

La estabilidad como mecanismo de selección natural sobre redes de interacción en sistemas biológicos

Stability as a natural selection mechanism on interaction
networks in biological systems

por Juan Ignacio Perotti

Presentado ante la Facultad de Matemática, Astronomía y Física
como parte de los requerimientos para la obtención del grado de
Doctor en Física de la

UNIVERSIDAD NACIONAL DE CORDOBA

Marzo de 2011

©FaMAF-UNC 2011

Director: Sergio Alejandro Cannas

a mis padres y hermana

Agradecimientos

A mi familia, por su apoyo incondicional.

A mi gran amor, Victoria.

A Sergio Cannas por todo el apoyo que me brindó en estos últimos años y su natural talento para convertir ruido en física.

A las personas y excelentes investigadores que colaboraron en forma directa con este trabajo: Orlando Billoni, Francisco Tamarit, Dante Chialvo.

A Pablo Gleiser y su invaluable predisposición a discutir ideas.

A todos los amigos y colegas que me acompañaron en esta etapa: Santiago P., Alejandro, Rodrigo, Santiago G., Ezequiel, Tito, Emilce, Gonzalo, Germán, Caro, Elena, Eugenia, Alfredo, Miguel, Nico, Roberto, Carlitos, Pablo y Noelia.

A María E. Boixadós por su invaluable ayuda.

A los compañeros de almuerzo.

Al grupo de Teatro.

Al grupo de Teoría de la Materia Condensada.

A las instituciones que me permitieron llevar a cabo este proyecto: al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) y a la Facultad de Matemática, Astronomía y Física (FaMAF-UNC).

Resumen

En la naturaleza abundan sistemas que exhiben una estructura o un comportamiento que no puede entenderse ni como un patrón ordenado, ni como un fenómeno altamente aleatorio que da lugar a explicaciones estadísticamente simples. A dichos sistemas se los suele denominar complejos, y al menos en la mayoría de los casos se los puede pensar como un conjunto de muchos componentes de cuya interacción emerge la complejidad. Las interacciones en un sistema de muchos componentes puede caracterizarse por el concepto de red y es de esperar que las redes que representan a los sistemas complejos puedan heredar tal propiedad. De hecho la mayoría de las redes en la naturaleza exhiben una topología compleja, y aunque se han realizado importantes avances, los orígenes de dicha complejidad no están del todo claros. Existe la posibilidad de que no haya una forma unificada de explicar las distintas formas en que las redes complejas emergen en la naturaleza *. Sin embargo se han podido descubrir un conjunto de principios que combinados de distintas formas proveen de explicaciones para muchas de las propiedades observadas en redes complejas de origen natural. Muchos de estos principios adoptan la forma de mecanismos a partir de los cuales se generan redes cuyas propiedades son luego contrastadas con las propiedades de las redes observadas en la naturaleza. En este trabajo de investigación se propone un mecanismo evolutivo de carácter muy general que resulta en una posible explicación a la ubicuidad de ciertas propiedades de las redes de interacción en una gran variedad de sistemas biológicos. Este mecanismo está basado en la estabilidad de la dinámica del sistema que subyace a la red. Dicho mecanismo es incorporado en un modelo de redes de agentes interactuantes, cuyas topologías evolucionan en el tiempo por medio de relacionar la pertenencia de cada agente a la red con el efecto que tienen los mismos sobre la estabilidad del sistema. Del estudio del modelo se concluye que dicho mecanismo da lugar a la emergencia de redes de topologías complejas, constituyendo así una posible explicación a la ubicuidad observada respecto de ciertas topologías complejas en las redes biológicas [165, 164]. En particular, permite encontrar una relación entre redes biológicas de sistemas microscópicos tales como las redes de proteínas o las redes metabólicas, hasta redes de sistemas macroscópicos como las redes ecológicas, cubriendo así un amplio rango de escalas.

PACS: 89.75Hc, 89.75Da, 89.75Fb

Palabras Clave: Redes complejas, Estabilidad en sistemas dinámicos, Simulaciones numéricas.

*Incluyo en naturaleza las construidas por el hombre, es decir, las comúnmente llamadas artificiales.

