring [117] estudiaron y compararon 13 índices de centralidad en redes ecológicas y determinaron que cuando los enlaces no están ponderados, *betweenness* es el mejor índice para identificar la importancia posicional de las especies. Además, las especies clave identificadas en redes con dos o más tipos de interacción pueden diferir notablemente respecto a aquellas encontradas en redes bipartitas limitadas a un tipo de interacción.

En nuestro caso específico identificamos especies clave dentro tanto de redes bipartitas mutualistas como antagonistas: las primeras incluyen a *Ribes magellanicum* y *Embothrium coccineum* (Tabla A.1; Figura A.1), mientras que las segundas incluyen a *Ribes cucullatum* y *Nothofagus antarctica* (Tabla A.2; Figura A.2). Esto demuestra que las especies clave dentro del sistema no son necesariamente las mismas que aquellas identificadas únicamente dentro del contexto bipartito; claramente podríamos perder detalles importantes del ecosistema si sólo consideramos un tipo de interacción dentro del análisis propuesto.

Aunque este capítulo se ha centrado exclusivamente en redes compuestas por polinizadores-plantas-gremios de herbívoros, el método desarrollado puede aplicarse a otros tipos variados de redes ecológicas donde existan amplias posibilidades para análisis posteriores; esto podría resultar útil para identificar características comunes entre especies dentro de una misma comunidad. Como enfoque matemático generalizado, nuestro algoritmo podría aplicarse a diversas disciplinas donde identificar nodos clave permita determinar qué elementos desempeñan roles centrales dentro del sistema analizado. De este modo, los investigadores podrán verificar si su red contiene información valiosa e invitar a futuros estudios a aplicar el método propuesto para realizar un análisis exhaustivo antes de extraer conclusiones anticipadas.

# Capítulo 3

## Medidas de Complejidad de redes ecológicas bipartitas

### 3.1. Introducción

En el área de ecología, la teoría de redes se ha utilizado para analizar sistemas ecológicos debido a su capacidad de proporcionar un potente marco estadístico-matemático que permite capturar la complejidad y estructura de las comunidades ecológicas. El uso de redes ha permitido a los investigadores identificar cómo la estructura topológica de los sistemas ecológicos está vinculada a sus propiedades y procesos ecológicos. Como resultado, se han propuesto numerosas métricas para capturar diferentes aspectos de la estructura de las redes y su relación con los sistemas ecológicos [118]. La mayoría de los estudios se han centrado en redes bipartitas que describen un único tipo de interacción, siendo los sistemas mutualistas y antagonistas los más comúnmente estudiados [119]. Los investigadores han examinado diversos índices de redes para identificar sus propiedades estructurales. Sin embargo, el grado de complejidad de estas redes no se ha explorado a fondo, y es esencial comparar y contrastar diferentes valores de complejidad [74, 120].

La palabra complejidad se utiliza a menudo de distintas formas según el ámbito en el que se estudie, en el área de la ecología puede emplearse para describir el estado de un ecosistema. La complejidad ecológica se ha relacionado con conceptos como diversidad, resiliencia y se ha propuesto como una guía en el área ecológica. Otras medidas que se pueden encontrar dentro del área de la ecología son: medidas temporales de complejidad, éstas se utilizan para caracterizar series temporales de diferentes variables que describen la dinámica de un sistema; medidas espaciales de complejidad, éstas no describen la dinámica o función de un sistema, sino que describen su configuración en el espacio en un momento del tiempo. En el análisis de redes, las medidas de complejidad estructural describen la organización y las relaciones entre los componentes del sistema

[121]. Landi [122] menciona diferentes formas de medir la complejidad desde un punto de vista ecológico, por ejemplo, la conectividad fue utilizada por May [123, 124] para describir la complejidad de la red. La palabra complejidad es muy amplia y puede adoptar diferentes significados dependiendo del contexto. En Montoya [125] las métricas de complejidad se definieron como un conjunto de índices, incluyendo especies, enlaces, enlaces por especie, grado y distribución de grado de la red. Sin embargo, dependiendo de la pregunta de investigación, la complejidad de una red puede definirse de diferentes maneras. Según Atmanspacher [126], las medidas de complejidad tienden a ser de dos tipos: las medidas del primer tipo aumentan linealmente con el incremento del desorden en el sistema, y las medidas del tipo 2 son una función convexa, atribuyendo sus valores más altos a sistemas de regularidad intermedia.

Recientemente, los investigadores se han interesado cada vez más por las redes que contienen múltiples tipos de interacción [127–130], y se han estudiado varios índices para redes bipartitas y multicapa [131].

Se ha desarrollado una herramienta para determinar si una red codifica información relevante en su sistema. La complejidad de las redes se ha estudiado utilizando diferentes índices, entre ellos el índice de *singular value decomposition entropy* (SVD). Muchos índices de complejidad utilizados actualmente en ecología han sido adaptados de otras áreas como la física y la teoría de la información [121].

El enfoque propuesto en este capítulo se basa en la teoría de la información y es una adaptación que puede aplicarse a distintos tipos de redes. Se propone una nueva metodología que pretende comparar los valores iniciales de los índices de complejidad con los valores obtenidos a partir de sus versiones aleatorizadas, un aspecto que no se ha explorado en profundidad.

Aunque algunos resultados derivados de estudios basados en redes ayudan a caracterizar su estructura, en este capítulo se pretende ir más allá del paradigma tradicional basado en redes. El procedimiento propuesto puede aplicarse a todo tipo de redes, con el objetivo de determinar si contienen información relevante codificada en su estructura.

Se pueden utilizar varios índices para aproximar la complejidad de una red, pero es importante tener en cuenta el tipo de red y la herramienta utilizada. Entre los índices que pueden utilizarse se encuentran la entropía y el desequilibrio. El desequilibrio está definido como:

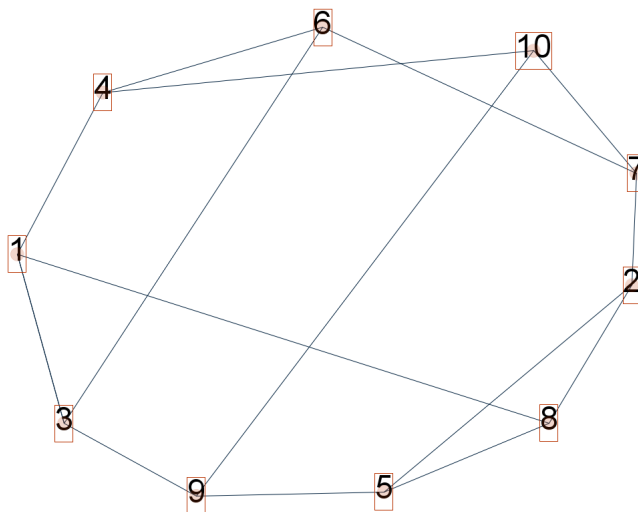
$$D = \sum_{i=1}^N \left( p(x_i) - \frac{1}{N} \right)^2. \quad (3.1)$$

En este capítulo  $p(x_i) = \frac{\text{Nodos en la comunidad } i}{\text{nodos totales}}$ ,  $N$  es el número de comunidades

en la red y la entropía está definida por

$$H(X) = - \sum_{i=1}^N P(x_i) \cdot \log(P(x_i)), \quad (3.2)$$

siendo  $X = \{x_1, \dots, x_N\}$  la variable aleatoria y  $H(x)$  es calculada usando la fórmula de Shannon. La complejidad puede ser definida de varias maneras, tanto como el producto de la entropía y el desequilibrio o como un conjunto de índices como *Singular Value Decomposition* (SVD) entropía y K-complejidad. Cuando la complejidad es aproximada usando el producto de la entropía por el desequilibrio, la forma de calcular la entropía debe ser tenida en cuenta, ya que ésta puede variar dependiendo de la variable aleatoria usada. Consideremos un grafo 3-regular (Figura 3.1).



**Figura 3.1:** Grafo regular de grado 3 con 10 vértices

Este grafo puede ser descrito usando la siguiente matriz de adyacencia:

$$\begin{bmatrix} 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

Por ejemplo, si consideramos una matriz de  $10 \times 10$  y su proporción de unos y ceros, cada vector de esta matriz puede tratarse como una variable aleatoria con dos resultados posibles, 0 y 1. Contando el número de ocurrencias de estos resultados, la variable aleatoria  $X = \{x_0=0.7, x_1=0.3\}$  y su entropía  $H(X) = 1$ . Este valor fue calculado usando la fórmula mencionada previamente,  $H(X) = -0.7 \log_2(0.7) - 0.3 \log_2(0.3)$ . Como una alternativa, este grafo puede ser descrito usando su grado  $D = \langle 3, 3, 3, 3, 3, 3, 3, 3, 3, 3 \rangle$  el cual puede ser tratado como una variable aleatoria con entropía  $H(D) = 0$ . Otra posible variable aleatoria que puede ser derivada de este grafo es la distribución de grado  $PD = \{d_0 = 0, d_1 = 0, d_2 = 0, d_3 = 1\}$  con una entropía de  $H(PD) = 0$  [132].

En el estudio de la teoría de la información y la complejidad de redes, es fundamental comprender que existen múltiples métodos para definir una variable aleatoria. Los valores de entropía derivados de estas variables pueden variar significativamente, dependiendo de la naturaleza específica de la variable aleatoria en cuestión. En este capítulo, además de explorar la entropía, se abordará la complejidad de Kolmogorov, un concepto que mide la aleatoriedad intrínseca y la información estructural de una red, sin depender de ninguna distribución de probabilidad particular. Esta característica hace que la complejidad de Kolmogorov sea más robusta y fiable en comparación con la entropía, que puede fluctuar en función de cómo se elija la distribución.

La complejidad de Kolmogorov proporciona una perspectiva única sobre la estructura de las redes, capturando aspectos que pueden no ser evidentes a través de medidas basadas en entropía. Este índice puede ser aproximado mediante diversos métodos, como el algoritmo probabilístico propuesto por Solomonoff y Levin [59, 60], o a través del Método de Descomposición en Bloques, que es particularmente útil para cadenas o matrices más grandes [59]. La K-complejidad ha encontrado aplicaciones prácticas en campos como el aprendizaje automático [133, 134] y la teoría general de la computación [135], donde su capacidad para cuantificar la complejidad se traduce en mejoras significativas en el rendimiento y la eficiencia.

Recientemente, se ha introducido un enfoque novedoso que considera la comple-

alidad a través de la descomposición en valores singulares (SVD) [131]. Este método combina la medida de entropía de Shannon con los valores singulares distintos de cero obtenidos a partir del SVD, ofreciendo un nuevo índice que será detallado en la sección metodológica. Sin embargo, es crucial considerar estos índices dentro del contexto más amplio proporcionado por un modelo nulo. La comparación con un modelo nulo permite evaluar cómo varían estos índices al aleatorizar las redes, proporcionando una visión más completa del fenómeno estudiado.

Para alcanzar los objetivos centrales de este capítulo, se emplean métodos sistemáticos para caracterizar redes utilizando datos extraídos de sitios web. Se examinan métricas clave como modularidad, entropía, entropía SVD y K-complejidad, con el fin de identificar las herramientas más efectivas para distinguir las complejidades inherentes a diferentes redes. Presentamos dos enfoques sistemáticos para incrementar gradualmente la aleatoriedad dentro de una red y analizamos cómo estos cambios impactan las métricas mencionadas. Para ello, utilizamos dos algoritmos diseñados específicamente para generar redes aleatorias. Los objetivos específicos de investigación son: 1) Determinar el nivel de complejidad presente en las redes analizadas; 2) Identificar las herramientas más adecuadas según el tipo específico de red que se esté caracterizando.

La sección metodológica detalla la implementación del algoritmo, así como los criterios utilizados para seleccionar las herramientas y los conjuntos de datos empleados. Posteriormente, en la sección dedicada a los resultados, se presentan nuestros hallazgos e interpretaciones. La sección 3.4 se reserva para una discusión profunda sobre las implicaciones derivadas de nuestros resultados.

El análisis exhaustivo de estas métricas no sólo proporciona una comprensión más profunda sobre las características estructurales de las redes ecológicas, sino que también permite identificar patrones que podrían ser indicativos del estado general del ecosistema. Al comparar los índices obtenidos con sus modelos nulos correspondientes, podemos discernir si los cambios observados son realmente significativos o si pueden atribuirse a variaciones aleatorias inherentes al sistema.

Además, al aplicar estos métodos a diferentes tipos de redes ecológicas, se abre un campo amplio para futuras investigaciones que busquen identificar características comunes entre especies dentro de una misma comunidad. Este enfoque matemático tiene el potencial no sólo para enriquecer nuestro entendimiento teórico, sino también para ser útil en diversas disciplinas donde se requiera determinar qué elementos desempeñan roles centrales dentro del sistema analizado.

Esperamos que los resultados y metodologías presentados en este capítulo sirvan como base para estudios futuros que busquen aplicar el método propuesto para realizar un análisis cuidadoso antes de extraer conclusiones anticipadas sobre las interacciones dentro de las redes ecológicas. La capacidad para identificar y cuantificar la complejidad en estos sistemas es esencial para avanzar en nuestra comprensión del funcionamiento

ecológico y contribuir a estrategias efectivas para su conservación y manejo sostenible.

## 3.2. Metodología

### 3.2.1. Datos analizados

Para evaluar la metodología propuesta se han seleccionado un conjunto de redes ecológicas, 23 redes de planta-polinizador y 19 redes hospedador-parásito de la base de datos <https://www.web-of-life.es/>. Esta base de datos incluye redes de interacción entre diferentes especies y de diversos hábitats, organismos, años de muestreo y metodologías. Las redes a estudiar contienen entre 27 y 997 nodos y entre 32 a 2933 enlaces. La anidación se calculó revelando que las redes polinizador-planta están más anidadas que las redes hospedador-parásito. La densidad en estas redes es inferior a 0.3.

### 3.2.2. Medidas estadísticas frecuentemente usadas en la caracterización de las redes

Se pueden utilizar diversos parámetros para caracterizar una red, tales como el número de enlaces, la cantidad de nodos, la asortatividad, el anidamiento, la distribución de grado, la densidad, la modularidad y la entropía. En nuestro estudio, nos hemos centrado en calcular el número de enlaces y nodos, el anidamiento, la densidad, la modularidad y la entropía. El número de nodos representa la cantidad de especies en la red, mientras que el número de enlaces denota las interacciones entre estas especies.

El anidamiento es una propiedad que establece que, para dos nodos cualesquiera  $i$  y  $j$ , si el grado de  $i$  es menor que el grado de  $j$ , entonces la vecindad de  $i$  está contenida en la vecindad de  $j$ . La densidad se refiere a la relación entre los enlaces existentes en la red y los posibles. Por otro lado, la modularidad mide la intensidad de los enlaces intracomunitarios en comparación con los enlaces intercomunitarios, lo que ayuda a caracterizar la estructura de la comunidad. Las comunidades se definen como conjuntos disjuntos de nodos que comparten propiedades y/o funciones comunes dentro de la red.

Sabemos que la modularidad puede ser calculada utilizando diferentes algoritmos [61, 62, 70, 78, 79]. En este capítulo utilizamos el algoritmo de Louvain, que agrupa los nodos con el objetivo de maximizar la función de modularidad considerando previamente varios grupos de nodos y seleccionando aquellos que maximizan dicha función. Además, se ha calculado el número de comunidades y la modularidad utilizando el paquete `lpbrim` en R [136]; los resultados obtenidos fueron análogos a los generados con el algoritmo de Louvain.

El algoritmo de módulos bipartitos inducidos recursivamente (BRIM) fue propuesto

por Barber en 2007 [62] y busca optimizar la modularidad asignando nodos a diferentes comunidades de manera sucesiva. Este proceso tiene como objetivo maximizar la contribución individual de cada nodo a la modularidad, considerando las asignaciones previas. En este contexto, cada conjunto de nodos induce recursivamente al otro conjunto.

El algoritmo BRIM puede implementarse utilizando dos enfoques heurísticos diferentes: el BRIM adaptativo y una combinación del método de Propagación de Etiquetas (LP) con el algoritmo BRIM. El enfoque adaptativo utiliza el algoritmo BRIM junto con una estrategia para identificar divisiones óptimas en comunidades dentro de la red. Esta estrategia consiste en duplicar el número de comunidades en factores de dos hasta que la modularidad (Q) (2.1) comience a disminuir. Luego se aplica un método de bisección para encontrar el valor máximo de Q entre las particiones actuales y anteriores.

Por otro lado, el método LP.BRIM combina Propagación de Etiquetas (LP) con el algoritmo BRIM para realizar particiones en comunidades [137]. Este proceso consta de dos fases: durante la primera fase, los nodos vecinos intercambian etiquetas que representan su comunidad; cada nodo adopta la etiqueta más común entre sus vecinos. Este proceso se repite hasta que los grupos densamente conectados alcanzan un consenso sobre su etiqueta más representativa, lo cual se indica cuando la modularidad deja de aumentar. En la segunda fase, el algoritmo BRIM refina las particiones obtenidas mediante Propagación de Etiquetas.

La entropía es calculada usando una variable aleatoria. En este capítulo hemos definido una variable aleatoria discreta que toma  $K$  valores. Las comunidades fueron calculados usando el algoritmo de Louvain. Para cada comunidad, se calculó  $P(k) = \frac{n_k}{n}$  donde  $n_k$  = número de nodos en esa comunidad y  $n$  = número total de nodos de la red,  $k$  = número de comunidades. La entropía fue definida como la define Meila [80]

$$H(c) = - \sum_{k=1}^K P(k) \cdot \log(P(k)). \quad (3.3)$$

### 3.2.3. Herramientas para estimar la complejidad de una red: *Singular value decomposition* y K-complejidad

Existe una variedad de herramientas para estimar la complejidad de una red, en este capítulo presentaremos dos de ellas, el *singular value decomposition* (SVD) entropía y la complejidad de Kolmogorov (K-complejidad). En el caso del método SVD entropía se factoriza una matriz  $A$  (donde  $A_{m,n} \in \mathbb{R}$ ) en  $U \cdot D \cdot V^T$ .  $U \in \mathbb{R}^{m \times m}$  y  $V \in \mathbb{R}^{n \times n}$  son matrices ortogonales.  $D \in \mathbb{R}^{m \times n}$  es una matriz que sólo contiene valores no-negativos  $d$  en su diagonal y todas las otras entradas son cero. SVD entropía es calculada usando

$s_i = \frac{d_i}{\sum_i(d_i)}$ , donde  $d_i = D_{ii}$  son elementos de la diagonal, estos valores son conocidos como los valores singulares de  $A$ . Intuitivamente, el valor singular ( $d_i$ ) mide qué parte del conjunto de datos es explicado (proporcionalmente) por cada vector, por lo tanto, es posible medir la entropía siguiendo la fórmula de Shannon [138]. Los valores altos de SVD entropía reflejan que todos los vectores son igual de importantes, es decir, que la estructura de la red ecológica no se puede comprimir de forma eficiente y, por lo tanto, indicaría una alta complejidad [139]. Dado que las redes tienen diferentes dimensiones se normalizan los valores para garantizar que los valores sean inferiores a 1, y se cuantifica el índice SVD entropía definiendo  $J$  como

$$J = -\frac{1}{\ln(k)} \sum_{i=1}^k s_i \ln(s_i), \quad (3.4)$$

donde  $k = rk(A)$ . El rango de  $A$  ( $k = rk(A)$ ) es la dimensión del espacio vectorial que tiene la matriz y corresponde al número de filas o columnas linealmente independientes (l.i). Se dice que una matriz es de rango completo si  $k = M$ , es decir, todas sus filas/columnas son linealmente independientes. Las matrices que no son de rango completo se denominan de rango deficiente, y podemos medir la deficiencia de rango utilizando  $d = M - k$ . Una forma para controlar la diferencia en la riqueza de especies de las distintas redes en el área de ecología es informar de la deficiencia relativa de rango, es decir, expresada como relación entre la deficiencia de rango y el rango máximo:

$$D = 1 - \frac{k}{M}. \quad (3.5)$$

Esta medida devuelve valores entre 0 (la matriz tiene rango completo) y  $1 - M^{-1} \approx 1$  (la matriz tiene rango 1). Esto sirve como estimación aproximada de la complejidad, ya que cuantas menos columnas/filas l.i haya en la matriz, mayor será ese valor. Sin embargo, también puede carecer de sensibilidad, porque impone una prueba estricta de unicidad, lo que exige enfoques más cuantitativos de la complejidad [131].

Aunque la complejidad de Kolmogorov (K-complejidad) es más confiable y robusta que SVD entropía porque la complejidad  $K$  mide la aleatoriedad intrínseca o el contenido de información de una configuración de una red, es incomputable. La definición formal de K-complejidad es  $K_T(s) = \min\{|P|, T(P) = s\}$ , donde  $s$  es una cadena,  $P$  es un programa que produce las cadenas  $s$  cuando se ejecuta en una máquina de Turing universal  $T$  y  $|P|$  es la longitud del programa  $P$ , es decir, el número de bits necesarios para representar  $P$  [132]. La máquina de halting de Turing, derivada del problema de parada planteado por Alan Turing en 1936, establece que no existe un algoritmo general que pueda determinar si un programa se detendrá o continuará indefinidamente. Este resultado, que demuestra la indecidibilidad del problema, es crucial en la teoría de

la computación, ya que define los límites de lo que es computable. La máquina de halting subraya que existen problemas que no pueden ser resueltos por ninguna máquina de Turing, marcando una frontera teórica en la computación [140]. Afortunadamente, K-complejidad puede aproximarse utilizando diferentes algoritmos. Una forma de aproximar el verdadero valor de  $K$  es utilizar la probabilidad algorítmica introducida por Solomonoff y Levin [59, 60]. La probabilidad algorítmica utilizada para calcular la complejidad  $K$  tiene como base al Teorema de Codificación [59]. El teorema de codificación aproxima  $K$  de la siguiente manera  $K_T(s) \approx -\log_2 p^a(s)$  con  $p^a(s) = \sum_{P:T(P)=s} \frac{1}{2^{|P|}}$  donde  $P$  es un programa aleatorio que da como resultado la cadena  $S$ . La medida de probabilidad algorítmica  $p^a(s)$  está relacionada con la K-complejidad, ya que  $p^a(s)$  es al menos el término máximo en la suma de programas, dado que el programa más corto tiene el mayor peso en la suma. La probabilidad algorítmica se puede utilizar para calcular K-complejidad utilizando el Teorema de la Codificación [59] que establece que la diferencia entre la probabilidad algorítmica y K-complejidad es menor que  $c$  (3.6). La consecuencia del Teorema de Codificación es que asocia la frecuencia de la cadena del string  $s$  con su complejidad.

$$|\log_2 P^a(s) - K_T(s)| < c. \quad (3.6)$$

Esta fórmula ha inspirado el desarrollo del CTM (Coding Theorem Method). Siendo

$$CTM(s) = \frac{|\{T \in \mathcal{T}(n, m) : T(P) = s\}|}{|\{T \in \mathcal{T}(n, m) : T(P) : halts\}|}, \quad (3.7)$$

donde  $\mathcal{T}(n, m)$  es el espacio de todas las máquinas de Turing universales con  $n$  estados y  $m$  símbolos. La función  $D(n, m, s)$  calcula el cociente en todas las máquinas halting con  $n$  estados  $m$  símbolos que producen la cadena  $s$  [141]. El problema es que el CTM sólo puede aplicarse a cadenas cortas de 12 caracteres o menos. Para cadenas y matrices más grandes, se debe utilizar el BDM (método de descomposición en bloques). El BDM utiliza la descomposición de la cadena  $s$  en bloques  $\{b_1, b_2, \dots, b_k\}$  (posiblemente superpuestos) [132, 142], para los que el CTM proporciona una estimación y luego junta los valores basándose en la teoría clásica de la información. Sea  $s$ , una cadena larga,  $BDM(s)$  se calcula como

$$BDM(s) = \sum_{i=1}^k CTM(b_i) + \log_2 |b_i|, \quad (3.8)$$

donde  $CTM(b_i)$  denota el número de veces que el bloque  $b_i$  aparece en  $s$ . BDM básicamente se divide en dos partes. Por un lado, las aproximaciones de K-complejidad se realizan mediante CTM, cuyo orden de complejidad es 1. Por otro lado, BDM descompone los datos originales en fragmentos para los cuales CTM proporciona una

estimación y luego junta los valores basándose en la teoría clásica de la información. BDM es una medida híbrida de complejidad que combina la Entropía de Shannon pero proporciona estimaciones locales de la complejidad algorítmica. Está diseñada para mejorar las propiedades de la Entropía de Shannon que en la práctica se reducen a encontrar regularidades estadísticas y para extender la potencialidad de CTM. Consiste en descomponer los strings en piezas más pequeñas para las cuales se han estimado numéricamente las aproximaciones de complejidad algorítmica utilizando CTM, luego se aproxima la complejidad de Kolmogorov para el string más grande sumando la complejidad de los componentes individuales del string [143].