Abstract

Systems abound in nature that exhibit a structure or behavior that cannot be understood as a regular pattern, or else as a highly random phenomena leading to a simple statistical explanation. Such systems are often called complex, and at least in most cases they can be conceived of as a collection of many components from whose interactions complexity emerges. The interactions arising in a system of many components may be characterized by the concept of network, and it is hoped that networks associated to complex systems could inherit such property. In fact most networks in nature exhibit complex topology and although significant progress has been made the origins of such complexity are not entirely clear. The possibility exists that there is no unified way of explaining the emergence of complexity in natural networks. *. However, a set of principles have been discovered such that if combined in different ways provide explanations for many of the observed properties of natural complex networks. Many of these principles take the form of mechanisms through which new networks are generated. The properties of these networks are later on compared with the properties of networks observed in nature. In this research study a general evolutionary mechanism is proposed as a possible explanation for the ubiquity of certain properties of interaction networks in biological systems. This mechanism, which is based on the system dynamic stability underlying the network, is incorporated into a model of a growing network of interacting agents. In this model each new agent's membership in the network is determined by the agent's effect on the network's global stability. From the study of this model it has been found that this mechanism leads to the emergence of complex network topologies, offering a possible explanation of the observed ubiquity of certain complex topologies in biological networks [165, 164]. In particular, it establishes a relationship between biological systems, from microscopic scales such as proteins or metabolic networks to macroscopic scales such as ecological networks. Therefore, the mechanism covers a wide range of scales.

PACS: 89.75Hc, 89.75Da, 89.75Fb

Keywords: Complex networks, Dynamical system stability, Numerical simulations.

*Here I also consider networks that emerge from human activity.

Índice general

1. Introducción	19
2. Marco teórico	23
2.1. La distribución de grados de los nodos de una red	23
2.1.1. Las redes aleatorias: El modelo nulo	24
2.1.2. La distribución de grados de las redes empíricas	25
2.1.3. El modelo de Barabási y Albert: El mecanismo de crecimiento y conexión preferencial	28
2.2. Las distancias topológicas entre nodos de la red	36
2.2.1. La distancia topológica de un grafo aleatorio	37
2.3. El concepto de clusterización en redes	38
2.3.1. La clusterización de las redes aleatorias	39
2.3.2. La clusterización en redes reales	39
2.3.3. La clusterización en el modelo de Barabási–Albert	40
2.4. El concepto de Mundo Pequeño	41
2.4.1. El modelo de Watts–Strogatz	42
2.4.2. La distancia topológica en redes reales	47
2.5. La clusterización en función del grado de los nodos	48
2.5.1. La organización jerárquica de Ravasz <i>et. al.</i>	49
2.6. La correlación entre los grados de nodos vecinos: Asortatividad	51
2.6.1. La asortatividad/disasortatividad en redes reales	54
2.6.2. Algunas consideraciones teóricas sobre la asortatividad	57
2.7. Resistencia ante fallas aleatorias o ataques dirigidos en una red	58
2.8. La densidad de conexiones en una red: “Conectividad”	60
2.8.1. La conectividad y otras propiedades en redes ecológicas	63
2.8.2. Mundo pequeño y clusterización en redes ecológicas	66
2.8.3. La distribución de grados en las redes ecológicas	67
2.8.4. El ataque dirigido y las fallas aleatorias en redes ecológicas	68
2.9. La estabilidad en las redes de interacción: El modelo de R. May	68

2.9.1. Analizando el concepto de estabilidad utilizado por May: Extensiones y alternativas	70
2.9.2. Espectro de autovalores en redes complejas	72
3. Un mecanismo de presión selectiva basado en la estabilidad	75
3.1. El modelo	79
4. Resultados	83
4.0.1. Detalles técnicos	83
4.0.2. Las distintas variaciones del modelo.	84
4.1. Propiedades topológicas	85
4.1.1. Conectividad	85
4.1.2. La distribución de grados	89
4.1.3. Propiedades de mundo pequeño	100
4.1.4. Organización jerárquica	107
4.1.5. Correlación en el grado de los nodos vecinos	107
4.1.6. Tolerancia a las fallas topológicas	112
4.1.7. Límites al tamaño máximo	113
4.2. Propiedades dinámicas	117
4.2.1. Interacciones anticorrelacionadas	119
4.2.2. Forzado de la anticorrelación	122
5. Conclusiones	127
5.1. Posibles extensiones al presente trabajo	129
A. Tolerancia a ataques dirigidos y fallas aleatorias en redes tróficas	137
A.1. Modularidad jerárquica en redes tróficas	139
B. Sobre los datos obtenidos de la <i>home page</i> de A.G. Rossberg	147