### 3.2.4. Modelos nulos

En este capítulo, empleamos dos modelos nulos para generar versiones aleatorizadas de la red original. Estos modelos son fundamentales para evaluar la complejidad de la red, ya que permiten establecer un punto de referencia al comparar las métricas de la red real con las de sus versiones aleatorizadas.

Los detalles exhaustivos sobre ambos modelos se presentan a continuación para garantizar una comprensión completa del proceso utilizado para generar redes aleatorizadas.

El primer modelo nulo se centra en preservar la distribución de grado de cada nodo, tal como se detalla en la Figura 3.2B. Este enfoque asegura que, aunque se alteren las conexiones entre nodos, la cantidad total de enlaces conectados a cada nodo se mantenga constante. Para lograr esto, el procedimiento consiste en seleccionar aleatoriamente dos pares de nodos (es decir, cuatro nodos en total) y luego intercambiar los enlaces entre ellos. Este método fue desarrollado por Huaylla et al. [130] y permite desordenar la red sin alterar su estructura fundamental.

El segundo modelo nulo tiene como objetivo preservar el número total de enlaces en la red, aunque no garantiza que la distribución de grado permanezca inalterada (Figura 3.2C). Este modelo se basa en un enfoque descrito por Watts y Strogatz [68, 144], aunque es la primera vez que se aplica específicamente a redes ecológicas.

#### Descripción detallada del algoritmo

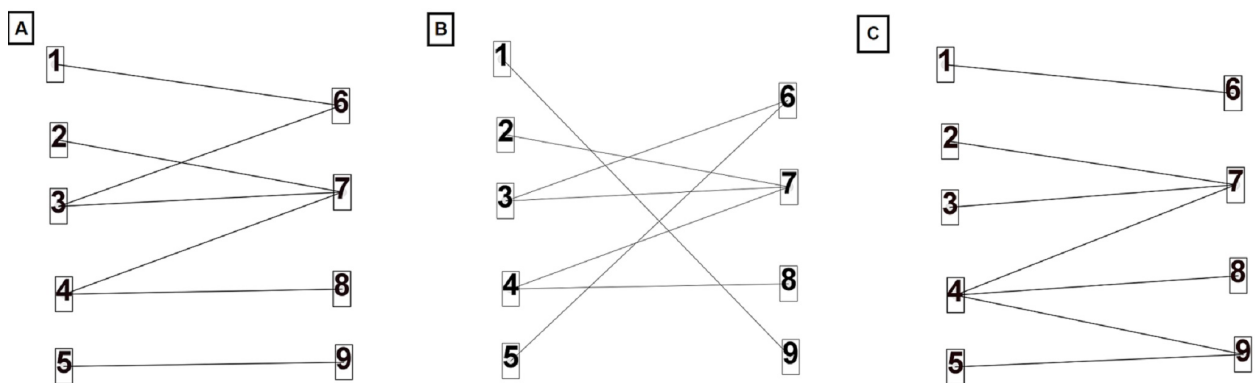
Paso 1: Armar una lista con la posición de los unos de la matriz original (número de fila, número de columna), la cual llamaremos lista 1.

Paso 2: Armar una lista con la posición de los ceros de la matriz (número de fila, número de columna), la llamaremos lista 2.

Paso 3: Elegir una fila aleatoria de la lista 1.

- Paso 4: Verificar si el nodo elegido tiene más de un vecino.
- Paso 5: Si el nodo tiene más de un vecino luego elegimos una fila aleatoria de la lista 2.
- Paso 6: Reemplazar por 1 en la matriz original usando el número de fila y de columna elegida de la lista 2 y reemplazar por 0 en la matriz original en la posición elegida de la lista 1.
- Paso 7: Calcular el número de diferentes enlaces entre la original y la red nueva. Para esto, se puede comparar la matriz final de adyacencia con la matriz original.
- Paso 8: Calcular la modularidad, la entropía, SVD entropía y K-complejidad de la nueva red.
- Paso 9: Repetir los pasos 1 hasta el 8, varias veces hasta obtener una nueva red con un grado de desorden. Una vez que los 4 índices alcanzaron un valor estable, terminaremos con las repeticiones.

En ambos modelos nulos, es fundamental tener especial cuidado para evitar la creación de enlaces dobles o nodos desconectados, lo que podría comprometer la validez del análisis posterior.

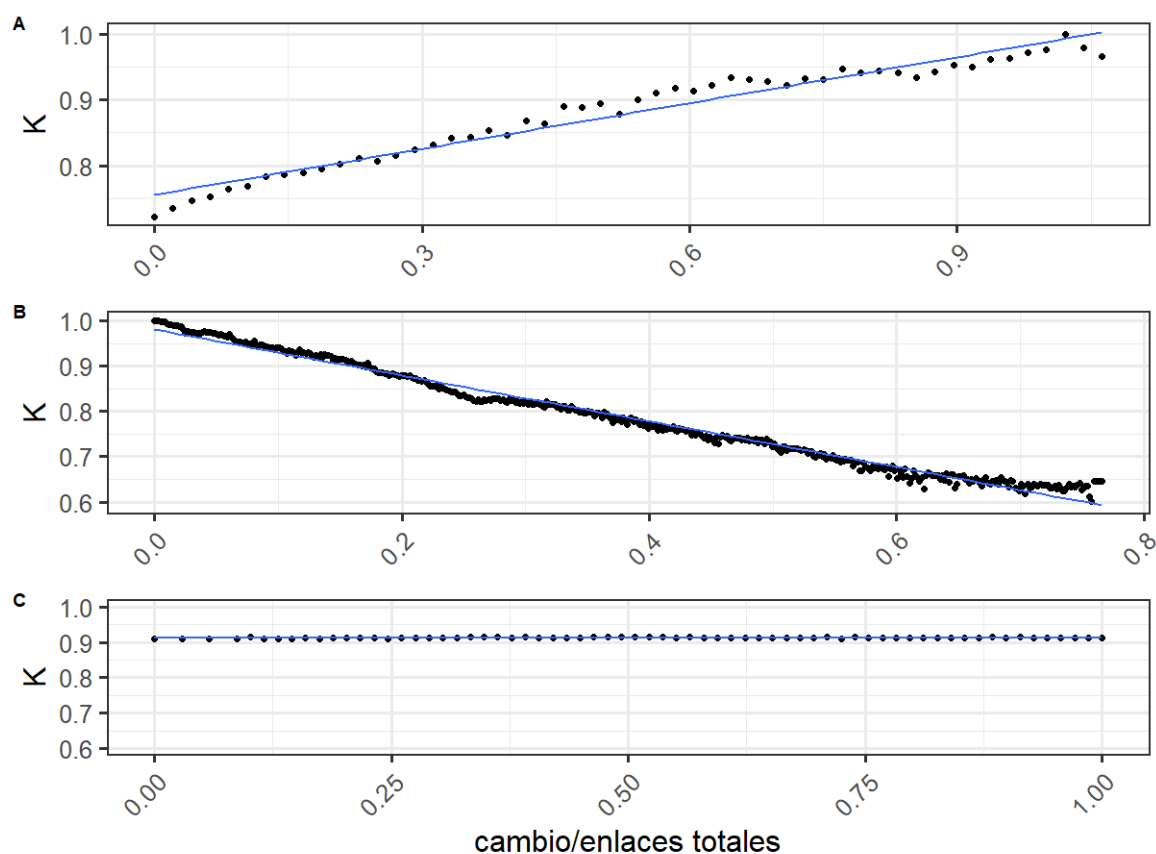


**Figura 3.2:** Representación simplificada de un cambio, donde  $A_{ij}$  es el elemento de la fila  $i$  y la columna  $j$  de la matriz de adyacencia y  $D_i$  es el grado del nodo  $i$ . (A) Grafo original.  $A_{16} = 1$ ,  $A_{59} = 1$ ,  $A_{56} = 0$ ,  $A_{19} = 0$ ,  $D_4 = 2$ ,  $D_6 = 2$ . (B) Grafo con un cambio.  $A_{16} = 0$ ,  $A_{59} = 0$ ,  $A_{56} = 1$ ,  $A_{19} = 1$ , donde la distribución de grado es la misma de la red original. (C) Grafo con un cambio, donde la distribución de grado ya no es la misma.  $D_4 = 3$ ,  $D_6 = 1$ .

### 3.2.5. Comparando medidas de complejidad

Para evaluar la eficacia del protocolo propuesto en este capítulo, utilizamos 42 redes bipartitas, donde 19 de ellas son redes de hospedador-parásito (Tabla B.1) y 23 redes planta-polinizador (Tabla B.2). Estos conjuntos de datos se obtuvieron del sitio web <https://www.web-of-life.es/>, el cual contiene una amplia variedad de interacciones

entre especies de varios continentes, diferentes entornos y metodologías a lo largo de numerosos años de muestreo. Esta metodología propone comparar los valores iniciales de entropía, modularidad, SVD entropía y K-complejidad de cada red con los valores de sus versiones aleatorizadas. Ya que considerar solamente los valores iniciales no es suficiente para afirmar que la red original tiene información codificada, es por ello que proponemos aleatorizarlas. Si la red original tiene información codificada en su sistema, los parámetros medidos deberían ser distintos a los de una red aleatoria. Por lo tanto, realizamos cambios controlados en las redes para aumentar el desorden y alterar las estructuras existentes con el fin de hacer la transición a sus versiones aleatorizadas. Mediante el seguimiento de estos índices a través de estos cambios, sus valores deberían desviarse de los iniciales a los de una red aleatoria con topologías similares. Se propusieron dos modelos: el primero conservaba la distribución de grado, mientras que el segundo modelo conserva el número de enlaces pero modifica su distribución de grado.



**Figura 3.3:** Estas figuras representan tres posibles escenarios para el comportamiento de los diferentes índices que consideramos para el estudio de la complejidad. Consideraremos tres escenarios posibles para el comportamiento del índice de K-complejidad versus los cambios/total de enlaces y luego ajustamos esos puntos utilizando un modelo lineal para identificar el signo del coeficiente principal. Tendremos tres escenarios posibles para ese coeficiente principal, replicamos esto para todos los índices calculados. A)  $\beta_k > 0$ . K-complejidad versus cambios/total de enlaces. Cada valor es un promedio de 10 redes excepto la original. B)  $\beta_k < 0$ . K-complejidad versus cambios/total de enlaces en la red original. Cada valor es un promedio de 10 redes excepto la original. C)  $\beta_k \approx 0$ . K-complejidad versus cambios/total de enlaces en la red original. Cada valor es un promedio de 10 redes, excepto la original.

Después de calcular SVD entropía, K-complejidad, la modularidad y la entropía a medida que realizamos los cambios, graficamos SVD entropía versus los cambios/total de enlaces, luego ajustamos esos puntos utilizando un modelo lineal para identificar el signo del coeficiente principal ( $\beta_{svd}$ ). Utilizando este método obtenemos un valor para cada red y podemos comparar el comportamiento con redes del mismo grupo. Lo mismo se hizo para K-complejidad versus cambios/total enlaces ( $\beta_k$ ), Entropía versus cambios/total enlaces ( $\beta_{entropía}$ ) y modularidad versus cambios/total enlaces ( $\beta_{modularidad}$ ). Una representación de esto se puede ver en la Figura 3.3. Nos interesa el comportamiento después de los cambios. Para cada red y para cada modelo, se calcularon cuatro coeficientes principales,  $\beta_{entropía}$ ,  $\beta_{modularidad}$ ,  $\beta_{svd}$  y  $\beta_k$ . Por último, comparamos estas medidas.

### 3.2.6. Medidas de complejidad aplicadas en 3 redes aleatorias conocidas

Generamos redes aleatorias para determinar qué índice era más sensible frente a los cambios realizados en las redes. Se generaron 100 redes aleatorias con las mismas características, éstas son de 100, 200, 500 y 1000 nodos y una densidad entre 0.01 y 0.5. Se calcularon cuatro índices para cada red, modularidad, entropía, SVD-entropía y K-complejidad, y se estudió el comportamiento de estos índices a medida que aumentaba la densidad de la red. Las redes fueron generadas usando el método de Erdos-Renyi, Watts - Strogatz y libre de escala. Para la generación de estas redes se crearon funciones en R usando las librerías *igraph* y *Matrix*.

#### Redes Erdős Renyi

Dado un conjunto de nodos  $N = \{1, \dots, n\}$  suponemos que la probabilidad de formar un enlace entre los nodos  $i$  y  $j$  es  $p$  y que la formación de enlaces es independiente. Generalmente para simular una red con este modelo aleatorio se enumeran todas las parejas posibles (sin importar el orden) y después se recorre la lista de parejas formando enlaces de forma independiente con probabilidad  $p$ .

Suponemos que tenemos  $n$  nodos y queremos una red de  $m$  enlaces, la probabilidad de que se forme una red específica de  $m$  enlaces es:

$$P(G) = p^m \cdot (1 - p)^{\binom{n(n-1)}{2} - m}. \quad (3.9)$$

Sea  $G(n)$  el conjunto de todas las redes con  $n$  nodos y cada red con  $m$  enlaces se selecciona con probabilidad dada por la ecuación (3.9).

Al modelo de Erdos-Renyi se lo suele notar como  $G(n, p)$ . En este modelo no fijamos el número sino la probabilidad de enlaces entre nodos. Conservamos  $n$  nodos, pero ahora colocamos un enlace entre cada par distinto con probabilidad independiente  $p$ .  $G(n, p)$  es el conjunto de grafos con  $n$  nodos en los que cada grafo simple  $G$  aparece con probabilidad  $P(G)$ .

Es posible calcular ciertos índices para estas redes, como por ejemplo, el grado de cada nodo, esto es el número de nodos con los cuales se tiene un enlace. Además, la probabilidad de que un nodo  $i$  tenga exactamente  $d$  enlaces es:

$$\binom{n-1}{d} p^d (1-p)^{n-1-d}. \quad (3.10)$$

Cuando  $n$  es grande y  $p$  es pequeño la ecuación (3.10) se puede aproximar por una distribución de Poisson,

$$\frac{e^{(n-1)p}((n-1)p)^d}{d!}. \quad (3.11)$$

Aunque el número de enlaces no está fijo es posible calcular su media o valor esperado. El número de redes que tienen exactamente  $n$  nodos y  $m$  enlaces es igual a elegir  $m$  enlaces de  $\binom{n}{2}$ , siendo éste la cantidad total de enlaces que tiene la red dado que tenemos  $n$  nodos.

La probabilidad de tener un grafo con  $m$  enlaces es

$$P(m) = \binom{\binom{n}{2}}{m} p^m (1-p)^{\binom{n}{2}-m}. \quad (3.12)$$

La media de  $m$  es

$$\langle m \rangle = \sum_{m=0}^{\binom{n}{2}} m P(m) = \binom{n}{2} p. \quad (3.13)$$

El número esperado de enlaces entre un par de nodos es igual a la probabilidad  $p$  de un enlace entre los mismos nodos, como se muestra en la ecuación (3.13).

Se deduce del resultado anterior el grado medio de un nodo en la red de Erdos-Renyi.

El grado medio en un grafo con exactamente  $m$  aristas es  $\langle k \rangle = \frac{2m}{n}$ , y por tanto el grado medio en  $G(n, p)$  es

$$\langle k \rangle = \sum_{m=0}^{\binom{n}{2}} \frac{2m}{n} P(m) = \frac{2}{n} \binom{n}{2} p = (n-1)p. \quad (3.14)$$

Para este resultado hemos usado la ecuación (3.13) y asumimos que  $n$  es constante. Esto nos dice que el número esperado de enlaces conectados a un nodo es igual al número esperado  $p$  entre el nodo y cualquier otro nodo, multiplicado por  $n-1$  porque se puede conectar con  $n-1$  nodos, ya que la red tiene  $n$  nodos en total.

Podemos calcular la probabilidad de que un nodo sea conectado a  $k$  nodos como  $p^k(1-p)^{n-1-k}$ . Además como el número total de nodos es  $n$  entonces tenemos  $\binom{n-1}{k}$  maneras de elegir  $k$  nodos. Luego la probabilidad de que un nodo sea conectado a  $k$  nodos está dada por la ecuación (3.15), la cual sigue una distribución binomial.

$$p_k = \binom{n-1}{k} p^k (1-p)^{n-1-k}. \quad (3.15)$$

Para generar las redes de Erdos-Renyi creamos una función usando R donde tenemos como argumento la cantidad de nodos, enlaces, la semilla que será usada y el método *cluster* a utilizar. Esta función fue pensada para usar el método de Louvain, que fue el método usado para detección de comunidades de la sección anterior, pero además agregamos a la función el método como un argumento ya que la idea es usar otros

métodos para decidir. A la función la llamaremos *generar\_redes*, donde se usaron 100 semillas  $seq(10, 1000, 10)$ , de esta forma nos garantizamos tener 100 redes originales diferentes. Además, estas redes tienen distintas densidades, en este caso usamos las siguientes densidades: 0.010, 0.025, 0.035, 0.05, 0.075, 0.1, 0.2, 0.3, 0.4 y 0.49.

En este caso en particular, para corroborar los resultados obtenidos mediante el algoritmo de Louvain, se utilizó el algoritmo de Leiden, que ofrece mejoras en la calidad y estabilidad de las particiones detectadas. La comparación entre ambos resultados permitió verificar la consistencia de las comunidades detectadas y asegurar que las particiones fueran más robustas. En R la función para usar el algoritmo de Leiden tiene como argumento un parámetro de resolución, en consecuencia tuvimos que determinar qué parámetro de resolución usar. Para esto generamos una función llamada *buscando\_el\_parametro*, donde los argumentos son los nodos, los enlaces, la semilla, el método de *cluster* que voy a usar y el parámetro es el parámetro de resolución. Usamos un vector para el parámetro  $seq(0.20, 1, 0.010)$  para cada tipo de red y de esta forma identificamos donde se produce algún cambio, es decir, al principio hay un número de comunidades y después se produce un cambio brusco en el medio y de esa forma determinamos el parámetro.

## Redes Watts-Strogatz

Las redes de Watts-Strogatz (W-S) contienen  $n$  nodos y cada uno de ellos tiene grado  $c$ , es decir, tenemos  $c$  enlaces que se conectan con cada nodo. Se recorre cada uno de los enlaces sucesivamente y con cierta probabilidad  $p$  se elimina ese enlace y se reemplaza por otro que conecte dos nodos elegidos uniformemente al azar. Si  $p = 0$  no se cambian los enlaces y conservamos el círculo original. Cuando  $p = 1$  todas las aristas se reordenan en lugares aleatorias y tenemos una red aleatoria. Para valores intermedios de  $p$  generamos redes intermedias.

El punto crítico sobre el modelo es que a medida que  $p$  aumenta, las comunidades se mantienen hasta valores bastante grandes de  $p$ . Como el modelo de W-S tiene forma de anillo, al principio podemos tener un camino largo para llegar de un nodo a otro. Una vez que se agregaron los enlaces para llegar a un nodo desde otro nodo, el grado de ese nodo es  $c$  más la cantidad de enlaces que se agregaron a él. La definición de este tipo de redes nos dice que para cada uno de los enlaces originales, de los cuales son  $\frac{1}{2} \cdot n \cdot c$  se añade un camino más corto con probabilidad  $p$  entre dos nodos aleatorios. De esta forma ahora tenemos  $\frac{1}{2} \cdot n \cdot c \cdot p$  caminos más cortos entre dos nodos, es decir, que en promedio tenemos  $c \cdot p$  formas de llegar de un nodo a otro mediante un camino más corto. En particular, el número de caminos más cortos ( $s$ ) a cualquier nodo sigue una distribución de Poisson con media  $c \cdot p$ .

$$p_s = e^{-cp} \frac{(cp)^s}{s!}. \quad (3.16)$$

Si el grado total de un nodo es  $k = s + c$ , resulta  $s = k - c$  en la ecuación (3.16) por lo tanto la distribución de grado de una red de W-S es:

$$p_k = e^{-c.p} \frac{(c.p)^{k-c}}{(k-c)!}, \quad (3.17)$$

para  $k \geq c$  y  $p_k = 0$  si  $k < c$ .

Por otro lado, el modelo de W-S nunca pretendió simular las distribuciones de grado del mundo real. Lo que sí hace bien es reflejar los valores de *clustering* y de las distancias que se ven en las redes reales.

Para generar redes de W-S armamos una función a la que llamamos *generar\_redes2* que tiene como argumento los nodos, el número de vecinos, la probabilidad, la semilla y el método *cluster*. En este caso a diferencia del método de Erdos-Renyi tuvimos que agregar como argumento la probabilidad ya que ese argumento es importante para generar las redes de W-S. Al usar la función *generar\_redes2* usamos el método de Louvain.

El método de W-S parte de una red ordenada, luego recorre cada uno de los vínculos de la red y se decide con cierta probabilidad si cambio ese vínculo o no, esa probabilidad es la que agregamos como argumento de la función. Como se menciona previamente para valores de  $p$  cercanos a 0 se cambian pocos vínculos, pero para valores altos de  $p$  se cambian más vínculos y hay más desorden en la red. Sin embargo, la densidad no cambia porque se mantiene la cantidad de vínculos de la red original. En la simulación usamos diferentes valores de probabilidad, los valores usados fueron: 0.01, 0.05, 0.1, 0.2, 0.5 y para cada valor de probabilidad generamos 100 redes. En la función anterior usamos el algoritmo de Louvain para generar la división de comunidades.

También usamos el algoritmo de Leiden y directamente armamos una función para generar redes y determinar las comunidades con los parámetros de resolución seleccionados (0.5, 0.6, 0.7) previamente con el modelo anterior. La función se llama *generar\_redes4* y tiene como argumento los nodos, vecinos, probabilidad, la semilla, el método *cluster* y el parámetro.

## Redes Libres de escala

El nombre libre de escala (Scale free) captura la falta de una escala interna, una consecuencia de esto es que en la misma red coexisten nodos con grados muy diferentes, es decir, algunos nodos están conectados con muchos nodos y otros con pocos. Esta característica es la que la distingue a las redes sin escala de las redes que tienen el mismo grado para todos sus nodos, o de las redes aleatorias, cuyos grados varían en un

rango acotado.

En las redes libres de escala, los grados de los nodos más pequeños y más grandes son ampliamente diferentes. En este tipo de redes el  $n$ -ésimo momento de la distribución de grado, está dado por la siguiente ecuación:

$$k^n = \int_{K_{min}}^{k_{max}} K^n p(k) dk = C \frac{k_{max}^{n-\gamma+1} - K_{min}^{n-\gamma+1}}{n - \gamma + 1}, \quad (3.18)$$

donde  $k_{min}$  está fijo y  $k_{max}$  es el grado máximo de los nodos.

Los nodos de estas redes tienen grados comparables en una red aleatoria. Dado que el valor del exponente de grado desempeña un papel importante en la predicción de diversas propiedades de la red, necesitamos herramientas para ajustar la distribución  $p(k)$  y estimar  $\gamma$ . Esto nos lleva a abordar varios temas relacionados con la representación gráfica y el ajuste de leyes de potencia. Cuando se tienen nodos con grados muy distintos, una representación lineal resulta incapaz de exhibirlos todos de manera adecuada. Para obtener las distribuciones de grado nítidas y claras se usa un gráfico con una doble escala logarítmica, a lo que se le suele llamar un gráfico log-log. Las redes de escala libre están presente en muchas redes de interés y de amplias disciplinas.

Se generaron redes libres de escala, ya que éstas nos pueden dar información sobre el comportamiento de las redes reales. En este trabajo se ha creado una función para generar esas redes y se la denominó *generar\_redes5*, la cual tiene como argumentos los nodos, parampl, semilla, método *cluster* y parámetro. Siendo parampl la cantidad de nodos que tenemos inicialmente en la red, de esta forma se generaron redes con cierta densidad. El argumento de la función llamado parámetro, es el parámetro de resolución de Leiden, en este caso usamos los parámetros elegidos previamente, 0.5, 0.6 y 0.7.

### 3.3. Resultados

#### 3.3.1. Modelos nulos

##### Conservando la distribución de grado

Se aleatorizó la red original preservando el grado de cada nodo y se calcularon los 4 índices: Entropía, modularidad, SVD entropía y K-complejidad.

Luego calculamos  $\beta_{entropía}$ ,  $\beta_{modularidad}$ ,  $\beta_{SVD}$  y  $\beta_k$  para cada red seleccionada, con el fin de identificar cuáles comienzan a crecer o decrecer a medida que realizamos los cambios.