Índice de figuras

2.1. Distribuciones de grados de los modelos de Amaral et. al. [17]	36
2.2. Comparación de $Cc^{(WS)}/\langle k \rangle$ entre redes empíricas y redes aleatorias	41
2.3. Redes <i>small-world</i> de Watts–Strogatz [204]	44
2.4. $Cc(p)$ y $L(p)$ para el modelo de Watts–Strogatz	44
2.5. $P(k)$ del modelo de Watts–Strogatz	46
2.6. Comparación de L entre redes empíricas y redes aleatorias	49
2.7. Red con organización jerárquica	51
2.8. $Cc(k)$ para redes de proteínas y metabólicas	52
2.9. $k_{nn}(k)$ para una red de interacciones proteína–proteína	56
2.10. Asortatividad local para la red de proteínas del <i>H. Sapiens</i>	56
2.11. Ataques dirigidos y fallas aleatorias	61
2.12. Los distintos tipos de interacción posibles	64
2.13. $P(\lambda)$ en matrices de May sobre topologías libres de escala	74
3.1. Representación esquemática de la dinámica del sistema bajo presión selectiva	79
4.1. $C(n)$ en el presente modelo para diferentes valores de b	86
4.2. $C(n)$ en el presente modelo para diferentes valores de q	87
4.3. Comparación de $C(n)$ entre el presente modelo y el de May[125]	88
4.4. $C(n)$ correspondiente a diferentes redes empíricas	89
4.5. Escaleo de $P(k)$ con n	91
4.6. Dependencia de $P(k)$ con Δ	93
4.7. Dependencia de $P(k)$ con R	94
4.8. Dependencia de $n_s(n)$ con R	95
4.9. Dependencia de $n_s(n)$ con b	96
4.10. Dependencia de $n_s(n)$ con q	96
4.11. $P(k_{in})$ y $P(k_{out})$	97
4.12. Comportamiento de $P(k)$ ante cooperación/competición	98
4.13. El conexionado preferencial en el presente modelo	99
4.14. Comparación de L entre el presente modelo, redes aleatorias, y redes empíricas	100
4.15. Dependencia de $Cc^{(WS)}(n)$ con b	103

4.16. Dependencia de $Cc^{(Tr)}(n)$ con b	103
4.17. Dependencia de $Cc^{(WS)}(n)$ con q	103
4.18. Dependencia de $Cc^{(Tr)}$ con q	103
4.19. Comparación de $Cc^{(WS)}(n)$ entre redes del presente modelo, aleatorias, y empíricas	104
4.20. Comparación de $Cc^{(Tr)}(n)$ entre redes del presente modelo, aleatorias, y empíricas	104
4.21. Comparación de $Cc^{(WS)}(n)/\langle k \rangle(n)$ entre redes del presente modelo, aleatorias, y empíricas	104
4.22. Comparación de $Cc^{(Tr)}(n)/\langle k \rangle(n)$ entre redes del presente modelo, aleatorias, y empíricas	104
4.23. Comparación de $\langle k \rangle$ del presente modelo con redes empíricas	106
4.24. Evolución temporal del tamaño de las redes	106
4.25. Evolución temporal de $Cc^{(WS)}$ y L	106
4.26. $Cc(k)$ del presente modelo	108
4.27. Dependencia de $Cc(k)$ con b	108
4.28. Dependencia de $Cc(k)$ con q	108
4.29. Grado medio $k_{nn}(k)$ de los nodos vecinos a nodos de grado k	109
4.30. Dependencia de $k_{nn}(k)$ con q	111
4.31. Comparación de $\langle r \rangle(n)$ entre redes del presente modelo, aleatorias y empíricas	111
4.32. Tolerancia a fallas y ataques dirigidos	113
4.33. Grado total promedio de los nodos ingresantes a la red $k_n^{(t)}$ en función del tamaño n	116
4.34. Promedio de la máxima parte real λ_{max} de entre todos los autovalores en función de n	117
4.35. Distribución de las partes reales de los autovalores $P(\lambda)$.	118
4.36. Densidad de probabilidades $P(a_{i,j})$ de las entradas de las matrices aceptadas por la selección	119
4.37. Correlación de las magnitudes de las interacciones dobles $\langle a_{i,j}a_{j,i} \rangle$ en función de n	122
4.38. Correlación de las magnitudes de las interacciones dobles $\langle a_{i,j}a_{j,i} \rangle$ en función de k	123
4.39. Ejemplo de la estructura de conexiones dirigidas de la red	124
4.40. Conectividad $C(n)$ para distintos grados de forzamiento de la anticorrelación h	125
4.41. Distribución de grados $P(k)$ para distintos grados de forzamiento de la anticorrelación h	125
A.1. $S(f)$ y $\langle s \rangle(f)$ para la red trófica <i>Ythan Estuary (96)</i> . .	140
A.2. $S(f)$ y $\langle s \rangle(f)$ en redes tróficas. <i>Benguela Current, Bridge Brook Lake, British Grassland y Canton Creek</i>	141
A.3. $S(f)$ y $\langle s \rangle(f)$ en redes tróficas. <i>Caribbean Reef, Chesapeake, Coachella Valley y El Verde Rainforest</i>	142
A.4. $S(f)$ y $\langle s \rangle(f)$ en redes tróficas. <i>Little Rock Lake, Northeast US Shelf, Scotch Broom y Skipwith Pond</i>	143