Los resultados obtenidos revelaron que en el caso de las redes de planta-polinizador, la modularidad decrece y la entropía aumenta a medida que los cambios se incrementan, esto se puede observar en el signo de  $\beta_{modularidad}$  y  $\beta_{entropía}$  (Figura 3.4 C- Figura 3.4 D). Además, el desorden de la red crece a medida que los valores de entropía cam-

bian hacia los valores de una red aleatoria. El cambio en la entropía es un indicador de que en la red se perdió información codificada. Sin embargo, en el caso de las redes hospedador-parásito el desorden no es reflejado utilizando la modularidad, ni la entropía. Observamos que a veces  $\beta_{modularidad} > 0$  y otras veces  $\beta_{modularidad} < 0$ . La densidad de las redes hospedador-parásito es mayor que la de la red polinizador-planta. La densidad juega un papel importante y puede ser la razón de que ciertas herramientas no reflejen lo esperado. Es decir, ciertas herramientas pueden proporcionar información relevante sobre la red mientras que otras no porque son sensibles a la densidad de la red. En el caso del estudio de la complejidad, deben seleccionarse determinadas herramientas para explicar la complejidad en función de su densidad.

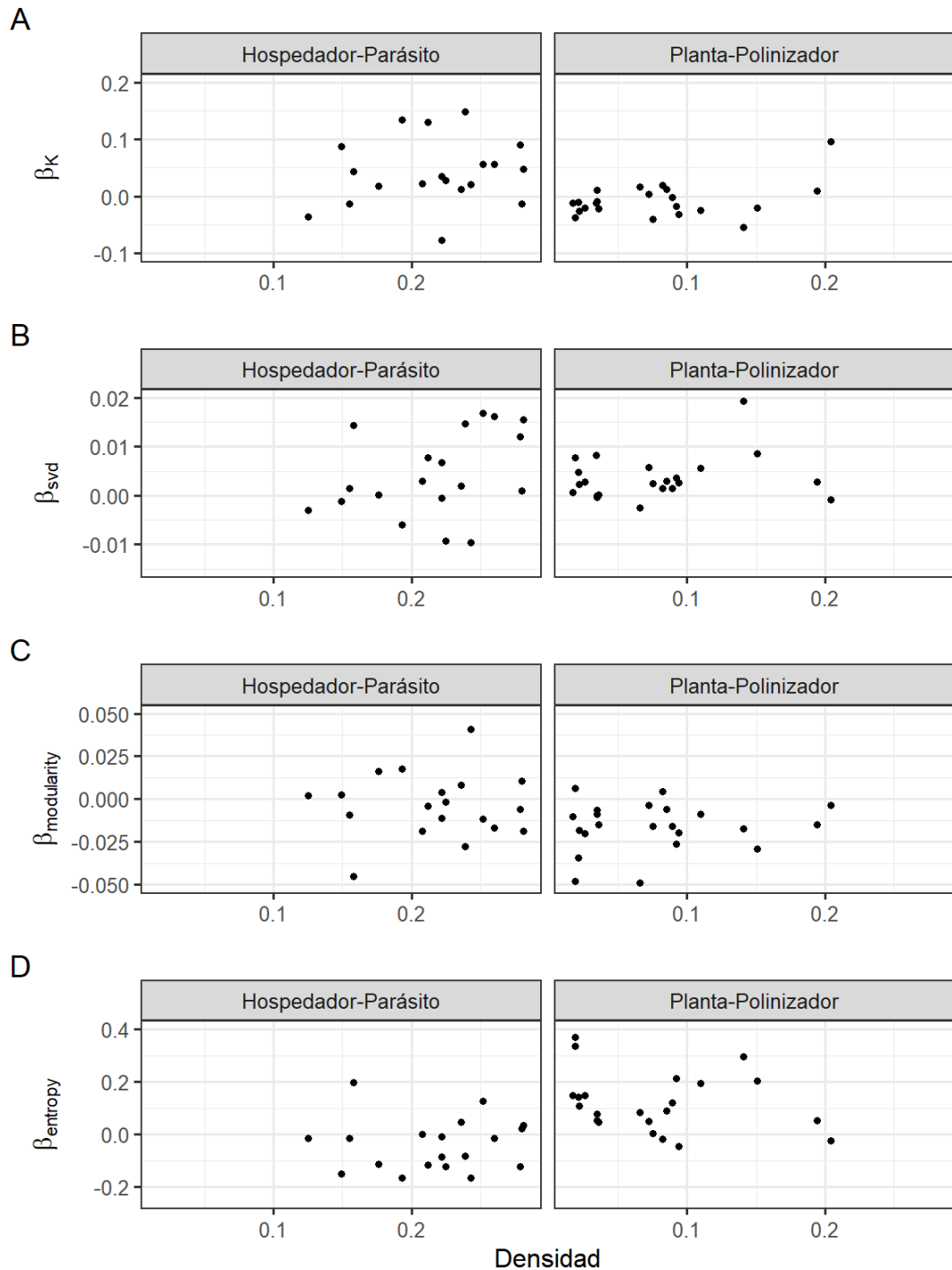
En este capítulo, las redes polinizador-planta tienen una densidad inferior a 0.2. Si las redes tienen una densidad superior a 0.15, la entropía puede no ser un indicador para medir la aleatoriedad. La entropía fue calculada utilizando la división de comunidades pero este índice no refleja el cambio que se hizo originalmente para las redes que tienen una densidad mayor a 0.20. Al hacer los cambios, hay nodos que no cambian de comunidad porque no tienen mucho margen de cambio y por lo tanto la entropía no varía cuando se generan redes aleatorias en redes con una densidad alta y como consecuencia la entropía no se puede utilizar para determinar la complejidad de una red si tiene una densidad mayor a 0.2.

Además, se calcularon otros índices para determinar la complejidad de las redes. SVD-entropía no mostró ningún cambio significativo en el caso de las redes polinizador-planta a medida que realizábamos cambios y casi ninguna variación, la variación se situó entre 0 y 0.1. En el caso de las redes hospedador-parásito, observamos que para algunas redes había un cambio más pronunciado que en otros casos.

K-complejidad reveló que las redes polinizador-planta son más complejas y que la información codificada se pierde a medida que aumenta el nivel de aleatorización. En este caso  $\beta_k$  generalmente toma un valor negativo y para la red hospedador-parásito muchas veces ese valor es positivo. Esto significa que en el caso de las redes polinizador-planta los valores disminuyen al aumentar los cambios.

### Preservando el número de enlaces

Utilizando el segundo método para realizar los cambios, el cual preserva el número de enlaces pero cambia su distribución de grado, se calculó SVD entropía y K-complejidad a medida que realizábamos los cambios para determinar el grado de complejidad de las redes hospedador-parásito. Se observó que en las redes polinizador-planta, el valor de K-complejidad disminuye a medida que realizamos los cambios, pero no ocurre lo mismo cuando utilizamos las redes hospedador-parásito, es decir, estas redes no presentan información codificada en su sistema. Esto se puede obser-



**Figura 3.4:** A) Relación entre  $\beta_k$  y la densidad para cada red, vale la pena resaltar que el grado de cada nodo se conserva en estas redes. B) Relación entre  $\beta_{svd}$  y densidad para cada red. C) Relación entre  $\beta_{modularidad}$  y densidad para cada grafo. D) Relación entre  $\beta_{entropía}$  y densidad para cada red.

var en la Figura 3.5 A) cuando vemos que una gran proporción de las pendientes toman valores negativos en el caso de las redes polinizador-planta frente a las redes hospedador-parásito. Utilizando SVD-entropía vimos que, al aumentar los cambios, se convierten en redes más complejas, esto puede deberse a que su densidad es superior a 0.1. SVD entropía sigue sin ser una herramienta confiable para determinar la complejidad de una red con una densidad superior a 0.2 porque en estos casos la pendiente puede ser negativa o positiva.

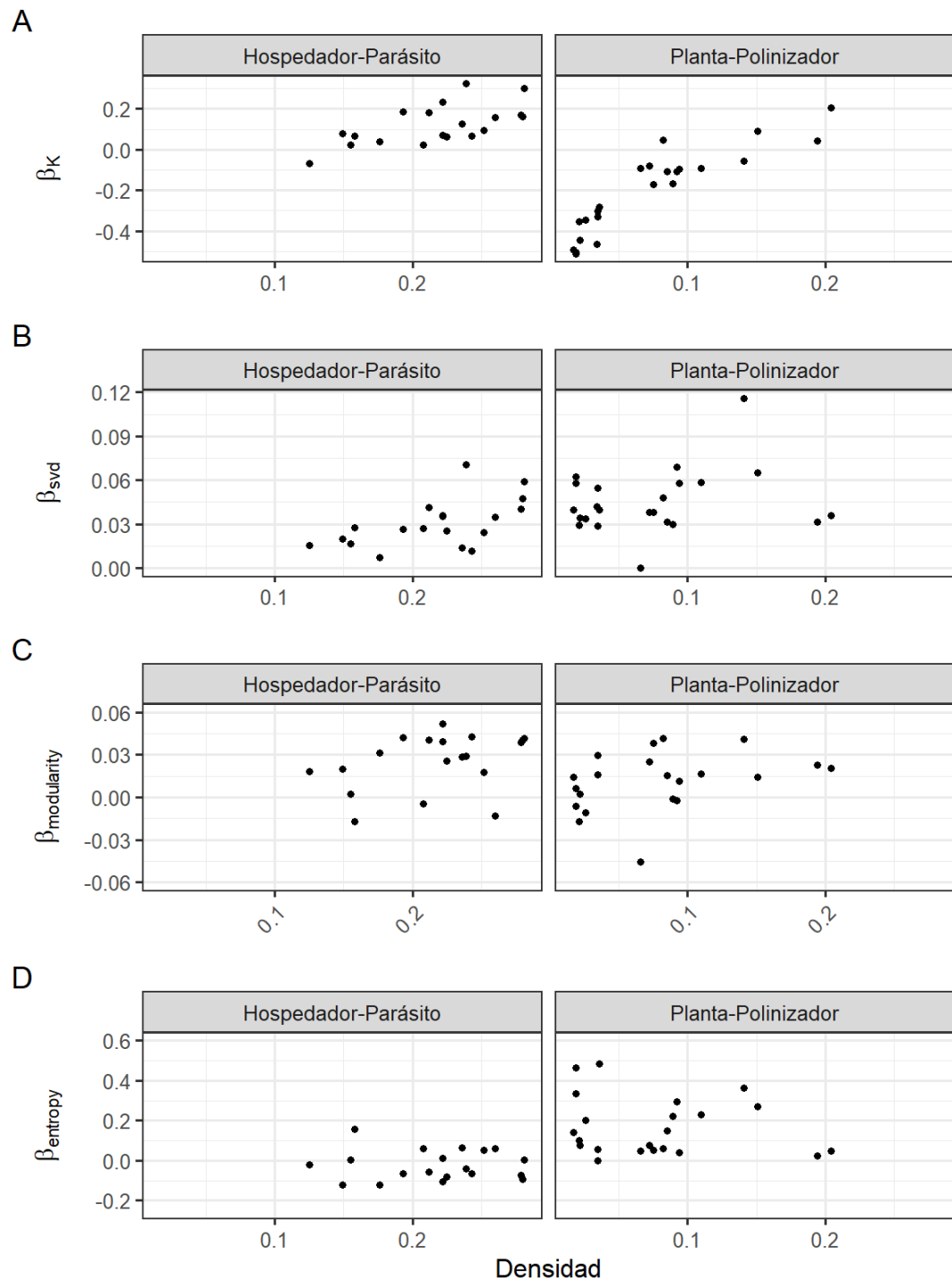
Para las redes planta-polinizador,  $\beta_{entropía} > 0$ , lo que significa que la red original tenía información codificada en su sistema. Sin embargo, en las redes hospedador-parásito, a veces  $\beta_{entropía}$  es negativo y otras veces es positivo. Esto podría ser un indicio de que algunas redes están más cerca que otras de su complejidad máxima dada su densidad. La distribución de grado se modificó y la modularidad se vio influenciada por este cambio. La Figura 3.5.C muestra que los valores de  $\beta_{modularidad}$ , en general, son mayores que cero porque al cambiar la distribución de grado, también se modificaron las comunidades, lo que produjo cambios notables en este índice.

Es importante destacar que la distribución de grado probablemente está asociada con características de la red que responden a características del sistema y que cuando se realizan cambios aleatorios, esto se pierde. El segundo método deja claro que la distribución de grado es relevante en el análisis de redes, es decir, juega un papel clave en su análisis. Es razonable que el valor disminuya en las redes planta-polinizador porque la red original tiene información codificada en su sistema, por eso el valor del índice de complejidad es alta y cuando se aleatoriza la red, pierde esa información codificada en su sistema.

Vale la pena señalar que el subconjunto de redes seleccionadas para el estudio refleja lo que ya se sabía, que las redes de polinizadores son más anidadas que las redes de hospedador-parásito (ver apéndice Figura B.1). No obstante, fue posible avanzar en el estudio de la complejidad de una red en el área de ecología utilizando estos dos conjuntos de datos.

Se calcularon los valores iniciales de SVD entropía (ver apéndice Figura B.2) y en todos los casos los resultados obtenidos indicaron que los valores de entropía eran elevados.

Los detalles de la cantidad de valores de  $\beta$  cuyo intervalo de confianza incluye el cero, la cantidad de valores de  $\beta$  positivos cuyo intervalo de confianza no incluye al cero y la cantidad de valores de  $\beta$  negativos donde su intervalo de confianza no incluye al cero se puede ver en la Tabla B.3.



**Figura 3.5:** A) Relación entre  $\beta_k$  y la densidad de cada red; el número de enlaces se mantiene en estas redes pero su distribución de grado no. B) Relación entre  $\beta_{svd}$  y la densidad para cada red. C) Relación entre  $\beta_{modularidad}$  y la densidad para cada red. D) Relación entre  $\beta_{entropía}$  y la densidad para cada red.

## 3.4. Discusión

En este estudio, nuestro objetivo principal fue caracterizar diversos aspectos de la complejidad de las redes ecológicas, utilizando un conjunto representativo de datos. Se analizaron 23 redes de interacción planta-polinizador y 19 redes de interacción hospedador-parásito extraídas de la base de datos web-of-life.es. Para llevar a cabo esta caracterización, se calcularon varios índices fundamentales que incluyen el número de enlaces, el número de nodos, la asortatividad, el anidamiento, la densidad, la modularidad y la entropía. Además, se introdujeron dos índices específicos para estimar la complejidad de las redes: la entropía de descomposición en valores singulares (SVD) y la complejidad de Kolmogorov (K-complejidad). La entropía SVD permite cuantificar la factorización de una matriz en matrices ortogonales, mientras que la K-complejidad proporciona una aproximación del valor real de la complejidad mediante diferentes algoritmos.

Para evaluar adecuadamente las características estructurales de las redes, se implementaron dos modelos nulos. El primer modelo preserva el grado de cada nodo al intercambiar aleatoriamente pares de nodos conectados. El segundo modelo mantiene constante el número total de enlaces al desconectar un nodo y volverlo a conectar a otro nodo aleatorio. A pesar del aumento en la aleatoriedad introducido por estos modelos, ciertas características fundamentales de la red permanecen constantes: en el primer modelo se conserva la distribución de grado y en el segundo se mantiene el número total de enlaces. Esta metodología nos permitió evaluar los cambios en la red original a medida que se incrementaba la aleatoriedad.

Uno de los hallazgos más significativos fue que la K-complejidad disminuyó monótonamente hasta alcanzar un estado completamente aleatorio en las redes planta-polinizador cuando se aplicó el modelo que preserva el número total de enlaces. Asimismo, descubrimos que tanto la entropía como la modularidad son herramientas confiables para analizar redes con una densidad inferior a 0.1. A medida que se introducen cambios en la red, observamos que el valor de entropía tiende a aumentar, lo que indica un incremento en el desorden y una pérdida progresiva de información respecto a la estructura original. La modularidad complementa estos resultados al evidenciar una pérdida en la estructura comunitaria dentro de las redes analizadas.

En función de los resultados obtenidos, hemos establecido criterios claros para seleccionar entre los dos modelos nulos. En situaciones donde la densidad es menor a 0.1, se recomienda utilizar el primer modelo; en estos casos, tanto la entropía como la modularidad son suficientes para determinar si una red presenta una estructura significativamente diferente a aquella observada en redes aleatorias. Por otro lado, si la densidad supera 0.1, es más apropiado aplicar el segundo modelo, que permite observar cómo otros índices como SVD entropía y K-complejidad responden ante cambios en la

red.

Es importante destacar que aunque este análisis se centró exclusivamente en redes bipartitas, nuestra metodología es versátil y puede aplicarse a otros tipos de redes. Este enfoque es particularmente útil para identificar cuándo una red carece de información codificada en su sistema. La ausencia de información puede ser indicativa de enlaces faltantes o estructuras subyacentes no detectadas. Además, al estudiar redes sin información codificada, podemos hacer predicciones sobre enlaces futuros entre nodos. La predicción de enlaces es un área emergente que tiene el potencial de ayudar a identificar conexiones irrelevantes dentro de una red [145, 146]. Comprender cómo los cambios afectan la estructura de las redes es esencial para avanzar en nuestra comprensión ecológica.

Finalmente, nuestros algoritmos e índices seleccionados tienen aplicaciones potenciales en diversas disciplinas más allá del ámbito ecológico. Esta flexibilidad permite verificar si diferentes tipos de redes contienen información valiosa y evaluar su nivel general de complejidad. Así, esperamos que nuestros hallazgos no sólo contribuyan al entendimiento actual sobre las interacciones ecológicas sino que también sirvan como base para futuras investigaciones que exploren nuevas dimensiones en el estudio de las redes complejas.

# Capítulo 4

## Hipergrafos: Un enfoque complementario

### 4.1. Introducción

La teoría de grafos se ha transformado en una herramienta fundamental en diversas disciplinas, abarcando desde la informática hasta la biología. Su capacidad para representar sistemas complejos mediante redes, donde las entidades se modelan como nodos y las interacciones como aristas, ha llevado a su adopción en tareas de reconocimiento de patrones y aprendizaje automático. En los últimos años, se ha intensificado la investigación en el ámbito ecológico, donde los grafos permiten describir la estructura de sistemas complejos al especificar las conexiones entre pares de entidades. Sin embargo, esta representación binaria presenta limitaciones cuando las interacciones involucran a más de dos entidades simultáneamente, lo que ha llevado al desarrollo de herramientas más sofisticadas como los hipergrafos.

Los hipergrafos son una generalización de los grafos que permiten que sus aristas, denominadas hiperaristas, conecten múltiples nodos. Esta característica es crucial para representar interacciones complejas y simultáneas entre varias entidades, ofreciendo una representación más rica en contextos como la biología. A pesar de su potencial, el uso de hipergrafos en biología es menos conocido que el de los grafos tradicionales; a menudo se aplican sin ser explícitamente identificados. Sin embargo, su estructura permite superar limitaciones conceptuales y proporciona descripciones más completas de los sistemas reales en estudio [147]. La capacidad de los hipergrafos para modelar interacciones de orden superior es particularmente valiosa en el análisis de redes ecológicas. Por ejemplo, las redes alimentarias pueden representarse mediante grafos dirigidos, donde los nodos representan especies y las aristas indican relaciones tróficas. Sin embargo, esta representación no captura completamente la complejidad de las interacciones entre grupos de especies competidoras. Para abordar esta limitación, se

propone el uso de un hipergrafo de competencia, donde los nodos representan especies y los hiperenlaces agrupan aquellas que compiten por recursos comunes [148]. Investigaciones han demostrado que los hipergrafos de competencia ofrecen una descripción más detallada de las relaciones depredadoras en comparación con los grafos convencionales.

Las interacciones de orden superior poseen un gran poder explicativo, pero su aplicación directa en modelos empíricos es escasa debido a la falta de datos adecuados [149]. En ciertos casos, como en redes de coautoría o ecosistemas específicos, es posible observar directamente estas interacciones. Sin embargo, para muchos sistemas —como la actividad cerebral o las interacciones sociales— esta información no está disponible. En el contexto ecológico, incorporar interacciones de orden superior ha demostrado estabilizar dinámicas poblacionales y facilitar la coexistencia entre especies frente a perturbaciones [150]. Estudios recientes han explorado cómo medir la influencia de cada especie dentro de un hipergrafo mediante la entropía de sus hiperenlaces, analizando cómo esta medida se relaciona con propiedades críticas como el grado de nodos y coeficientes de agrupamiento [151]. Este enfoque tiene implicaciones significativas para caracterizar la complejidad de las redes ecológicas usando otras herramientas [152] y permite obtener interacciones indirectas [64].

El objetivo principal del presente capítulo es comparar la información sobre comunidades obtenida en el Capítulo 2 con los nodos que componen cada hiperenlace del hipergrafo construido a partir de datos empíricos. Se busca resaltar la importancia del uso de hipergrafos como herramientas efectivas para representar y estudiar redes ecológicas. En particular, nos proponemos: (1) determinar la relación entre las comunidades identificadas y los nodos que componen los hiperenlaces; (2) identificar los nodos más influyentes dentro del hipergrafo. Los detalles metodológicos relacionados con esta identificación se presentan en la Sección 4.2, mientras que los resultados y su interpretación se discuten en la Sección 4.3. Finalmente, la Sección 4.4 aborda las implicancias del trabajo realizado y su relevancia para futuras investigaciones en ecología y otras áreas relacionadas.

Como referencia en este capítulo se usaron los datos del capítulo 2. La red contiene 3 tipos de nodos y dos tipos de interacción, la misma está compuesta por polinizadores, especies de plantas y gremios de herbívoros. Esta es una red multicapa y contiene tres capas.

## 4.2. Metodología

### 4.2.1. Construcción del hipergrafo a estudiar

Varios estudios en redes sociales y en diferentes disciplinas [153–156] sugieren que las interacciones entre nodos pueden explicar la densidad local de las redes. Según

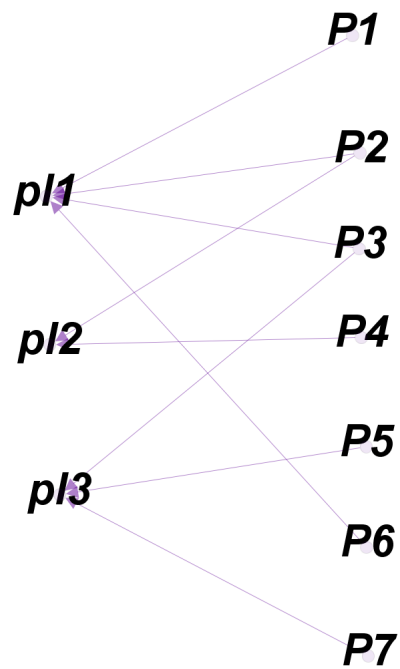
estas investigaciones, los nodos o entidades están conectados debido a un contexto compartido, lo que permite la formación de conexiones de orden superior que incluyen relaciones más complejas entre esos nodos.

Para identificar interacciones de orden superior en datos de redes, a menudo se utilizan cliques [157–159]. Los cliques son subgrafos completamente conectados, lo que significa que todos los nodos dentro de una clique son vecinos entre sí. Esta propiedad permite detectar interacciones de orden superior. Sin embargo, Young [160] señala que la descomposición en cliques no resuelve completamente el problema de recuperar todas las interacciones relevantes. Determinar qué tipo de descomposición utilizar puede ser complicado; por ejemplo, un triángulo puede descomponerse en una única 2-clique o en tres 1-cliques. En teoría de grafos, se busca el conjunto mínimo de cliques que cubra toda la red.

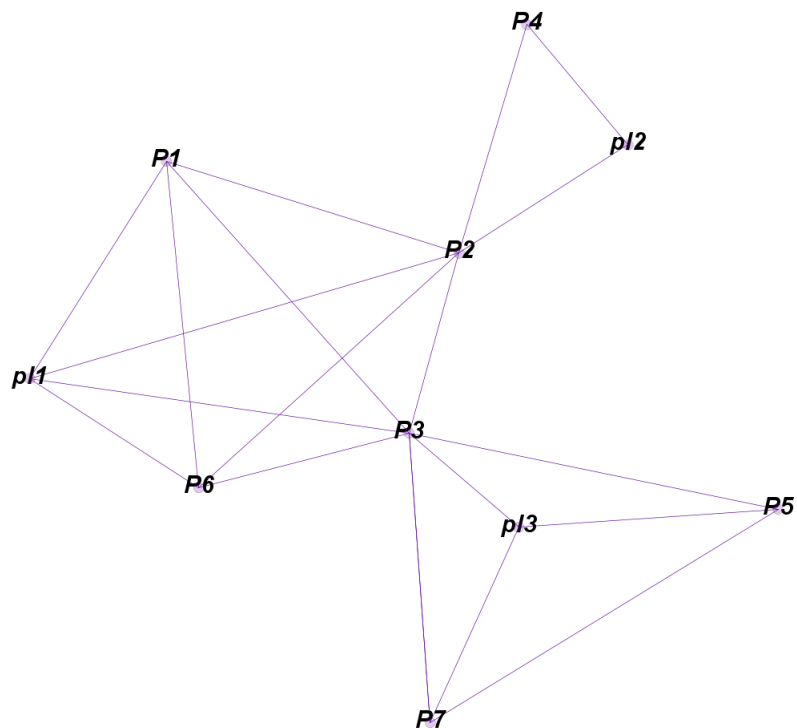
Además, existen enfoques que incorporan elementos o procesos aleatorios para abordar este problema. Algunos modelos simulan cómo se generan los datos, lo que ayuda a estabilizar las soluciones y restringirlas para evitar que sean excesivamente complejas o que sobreajusten los datos. Esto es especialmente útil en problemas mal planteados o aquellos con múltiples soluciones posibles.