A.5. $S(f)$ y $\langle s \rangle(f)$ en redes tróficas. <i>St. Marks Seagrass, St. Martin Island, Stony Stream e Ythan Estuary (91)</i>	144
A.6. $P_{\geq}(k)$ en redes tróficas. <i>Northeast US Shelf, Caribbean Reef, Benguela Current y British Grassland</i>	145
A.7. $P_{\geq}(k)$ para la red trófica <i>El Verde Rainforest</i>	145
A.8. $Cc(k)$ en redes tróficas	146

Índice de cuadros

2.1. Redes empíricas libres de escala	27
2.2. Comparación de L entre redes empíricas y redes aleatorias	48
2.3. Asortatividad de los grados en redes reales	58
4.1. Asortatividad de las redes tróficas	112
4.2. Tamaños máximos	116
4.3. Valores de ϵ y γ ajustados en las figuras (4.40) y (4.41) .	126
A.1. Robustéz de las redes tróficas ante ataques o fallas	139

Capítulo 1

Introducción

Existen en la naturaleza innumerables sistemas compuestos por unidades simples, fáciles de describir y predecibles, las cuales sin embargo dan lugar a comportamientos colectivos altamente complejos y difíciles de predecir. El origen de esta *complejidad emergente* puede ser buscado en los intrincados mecanismos de interacción entre los constituyentes. Tal es el caso de sistemas que nos son tan familiares como un vidrio de botella, o tan ubicuos como las proteínas de las que estamos hechos o, en el extremo de la complejidad, tan enigmáticos como el cerebro responsable de nuestros estados mentales. Un subconjunto muy especial de estos *Sistemas Complejos* deben la no trivialidad de su comportamiento a la *topología* de la red de interacciones entre sus elementos, y han dado lugar en estas últimas décadas a una nueva y promisoría disciplina de estudio dentro de la Física Estadística denominada *Redes Complejas* (algunos *Reviews* sobre distintos tópicos en el área: [10], [146],[58],[59], [57],[34],[54]). Sin duda lo que más ha llamado la atención a partir de los resultados de estos primeros años de estudio, es la aparente *ubiquidad* de algunas propiedades de estas redes. Fenómenos tan disímiles como las redes metabólicas que regulan el metabolismo de los seres vivos [97, 200], Internet [71, 163], la *World-Wide-Web* [7, 11], diferentes relaciones sociales entre individuos [23, 17, 117] o incluso redes derivadas a partir de estudios de resonancia magnética en la actividad cerebral [68], parecen compartir ciertos elementos topológicos básicos que dan lugar a comportamientos muy similares desde el punto de vista cualitativo *. Incluso la definición de una red de interacciones para un dado sistema frecuentemente no es única (ver por ejemplo los siguientes casos en donde distintos tipos de redes pueden definirse a partir de proteínas [16, 209, 53]), la misma depende del nivel de *coarse graining* que se utiliza. Sin embargo, varias propiedades topológicas resultan ser independientes de la definición precisa de la red.