Una vez identificadas las cliques, se procede a crear un hipergrafo. En este contexto particular, se buscó construir un hipergrafo a partir de un grafo bipartito. Young et al. [160] proponen un método basado en el principio de parsimonia y utilizan una estructura de orden superior fundamentada en evidencia estadística para determinar el hipergrafo. Para aplicar el método propuesto por Young, se realizó una proyección de la red con el fin de obtener cliques a partir del conjunto de datos. En redes bipartitas, solo existen enlaces entre dos tipos de nodos (es decir, 1-clique). En nuestro caso de estudio, se analiza una red tripartita compuesta por polinizadores, especies de plantas y gremios de herbívoros. El desafío radica en que no es posible determinar cliques directamente a partir de la red estudiada. Por ello, se llevó a cabo una proyección donde se considera que si dos polinizadores visitan la misma especie de planta, esos nodos están conectados entre sí (Figura 4.1). De esta manera, obtenemos un grafo (Figura 4.2) del cual es posible hallar cliques y posteriormente construir un hipergrafo.

A modo de ejemplo, a continuación representamos una red bipartita, donde tenemos 7 polinizadores y 3 especies de plantas.



**Figura 4.1:** Red bipartita de 10 nodos. 3 plantas (*pl1*, *pl2* y *pl3*) y 7 polinizadores (*P1*, *P2*, *P3*, *P4*, *P5*, *P6* y *P7*).



**Figura 4.2:** Grafo de 10 nodos construido a partir de la red bipartita.

Otro método para crear un hipergrafo es el propuesto por Estrada [161]. En ese



índice adecuado para comparar comunidades con hiperenlaces. El índice Rand [162, 163] es una herramienta que evalúa la similitud entre dos particiones. Cada partición se representa comúnmente mediante dos matrices ( $R$  y  $Q$ ). Asumiendo que la partición de referencia es  $R$ , esta matriz divide los datos en  $k$  clases, mientras que la otra partición clasifica los datos en diferentes categorías o *clusters*.

El índice Rand evalúa cuántos pares de nodos están presentes en ambas particiones, es decir, cuántos nodos comparten el mismo grupo y cuántos están en grupos distintos. Sin embargo, este índice no es adecuado para escenarios donde un elemento puede pertenecer a más de un grupo, como ocurre en los hipergrafos, donde los hiperenlaces pueden compartir nodos.

Para abordar esta limitación, Campello [164] extiende una medida que permite realizar comparaciones entre un grupo de nodos disjuntos y otro grupo que puede tener intersecciones. Esta extensión posibilita que los elementos pertenezcan a múltiples grupos simultáneamente. El índice propuesto no solo considera la coincidencia exacta de pares de nodos, sino que también evalúa el grado en el cual estos pares comparten su pertenencia en ambas particiones.

El índice Rand [162] es definido como sigue :

$$oAri = \frac{|V \cap Y| + |X \cap Z|}{|V \cap Y| + |V \cap Z| + |X \cap Y| + |X \cap Z|}, \quad (4.1)$$

donde  $V, X, Y$ , y  $Z$  se describen como sigue:

$V$ : es el conjunto de pares de datos que pertenecen a la misma comunidad  $R$ .

$X$ : es el conjunto de pares de datos que pertenecen a diferentes comunidades en  $R$ .

$Y$ : es el conjunto de pares de datos que pertenecen al mismo hiperenlace en  $Q$ .

$Z$ : es el conjunto de pares de datos que pertenecen a diferentes hiperenlaces en  $Q$ .

Llamando

$$a = |V \cap Y|$$

$$b = |V \cap Z|,$$

$$c = |X \cap Y|,$$

$$d = |X \cap Z|$$

$$oARI = \frac{a + d}{a + b + c + d}. \quad (4.2)$$

Para tener en cuenta las particiones aleatorias, se normalizó el valor del índice anterior. De esta forma, su valor (esperado) es 0 cuando las particiones se seleccionan al azar, lo que es equivalente a decir que las particiones son muy distintas y 1 cuando se consigue una coincidencia perfecta.

El índice corregido es:

$$oARI_{adj} = \frac{a - \frac{(a+c)(a+b)}{d}}{\frac{(a+c)+(a+b)}{2} - \frac{(a+c)(a+b)}{d}}. \quad (4.3)$$

Este índice fue programado en R ya que no hay un paquete que lo calcule directamente. La mayor parte de los paquetes realizan la comparación entre conjuntos disjuntos. Después de calcular el índice nos interesaba determinar qué tan diferente era ese valor cuando aleatorizábamos la red original. Para esto generamos hipergrafos de esas redes aleatorizadas y comparamos las comunidades de cada red con sus hiperenlaces.

### 4.2.3. La entropía como herramienta para determinar nodos influyentes

En el Capítulo 3 usamos la entropía como una herramienta para medir la complejidad de la red. Este índice es ampliamente usado en diferentes disciplinas [138, 165], es usado para procesos de toma de decisiones y también se lo utiliza para medir la complejidad estructural en grafos. En este capítulo usamos la entropía para el análisis de nuestro hipergrafo.

Si la probabilidad de que ocurra un evento es representada por  $p$ , el valor total de la entropía usando la probabilidad de  $n$  eventos es calculado por la siguiente ecuación:

$$h(p_1, p_2, \dots, p_n) = - \sum_{i=1}^n p_i \cdot \log(p_i). \quad (4.4)$$

Podemos calcular el valor de la entropía de cada nodo o de cada hiperenlace, para lo cual se utilizan los valores de la matriz  $(I_{ij})$ , que representa al hipergrafo, donde  $e$  representa a los hiperenlaces y  $v$  representa a los nodos. Además es posible hallar el grado de cada hiperenlace,  $d(v)$  indica el grado del nodo y  $d(e)$  indica el grado del hiperenlace. En el caso de los hipergrafos pesados, se suman los pesos.

**Tabla 4.1:** Matriz de incidencia y grados

	$e_1$	$e_2$	$e_3$	$e_4$	$d(v)$
$v_1$	1	1	0	0	2
$v_2$	0	1	0	1	2
$v_3$	1	1	1	0	3
$v_4$	1	0	1	1	3
$v_5$	0	0	0	1	1
$d(e)$	3	3	2	3	

Utilizando los valores de la Tabla 4.1, se pueden calcular los valores de entropía de los nodos como así también de los hiperenlaces usando las siguientes ecuaciones. Estos

valores indican la influencia de los nodos o hiperenlaces en el sistema.

$$h(v_i) = - \sum_{j=1}^n \frac{|I_{ij}|}{d(v_i)} \log \left( \frac{|I_{ij}|}{d(v_i)} \right), \quad (4.5)$$

siendo  $n$  la cantidad de nodos del hipergrafo a estudiar que tienen asignado un valor distinto de cero en la fila  $v_i$ .

$$h(v_1) = - \sum_{j=1}^n \frac{|I_{1j}|}{d(v_1)} \log \left( \frac{|I_{1j}|}{d(v_1)} \right) = - \left( \frac{1}{2} \log \left( \frac{1}{2} \right) + \frac{1}{2} \log \left( \frac{1}{2} \right) \right) = 1$$

De manera similar podemos definir la entropía para los hiperenlaces.

$$h(e_j) = - \sum_{i=1}^m \frac{|I_{ij}|}{d(e_j)} \log \left( \frac{|I_{ij}|}{d(e_j)} \right), \quad (4.6)$$

siendo  $m$  la cantidad de filas de esa columna que son distintas de cero.

Calculamos  $h(e_1)$  como

$$h(e_1) = - \sum_{i=1}^m \frac{|I_{i1}|}{d(e_1)} \log \left( \frac{|I_{i1}|}{d(e_1)} \right) = - \left( \frac{1}{3} \log \left( \frac{1}{3} \right) + \frac{1}{3} \log \left( \frac{1}{3} \right) + \frac{1}{3} \log \left( \frac{1}{3} \right) \right) = 1,585$$

Se creó una función en R para poder calcular estos valores para cualquier matriz de incidencia.

## 4.3. Resultados

### 4.3.1. Hipergrafo del caso de estudio

En esta sección construimos un hipergrafo usando las interacciones entre las 11 especies de plantas con los 11 gremios de herbívoros y las 25 especies distintas de polinizadores. La red original del Capítulo 2 contiene 2 tipos de interacción y 111 enlaces.

En la sección 4.2.1 mencionamos dos formas diferentes de construir un hipergrafo. Nosotros realizamos lo propuesto por Young [160] usando nuestro conjunto de datos pero al construir el grafo e identificar las cliques nos encontramos que una gran proporción de nodos estaban conectados entre sí. Por lo tanto, las cliques identificadas eran cliques que contenían la mayor cantidad de nodos y llegábamos a hipergrafos con dos hiperenlaces. Concluimos que el método propuesto por Young [160] no es recomendable en redes bipartitas.

Usando el método propuesto por Estrada [161] construimos un hipergrafo con 11 hiperenlaces (Figura 4.3), donde cada hiperenlace está formado por una especie de planta, por los herbívoros y polinizadores que interactúan con esa planta.

Detalles de los hiperenlaces:

- $e_1 = \{\text{Ants, Bee1, Bee5, Fly1, Fly15, Fly5, Herb8, } B.darwinii\}$
- $e_2 = \{\text{Ants, Bee1, Bee2, Bee3, Bee4, Bee5, Bee6, Bee7, Beetle1, Beetle2, Beetle3, Fly4, Fly7, Fly8, Unknown, Wasp1, Herb2, Herb8, Herb7, } E.coccineum \}$
- $e_3 = \{\text{Ants, Bee1, Bee2, Bee3, Bee4, Bee5, Bee6, Fly1, Fly7, Fly9, Unknown, Herb2, Herb10, Herb4, Herb9, Herb6, } L.hirsuta\}$
- $e_4 = \{\text{Fly1, Fly7, Wasp1, Herb2, Herb4, Herb7, Herb9, } M.boaria\}$
- $e_5 = \{\text{Ants, Beetle1, Butterfly2, Fly1, Fly9, Wasp1, Herb2, Herb7, } M.chubutensis\}$
- $e_6 = \{\text{Fly1, Herb2, Herb8, Herb4, Herb11, Herb3, Herb7, Herb9, Herb6, } R.cucullatum\}$
- $e_7 = \{\text{Ants, Bee1, Bee14, Bee5, Bee6, Fly1, Fly11, Fly13, Fly15, Fly4, Fly5, Fly7, Fly8, Fly9, Wasp5, Herb8, Herb11, Herb7, } R.magellanica\}$
- $e_8 = \{\text{Ants, Bee1, Bee2, Bee5, Beetle1, Beetle3, Fly1, Fly5, Fly7, Fly8, Fly9, Herb2, Herb7, Herb9, Herb5, Herb6, } S.patagonicus\}$
- $e_9 = \{\text{Herb11, } A.microphylla\}$
- $e_{10} = \{\text{Herb7, Herb5, } G.mucronata\}$
- $e_{11} = \{\text{Herb2, Herb8, Herb10, Herb4, Herb7, Herb1, Herb5, Herb6, } N.antarctica\}$

Detalle de las comunidades obtenidas en el Capítulo 2. La comunidad 1 está formada por los nodos de color turquesa:

$$C_1 = \{\text{Bee1, Bee5, Fly5, Fly7, Beetle3, } B.darwinii, S.patagonicus\}$$

La comunidad 2 por los nodos de color violeta:

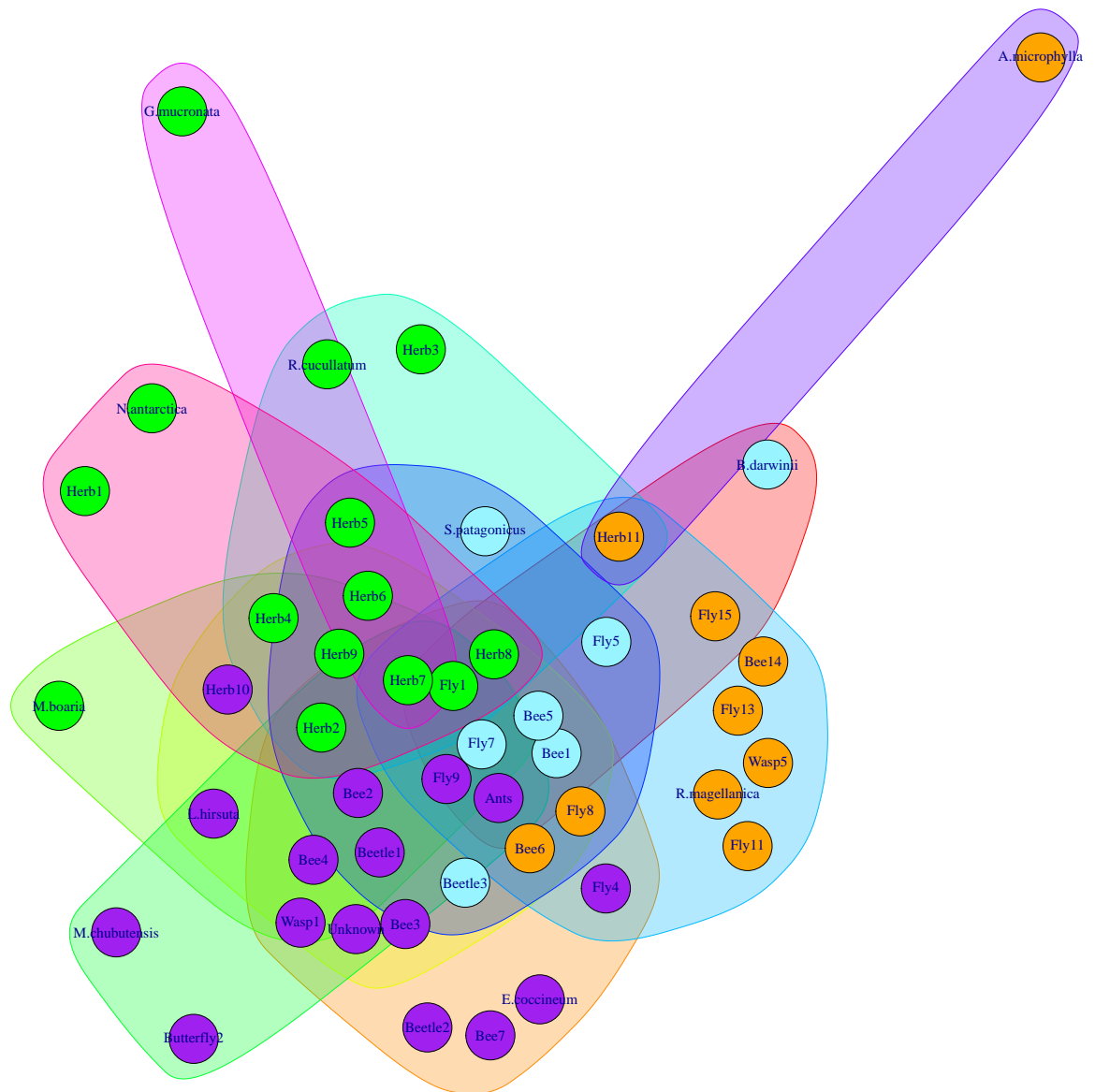
$$C_2 = \{\text{Ants, Bee2, Bee3, Bee4, Bee7, Beetle1, Beetle2, Butterfly2, Fly4, Fly9, Unknown, Wasp1, Herb10, } E.coccineum, L.hirsuta, M.chubutensis\}$$

La comunidad 3 por los nodos de color verde:

$$C_3 = \{\text{Fly1, Herb1, Herb2, Herb3, Herb4, Herb5, Herb6, Herb7, Herb8, Herb9, } R.cucullatum, G.mucronata, N.antarctica, M.boaria\}$$

La comunidad 4 por los nodos de color naranja:

$$C_4 = \{\text{Bee14, Bee6, Fly11, Fly13, Fly15, Fly8, Wasp5, Herb11, } R.magellanica, A.microphylla\}$$

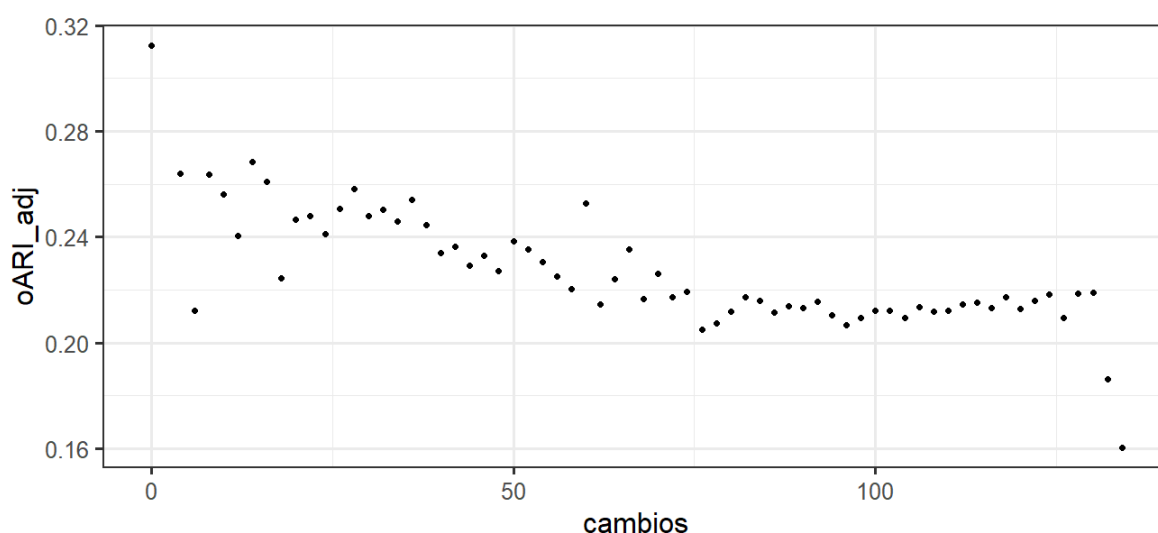


**Figura 4.3:** Hipergrafo con 11 hiperenlaces, donde sus nodos están coloreados dependiendo de la comunidad a la que pertenecen.

En la Figura 4.3 podemos observar que hay hiperenlaces que están compuestos por nodos de diferentes comunidades, es decir, esos hiperenlaces atraviesan todas las comunidades, sin embargo, tenemos 2 hiperenlaces que están completamente contenidos en una comunidad. El hiperenlace  $e_{10}$  está completamente contenido en la comunidad 3 y el hiperenlace  $e_9$  está completamente contenido en la comunidad  $C_4$ .

### 4.3.2. Comparando comunidades versus hiperenlaces

Se aleatorizó la red original usada en el Capítulo 2 preservando el grado de cada nodo y con un grado de aleatoriedad incorporado a la red original intercambiando los nodos adyacentes de dos enlaces elegidos al azar. Se calcularon las comunidades de cada red usando el algoritmo de Louvain. De cada red que se generaba a medida que se realizaban los cambios se generaba un hipergrafo y se calculó el índice  $oARI_{adj}$  entre esa comunidad obtenida por el algoritmo de Louvain y los hiperenlaces de cada hipergrafo. Además, para cada número de cambios, generamos diez redes diferentes y calculamos el promedio  $oARI_{adj}$ . En la Figura 4.4 se observa que el índice  $oARI_{adj}$  disminuye a medida que aumenta el número de cambios. Este cambio monótono es un indicador de que el valor obtenido es significativo ya que la red original tenía cierta información codificada que se destruyó al incorporar la aleatoriedad. El valor obtenido de  $oARI_{adj}$  entre las comunidades de la red original y los hiperenlaces de la red original es 0.3125.



**Figura 4.4:** Valores de  $oARI$  ajustado vs. el número de cambios (*swappings*) en la red original. Cada valor es un promedio de 10 redes excepto por la red original.

### 4.3.3. Determinando los nodos influyentes usando la entropía

Se llevó a cabo un cálculo exhaustivo del valor de la entropía para cada hiperenlace, como se detalla en la Tabla 4.2. En este contexto, cada hiperenlace representa

una especie de planta específica. De manera similar, es posible calcular los valores de entropía para la lista de polinizadores y los gremios de herbívoros asociados.

El método que hemos propuesto establece que la especie de planta más influyente es aquella que presenta un valor de entropía superior al de las demás. En nuestro análisis, hemos identificado que la especie con el mayor valor de entropía es *E.coccineum*, seguida por *R.magellanicum*, que ocupa el segundo lugar en términos de influencia. Este hallazgo refuerza las conclusiones a las que llegamos en el Capítulo 2.

Además, los hiperenlaces formados por estas dos especies de plantas son cruciales, ya que albergan una amplia variedad de polinizadores y gremios de herbívoros. Esta diversidad es fundamental para la estabilidad del ecosistema en el que se encuentran. El hiperenlace correspondiente a *E.coccineum* destaca por contener una mayor cantidad de nodos pertenecientes a la comunidad 2, aunque también incluye nodos de otras comunidades. Por otro lado, el hiperenlace asociado a *R.magellanicum* presenta una mayor proporción de nodos pertenecientes a la comunidad 4, con el resto de los nodos distribuidos entre las demás comunidades.

**Tabla 4.2:** Especies de plantas con su valor de entropía y con su grado

Especie de planta	H(planta)	Grado
<i>E.coccineum</i>	2.944	19
<i>R.magellanicum</i>	2.890	18
<i>L.hirsuta</i>	2.772	16
<i>S.patagonicus</i>	2.772	16
<i>R.cucullatum</i>	2.197	9
<i>M.chubutensis</i>	2.079	8
<i>N.antarctica</i>	2.079	8
<i>B.darwinii</i>	1.946	7
<i>M.boaria</i>	1.945	7
<i>G.mucronata</i>	0.6931	2
<i>A.microphylla</i>	0	1

## 4.4. Discusión

En este capítulo se presentó un método para construir hipergrafos que puede ser aplicado a redes bipartitas o a redes con más de dos tipos de interacción. Este enfoque permite modelar interacciones complejas donde más de dos entidades están involucradas simultáneamente.

Evaluamos cuantitativamente la similitud entre las comunidades obtenidas por el algoritmo de Louvain y los hiperenlaces del hipergrafo en cada paso que se aleatoriza la red. Al aleatorizar la red se borra cualquier información valiosa almacenada pero

al mismo tiempo se conservan intactas algunas características como su distribución de grado. Los resultados mostraron que las comunidades y los hiperenlaces no tienen una relación unívoca, ni tampoco es posible reconstruir un hiperenlace realizando la unión entre comunidades. Esto demuestra que ambos enfoques, análisis de redes y análisis de hipergrafos, son complementarios y proporcionan información diferente.

Las similitudes entre las comunidades obtenidas y los hiperenlaces fueron calculadas usando el índice  $oARI_{adj}$ . A pesar de aleatorizar la red se preservó la distribución de grado. Cabe destacar que el índice  $oARI_{adj}$  disminuyó monótonamente con respecto al primer valor que fue calculado usando la red y el hipergrafo original. Por último, se calculó el grado de los hiperenlaces y la entropía para cada especie de planta y se identificaron las especies más influyentes o también denominadas especies clave. Las especies identificadas fueron: *R. magellanicum* y *E. coccineum*. Este resultado coincide con lo obtenido usando la metodología del Capítulo 2, lo que resalta la robustez del enfoque propuesto.

Nuestros análisis revelaron que la construcción de un hipergrafo a partir de una red bipartita depende de la compatibilidad del método con los datos, ya que no todos los enfoques son aplicables a cualquier conjunto de datos.

Recientemente, los hipergrafos comenzaron a ser utilizados en estudios de ecología, principalmente para modelar sistemas ecológicos complejos que involucran interacciones de más de dos especies. A diferencia de los grafos tradicionales, que sólo permiten relaciones entre pares de nodos, los hipergrafos pueden capturar relaciones de grupo, lo que es esencial para comprender otro tipo de dinámicas.