Todo parece indicar que cuando los constituyentes de un sistema evolucionan en base a *procesos adaptativos*, dan lugar a un proceso *auto*

*La ubiquidad en los ejemplos aquí citados, corresponde a que todos presentan una distribución de grados libre de escala a diferencia de las Poissonianas correspondientes a las redes aleatorias.

organizado que permite construir redes topológicamente complejas en el sentido de que no presentan ni una organización regular, ni una completamente aleatoria. Este trabajo se centrará en el estudio de un nuevo concepto propuesto, un mecanismo evolutivo de generación de redes basado en la estabilidad de los sistemas. La idea es que muchas redes corresponden a sistemas abiertos los cuales están sujetos a un flujo constante de nuevos componentes (inmigración, mutación, etc.). Supondremos que en ausencia de dicho flujo los sistemas evolucionan a una dinámica de equilibrio estable. Sin embargo, la incorporación de nuevas componentes eventualmente desestabilizan los sistemas obligándolos a reestructurarse hasta alcanzar un nuevo equilibrio. Comprobaremos a lo largo de la tesis que dicho proceso de continua reestructuración sujeto al flujo inmigratorio provoca que los sistemas crezcan adquiriendo la estructura de redes no triviales o complejas.

En el Capítulo 2 se introducirán los conceptos teóricos y empíricos que serán de utilidad para poner en contexto los resultados obtenidos en esta tesis. Se discutirán los conceptos de redes libres de escala con sus distribuciones de grados tipo leyes de potencia, la extendida aparición de las mismas en la naturaleza y en diferentes modelos que la explican [10]. También se discutirán brevemente excepciones a esta regla [17] y algunas de las consecuencias de la diferencia [25, 12, 51]. Se discutirá el importante concepto de mundo pequeño [184] y su relación con la alta clustervización que presentan las redes reales [204]. En relación a la alta clustervización revisaremos los conceptos en torno a la organización jerárquica [176, 175]. El concepto de correlación asortativa/disasortativa será de utilidad por lo que también será discutido en este capítulo [144, 145, 148]. Además se tratará el concepto de estabilidad de May [125] y su relación con el dilema complejidad–estabilidad en redes tróficas [166, 63, 64]. En el Capítulo 3 se introducirá un modelo a través del cual se estudiará el concepto de un mecanismo evolutivo de generación de redes basado en la estabilidad de los sistemas. Se discutirán en él algunos aspectos técnicos y se intentará justificar a priori las distintas suposiciones/aproximaciones realizadas en el modelo. En el Capítulo 4 se presentarán los resultados obtenidos del estudio y análisis del modelo. Particularmente se estudiarán algunas variaciones del mismo para testear la robustez de las aproximaciones realizadas encontrándose resultados positivos. Se verá que el modelo es capaz de generar redes con distribuciones de grados tipo leyes de potencias, las propiedades de mundo pequeño, correlaciones disasortativas en el grado de los nodos, y organización jerárquica. Se compararán las respuestas ante daños topológicos entre las redes generadas por el modelo y redes tróficas encontrándose una gran similitud. El estudio del comportamiento de la densidad de conexiones en función del tamaño permite aportar conclusiones al dilema de complejidad–estabilidad, sugiriendo que la cuestión trasciende más allá de las redes ecológicas abarcando un conjunto más amplio de sistemas biológicos. Finalmente, en el Capítulo 5 se interpretarán los resultados y presentarán las conclusiones.

Particularmente se hipotetizará que los mecanismos evolutivos en donde actúa una presión selectiva se adaptan mejor a los sistemas biológicos que otros mecanismos propuestos los cuales generalmente se adaptan mejor a las redes sociales (ej. red de contactos sexuales) o tecnológicas (ej. Internet). Como agregado se proponen algunas posibles extensiones al presente trabajo.

Capítulo 2

Marco teórico

El marco teórico será organizado de la siguiente manera. Se irán introduciendo de manera secuencial magnitudes/funciones que son útiles para caracterizar propiedades de las redes. Se mencionará que tipo de estructura intenta *cuantificar* cada una de las definiciones. Simultáneamente se mencionará el estudio de dichas magnitudes sobre redes de origen empírico así como en distintos modelos de red. En dicha circunstancia se aprovechará para identificar las características que distinguen a las distintas clases de redes, y los ingredientes básicos que son introducidos en los distintos modelos que dan lugar a explicaciones de las correspondientes características.

Sobre la nomenclatura: Formalmente un grafo no dirigido G consiste en un par ordenado $\mathcal{G} = (\mathcal{N}, \mathcal{L})$ en donde \mathcal{N} es un conjunto de nodos y $\mathcal{L} \subseteq \{\{a, b\} : a, b \in \mathcal{N}\}$ es un conjunto de pares no ordenados de nodos* que representan links no dirigidos. Un grafo dirigido es un par ordenado $\mathcal{D} = (\mathcal{N}, \mathcal{E})$ en donde $\mathcal{E} = \{(a, b) : a, b \in \mathcal{N}\}$ es un conjunto de pares ordenados que representan los links dirigidos†. Sin embargo aquí no hará falta entrar en una notación matemática tan sofisticada. En general, para identificar los distintos *entes* se utilizarán términos que fácilmente se asocian con sus significados gracias al contexto. Los términos *red* y *grafo* serán utilizados como sinónimos. Lo mismo ocurre con los términos *link*, *conexión* y eventualmente *interacción*. En general los links se asumirán no dirigidos salvo que se especifique explícitamente lo contrario.

2.1. La distribución de grados de los nodos de una red

El grado de un nodo de una red, es el número de conexiones (o links) que alcanzan a dicho nodo. La distribución de grados de una red se define como la probabilidad $P(k)$ de encontrar en dicha red un nodo de grado k . Mas precisamente, para un ensamble de redes:

* \mathcal{L} por link.

† \mathcal{E} por edge.

$$P(k) = \left\langle \frac{1}{n} \sum_i \delta_{k,k_i} \right\rangle \quad (2.1)$$

donde el promedio $\langle \cdot \rangle$ es sobre el ensamble de redes, n es el tamaño de las redes, y para cada una de dichas redes k_i representa el grado del nodo i -ésimo en la misma. Naturalmente, el grado promedio se relaciona a la distribución de grados de acuerdo a $\langle k \rangle = \sum_k kP(k)$.

Una aclaración importante: En general se van a introducir las definiciones (como la de la ecuación (2.1)) considerando ensambles de redes ya que en el estudio de modelos se utiliza frecuentemente. Sin embargo, al estudiar redes reales en general no se dispone de un ensamble. En dichas circunstancias las definiciones siguen siendo aplicables por medio de omitir la promediación.

Observación: Después del tamaño n y el grado promedio $\langle k \rangle$, la distribución de grados $P(k)$ es una de las medidas más simples que se pueden intentar determinar en una red. Sin embargo como se expondrá en las siguientes secciones, la distribución de grados es suficiente para descubrir propiedades no triviales en las redes de orígenes empíricos. De hecho, a través de $P(k)$ se revelarán los primeros signos de complejidad en algunas de dichas redes.

2.1.1. Las redes aleatorias: El modelo nulo

Al estudiar las redes de origen empírico, una de las primeras cosas que resalta es la evidente falta de regularidad que muchas de ellas suelen presentar. Principalmente las de origen social o biológico, pero en algunos casos también se observa irregularidad en redes de origen tecnológico. Resulta natural entonces, intentar comparar las propiedades de las redes empíricas “irregulares” con la de redes aleatorias. La evidente falta de correlaciones en una red aleatoria, la convierte en el modelo ideal de red sin estructura. Es por ello que las redes aleatorias se consideran la *hipótesis nula* o *modelo nulo* por defecto. En general, para tratar de identificar la estructura presente en un ensamble de redes de aproximadamente n nodos y M links, en un primer paso se las compara con redes aleatorias de igual cantidad de nodos y links. ¿Pero concretamente que se denomina red aleatoria? Originalmente Erdős y Rényi [70] definieron un *grafo aleatorio* como un conjunto de n nodos y M links distribuidos al azar entre los $n(n-1)/2$ pares posibles de nodos. En total hay:

$$\binom{n(n-1)/2}{M} \quad (2.2)$$

tales grafos que forman un espacio de probabilidades en donde todas las realizaciones son equiprobables [10]. Una definición alternativa y generalmente equivalente a la del grafo aleatorio* es denominada *grafo binomial*.