Los hallazgos de este capítulo se diferencian de estudios previos que utilizan a los hipergrafos para modelar interacciones entre especies de plantas y polinizadores donde más de dos especies están involucradas en la misma interacción [166] y de los estudios donde se los usó para interacciones indirectas [64, 65]. Sin embargo, no hay estudios sobre la comparación entre comunidades y los hiperenlaces obtenidos por el hipergrafo dentro del área de ecología, ya que hasta el momento en las técnicas de aprendizaje no supervisado se comparan conjuntos disjuntos.

#### 4.4.1. Implicancias de la comparación entre comunidades versus hiperenlaces

Al comparar las comunidades de la red original con los hiperenlaces del hipergrafo original, se observó que algunos hiperenlaces atraviesan múltiples comunidades, mientras que otros están contenidos dentro de una sola comunidad. Esto sugiere que las comunidades y los hiperenlaces capturan diferentes aspectos de la estructura de la red. Estos son los hiperenlaces formados por los polinizadores y gremios de herbívoros que tienen interacción con las siguientes especies de plantas *B. darwinii*, *E. coccineum*,

*L.hirsuta*, *S.patagonicus* y *N.antarctica*. El hiperenlace 4 que está formado por los polinizadores y gremios de herbívoros que interactúan con *M.boaria* sólo tiene nodos de 3 comunidades. Además tenemos 3 hiperenlaces que contienen nodos de 2 comunidades. Finalmente tenemos 2 hiperenlaces que están completamente contenidos dentro de una comunidad, uno de ellos es el hiperenlace 9 formado por  $\{Herb11, A.microphylla\}$ , el cual está completamente contenido en la comunidad 4 y el hiperenlace 10 formado por  $\{Herb7, Herb5, G.mucronata\}$  está completamente contenido en la comunidad 3.

Por todo lo anterior, se esperaba que al calcular de forma cuantitativa la similitud entre las comunidades obtenidas usando el algoritmo de Louvain con los hiperenlaces ese valor no diera 1. Si comparamos 2 grupos y ese valor da 1 es porque son similares. El valor obtenido usando el  $oARI_{adj}$  puede parecer suficiente pero es necesario compararlo con otros valores obtenidos de la misma manera. Por ese motivo se aleatorizó la red original y a medida que se aleatorizaba la misma se calculaba el hipergrafo obtenido y se comparó la similitud entre los hiperenlaces con las comunidades obtenidas para cada número de cambios. Dado que un valor solo no es suficiente se consideraron 10 redes para cada número de cambios realizados en la red original. Se determinó que los hiperenlaces con las comunidades no son similares, ni tampoco es posible reconstruir un hiperenlace realizando la unión entre comunidades. Con lo cual podemos concluir que al analizar los datos usando el análisis de redes obtenemos un tipo de información diferente al que se obtiene con los hipergrafos, demostrando que ambos enfoques son complementarios.

#### 4.4.2. Identificación de nodos influyentes

Identificar los nodos influyentes de un hipergrafo es fundamental para comprender la estructura y dinámica de los hiperenlaces. Al realizar un análisis de redes tradicional muchas veces los nodos influyentes suelen estar asociados a los valores del grado de cada nodo, pero ya vimos en el Capítulo 2 que no necesariamente los nodos de mayor grado serán los más influyentes. Sin embargo, en el caso de los hipergrafos las relaciones son más complejas y los métodos tradicionales pueden no capturar completamente el rol de los nodos en estructuras de orden superior.

Además, los nodos influyentes en hipergrafos tienen importantes implicancias prácticas, como por ejemplo, en el área de ecología y otras disciplinas donde las interacciones no se limitan a pares de elementos. Al considerar hiperenlaces, se obtiene una visión más amplia de las relaciones que influyen en la dinámica del sistema y es por eso que nos interesaba determinar qué hiperenlace es más relevante o influye más en el hipergrafo. Nuestro análisis demostró que las dos especies de plantas más influyentes en el hipergrafo son: *R.magellanicum* y *E.coccineum*. Usando dos métodos distintos, obtuvimos que los nodos influyentes o especies clave son las mismas en nuestro sitio de

estudio.

Aunque en este análisis consideramos únicamente un hipergrafo que fue construido a partir de una red bipartita, el método usado para generar el hipergrafo puede aplicarse a otros tipos de redes de diferentes disciplinas que involucren interacciones complejas. Además el enfoque del índice  $oARI_{adj}$  es útil para identificar si dos conjuntos de nodos son similares entre sí, aunque uno de esos conjuntos no necesariamente sea disjunto. Dentro del área de biología los hipergrafos nos permiten modelar interacciones entre proteínas, mientras que en el área de ciencias sociales este enfoque nos permitiría analizar diferentes dinámicas dentro de grupos que comparten ciertas características entre los individuos. Este enfoque también puede ser útil para la toma de decisiones o para personalizar sugerencias con respecto a algún producto. Finalmente, nuestros algoritmos e índices calculados pueden aplicarse a diversas disciplinas.



# Capítulo 5

## Discusión

El eje central de este trabajo es el análisis de la estructura de diferentes redes mediante índices de complejidad. Este objetivo está fuertemente inspirado en las necesidades de un grupo de investigación en el área de la ecología que buscaba caracterizar y entender cómo analizar la información codificada en la estructura de las redes de interacción construídas en base a datos de campo. La motivación detrás de este enfoque surge de la necesidad de optimizar las estrategias de recolección de datos en campo sobre la estructura ecosistémica, especialmente en el contexto de las interacciones entre insectos y plantas en ecosistemas como la Patagonia Norte. El objetivo final es implementar prácticas de manejo forestal sostenibles, un desafío persistente en muchos ecosistemas. Sin embargo, en los últimos años, ha habido un aumento en los ejemplos de enfoques que simultáneamente mejoran la productividad y conservan la biodiversidad [167–169].

Mediante los estudios asociados a esta tesis, se ha logrado un avance significativo en el análisis de la complejidad de las redes. Se han propuesto nuevas herramientas para abordar este desafío, destacándose por un enfoque innovador en la generación de redes aleatorias. A diferencia de las metodologías tradicionales, en este estudio, las redes aleatorizadas se construyen a partir de la red original mediante un proceso de aleatorización controlada. Este proceso permite que la red se aleje gradualmente de su estructura original, lo que facilita una evaluación más precisa de cómo las alteraciones estructurales afectan las propiedades de la red.

En contraste con estudios previos, como el de Dunne [170] donde se compara la densidad de las redes originales con las aleatorias sin un proceso controlado de aleatorización, nuestro enfoque permite evaluar cambios específicos en la estructura de la red a lo largo de la aleatorización.

Por otro lado, en el estudio de Stouffer y Bascompte [171], las redes aleatorias se utilizan para contrastar con redes tróficas reales, pero se centran en la modularidad sin generar las redes aleatorias a partir de modificaciones controladas de la red original. Nuestro enfoque permite un seguimiento más detallado y controlado de cómo las mo-

dificaciones estructurales progresivas afectan las propiedades de la red, algo que no se ha abordado en los estudios mencionados.

Comprender la estructura de las interacciones entre insectos y plantas es fundamental tanto a nivel ecológico como económico, ya que proporciona herramientas para gestionar de manera eficiente y sustentable los recursos de esta región. A través del análisis de la topología de las redes de interacción entre plantas y polinizadores, es posible prever la estabilidad y funcionalidad de dichas redes. Aunque la presencia de especies generalistas puede favorecer estas propiedades, las especies raras o especializadas podrían verse involucradas y asociadas a efectos no triviales, al mismo tiempo que estarían más comprometidas por la competencia con especies generalistas más dominantes. Es por eso que entender cómo los disturbios afectan las respuestas de la red de interacción planta-polinizador es clave para diseñar estrategias de manejo que mantengan la estabilidad y funcionalidad de estas redes.

Pese a que se conoce poco sobre el impacto del cambio ambiental y climático en el manejo de bosques, especialmente en las especies forestales nativas y su interacción con los herbívoros artrópodos, es esencial investigar cómo responden estas especies y sus herbívoros a diferentes métodos de manejo forestal bajo condiciones microclimáticas variables, una variable puede ser la temperatura y en este caso se puede crear un invernadero cambiando esas variables climáticas [172]. Dado que no existían datos recolectados en la Patagonia Norte, el trabajo de campo que originó los datos sobre los que trabajamos midió simultáneamente las interacciones planta-herbívoro y el rendimiento de especies forestales. Esta información y su análisis podrían ser cruciales para la gestión de estos bosques en un futuro con temperaturas más elevadas.

## 5.1. Contribuciones

### 5.1.1. Capítulo 2

La metodología propuesta en el Capítulo 2 constituye una herramienta valiosa para el análisis de redes ecológicas, ya que permite discernir si una red construida en base a datos recolectados contiene información relevante revelada, entre otras cosas, a través de su estructura de comunidades. Si bien durante muchos años se han realizado estudios de redes ecológicas bipartitas, en este caso se propone un método que tiene en cuenta más de un tipo de interacción, lo cual constituye un enfoque más completo que otros desarrollados anteriormente con interacciones herbívoro-planta o polinizador-planta [173–175]. Al realizar un análisis en conjunto, se tiene en cuenta el efecto que cada nivel tiene sobre el otro. En el Capítulo 2 se evidencia que no es suficiente calcular la modularidad de la red original, sino que es necesario comparar ese valor con los obtenidos en sus versiones aleatorizadas, proponiendo un modelo nulo. De esta forma,

se agrega desorden en la red original, y de esa manera se puede ponderar cómo es que se pierde la información codificada en el sistema cuando se le agregan perturbaciones aleatorias, es decir, ruido.

Para analizar la relevancia local de lo hecho en este capítulo destacamos que la metodología usada permitió realizar un aporte en el área de ecología identificando las especies clave del ecosistema. Si bien hay una amplia variedad de herramientas para la identificación de especies clave, muchas de ellas no tienen en cuenta el sistema en su totalidad. Debido a las características de las especies estudiadas y a estudios anteriores [97], se esperaba que algunas plantas como *N. antarctica*, *M. boaria* y *R. cucullatum* tuvieran vínculos más fuertes con los gremios de herbívoros que con los polinizadores [97]. Por otro lado, *E. coccineum* y *R. magellanicum* tienen flores muy atractivas y, al mismo tiempo, sus hojas son muy palatables. Estas especies pueden tener una alta interacción con herbívoros y polinizadores. Por último, *E. coccineum* fue muy abundante entre las especies de plantas leñosas muestreadas.

Como resultado de los análisis realizados, detectamos que las especies seleccionadas como clave, *R. magellanicum* y *E. coccineum*, desempeñan un papel estabilizador en el ecosistema. Estas dos especies tienen funciones relevantes en lo que respecta a las interacciones ecológicas, pero también como bienes forestales. Las flores de éstas son altamente preferidas por polinizadores como insectos y aves, y sus hojas soportan una alta diversidad de gremios de herbívoros. *E. coccineum* posee un notable valor maderable, especialmente para la producción de leña, postes y mobiliario. Además, su notable capacidad de rebrote ofrece interesantes posibilidades para implementar estrategias de manejo de cosecha más sostenibles y diversificadas [109]. Además, tienen varias aplicaciones médicas y el fruto de *R. magellanicum* proporciona un valioso doble servicio como alimento para los animales del bosque y como producto potencial para el consumo humano. Por tanto, *E. coccineum* y *R. magellanicum* constituyen especies clave por su relación entre polinizadores, herbívoros y plantas leñosas de uso para actividades humanas, destacando que el seguimiento de estas especies y conocer su papel en los sistemas ecológicos, constituye una prioridad para la conservación y gestión de estos sitios de estudio.

Además, la estructura de comunidades hallada para las redes muestra resultados interesantes en cuanto a su composición. En la red del Capítulo 2 se encontraron diferencias interesantes entre las cuatro comunidades identificadas, que variaban en términos de polinizadores, especies de plantas y composición del gremio de herbívoros. En particular, aquellas comunidades cuyas plantas eran muy visitadas por los polinizadores eran menos preferidas por los herbívoros. Las especies vegetales con flores típicamente atractivas para los insectos polinizadores pertenecen a tres de las comunidades mencionadas [23]. Cabe destacar que estas plantas también tienen hojas menos palatables, excepto *R. magellanicum* [114]. Por otro lado, la comunidad que fue consumida en ma-

yor proporción por los diferentes gremios de herbívoros estuvo compuesta por plantas que fueron polinizadas con menor frecuencia por insectos. *G. mucronata*, *M. boaria*, *N. antarctica* y *R. cucullatum* tienen hojas de alta calidad [115].

Lo expuesto sugiere que el enfoque realizado en esta tesis podría convertirse en una estrategia útil para determinar si existe un equilibrio en los estratos de la red, proporcionando información para análisis posteriores. En general, los resultados obtenidos tienen sentido en términos biológicos y nos invitan a replicar el enfoque en otros sistemas de estudio de diferentes regiones o ampliando éste.

### 5.1.2. Capítulo 3

La metodología propuesta en el Capítulo 3 además de introducir una herramienta innovadora y útil ha contribuido significativamente al entendimiento de la complejidad de diferentes redes, seleccionando un conjunto de herramientas para caracterizarlas. Aunque esta metodología fue aplicada en 42 redes ecológicas entre ellas 23 redes de planta-polinizador y 19 redes hospedador-parásito de la base de datos web-of-life.es, tiene el potencial de extenderse a redes de otras áreas.

Para este estudio fue necesario caracterizar las redes, para lo cual se calcularon diferentes índices como: la cantidad de nodos, el número de enlaces, el anidamiento y la densidad de cada red. Además se estudió el comportamiento de 4 índices para abordar el análisis de la complejidad de las redes. Para esto se aleatorizaron las redes originales usando 2 modelos nulos, el primero de ellos conserva la distribución de grado y el segundo conserva la cantidad de enlaces. Se deja en evidencia que no es suficiente considerar un solo valor de la red original para hacer un estudio sobre la complejidad de la misma, sino que es necesario usar modelos nulos para contrastar esos valores originales.

El análisis realizado en este capítulo realizó un aporte con respecto a la utilización de ciertos índices dependiendo de la densidad de la red. Se tuvo presente que los dos conjuntos de redes variaban en cuanto a su densidad. En las redes de planta-polinizador, los valores oscilaron entre 0.017 y 0.204, mientras que en las redes de hospedador-parásito entre 0.125 y 0.28. Es importante tener en cuenta la densidad de la red para saber qué índice se ajusta mejor a esos datos.

Nos interesaba comparar el comportamiento de 4 índices: modularidad, entropía, SVD entropía y K-complejidad. Usando los dos modelos nulos se aleatorizó la red y además los 4 índices fueron calculados a medida que se realizaban los cambios en la red. Como nos interesaba identificar cómo era el comportamiento de esos dos tipos de redes dependiendo de la densidad se propuso ajustar los resultados de los índices a medida que se realizaban los cambios para cada red. Finalmente el método propone comparar las pendientes de los ajustes dependiendo de la densidad de la red.

Nuestro estudio reveló que la densidad juega un rol clave al momento de realizar un análisis de la red. En el caso de las redes con una densidad menor a 0.15 se observó que la complejidad de las mismas puede ser analizada usando el primer modelo, donde queda en evidencia que el comportamiento de la modularidad y la entropía ya nos dan información sobre la complejidad de las mismas. Sin embargo, nos interesaba ir más allá de estos índices es por eso que se calcularon SVD-entropía y K-complejidad y sus pendientes a medida que se realizaban los cambios. K-complejidad deja en evidencia que el comportamiento de las redes planta-polinizador se corresponde con los resultados obtenidos por la entropía y modularidad. Sin embargo SVD-entropía aún no marcaba ninguna diferencia entre redes de hospedador-parasito y planta-polinizador, se llegó a esta conclusión al comparar  $\beta_{SVD}$  de ambos conjuntos de redes. Para ver la influencia de este índice y de si efectivamente la densidad jugaba un rol clave se propone otro modelo.

El segundo modelo propuesto confirmó que la densidad influye en la toma de decisiones y que en conjunto con K-complejidad puede ser un instrumento para determinar si una red pertenece a cierto conjunto de redes y se puede predecir cómo será el comportamiento de la misma dependiendo del grupo al que pertenezca. Además se dejó en evidencia que SVD entropía no mostraba diferencias cuando comparábamos las redes, con lo cual queda demostrado que no es recomendable usar SVD entropía para el análisis de la complejidad de una red. Usando el segundo modelo se ve reflejado que la modularidad es sensible a la distribución de grado.

Se expuso que la entropía y la modularidad son herramientas útiles para determinar si una red contiene información codificada en su sistema. La modularidad complementa los resultados obtenidos por la entropía al mostrar que se está perdiendo la estructura de las comunidades. Cuando esto no ocurre, se pueden utilizar otras herramientas para determinar si la red contiene información relevante. Además, se ha determinado la importancia de otros índices como K-complejidad y se ha demostrado su robustez.

### 5.1.3. Capítulo 4

La metodología propuesta en el Capítulo 4 ofrece un aporte significativo para la comparación de conjuntos no necesariamente disjuntos, ya que hasta el momento en las técnicas de aprendizaje no supervisado se comparan conjuntos disjuntos o *clusters*. Además se presentaron dos formas distintas de generar un hipergrafo y se propone otra forma de identificar especies clave en un ecosistema. Los hipergrafos constituyen una herramienta valiosa para el análisis de interacciones ecológicas, ya que permiten modelar interacciones entre más de dos entidades. A pesar de su relevancia y sus beneficios aún no son tan utilizadas en el área de ecología.

Al estudiar la similitud entre dos conjuntos de nodos no disjuntos se uso el índice

$oARI_{adj}$  propuesto por Campello [164]. Se usó el modelo nulo propuesto a lo largo de toda la tesis para estudiar el comportamiento de ese índice a medida que se aleatorizaba la red. En cada cambio realizado se construyó un hipergrafo y se calculó la similitud entre las comunidades de esa red y los hiperenlaces del hipergrafo construido para ese número de cambio. Vale la pena resaltar que se generaron 10 redes para cada número de cambio y se construyeron 10 hipergrafos de esas 10 redes, de esta forma obtuvimos 10 valores para el índice  $oARI_{adj}$  y para cada número de cambio se realizó el promedio entre esos 10 valores.

Otro hallazgo relevante fue la construcción del hipergrafo ya que nuestra red original está compuesta por 2 tipos de interacción y donde es necesario explorar qué método se ajusta mejor a la construcción del hipergrafo. Vale la pena recordar que las interacciones provienen de redes bipartitas y no todos los enfoques nos llevarán a la construcción de un hipergrafo óptimo. Al construir el hipergrafo usando el método de Young et al [160] y al identificar los cliques del grafo proyectado se obtuvo que el hipergrafo contenía 21 nodos en un solo hiperenlace, esto se debe a que la mayor proporción de nodos estaban conectados entre sí. Por ese motivo para la construcción del hipergrafo se implementó la metodología propuesta por Estrada [161]. Los resultados para la construcción del hipergrafo nos llevaron a concluir que no todos los enfoques son adecuados para todas las redes.

Una vez que construimos el hipergrafo identificamos los nodos influyentes. Los nodos influyentes dentro de un ecosistema pueden ser denominados especies clave y jugar un rol clave para la estructura del ecosistema. En el Capítulo 4 otro de los desafíos fue abordar este tema desde una perspectiva de los hipergrafos, es por eso que nos interesaba estudiar un índice que sea capaz de identificar este tipo de nodos. La entropía fue calculada para cada hiperenlace, recordemos que por como construimos el hipergrafo, cada hiperenlace representa una especie de planta. También es posible calcular los nodos influyentes entre los conjuntos de nodos de los polinizadores y el gremio de herbívoros. Las dos especies de plantas más influyentes fueron: *R. magellanica* y *E. coccineum*. Estos hallazgos se corresponden con los resultados obtenidos usando la metodología del Capítulo 2, lo cual hace más confiable nuestro método usado en los hipergrafos.

El cálculo del índice  $oARI_{adj}$  fue implementado mediante una función en R, ya que, al tratarse de un índice novedoso, no existía una implementación previa para su uso en diferentes redes. Esta función, que ahora permite aplicarlo a una amplia variedad de redes, tiene el potencial de convertirse en un paquete de R en el futuro. Además, se desarrollaron funciones adicionales para la aleatorización de la red original, la creación de hipergrafos en paralelo y el cálculo de la entropía, todas programadas en R.

## 5.2. Alcances de este trabajo

A pesar de que en los análisis del Capítulo 3 nos centramos únicamente en redes bipartitas, es importante destacar que nuestro método puede extenderse a otros tipos de redes, incluso mediante proyecciones. Este enfoque metodológico resulta útil para identificar cuándo una red carece de información codificada en su sistema. Las razones detrás de esta falta de información pueden ser diversas, como la ausencia de enlaces no detectados o deficiencias en la forma en que se han recopilado los datos. Además, al investigar redes sin información codificada, se podrían predecir enlaces futuros entre nodos y determinar qué conexiones no son esenciales.

La predicción de enlaces es un campo amplio y representa un desafío científico de larga data con numerosas aplicaciones prácticas. Su utilidad abarca desde la identificación de enlaces faltantes hasta la anticipación de conexiones que podrían surgir con la evolución de las redes [93, 176, 177]. En el ámbito de las redes biológicas, el descubrimiento de interacciones suele ser un proceso costoso; por lo tanto, una predicción precisa resulta más eficiente que la búsqueda aleatoria de posibles interacciones latentes.

Asimismo, la detección de enlaces en redes sociales puede optimizar el rendimiento de algoritmos de clasificación basados en enlaces [178]. En este contexto, es posible recomendar enlaces futuros, es decir, conexiones entre pares de nodos no conectados a usuarios relevantes, lo que facilitaría la formación de nuevas relaciones [179]. Desde una perspectiva teórica, la predicción de enlaces también puede ofrecer metodologías valiosas para el modelado de redes [145]. Es importante destacar que en este trabajo se exploró una familia de algoritmos para la detección de comunidades, cuyos resultados han sido consistentes. No obstante, podrían obtenerse otros resultados utilizando diferentes algoritmos en los que esto no suceda.

La metodología y las herramientas empleadas en el Capítulo 4 pueden ser implementadas en diferentes áreas. De hecho puede resultar interesante aplicar el análisis de hipergrafos a redes que fueron analizadas con los métodos clásicos de teoría de grafos, lo cual nos permitiría explorar nuevas estructuras y relaciones que no son capturadas por los grafos convencionales. Este enfoque ofrece otra perspectiva, especialmente donde las interacciones involucran más de dos entidades a la vez, como en estudios ecológicos, biológicos y sociales.

## 5.3. Limitaciones y trabajos futuros

A pesar de los avances significativos en la detección de especies clave y el análisis de la complejidad de las redes, persisten preguntas intrigantes que podrían ser abordadas en investigaciones futuras. Algunas de estas preguntas surgen de los resultados

obtenidos y otras representan ampliaciones de lo realizado.

En primer lugar, sería interesante investigar qué ocurriría si se estudiaran redes en sitios que han sufrido modificaciones en el suelo a lo largo de los años, donde se han removido especies de plantas. Sería relevante determinar si las especies clave identificadas en nuestro estudio se mantienen tras tales alteraciones en el hábitat. También nos gustaría conocer si los polinizadores son los mismos en cada sitio modificado y si la presencia de herbívoros se mantiene constante en el área de estudio. Para este análisis, se podría implementar un enfoque diferente para la recolección de datos, recomendando una mejora en la clasificación de los herbívoros. En esta tesis, se utilizaron grupos de gremios de herbívoros, pero una clasificación taxonómica podría ofrecer una visión más detallada.

Si bien se han realizado algunos estudios sobre cómo cambian las interacciones ecológicas a lo largo del tiempo, hay una falta de investigaciones que analicen estas dinámicas utilizando enfoques estadísticos avanzados, como modelos de series temporales o métodos de machine learning. Estos enfoques podrían predecir cambios en la estructura de las redes debido a factores como el cambio climático o la pérdida de hábitat, proporcionando información valiosa sobre cómo las redes ecológicas evolucionan y se adaptan con el tiempo. Además, sería interesante plantar las especies seleccionadas como clave debido a su relación con polinizadores y herbívoros en diferentes localidades de la Patagonia Norte. Esto permitiría estudiar su impacto a nivel social y ecológico, dado que estas plantas ofrecen diversos beneficios. Por ejemplo, entre las especies forestales leñosas, *R. magellanicum* y *E. coccineum* desempeñan funciones relevantes en las interacciones ecológicas y como bienes forestales. Las flores de ambas especies son altamente atractivas para polinizadores como insectos y aves, y sus hojas sustentan una gran variedad de herbívoros. Además, *E. coccineum* es de gran importancia socioeconómica, ya que se utiliza para leña, postes y fabricación de muebles. Desde el punto de vista ecológico, tiene una notable capacidad de rebrote, lo que la convierte en una opción viable para la restauración de bosques, así como en una alternativa para la gestión de la cosecha y diversas aplicaciones en medicina. Por su parte, el fruto de *R. magellanicum* es un alimento tanto para la fauna del bosque como para los seres humanos, siendo utilizado en la elaboración de mermeladas, conservas y licores. Desde un enfoque más estadístico, el trabajo publicado en 2021 [130] resalta una contribución valiosa al estudio de las perturbaciones en redes ecológicas.