*Cuando se construye numéricamente un grafo aleatorio es necesario evitar las

En un grafo binomial se dispone de n nodos de modo que cada par de ellos está conectado con probabilidad p . En otras palabras, en un grafo binomial el número de links M es una variable aleatoria de media $pn(n-1)/2$ *. En este trabajo se utiliza el término red aleatoria para referirse indistintamente a un grafo aleatorio o a un grafo binomial, los cálculos numéricos han sido realizados utilizando la definición de grafo aleatorio, y los cálculos analíticos la definición de grafo binomial.

Para una red aleatoria, la distribución de grados es representada con buena aproximación por una distribución binomial:

$$P(k) \simeq \binom{n-1}{k} p^k (1-p)^{n-1-k} \quad (2.3)$$

donde $k \in \{1, \dots, n-1\}$. A su vez para valores grandes de n ésta expresión tiende a una distribución de Poisson [10]:

$$P(k) \simeq e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!}, \quad \langle k \rangle = pn \quad (2.4)$$

siempre y cuando $\langle k \rangle$ tienda a una constante con el aumento de n .

Observación: Los links son distribuidos al azar en una red aleatoria. Esto en principio podría dar lugar a nodos con un número arbitrario de vecinos entre 1 y $n-1$. Sin embargo típicamente un grafo aleatorio es bastante homogéneo. En vistas de la ecuación (2.4), la mayoría de sus nodos poseen una cantidad de conexiones que varía muy poco entorno a la media $\langle k \rangle$.

2.1.2. La distribución de grados de las redes empíricas

A pesar de la aparente falta de regularidad en las redes empíricas, un simple vistazo a algunas de ellas revela que las mismas poseen algunos nodos en donde $k \gg \langle k \rangle$. Dichos nodos con muchas conexiones se los denomina *hubs*. Si la organización de las redes empíricas tuviese un origen puramente aleatorio, la aparición de hubs sería algo altamente improbable en vistas de la ecuación (2.4). Esta discrepancia motivó la profundización del estudio de las propiedades de las redes empíricas.

Consistentemente con la observación de hubs, mucha de las redes empíricas exhiben una distribución de grados cuya cola es adecuadamente ajustada por una ley de potencias:

$$P(k) \sim k^{-\gamma} \quad , \text{para } k \gg 1 \quad (2.5)$$

conexiones múltiples, es decir, que exista más de un link que conecte a un dado par de nodos. Evitar las conexiones múltiples puede introducir pequeñas correlaciones.

*Las conexiones múltiples en un grafo binomial son inexistentes por definición, por lo que se evitan a priori las correlaciones que pueden aparecer en la construcción de los grafos aleatorios. El costo que hay que pagar por ello es que M se vuelve una variable aleatoria.

con γ típicamente entre 1 y 3 [10, 58, 146, 68]. Si el comportamiento tipo ley de potencias se cumple (al menos aproximadamente) para todo el rango de k , se suele decir que la correspondiente red es *libre de escala*. Algunas redes son demasiado pequeñas para tener una distribución de grados bien definida [189], sin embargo claramente exhiben hubs indicando una topología distinta a la de una red aleatoria. Hay otras formas en que la distribución de grado de las redes empíricas difieren de una distribución de Poisson. Tal es el caso de aquellas que exhiben una distribución tipo ley de potencias truncada[53, 132]:

$$P(k) \sim k^{-\gamma} f(k/\kappa) \quad , \text{ donde } f(x) = \begin{cases} \text{const.} & \text{si } x \ll 1 \\ \rightarrow 0 & \text{si } x \gtrsim 1 \end{cases} \quad (2.6)$$

En un ejemplo típico $f(x)$ decae exponencialmente o como una gaussiana. En la ecuación (2.6) el comportamiento tipo ley de potencias predomina mientras $k \ll \kappa$, pero cuando $k \gtrsim \kappa$ predomina el decaimiento rápido de $f(x)$ provocando una caída abrupta de la curva. Es por ello que κ da una idea del máximo grado que los nodos de la red en cuestión pueden presentar. De todas maneras, este tipo de comportamientos son consistentes con la aparición de hubs. También es importante remarcar que hay redes empíricas que son excepciones a esta regla (no poseen hubs). Tal es el caso de la red de distribución de energía eléctrica del sur de California de los EE.UU., o una red de amigos entre estudiantes en una preparatoria* también en EE.UU. En dichos casos las distribuciones de grados decaen como una exponencial o como una Gaussiana respectivamente[17]. En estos casos podría ocurrir que κ es demasiado pequeño apantallando una posible ley de potencias.