Existen áreas adicionales de investigación que podrían enriquecer nuestra comprensión sobre cómo estas redes responden a entornos cada vez más complejos y cambiantes. Por ejemplo, se podrían analizar perturbaciones de origen natural, explorando cómo las perturbaciones humanas, como la deforestación o el cambio climático, interactúan con perturbaciones naturales, como incendios o sequías, para afectar la estructura y funcionalidad de las redes ecológicas. Este enfoque podría revelar nuevas dinámicas

sobre cómo las redes responden a la creciente influencia humana. En este caso, el uso de simulaciones estadísticas y modelos de redes complejas podría ofrecer una visión más profunda de la robustez de estas redes en escenarios donde interactúan múltiples factores. Además, explorar estos aspectos mediante enfoques estadísticos avanzados y técnicas de análisis de redes podría abrir nuevas oportunidades para comprender mejor la dinámica de las redes ecológicas, contribuyendo a su conservación y gestión sostenible. El estudio de la complejidad de las redes ha generado una creciente demanda de conocimientos estadísticos en esta área. El principal reto radica en encontrar un equilibrio adecuado entre enfoques de análisis de redes que sean excesivamente complejos e inaccesibles y aquellos que sean accesibles pero demasiado simplistas. Consideramos que el progreso hacia este objetivo requerirá una colaboración estrecha y un compromiso firme entre las disciplinas de estadística y ecología.

El trabajo presentado podría complementarse con estudios comparativos en diferentes ecosistemas para entender cómo varían las interacciones ecológicas bajo diferentes condiciones ambientales. También se podrían incorporar datos genómicos en el análisis de redes para identificar variaciones en las especies y sus interacciones.

Un objetivo más ambicioso es el de desarrollar modelos predictivos que integren variables ambientales y biológicas para anticipar cambios en las redes ecológicas ante escenarios de cambio climático. Desarrollar estos modelos podría abrir nuevas oportunidades para entender mejor la dinámica de las redes ecológicas, contribuyendo a la conservación y gestión sostenible de los ecosistemas.

Desde el enfoque de los hipergrafos se pueden seguir explorando métricas alternativas o enfoques bayesianos para construir hipergrafos desde redes bipartitas. Además se puede proponer un análisis temporal para estudiar como evolucionan las comunidades y los hiperenlaces en el tiempo.

El código implementado en R usado para esta tesis estará disponible en un repositorio de github ([código Tesis.Huaylla](#)) y el pseudocódigo de los algoritmos principales se encuentran en el apéndice C.



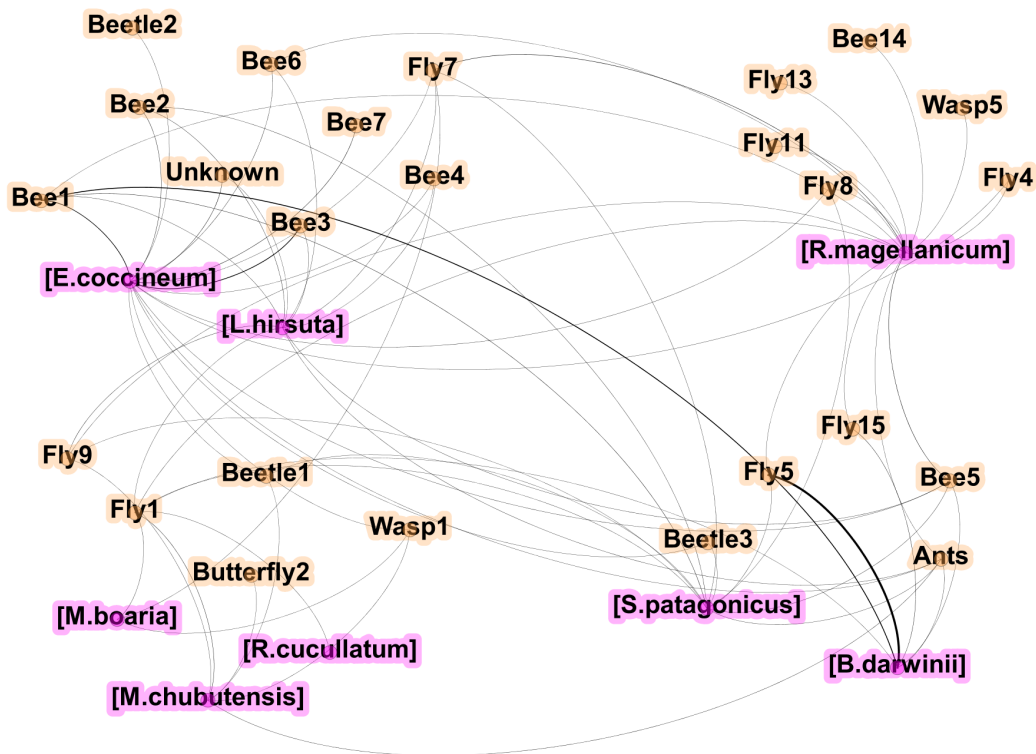
# Apéndice A

## Material complementario A

### Red de especies de planta-polinizador

**Tabla A.1:** Cambios en la estructura de la red después de eliminar una especie de planta. Las especies clave: *E.coccineum* y *R.magellanicum*

Especie de planta eliminada	Betweenness	Betweenness normalizada	Comunidades (antes)	Comunidades (después)	Componentes conectadas	Componentes conectadas (Después)
<i>R.magellanicum</i>	173.73	1	4	7	1	5
<i>E.coccineum</i>	159.28	0.92	4	7	1	3
<i>L.hirsuta</i>	58.85	0.34	4	4	1	1
<i>S.patagonicus</i>	53.66	0.31	4	4	1	1
<i>M.chubutensis</i>	42.77	0.25	4	5	1	1
<i>B.darwinii</i>	15.39	0.09	4	4	1	1
<i>M.boaria</i>	4.30	0.02	4	4	1	1
<i>R.cucullatum</i>	0	0	4	4	1	1

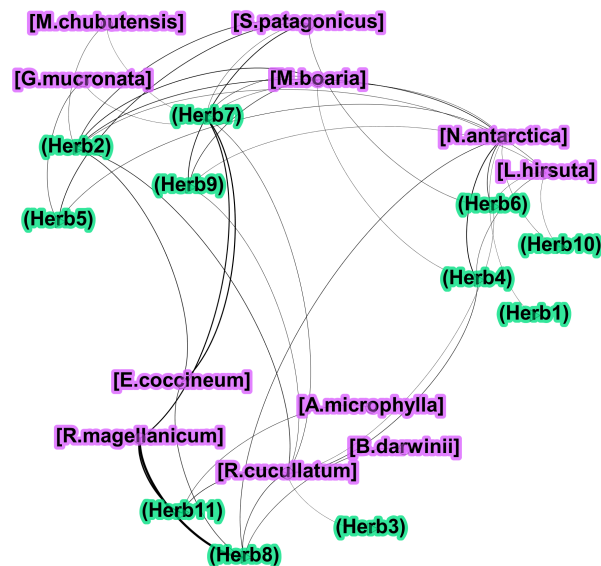


**Figura A.1:** Rede de interações entre planta-polinizador ordenada por comunidades. Quatro comunidades foram identificadas. Seus nós são apresentados como segue: As plantas em violeta e os polinizadores em color naranja.

## Red de herbívoro-planta

**Tabla A.2:** Cambios en la estructura de la red después de eliminar una especie de planta. Especies clave: *R.cucullatum* and *N.antarctica*

Especie de planta eliminada	Betweenness	Betweenness normalizada	Comunidades (antes)	Comunidades (después)	Componentes conectadas	Componentes conectadas (Después)
<i>R.cucullatum</i>	63.80	1	3	5	1	2
<i>N.antarctica</i>	54.17	0.85	3	4	1	2
<i>S.patagonicus</i>	11.30	0.18	3	4	1	1
<i>R.magellanicum</i>	9.91	0.16	3	4	1	1
<i>L.hirsuta</i>	9.10	0.14	3	4	1	1
<i>M.boaria</i>	4.81	0.08	3	4	1	1
<i>E.coccineum</i>	3.79	0.06	3	4	1	1



**Figura A.2:** Red de interacción entre herbívoros-plantas ordenadas por comunidades. Se identificaron tres comunidades. Sus nodos se representan de la siguiente manera: plantas en color violeta y gremios de herbívoros en verde.

## Proyección de la red planta-polinizador

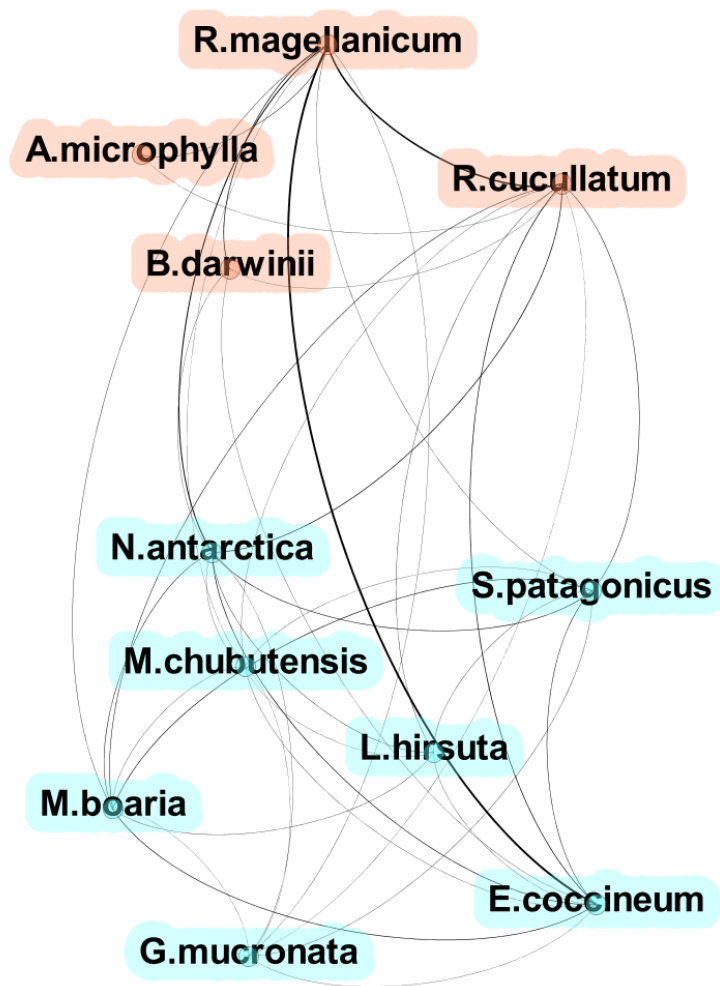
**Tabla A.3:** Cambios en la estructura de la red después de eliminar una especie de planta.

Especie de planta eliminada	Betweenness	Betweenness normalizada	Comunidades (antes)	Comunidades (después)	Componentes conectadas	Componentes conectadas (Después)
<i>B.darwinii</i>	0.16	1	1	1	1	1
<i>L.hirsuta</i>	0.16	1	1	1	1	1
<i>M.chubutensis</i>	0.16	1	1	1	1	1
<i>R.magellanicum</i>	0.16	1	1	1	1	1
<i>S.patagonicus</i>	0.16	1	1	1	1	1
<i>M.boaria</i>	0.16	1	1	1	1	1
<i>E.coccineum</i>	0	0	1	1	1	1
<i>R.cucullatum</i>	0	0	1	1	1	1

## Proyección de la red herbívoro-planta

**Tabla A.4:** Cambios en la estructura de la red después de eliminar una especie de planta.

Especie de planta eliminada	Betweenness	Betweenness normalizada	Comunidades (antes)	Comunidades (después)	Componentes conectadas	Componentes conectadas (Después)
<i>R.cucullatum</i>	6.16	1	2	2	1	1
<i>R.magellanicum</i>	4.5	0.73	2	2	1	1
<i>E.coccineum</i>	1.66	0.27	2	2	1	1
<i>N.antarctica</i>	1.66	0.27	2	2	1	1
<i>M.boaria</i>	0.33	0.05	2	2	1	1
<i>M.chubutensis</i>	0.33	0.05	2	2	1	1
<i>S.patagonicus</i>	0.33	0.05	2	2	1	1
<i>L.hirsuta</i>	0	0	2	2	1	1
<i>B.darwinii</i>	0	0	2	2	1	1
<i>A.microphylla</i>	0	0	2	2	1	1
<i>G.mucronata</i>	0	0	2	2	1	1



**Figura A.3:** Red de las especies de plantas ordenada por comunidades. Dos comunidades fueron identificadas.

A continuación (Tabla A.5) se presenta la lista de nodos de la red presentada en el Capítulo 2.

Nodo	Nombre
1	Ants
2	Bee1
3	Bee14
4	Bee2
5	Bee3
6	Bee4
7	Bee5
8	Bee6
9	Bee7
10	Beetle1
11	Beetle2
12	Beetle3
13	Butterfly2
14	Fly1
15	Fly11
16	Fly13
17	Fly15
18	Fly4
19	Fly5
20	Fly7
21	Fly8
22	Fly9
23	Unknown
24	Wasp1
25	Wasp5
26	Herb2
27	Herb8
28	Herb10
29	Herb4
30	Herb11
31	Herb3
32	Herb7
33	Herb1
34	Herb9
35	Herb5
36	Herb6
37	B.darwinii
38	E.coccineum
39	L.hirsuta
40	M.boaria
41	M.chubutensis
42	R.cucullatum
43	R.magellanica
44	S.patagonicus
45	A.microphylla
46	G.mucronata
47	N.antarctica

**Tabla A.5:** Lista de nodos

Conjunto de enlaces de la red [2.7](#)

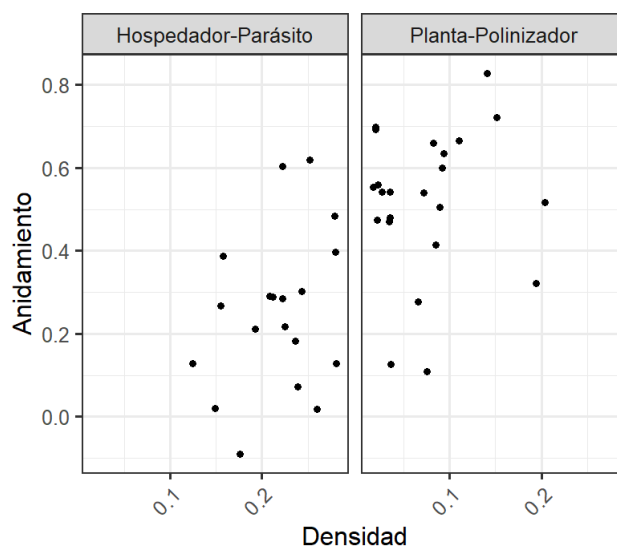
$$\begin{aligned}
 E = \{ & (1, 37), (1, 38), (1, 39), (1, 41), (1, 43), (1, 44), (2, 37), (2, 38), (2, 39), (2, 43), (2, 44), \\
 & (4, 38), (4, 39), (4, 44), (5, 38), (5, 39), (6, 38), (6, 39), (7, 37), (7, 38), (7, 39), (7, 43), \\
 & (7, 44), (8, 38), (8, 39), (8, 43), (10, 38), (10, 41), (10, 44), (12, 38), (12, 44), (14, 37), (14, 39), \\
 & (14, 40), (14, 41), (14, 42), (14, 43), (14, 44), (17, 37), (17, 43), (18, 38), (18, 43), (19, 37), (19, 43), \\
 & (19, 44), (20, 38), (20, 39), (20, 40), (20, 43), (20, 44), (21, 38), (21, 43), (21, 44), (22, 39), (22, 41), \\
 & (22, 43), (22, 44), (23, 38), (23, 39), (24, 38), (24, 40), (24, 41), (26, 38), (26, 39), (26, 40), (26, 41), \\
 & (26, 42), (26, 44), (26, 47), (27, 37), (27, 38), (27, 42), (27, 43), (27, 47), (28, 39), (28, 47), (29, 39), \\
 & (29, 40), (29, 42), (29, 47), (30, 42), (30, 43), (30, 45), (31, 42), (32, 38), (32, 40), (32, 41), (32, 42), \\
 & (32, 43), (32, 44), (32, 46), (32, 47), (33, 47), (34, 39), (34, 40), (34, 42), (34, 44), (35, 44), (35, 46), \\
 & (35, 47), (36, 39), (36, 42), (36, 44), (36, 47) \}
 \end{aligned}$$

# Apéndice B

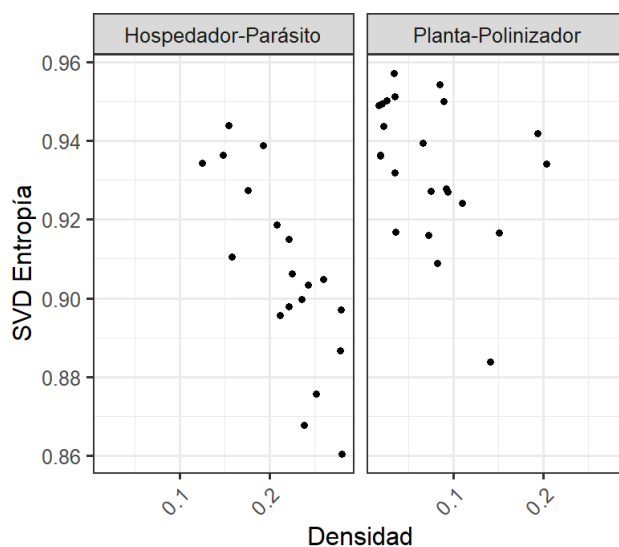
## Material complementario B

En el momento de estudiar los índices de complejidad, usamos 23 redes de planta-polinizador y 19 redes de hospedador-parásito del conjunto de datos que se encuentra disponible en [web-of-life.es](http://web-of-life.es).

Los valores de anidamiento y SVD entropía de las redes seleccionadas fueron calculadas y graficadas en la [Figura B.2](#).



**Figura B.1:** Relación entre anidamiento y densidad en la red hospedador-parásito (izquierda) y en la red planta-polinizador (derecha).



**Figura B.2:** Relación entre SVD-entropía y densidad en la red hospedador-parásito (izquierda) y en la red planta-polinizador (derecha).

Las tablas B.1-B.2 muestran el nombre de la red, número de nodos, enlaces, densidad y el sitio de estudio donde se tomaron los datos que fueron extraídos de la página web web-of-life.es.

**Tabla B.1:** Lista de redes de hospedador-parásito que se usaron en el capítulo 3

Red	Nodos	Cantidad de enlaces	Densidad	Anidamiento	Sitio de estudio
<i>A_HP_002</i>	42	96	0.222	0.284	Akmolinsk
<i>A_HP_006</i>	53	123	0.208	0.289	Armenia
<i>A_HP_008</i>	32	37	0.193	0.212	Chimkent
<i>A_HP_010</i>	49	88	0.158	0.388	East Balkhash
<i>A_HP_019</i>	27	36	0.222	0.603	Krasnojarsk
<i>A_HP_022</i>	34	68	0.236	0.182	Kustanai
<i>A_HP_024</i>	27	39	0.279	0.484	Mongolia, Central Khangay
<i>A_HP_025</i>	58	107	0.149	0.019	Mongolia, North Western Khangay
<i>A_HP_027</i>	47	108	0.212	0.288	Moyyunkum
<i>A_HP_029</i>	49	79	0.155	0.268	Kyrgyz Republic
<i>A_HP_030</i>	29	51	0.243	0.301	North Russian Far East
<i>A_HP_031</i>	56	134	0.225	0.218	Novosibirsk
<i>A_HP_032</i>	27	32	0.176	-0.090	Pavlodar
<i>A_HP_037</i>	38	90	0.252	0.619	Slovakia
<i>A_HP_038</i>	30	52	0.26	0.018	Southwestern Azerbaijan
<i>A_HP_042</i>	53	84	0.125	0.127	Tarbagatai
<i>A_HP_043</i>	38	73	0.28	0.396	Terskey Alatau
<i>A_HP_044</i>	27	197	0.281	0.129	Tomsk, Tumen
<i>A_HP_050</i>	62	226	0.239	0.072	Volga, Kama

**Tabla B.2:** Lista de redes de planta-polinizador que se usaron en el capítulo 3

Red	Nodos	Cantidad de enlaces	Densidad	Anidamiento	Sitio de estudio
<i>M.PL_005</i>	371	923	0.035	0.542	Pikes Peak, Colorado, USA
<i>M.PL_006</i>	78	146	0.141	0.828	Hickling, Norfolk, UK
<i>M.PL_009</i>	142	242	0.085	0.415	Latnjajaure, Abisko, Sweden
<i>M.PL_010</i>	107	456	0.194	0.321	Zackenbergl
<i>M.PL_013</i>	65	103	0.204	0.516	KwaZulu, Natal region, South Africa
<i>M.PL_015</i>	797	2933	0.034	0.471	Daphní, Athens, Greece
<i>M.PL_016</i>	205	412	0.089	0.506	Doñana Nat. Park, Spain
<i>M.PL_017</i>	104	299	0.151	0.721	Bristol, England
<i>M.PL_021</i>	768	1193	0.019	0.698	Ashu, Kyoto, Japan
<i>M.PL_023</i>	95	125	0.075	0.108	Rio Blanco, Mendoza, Argentina
<i>M.PL_026</i>	159	204	0.036	0.126	Galapagos
<i>M.PL_031</i>	97	156	0.066	0.276	Canaima Nat. Park, Venezuela
<i>M.PL_034</i>	154	312	0.094	0.633	Chiloe, Chile
<i>M.PL_044</i>	719	1125	0.017	0.552	Amami, Ohsima Island, Japan
<i>M.PL_047</i>	205	425	0.11	0.664	Isenbjerg
<i>M.PL_048</i>	266	671	0.092	0.599	Denmark
<i>M.PL_053</i>	393	589	0.021	0.474	Mt. Yufu, Japan
<i>M.PL_054</i>	431	773	0.022	0.559	Kyoto City, Japan
<i>M.PL_055</i>	259	431	0.035	0.4805	Nakaikemi marsh, Fukui Prefecture, Japan
<i>M.PL_056</i>	456	871	0.026	0.542	Mt. Kushigata, Yamanashi Pref, Japan
<i>M.PL_057</i>	997	1920	0.019	0.692	Kibune, Kyoto, Japan
<i>M.PL_072_04</i>	94	192	0.072	0.539	Difuntos, Pampas, Argentina
<i>M.PL_072_05</i>	84	138	0.082	0.659	El Morro, Pampas, Argentina

La siguiente tabla B.3 es un resumen de la cantidad de valores de  $\beta$  cuyo intervalo de confianza incluye el cero, la cantidad de valores de  $\beta$  positivos (sin incluir el cero) y la cantidad de valores de  $\beta$  negativos (sin incluir el cero).

		Modelo 1			Modelo 2		
		0	<0	>0	0	<0	>0
<b>Hospedador-parásito</b>	$\beta_k$	2	4	13	0	1	18
	$\beta_{svd}$	3	5	11	0	0	19
	$\beta_{modularidad}$	5	9	5	1	3	15
	$\beta_{entropía}$	8	8	3	4	10	5
<b>Planta-Polinizador</b>	$\beta_k$	2	15	6	0	19	4
	$\beta_{svd}$	2	4	17	1	0	22
	$\beta_{modularidad}$	0	22	1	1	6	16
	$\beta_{entropía}$	3	1	19	2	0	21

**Tabla B.3:** Tabla resumen

# Apéndice C

## Pseudocódigo de los algoritmos principales

Pseudocódigo para generar una matriz de adyacencia con ceros y unos

**Notación:**

- El archivo original contiene las interacciones entre nodos con sus pesos y los nombres de cada nodo.
- El archivo tiene en la primera columna el nombre de los nodos por ese motivo  $j$  comienza en 2.