En este trabajo nos interesa concentrarnos en los casos en que la distribución de grados presenta una gran variabilidad. En otras palabras, una *distribución de grados de cola larga* como en los casos de las ecuaciones (2.5) y (2.6) para $\kappa \gg 1$. La tabla (2.1) muestra un listado de redes de diferentes tamaños que exhiben una distribución de grados de cola larga. Es evidente que las redes que difieren de las redes aleatorias respecto de la distribución de grados cubren un amplio espectro de la naturaleza y la tecnología.

Observaciones: La evidente presencia de hubs en muchas de las redes empíricas, implica que las redes aleatorias no son un modelo adecuado. En otras palabras, a pesar de la irregularidad que presentan las redes reales, poseen algún tipo de organización que las diferencian de la pura aleatoriedad. En particular distribuciones tipo leyes de potencia (o de cola larga) están asociadas a grandes fluctuaciones, y altas correlaciones espaciales y temporales en la dinámica. Esto hace pensar que no es casualidad que muchas de las redes empíricas que exhiben distribuciones de grados de cola larga correspondan a sistemas evidentemente

*Del inglés *high school*.

Red	n	$\gamma_{in} / \gamma_{out}$	κ	referencia
WWW	325729	2.45 / 2.1	900	[11]
WWW	2×10^8	2.72 / 2.1	4000	[36]
WWW	4×10^7	2.38 / 2.1	1	[111]
WWW, site	260000	/ 1.94	-	[89]
Internet, domain	4389	2.2	40	[71]
Internet, router	3888	2.48	30	[71]
Internet, router	150000	2.4	60	[83]
Movie actors	212250	2.3	900	[23]
Coauthors, SPIRES	56627	1.2	1100	[143, 142, 141]
Coauthors, neuro.	209293	2.1	400	[24]
Coauthors, math.	70975	2.5	120	[24]
Sexual contacts	2810	3.4	-	[117]
Metabolic, E. Coli	778	2.2	110	[97]
Food web, Ythan estuary	134	1.05	35	[133]
Food web, Silwood Park	154	1.13	27	[133]
Citation	783339	3 /	-	[178]
Phone call	53×10^6	2.1	-	[9]
Words, cooccurrence	460902	2.67	3.03	[92]
e-mail messages	59912	1.5 / 2	-	[67]
Software packages	1439	1.4 / 1.6	-	[145]
Electronic circuits	24097	3	-	[91]
Peer to peer	880	2.1	-	[8, 124]
Protein Interaction	2115	2.4	-	[95]

Cuadro 2.1: Redes con distribuciones de grados tipo ley de potencias $P(k) \sim k^{-\gamma}$. Para cada red se especifica el tamaño n . En algunos casos se especifica γ_{in} y/o γ_{out} , los exponentes de las distribuciones de grados entrantes ó salientes respectivamente. En otros casos en que que aparece un sólo valor, simplemente se está especificando γ . También en algunos de los casos está especificado κ que, o bien representa un cut-off exponencial a la ley de potencias $P(k) \sim k^{-\gamma} e^{-k/\kappa}$, o bien es simplemente el máximo valor k_{max} de k que cada red presenta. Es importante resaltar que comúnmente γ oscila entre 1 y 3. Esta tabla es un resumen de la tabla II de [10] y de la tabla 3.1 de [146].

complejos como los sociales o los biológicos. El desafío es entonces explicar como se organizan, y que da lugar a dicha organización en estas redes evidentemente irregulares pero no completamente aleatorias.

Distribución de grados “libres de escala” en redes metabólicas y de proteínas

Particularmente en este trabajo, es importante resaltar que muchas redes de interacción de proteínas o redes metabólicas presentan distribuciones de grado libres de escala, o al menos leyes de potencias con truncamiento exponencial.

En [95] se reporta que la red de interacciones de proteínas de la *Saccharomyces cerevisiae* con $n = 1870$ y $M = 2240$, posee una distribución de grados ley de potencias con $\gamma = 2.4$ y un truncamiento exponencial en $\kappa \simeq 20$. En [209] también se estudia la caracterización de la red de proteínas de la *