**Entrada:** Archivo con datos de interacciones, con  $n$  filas y  $m$  columnas

**Salida :** Matriz de adyacencia binaria guardada en un archivo

- Leer el archivo y almacenarlo en la variable `salida`
- Obtener el número de filas  $n$  y el número de columnas  $m$  de `salida`
- Inicializar una matriz de ceros `matrizdeceros` de tamaño  $n \times n$

**for**  $j \leftarrow 2$  *hasta*  $m$  **do**

```
    for  $i \leftarrow 1$  hasta  $n$  do
        if salida[ $i, j$ ]  $\neq 0$  then
            matrizdeceros[ $i, j - 1$ ] = 1
        end
    end
```

**end**

- Obtener los nombres de las columnas desde la segunda hasta la última columna y asignarlos como etiquetas de filas y columnas de `matrizdeceros`
- Guardar `matrizdeceros` en un nuevo archivo llamado `matrizady`

**Algoritmo 1:** Generación de Matriz de Adyacencia

### Pseudocódigo para seleccionar región de interés

El algoritmo 2 calcula la región de interés, ya que nos interesa la región donde ninguna fila o columna es cero.

**Entrada:** Archivo con datos de una matriz, con  $n$  filas y  $m$  columnas

**Salida :** Vector con los valores mínimos y máximos de filas y columnas con datos no nulos

- Leer el archivo y almacenarlo en la variable `salida`
- Obtener el número de filas  $n$  y el número de columnas  $m$  de `salida`
- Inicializar listas vacías `filas` y `columnas`
- Inicializar contador `contador`  $\leftarrow 0$

**Paso 1: Identificar las filas con valores no nulos**

for  $i \leftarrow 1$  hasta  $n$  do

```

|   for  $j \leftarrow 1$  hasta  $m$  do
|       if salida[ $i, j$ ]  $\neq 0$  then
|           Incrementar contador
|           Almacenar  $i$  en filas en la posición contador
|       end
|   end
end

```

end

- Inicializar contador `conta`  $\leftarrow 0$

**Paso 2: Identificar las columnas con valores no nulos**

for  $j \leftarrow 1$  hasta  $m$  do

```

|   for  $i \leftarrow 1$  hasta  $n$  do
|       if salida[ $i, j$ ]  $\neq 0$  then
|           Incrementar conta
|           Almacenar  $j$  en columnas en la posición conta
|       end
|   end
end

```

end

- Calcular  $min_{fil} = \text{mín}(\text{filas})$
- Calcular  $max_{fil} = \text{máx}(\text{filas})$
- Calcular  $min_{col} = \text{mín}(\text{columnas})$
- Calcular  $max_{col} = \text{máx}(\text{columnas})$
- Definir el vector de dimensiones: `dimension` =  $(min_{fil}, max_{fil}, min_{col}, max_{col})$
- Devolver `dimension`

**Algoritmo 2:** Determinación de la Región No Nula en una Matriz

## Pseudocódigo para desordenar una red conservando el grado de los nodos

El algoritmo 3 realiza cambios en la matriz de la red original.

- $A$ : Matriz de adyacencia de la red original.

**Entrada:** Matriz de adyacencia de la red original  $A$ , número de cambios  $\Delta$

**Salida :** Matriz de adyacencia modificada conservando el grado de cada nodo, valores de la modularidad

### Paso 1: Calcular el grado de cada nodo

Calcular el grado de cada nodo a partir de la matriz de adyacencia  $A$

### Paso 2: Realizar $\Delta$ cambios en la red

**for**  $k = 1$  *hasta*  $\Delta$  **do**

    | Seleccionar dos nodos  $i$  e  $i'$  aleatoriamente

    | Seleccionar un vecino  $j$  de  $i$  tal que  $j$  no sea vecino de  $i'$

    | Seleccionar un vecino  $j'$  de  $i'$  tal que  $j'$  no sea vecino de  $i$

    | Intercambiar los enlaces:  $A_{ij} \leftarrow 0$ ,  $A_{i'j'} \leftarrow 0$ ,  $A_{ij'} \leftarrow 1$ ,  $A_{i'j} \leftarrow 1$

**end**

### Paso 3: Calcular el número de cambios

Comparar la nueva matriz de adyacencia con la original para determinar el número de enlaces distintos

### Paso 4: Calcular la modularidad de la nueva red

Calcular el valor de la modularidad en la matriz de adyacencia modificada

### Paso 5: Repetir hasta que haya una estabilidad en la modularidad

**while** *la modularidad no sea estable* **do**

    | Repetir pasos 2, 3 y 4 aumentando el grado de desorden

**end**

**Algoritmo 3:** Algoritmo para desordenar una red conservando el grado de los nodos

### Pseudocódigo para desordenar una red cambiando el grado de los nodos

**Entrada:** Archivo con matriz de adyacencia (*archivo*),  
 número total de cambios (*cambios totales*),  
 límites de submatriz (*filmin, filmax, colmin, colmax*)

**Salida :** Lista de matrices con la distribución modificada

- Leer el archivo de la matriz y almacenarlo en **salidap**
- Extraer la submatriz de interés **matrizcut** dentro de los límites dados
- Inicializar una lista vacía **multiarray** para almacenar las matrices modificadas
- Inicializar un contador **contamos** en 0

```

while contamos < cambios totales do
  | Obtener las posiciones de los valores 1 en matrizcut y almacenarlas en
  | posicion_unos
  | Seleccionar aleatoriamente una posición (f1, c1) de posicion_unos
  | Obtener las posiciones de los valores 0 en matrizcut y almacenarlas en
  | posicion_ceros
  | if La fila f1 y la columna c1 tienen más de un 1 then
  | | Incrementar contamos
  | | Establecer matrizcut[f1, c1] = 0
  | | Seleccionar aleatoriamente una posición (f0, c0) de posicion_ceros
  | | Establecer matrizcut[f0, c0] = 1
  | | Guardar la matriz modificada en multiarray[contamos]
  | end
end

```

- Devolver **multiarray** con las matrices modificadas

**Algoritmo 4:** Algoritmo para modificar la distribución de grado

## Pseudocódigo para calcular el índice $oARI_{adj}$

Se uso el algoritmo 5 para calcular la diferencia entre los nodos de las comunidades y los nodos de los hiperenlaces.

### Notación:

- $n_{11}$ : Ambos pares en la misma comunidad en ambas particiones
- $n_{00}$ : Ambos pares en comunidades diferentes en ambas particiones
- $n_{10}$  : Mismo en la primera, diferente en la segunda
- $n_{01}$ : Diferente en la primera, mismo en la segunda
- $n\_nodes$ : Número total de nodos de la red.
- `louvain_partition`: Comunidades obtenidas con el algoritmo de Louvain/
- `hyperedges_partition`: hiperenlaces del hipergrafo.
- `n_nodes`: número de nodos
- $oARI_{adj}$ : Índice ajustado de Rand

### Fórmula de $oARI_{adj}$ :

$$oARI_{adj} = \frac{a - \frac{(a+c)(a+b)}{d}}{\frac{(a+c)+(a+b)}{2} - \frac{(a+c)(a+b)}{d}}.$$

**Entrada:** *louvain\_partition*, *hyperedges\_partition*, *n\_nodes*

**Salida** :  $oARI_{adj}$

- Inicializar  $n_{11} \leftarrow 0$ ,  $n_{00} \leftarrow 0$ ,

$n_{10} \leftarrow 0$ ,  $n_{01} \leftarrow 0$

- Calcular las combinaciones de todos los pares de nodos: **node\_pairs**

$\leftarrow \text{combn}(1 : n\_nodes, 2)$

**foreach**  $p \in 1$  hasta  $n\text{col}(\text{node\_pairs})$  **do**

- Definir  $i \leftarrow \text{node\_pairs}[1, p]$  y  $j \leftarrow \text{node\_pairs}[2, p]$

- Verificar si  $i$  y  $j$  están en la misma comunidad:

$\text{in\_same\_community\_louvain} \leftarrow \text{True}$  si  $\text{louvain\_partition}[i] ==$   
 $\text{louvain\_partition}[j]$ , sino **False**

- Verificar si  $i$  y  $j$  están en el mismo hiperenlace:

$\text{in\_same\_community\_hyperedges} \leftarrow \text{True}$  si existe una comunidad tal que  
 $i$  y  $j$  este en el mismo hiperenlace, sino **False**

**if**  $\text{in\_same\_community\_louvain}$  y  $\text{in\_same\_community\_hyperedges}$  **then**

    Incrementar  $n_{11} \leftarrow n_{11} + 1$

**end**

**else if** *no* ( $\text{in\_same\_community\_louvain}$ ) y *no*

( $\text{in\_same\_community\_hyperedges}$ ) **then**

    Incrementar  $n_{00} \leftarrow n_{00} + 1$

**end**

**else if**  $\text{in\_same\_community\_louvain}$  y *no*

( $\text{in\_same\_community\_hyperedges}$ ) **then**

    Incrementar  $n_{10} \leftarrow n_{10} + 1$

**end**

**else**

    Incrementar  $n_{01} \leftarrow n_{01} + 1$

**end**

**end**

**Calcular:**  $d \leftarrow n_{11} + n_{10} + n_{01} + n_{00}$

- Definir los términos:  $a \leftarrow n_{11}$ ,  $b \leftarrow n_{10}$ ,  $c \leftarrow n_{01}$

- Calcular el numerador:  $\text{num} \leftarrow a - \frac{(a+c)(a+b)}{d}$

- Calcular el denominador:  $\text{denom} \leftarrow \frac{(a+c)+(a+b)}{2} - \frac{(a+c)(a+b)}{d}$

- Calcular el índice ajustado de Rand:  $oARI_{adj} \leftarrow \frac{\text{num}}{\text{denom}}$

- Devolver  $oARI_{adj}$

**Algoritmo 5:** Cálculo del Overlap Adjusted Rand Index

# Bibliografía

- [1] Newman, M. E. J. The structure and function of complex networks. *SIAM Review*, **45** (2), 167–256, 2003. URL <http://dx.doi.org/10.1137/S003614450342480>. 1
- [2] Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., Hwang, D. Complex networks: Structure and dynamics. *Physics Reports*, **424** (4), 175–308, 2006. URL <https://www.sciencedirect.com/science/article/pii/S037015730500462X>.
- [3] Estrada, E. The Structure of Complex Networks: Theory and Applications. Oxford University Press, 2011. 1
- [4] Albert, R., Barabási, A. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, **74**, 47–97, 2002. URL <https://link.aps.org/doi/10.1103/RevModPhys.74.47>. 1
- [5] Newman, M. Networks: An introduction. oxford university press. *Journal of Animal Ecology*, **18**, 2010. 7, 14, 20, 36
- [6] Barabási, A., Pósfai, M. Network science. Cambridge: Cambridge University Press, 2016. URL <http://barabasi.com/networksciencebook/>. 1
- [7] FAO. State Of The World's Forests 2012. Food and Agriculture Organization of The United Nations., 2012. 2, 4
- [8] Daily, G. C. Nature's services. societal dependence on natural ecosystems. *Animal Conservation*, **1** (1), 75–76, 1996. 2
- [9] Garibaldi, L., Kitzberger, T., Mazía, C., Chaneton, E. Nutrient supply and bird predation additively control insect herbivory and tree growth in two contrasting forest habitats. *Oikos*, **119** (2), 337–349, 2010.
- [10] Garibaldi, L., Kitzberger, T., Chaneton, E. Environmental and genetic control of insect abundance and herbivory along a forest elevational gradient. *Oecologia*, **167**, 117–129, 2011. URL <https://doi.org/10.1007/s00442-011-1978-0>. 33

- [11] Garibaldi, L., Casas, C., Goldenberg, M., Gowda, J., Heinemann, K., Kitzberger, T., *et al.* Silvicultura del ciprés de la cordillera: ¿una alternativa rentable y sustentable para la provincia de Rio Negro? Lavallol C.I. (ed.), 2016. 2
- [12] Echeverría, C., Newton, A., Coomes, A., Nahuelhual, L., Rey-Benayas, J. How landscapes change: integration of spatial patterns and human processes in temperate landscapes of southern Chile. *Applied Geography*, **32**, 822–831, 2007. 3
- [13] McIntyre, S., Hobbs, R. A framework for conceptualizing human effects on landscapes and its relevance to management and research models. *Conservation Biology*, **13**, 1282–1292, 1999. 3
- [14] Echeverría, C., Coomes, D., Salas, J., Rey-Benayas, J., Lara, A., Newton, A. Rapid deforestation and fragmentation of Chilean temperate forests. *Biological Conservation*, **130** (4), 481–494, 2006. URL <https://www.sciencedirect.com/science/article/pii/S0006320706000358>. 3
- [15] Steininger, M. K. Satellite estimation of tropical secondary forest above-ground biomass: Data from Brazil and Bolivia. *International Journal of Remote Sensing*, **21**, 1139–1157, 2000. 3
- [16] Hooper, D., Chapin III, F. S., Ewel, J. J., Hector, A., Inchausti, P., Lavorel, S., *et al.* Effects of biodiversity on ecosystem functioning: A consensus of current knowledge. *Ecological Monographs*, **75**, 1139–1157, 2005. 3
- [17] Steffan-Dewenter, I., Kessler, M., Barkmann, J., Bos, M., Buchori, D., Erasmí, S., *et al.* Tradeoffs between income, biodiversity, and ecosystem functioning during tropical rainforest conversion and agroforestry intensification. *Proceedings of the National Academy of Sciences of the United States of America*, **104** (12), 4973–4978, 2007. 4
- [18] Nylund, J., Gowda, J. Forest legislation: stimulating or discouraging forest owners? *Forest Ideas*, **19**, 100–106, 01 2011. 4
- [19] Chauchard, L. M., Bava, J., Castañeda, P., S. and Laclau, Loguercio, G. A., Pantanius, P. M., Rusch, V. Manual para las buenas prácticas forestales en bosques nativos de Norpatagonia. CIEFAP, 2008. 4
- [20] Rusch, V., Lantschner, M. V. La Biodiversidad. Manual de Buenas Prácticas para el manejo forestal sustentable de plantaciones, con énfasis en la conservación de la biodiversidad en la región de la Patagonia. Ministerio de Agricultura, Ganadería y Pesca de la Nación, Unidad para el Cambio Rural, Buenos Aires., 2015. 4

- [21] Jaksic, F., Feinsinger, P. Bird assemblages in temperate forests of north and south america: A comparison of diversity, dynamics, guild structure, and resource use. *Revista Chilena de Historia Natural*, **64**, 491–510, 1991. [4](#)
- [22] Veblen, T., Hill, R., Read, J. Ecology of southern Chilean and Argentinean Nothofagus forests. *The Ecology and Biogeography of Nothofagus Forests*. Yale University Press, 1996.
- [23] Raffaele, E., De Torres Curth, M., Morales, C., Kitzberger, T. Ecología e historia natural de la Patagonia Andina. Un cuarto de siglo de investigación en biogeografía, ecología y conservación. Fundación de Historia Natural Félix de Azara. Buenos Aires., 2014. [5](#), [41](#), [85](#)
- [24] Gönc, R., Casaux, R. J., Sulkin-Dolhatz, D. Effects of disturbances generated by different management strategies on the vegetation strata of nothofagus antarctica forests of chubut , argentina. *Ecología Austral*, **25**, 231–241, 2015. [4](#)
- [25] MEA. Ecosystems and human well-being: Millennium Ecosystem Assessment. Island Press, Washington, D.C., EEUU., 2015. [4](#)
- [26] Martínez Pastur, G., Soler, R., Lencinas, M., Cellini, J. M., Peri, P. L. Long-term monitoring of thinning for silvopastoral purposes in nothofagus antarctica forests of tierra del fuego, argentina. *Forest Systems*, **27** (1), e01S, 2018. [4](#)
- [27] Gowda, J., Kitzberger, T., Premoli, A. Landscape responses to a century of land use along the northern patagonian forest-steppe transition. *Plant Ecology*, **213**, 259–272, 2012. [5](#)
- [28] Bava, J. O., Loguercio, G. A., Salvador, G. ¿por qué plantar en patagonia? estado actual y el rol futuro de los bosques plantados. *Ecología Austral*, **25** (2), 101–111, 2015. [5](#)
- [29] Pretzsch, H., Forrester, D., J., B. Mixed-Species Forests. Springer Berlin Heidelberg., 2017. [5](#)
- [30] Pockock, M., Evans, D., Memmott, J. The robustness and restoration of a network of ecological networks. *Science*, **335** (6071), 973–977, 2012. [5](#)
- [31] Morales, C. L., Aizen, M. A. Invasive mutualisms and the structure of plant–pollinator interactions in the temperate forests of north-west patagonia, argentina. *Journal of Ecology*, **94** (1), 171–180, 2006. [5](#)

- [32] Tylianakis, J. M., Laliberté, E., Nielsen, A., Bascompte, J. Conservation of species interaction networks. *Biological Conservation*, **143** (10), 2270–2279, 2010. URL <https://www.sciencedirect.com/science/article/pii/S0006320709005126>.
- [33] Ceballos, S. J., Chacoff, N., Malizia, A. Interaction network of vascular epiphytes and trees in a subtropical forest. *Acta Oecologica*, **77**, 152–159, 2016. URL <https://www.sciencedirect.com/science/article/pii/S1146609X16300790>.
- [34] Emer, C., Jordano, P., Pizo, M. A., Ribeiro, M. C., Da Silva, F. R., Galetti, M. Seed dispersal networks in tropical forest fragments: Area effects, remnant species, and interaction diversity. *Biotropica*, **52**, 81–89, 2020. 5
- [35] Scherber, C., Eisenhauer, N., Weisser, W., Schmid, B., Voigt, W., Fischer, M., *et al.* Bottom-up effects of plant diversity on multitrophic interactions in a biodiversity experiment. *Nature*, **468**, 553–556, 2010. 5, 6
- [36] Garibaldi, L., Steffan-Dewenter, I., Winfree, R., Aizen, M. A., Bommarco, R., Cunningham, S. A., *et al.* Wild pollinators enhance fruit set of crops regardless of honey bee abundance. *Science*, **339** (6127), 1608–1611, 2013. URL <https://www.science.org/doi/abs/10.1126/science.1230200>. 5
- [37] Fontaine, C., Dajoz, I., Meriguet, J., Loreau, M. Functional diversity of plant-pollinator interaction webs and its effect on the stability of plant communities. *Nature*, **444** (7117), 1–4, 2006. 5
- [38] Chacoff, N. P., Monmany, A. C. Interacciones planta-insecto y sus implicancias para los sistemas agrícolas en el pedemonte del Alto Bermejo, págs. 345–365. Ediciones del Subtrópico, Fundación ProYungas, 2009. 5
- [39] Coley, P. D., Barone, J. A. Herbivory and plant defenses in tropical forests. *Annual Review of Ecology and Systematics*, **27**, 305–335, 1996. URL <http://www.jstor.org/stable/2097237>. 6
- [40] Turcotte, M. M., Thomsen, C. J. M., Broadhead, G. T., Fine, P. V. A., Godfrey, R. M., Lamarre, G. P. A., *et al.* Percentage leaf herbivory across vascular plant species. *Ecology*, **95**, 788–788, 2013. URL <https://doi.org/10.1890/13-1741.1>. 6
- [41] Maguire, D. Y., James, P. M., Buddle, C. M., Bennett, E. M. Landscape connectivity and insect herbivory: A framework for understanding tradeoffs among ecosystem services. *Global Ecology and Conservation*, **4**, 73–84, 2015. URL <https://www.sciencedirect.com/science/article/pii/S2351989415000542>. 6

- [42] Guyot, V., Castagneyrol, B., A., V., M., D., Jactel, H. Tree diversity reduces pest damage in mature forests across europe. *Biology Letters*, **12**, 2016. 6
- [43] Kearns, C., Inouye, D., Waser, N. Endangered mutualisms: The conservation of plant-pollinator interactions. *Annual Review of Ecology, Evolution, and Systematics*, **29**, 83–112, 2016. 6
- [44] Potts, S. G., Vulliamy, B., Dafni, A., Ne'eman, G., O'Toole, C., Roberts, S., *et al.* Response of plant-pollinator communities to fire: Changes in diversity, abundance and floral reward structure. *Oikos*, **101** (1), 103–112, 2003. URL <http://www.jstor.org/stable/3548348>. 6
- [45] Klein, A., Vaissière, B., Cane, J., Steffan-Dewenter, I., Cunningham, S., Kremen, C., *et al.* Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B*, **274**, 303–313, 2007. 6
- [46] Allen-Wardell, G., Bernhardt, P., Bitner, P., Burquez, A., Buchmann, S., Cane, J., *et al.* The potential consequences of pollinator decline on the biodiversity and stability of food crop yields. *Conservation Biology*, **12**, 8–17, 1998. 6
- [47] Kremen, C., Williams, N., Aizen, M., Gemmill-Herren, B., LeBuhn, G., Minckley, R., *et al.* Pollination and other ecosystem services produced by mobile organisms: a conceptual framework for the effects of land-use change. *Ecology Letters*, **10**, 299–314, 2007. 6
- [48] Tylianakis, J., Tschardtke, T., Lewis, O. Habitat modification alters the structure of tropical host–parasitoid food webs. *Nature*, **445**, 202–205, 2007. 6
- [49] Aizen, M. A., Feinsinger, P. Bees Not to Be? Responses of Insect Pollinator Faunas and Flower Pollination to Habitat Fragmentation, págs. 111–129. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003. URL [https://doi.org/10.1007/978-3-662-05238-9\\_7](https://doi.org/10.1007/978-3-662-05238-9_7). 6
- [50] Ghazoul, J. Impact of logging on the richness and diversity of forest butterflies in a tropical dry forest in thailand. *Biodiversity and Conservation*, **11**, 521–541, 2002. 6
- [51] Cottee-Jones, H., Whittaker, R. The keystone species concept: a critical appraisal. *frontiers of biogeography*, **4** (3), 117–127, 2012. URL <https://doi.org/10.21425/F5FBG12533>. 6
- [52] Mello, M., Rodrigues, F., da Fontoura Costa, L., Kissling, W., Şekercioğlu, C. H., Marquitti, F., *et al.* Keystone species in seed dispersal networks are mainly

- determined by dietary specialization. *Oikos*, **124**, 1031–1039, 2015. URL <https://api.semanticscholar.org/CorpusID:83672149>. 6
- [53] Paine, R. T. Food web complexity and species diversity. *American Naturalist*, **100** (910), 65–75, 1966. 6
- [54] Hooper, D. U., Vitousek, P. M. The effects of human disturbance on ecosystem processes and sustainability. *Ecological Applications*, **7** (3), 694–708, 1997. 6
- [55] Diestel, R. Graph Theory. Springer, 2000. 7
- [56] Boccaletti, S., Bianconi, G., Criado, R., del Genio, C., Gómez-Gardeñes, J., Romance, M., *et al.* The structure and dynamics of multilayer networks. *Physics Reports-Review Section of Physics Letters*, **544** (1), 1–122, 2014. URL <https://www.sciencedirect.com/science/article/pii/S0370157314002105>. 8, 22, 28, 31, 34
- [57] Kivela, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., Porter, M. A. Multilayer networks. *Journal of Complex Networks*, **2** (3), 203–271, 2014. URL <http://dx.doi.org/10.1093/comnet/cnu016>. 8, 22, 31
- [58] Almeida-Neto, M., Guimarães, P., Guimarães, P. A consistent metric for nestedness analysis in ecological systems: reconciling concept and measurement. *Oikos*, **117**, 1227–1239, 2008. 10
- [59] Levin, L. A. Laws on the conservation (zero increase) of information, and questions on the foundations of probability theory. *Akademii Nauk SSSR. Problemy Peredachi Informatsii*, **10**, 30–35, 1974a. 11, 46, 51
- [60] Solomonoff, R. A formal theory of inductive inference. part i. *Information and Control*, **7** (1), 1–22, 1964. 11, 46, 51
- [61] Blondel, V. D., Guillaume, J. L., Lambiotte, R., Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech.-Theory Exp.*, **10**, P10008, 2008. URL <http://dx.doi.org/10.1088/1742-5468/2008/10/P10008>. 11, 23, 48
- [62] Barber, M. J. Modularity and community detection in bipartite networks. *Phys. Rev. E*, **76**, 066102, Dec 2007. URL <https://link.aps.org/doi/10.1103/PhysRevE.76.066102>. 12, 48, 49
- [63] Farage, C., Edler, D., Eklöf, A., Rosvall, M., Pilosof, S. Identifying flow modules in ecological networks using infomap. *Methods in Ecology and Evolution*, **12** (5), 778–786, 2021. 12, 23

- [64] Golubski, A. J., Westlund, E. E., Vandermeer, J., Pascual, M. Ecological networks over the edge: Hypergraph trait-mediated indirect interaction (tmii) structure. *Trends in Ecology & Evolution*, **31** (5), 344–354, 2016. [13](#), [68](#), [79](#)
- [65] Gaucherel, C., Cosme, M., Noûs, C., Pommereau, F. A single changing hyper-network to represent (social-)ecological dynamics. *bioRxiv*, 2024. URL <https://www.biorxiv.org/content/early/2024/01/15/2023.10.30.564699>. [13](#), [79](#)
- [66] Dai, Q., Gao, Y. *Hypergraph Computation*. Springer, 2023. [14](#)
- [67] Berge, C. *Graphs and Hypergraphs*. Elsevier, New York, 1973. [15](#)
- [68] Watts, D. J., Strogatz, S. H. Collective dynamics of 'small-world' networks. *Nature*, **393**, 440–442, 1998. [15](#), [52](#)
- [69] Borgatti, S. P., Everett, M. G. Network analysis of 2-mode data. *Social Networks*, **19** (3), 243–269, 1997. URL <https://www.sciencedirect.com/science/article/pii/S0378873396003012>. [20](#)
- [70] Newman, M. E. J. Analysis of weighted networks. *Phys. Rev. E*, **70**, 056131, Nov 2004. URL <https://link.aps.org/doi/10.1103/PhysRevE.70.056131>. [23](#), [48](#)
- [71] Leicht, E. A., Holme, P., Newman, M. E. J. Vertex similarity in networks. *Phys. Rev. E*, **73**, 026120, Feb 2006a. URL <https://link.aps.org/doi/10.1103/PhysRevE.73.026120>. [20](#)
- [72] Bascompte, J. Networks in ecology. *Basic and Applied Ecology*, **8** (6), 485–490, 2007. URL <https://www.sciencedirect.com/science/article/pii/S1439179107000576>. [20](#)
- [73] Guimarães, P. R., Rico-Gray, V., Furtado dos Reis, S., Thompson, J. N. Asymmetries in specialization in ant–plant mutualistic networks. *Proc. R. Soc. B*, **273**, 2041–2047, 2006. [20](#)
- [74] Poulin, R. Network analysis shining light on parasite ecology and diversity. *Trends in Parasitology*, **26** (10), 492–498, 2010. URL <https://www.sciencedirect.com/science/article/pii/S1471492210001108>. [43](#)
- [75] Selva, N., M.A., F. The nested structure of a scavenger community. *Proceedings of the Royal Society of London B*, **274**, 1101–1108, 2007.
- [76] Jordano, P., Bascompte, J., Olesen, J. The ecological consequences of complex topology and nested structure in pollination webs. plant-pollinator interactions. from specialization to generalization. *University of Chicago Press, Chicago*, págs. 173–199, 2006. [20](#)

- [77] Newman, M. E. J., Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E.*, **69**, 026113, 2004. [23](#)
- [78] Girvan, M., Newman, M. Community structure in social and biological networks. *Proc. Natl. Acad. Sci.*, **99** (12), 7821–7826, 2002. [48](#)
- [79] Newman, M. E. J. Fast algorithm for detecting community structure in networks. *Phys. Rev. E.*, **69** (6), 066133, 2004b. [23](#), [48](#)
- [80] Meila, M. Comparing clusterings - an information based distance. *Journal of Multivariate Analysis*, **98**, 873–895, 2007. [24](#), [49](#)
- [81] Cover, T., Thomas, J. Elements of Information Theory. Wiley & Sons, New York, 2005. [24](#)
- [82] Fornito, A., Bullmore, E. T., Zalesky, A. Fundamentals of brain network analysis. *Elsevier Science: San Diego, CA, USA*, 2016. [25](#)
- [83] R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria, 2019. URL <https://www.R-project.org/>. [27](#)
- [84] Okuyama, T. Do mutualistic networks follow power distributions? *Ecological Complexity*, **5**, 59–65, 2008. [28](#)
- [85] Williams, R. J. Biology, methodology or chance? the degree distributions of bipartite ecological networks. *PLoS ONE*, **6** (3), e17645, 2011.
- [86] Marina, T. I., Saravia, L. A., Cordone, G., Salinas, V., Doyle, S. R., Momo, F. R. Architecture of marine food webs: To be or not be a ‘small-world.’. *PLoS ONE*, **13** (5), 1–13, 2018. [28](#)
- [87] Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., Barabási, A.-L. The large-scale organization of metabolic networks. *Nature*, **407** (6804), 651–654, 2000. URL <http://dx.doi.org/10.1038/35036627>. [28](#)
- [88] Jeong, H., Mason, S. P., Barabási, A. L., Oltvai, Z. N. Lethality and centrality in protein networks. *Nature*, **411** (6833), 41–42, 2001. URL <http://dx.doi.org/10.1038/35075138>.
- [89] Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., Barabási, A.-L. Hierarchical organization of modularity in metabolic networks. *Science*, **297** (5586), 1551–1555, 2002. URL <https://www.science.org/doi/abs/10.1126/science.1073374>. [28](#)

- [90] Sole, R., Goodwin, B. Signs of life: How complex pervades biology. Basic Books, New York., 2001. 28
- [91] Barabási, A.-L., Albert, R. Emergence of scaling in random networks. *Science*, **286** (5439), 509–512, 1999. URL <https://www.science.org/doi/abs/10.1126/science.286.5439.509>. 28, 36
- [92] Newman, M. Power laws, pareto distributions and zipf’s law. *Contemporary Physics*, **46** (5), 323–351, 2005. URL <http://dx.doi.org/10.1080/00107510500052444>. 29
- [93] Clauset, A., Shalizi, R., Cosma, Newman, M. E. J. Power-law distributions in empirical data. *SIAM Review*, **51** (4), 661–703, 2009. URL <http://dx.doi.org/10.1137/070710111>. 29, 89
- [94] Freeman, L. C. Centrality in social networks conceptual clarification. *Social Networks*, **1**, 215–239, 1978. URL <https://api.semanticscholar.org/CorpusID:751590>. 31
- [95] Dormann, C. How to be a specialist? quantifying specialisation in pollination networks. *Network biology*, **1**, 1–20, 2011. URL <https://api.semanticscholar.org/CorpusID:73589511>. 31
- [96] Coulin, C., Aizen, M., Garibaldi, L. Contrasting responses of plants and pollinators to wood-land disturbance. *Wiley Online Library. Austral Ecology*, **44** (6), 1040–1051, 2019. URL <http://dx.doi.org/10.1111/aec.12771>. 33, 34, 41
- [97] Nacif, M., Kitzberger, T., Garibaldi, L. Positive outcomes between herbivore diversity and tree survival: Responses to management intensity in a patagonian forest. *Forest Ecology and Management*, **458**, 117738, 2020. URL <https://www.sciencedirect.com/science/article/pii/S0378112719313192>. 33, 41, 85
- [98] Goldenberg, M., Oddi, F., Gowda, J., Garibaldi, L. Effects of firewood harvesting intensity on biodiversity and ecosystem services in shrublands of northern patagonia. *Forest Ecosystems*, **7**, 2020. URL <https://doi.org/10.1186/s40663-020-00255-y>. 34
- [99] Stanturf, J., Palik, B., Dumroese, R. K. Contemporary forest restoration: A review emphasizing function. *Forest Ecology and Management*, **331**, 292–323, 2014. URL <https://www.sciencedirect.com/science/article/pii/S0378112714004654>. 38

- [100] Gatica Saavedra, P., Echeverria, C., Nelson, C. Ecological indicators for assessing ecological success of forest restoration: a world review. *Restoration Ecology*, **25** (6), 850–857, 2017. 38
- [101] Bascompte, J., Jordano, P. Mutualistic Networks. Princeton University Press, 2013. 40
- [102] Burgos, E., Ceva, H., Perazzo, R., Devoto, M., Medan, D., Zimmermann, M., *et al.* Why nestedness in mutualistic networks? *Journal of Theoretical Biology*, **249** (2), 307–313, 2007. URL <https://www.sciencedirect.com/science/article/pii/S0022519307003670>.
- [103] Martín González, A., Dalsgaard, B., Olesen, J. Centrality measures and the importance of generalist species in pollination networks. *Ecological Complexity*, **7** (1), 36–43, 2010. URL <https://www.sciencedirect.com/science/article/pii/S1476945X09000294>.
- [104] Medan, D., Perazzo, R. P. J., Devoto, M., Burgos, E., Zimmermann, M. G., Ceva, H., *et al.* Analysis and assembling of network structure in mutualistic systems. *Journal of Theoretical Biology*, **246**, 510–521, 2007. URL <https://doi.org/10.1016/j.jtbi.2006.12.033>.
- [105] Memmott, J., Waser, N., Price, M. Tolerance of pollination networks to species extinctions. *Proceedings of the Royal Society*, **271**, 2605–2611, 2004. URL <http://doi.org/10.1098/rspb.2004.2909>.
- [106] Poisot, T., Stouffer, D., Gravel, D. Beyond species: Why ecological interaction networks vary through space and time. *Oikos*, **124** (3), 243–251, 2014. URL <https://doi.org/10.1111/oik.01719>.
- [107] Vázquez, D., Simberloff, D. Ecological specialization and susceptibility to disturbance: conjectures and refutations. *The American Naturalist*, **159** (6), 606–623, 2002. URL <https://doi.org/10.1086/339991>. 40
- [108] Strona, G., Nappo, D., Boccacci, F., Fattorini, S., San-Miguel-Ayanz, J. A fast and unbiased procedure to randomize ecological binary matrices with fixed row and column totals. *Nature Communications*, **5**, 4114, 2014. 40
- [109] Donoso Zegers, C. Las especies arbóreas de los bosques templados de Chile y Argentina: autoecología. Valdivia, Chile: Marisa Cuneo, ediciones., 2006. 41, 85
- [110] Rapoport, E., Ladio, A., Sanz, E. Plantas Nativas Comestibles de la Patagonia Andina, Argentino-Chilena. Parte I. San Carlos de Bariloche: Universidad Nacional del Comahue y National Geographic Society, 1999. 41

- [111] Itkin, S. Plantas de la Patagonia para la salud. San Carlos de Bariloche: Caleuche,, 2004. [41](#)
- [112] Gut, B. Trees in Patagonia. Springer Science & Business Media, 2008. [41](#)
- [113] Hansen, N., Codesal, P., Quinteros, P., Gallardo, C. Especies del sotobosque en ambientes de Ñire en Chubut. Ediciones INTA, 2013. [41](#)
- [114] Fajardo, A., Siefert, A. Phenological variation of leaf functional traits within species. *Oecologia*, **180**, 951–959, 2016. [41](#), [85](#)
- [115] Blackhall, M., Raffaele, E., Veblen, T. Is foliar flammability of woody species related to time since fire and herbivory in northwest patagonia, argentina? *Journal of Vegetation Science*, **23**, 931–941, 2012. [41](#), [86](#)
- [116] Nacif, M. E., Quintero, C., Garibaldi, L. A. Intermediate harvesting intensities enhance native tree performance of contrasting species while conserving herbivore diversity in a patagonian woodland. *Forest Ecology and Management*, **483**, 118719, 2021. URL <https://www.sciencedirect.com/science/article/pii/S0378112720314882>. [41](#)
- [117] Jordán, F., Scheuring, I. Searching for keystones in ecological networks. *Oikos*, **99**, 118719, 2002. [42](#)
- [118] Delmas, E., Besson, M., Brice, M., Burkle, L., Dalla Riva, G., Fortin, M., *et al.* Analysing ecological networks of species interactions. *Biol Rev.*, **94**, 16–36, 2019. [43](#)
- [119] Morrison, B. M. L., Brosi, B. J., Dirzo, R. Agricultural intensification drives changes in hybrid network robustness by modifying network structure. *Ecol. Lett.*, **23**, 359–369, 2020. [43](#)
- [120] Bascompte, J., Jordano, P., Melián, C., Olesen, J. The nested assembly of plant-animal mutualistic networks. *Proc. Natl. Acad. Sci. USA* **100**, **16**, 9383–9387, 2003. [43](#)
- [121] Parrot, L. Measuring ecological complexity. *Ecological Indicators*, **10** (6), 1069–1076, 2010. [44](#)
- [122] Landi, P., Minoarivelo, H., Brännström, A., Hui, C., Dieckmann, U. Complexity and stability of ecological networks: a review of the theory. *Population Ecology*, **60** (4), 319–345, 2018. [44](#)
- [123] May, R. Will a large complex system be stable? *Nature*, **238**, 413–414, 1972. [44](#)

- [124] May, R. *Stability and Complexity in Model Ecosystems*. Princeton University Press, New York., 1973. [44](#)
- [125] Montoya, J., Pimm, S., Solé, R. Ecological networks and their fragility. *Nature*, **442**, 259–264, 2006. [44](#)
- [126] Atmanspacher, H. A semiotic approach to complex systems. In: Mehler, A., Köhler, R. (Eds.), *Aspects of Automatic Text Analysis*. Springer, Berlin., 2007. [44](#)
- [127] Fontaine, C., Paulo, R., Guimarães, J., Kéfi, S., Loeuille, N., Memmott, J., *et al.* The ecological and evolutionary implications of merging different types of networks. *Ecology letters*, **14**, 1170–1181, 2011. [44](#)
- [128] Melián, C., Bascompte, J., Jordano, P., Krivan, V. Diversity in a complex ecological network with two interaction types. *Oikos*, **118** (1), 122–130, 2009.
- [129] Pilosof, S., Porter, M., Pascual, M., Kéfi, S. The multilayer nature of ecological networks. *Nat Ecol Evol*, **1**, 0101, 2017.
- [130] Huaylla, C., Nacif, M., Coulin, C., Kuperman, M., Garibaldi, L. Decoding information in multilayer ecological networks: The keystone species case. *Ecological Modelling*, **460**, 109734, 2021. [44](#), [52](#), [90](#)
- [131] Strydom, T., Dalla Riva, G. V., Poisot, T. Svd entropy reveals the high complexity of ecological networks. *Front. Ecol. Evol.*, **9**, 623141, 2021. [44](#), [47](#), [50](#)
- [132] Morzy, M., Kajdanowicz, T., Kazienko, P. On measuring the complexity of networks: Kolmogorov complexity versus entropy. *Complexity*, **2017**, 2017. [46](#), [50](#), [51](#)
- [133] Schmidhuber, J. Discovering neural nets with low kolmogorov complexity and high generalization capability. *Neural Networks*, **10** (5), 857–873, 1997. [46](#)
- [134] Faloutsos, C., Megalooikonomou, V. On data mining, compression, and kolmogorov complexity. *Data Mining and Knowledge Discovery*, **15** (1), 3–20, 2007. [46](#)
- [135] Li, M., Vitányi, P. *Kolmogorov complexity and its applications*. Springer New York, NY, 2014. [46](#)
- [136] Poisot, T., Stouffer, D. B. Optimization of bipartite modularity using lp-brim (label propagation followed by bipartite recursively induced modularity), 2021. R package version 1.0.0. [48](#)

- [137] Liu, X., Murata, T. An efficient algorithm for optimizing bipartite modularity in bipartite networks. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, **14**, 408–415, 2010. [49](#)
- [138] Shannon, C. E. A mathematical theory of communication. *Bell System Technical Journal*, **27**, 379–423, 1948. [50](#), [73](#)
- [139] Gu, R., Shao, Y. How long the singular value decomposed entropy predicts the stock market? — evidence from the dow jones industrial average index. *Physica A: Statistical Mechanics and its Applications*, **453**, 150–161, 2016. URL <https://www.sciencedirect.com/science/article/pii/S0378437116001965>. [50](#)
- [140] Turing, A. On computable numbers, with an application to the entscheidungsproblem. *Proceedings of the London Mathematical Society, Series 2*, **42**, 230–265, 1936. [51](#)
- [141] Soler-Toscano, F., Zenil, H., Delahaye, J., Gauvrit, N. a. Calculating kolmogorov complexity from the output frequency distributions of small turing machines. *PLOS ONE*, **9** (5), 2014. [51](#)
- [142] Zenil, H., Hernández-Orozco, S., Kiani, N., Soler-Toscano, F., Rueda-Toicen, A. A decomposition method for global evaluation of shannon entropy and local estimations of algorithmic complexity. *Entropy*, **20** (8), 605, 2016. [51](#)
- [143] Zenil, H., Soler-Toscano, F., Kiani, N. A., Hernández-Orozco, S. A decomposition method for global evaluation of shannon entropy and local estimations of algorithmic complexity. *Entropy*, **20** (3), 2018. [52](#)
- [144] Kuperman, M., Abramson, G. Small world effect in an epidemiological model. *Phys. Rev. Lett.*, **86**, 2909–2912, Mar 2001. [52](#)
- [145] Tan, F., Xia, Y., Zhu, B. Link prediction in complex networks: A mutual information perspective. *PLoS ONE*, **9** (9), e107056, 2014. [66](#), [89](#)
- [146] Li, J., Peng, J., Liu, S., Ji, X., Li, X., Hu, X. Link prediction in directed networks utilizing the role of reciprocal links. *IEEE Access*, **8**, 28668–28680, 2020. [66](#)
- [147] Klamt, S., Haus, U., Theis, F. Hypergraphs and cellular networks. *PLoS Computational Biology*, **5**, e1000385, 2009. [67](#)
- [148] Sonntag, M., Teichert, H.-M. Competition hypergraphs. *Discrete Applied Mathematics*, **143** (1), 324–329, 2004. URL <https://www.sciencedirect.com/science/article/pii/S0166218X04000976>. [68](#)

- [149] Battiston, F., Cencetti, G., Iacopini, I., Latora, V., Lucas, M., Patania, A., *et al.* Networks beyond pairwise interactions: Structure and dynamics. *Physics Reports*, **874**, 1–92, 2020. URL <https://www.sciencedirect.com/science/article/pii/S0370157320302489>. 68
- [150] Grilli, J., Barabás, G., Michalska-Smith, M., Allesina, S. Higher-order interactions stabilize dynamics in competitive network models. *Nature*, **548**, 210–213, 2017. 68
- [151] Tuğal, I., Pala, Z. Centrality with entropy in hypergraphs. *Rahva Teknik ve Sosyal Araştırmalar Dergisi*, **1**, 84–91, 2021. 68
- [152] Chen, C., Rajapakse, I. Tensor entropy for uniform hypergraphs. *IEEE Transactions on Network Science and Engineering*, **7**, 2889–2900, 2020. 68
- [153] Iacobucci, D., Wasserman, S. Social networks with two sets of actors. *Psychometrika*, **55**, 707–720, 1990. 68
- [154] Frank, K. A. Mapping interactions within and between cohesive subgroups. *Social Networks*, **18** (2), 93–119, 1996. URL <https://www.sciencedirect.com/science/article/pii/037887339500257X>.
- [155] Watts, D. J., Dodds, P. S., Newman, M. E. J. Identity and search in social networks. *Science*, **296** (5571), 1302–1305, 2002. URL <https://www.science.org/doi/abs/10.1126/science.1070120>.
- [156] Newman, M. E. J. Properties of highly clustered networks. *Phys. Rev. E*, **68**, 026121, Aug 2003. URL <https://link.aps.org/doi/10.1103/PhysRevE.68.026121>. 68
- [157] Agarwal, S., Branson, K., Belongie, S. Higher order learning with graphs. ICML '06. New York, NY, USA: Association for Computing Machinery, 2006. URL <https://doi.org/10.1145/1143844.1143847>. 69
- [158] Aksoy, S., Joslyn, C., Ortiz Marrero, C., Praggastis, B., Purvine, E. Hypernetwork science via high-order hypergraph walks. *EPJ Data Science*, **19** (9), 2020. URL <https://doi.org/10.1140/epjds/s13688-020-00231-0>.
- [159] Lotito, Q., Musciotto, F., Montresor, A. Higher-order motif analysis in hypergraphs. *Communications Physics*, **79** (5), 2022. URL <https://doi.org/10.1038/s42005-022-00858-7>. 69
- [160] Young, J., Petri, G., Peixoto, T. Hypergraph reconstruction from network data. *Communications Physics volume*, **135**, 135, 2021. 69, 74, 88

- [161] Estrada, E., Rodríguez-Velázquez, J. A. Subgraph centrality and clustering in complex hyper-networks. *Physica A: Statistical Mechanics and its Applications*, **364**, 581–594, 2006. URL <https://www.sciencedirect.com/science/article/pii/S0378437105012550>. 70, 75, 88
- [162] Rand, W. Objective criteria for the evaluation of clustering methods. *Pattern Recognition Letters*, **66** (336), 846–850, 1971. 72
- [163] Everitt, B., Landau, S., Leese, M. Cluster Analysis. Arnold, 2001. 72
- [164] Campello, R. A fuzzy extension of the rand index and other related indexes for clustering and classification assessment. *Pattern Recognition Letters*, **28** (7), 833–841, 2007. 72, 88
- [165] Bromiley, P., Thacker, N., Bouhova-Thacker, E. Shannon entropy, renyi entropy, and information. En: Shannon Entropy, Renyi Entropy, and Information. 2004. URL <https://api.semanticscholar.org/CorpusID:17565994>. 73
- [166] Bairey, E., Kelsic, E., Kishony, R. High-order species interactions shape ecosystem diversity. *Nature Communications*, **7**, 12285, 2016. 79
- [167] Chaudhary, A., Burivalova, Z., Koh, L., Hellweg, S. Impact of forest management on species richness: Global meta-analysis and economic trade-offs. *Scientific Reports*, **6** (23954), 2016. 83
- [168] Martínez Pastur, G., Soler, R., Lencinas, M. V., Cellini, J. M., Peri, P. L. Long-term monitoring of thinning for silvopastoral purposes in nothofagus antarctica forests of tierra del fuego, argentina. *Forest Systems*, **27** (1), e01S, May 2018. URL <https://fs.revistas.csic.es/index.php/fs/article/view/11928>.
- [169] Grosfeld, J., Chauchard, L., Gowda, J. H. Debates: ¿podemos manejar sustentablemente el bosque nativo de patagonia norte? *Ecología Austral*, **29** (1), 156–163, 2019. URL [https://ojs.ecologiaaustral.com.ar/index.php/Ecologia\\_Austral/article/view/775](https://ojs.ecologiaaustral.com.ar/index.php/Ecologia_Austral/article/view/775). 83
- [170] Dunne, J. A., Williams, R. J., Martinez, N. D. Food-web structure and network theory: The role of connectance and size. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, **99** (20), 12917–12922, 2002. 83
- [171] Stouffer, D. B., Bascompte, J. Compartmentalization increases food-web persistence. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)*, **108** (9), 3648–3652, 2011. 83

- [172] Nacif, M. E. Influencia de la intensidad de manejo sobre las interacciones planta-herbívoro y el desempeño de plantaciones multiespecíficas en un matorral norpatagónico. Tesis Doctoral, Universidad Nacional del Comahue, Centro Regional Universitario Bariloche, 2021. [84](#)
- [173] Kessler, A., Baldwin, I. Plant responses to insect herbivory: the emerging molecular analysis. *Annual Review of Plant Biology*, **53**, 299–328, 2002. URL <https://doi.org/10.1146/annurev.arplant.53.100301.135207>. [84](#)
- [174] Vázquez, D. P., Aizen, M. A. Null model analyses of specialization in plant-pollinator interactions. *Ecology*, **84** (9), 2493–2501, 2003. URL <http://www.jstor.org/stable/3450152>.
- [175] Bascompte, J., Jordano, P., Olesen, J. M. Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science*, **312** (5772), 431–433, 2006. URL <https://www.science.org/doi/abs/10.1126/science.1123412>. [84](#)
- [176] Liben-Nowell, D., Kleinberg, J. The link-prediction problem for social networks. *Journal of the American Society for Information Science and Technology*, **58**, 1019–1031, 2007. [89](#)
- [177] Guimerà, R., Sales-Pardo, M. Missing and spurious interactions and the reconstruction of complex networks. *Proceedings of the National Academy of Sciences*, **106** (52), 22073–22078, 2009. URL <https://www.pnas.org/doi/abs/10.1073/pnas.0908366106>. [89](#)
- [178] Wang, P., Xu, B., Wu, Y., Zhou, X. Link prediction in social networks: the state-of-the-art. *SCIENCE CHINA Information Sciences*, **58**, 2015. [89](#)
- [179] Zeng, Z., Chen, K.-J., Zhang, S., Zhang, H. A link prediction approach using semi-supervised learning in dynamic networks. En: 2013 Sixth International Conference on Advanced Computational Intelligence (ICACI), págs. 276–280. 2013. [89](#)

# Publicaciones asociadas

Los estudios desarrollados en los capítulos 2 y 3 de esta tesis se encuentran publicados en revistas indexadas.

1. Huaylla, C., Nacif, M., Coulin, C., Kuperman, M., Garibaldi, L. **Decoding information in multilayer ecological networks: The keystone species case.** Ecological Modelling, 460, 109734, 2021. <https://doi.org/10.1016/j.ecolmodel.2021.109734>
2. Huaylla, C. A., Kuperman, M.N., Garibaldi, L.A. **Comparison of two statistical measures of complexity applied to ecological bipartite networks,** Physica A: Statistical Mechanics and its Applications, 642, 129764, 2024. <https://doi.org/10.1016/j.physa.2024.129764>

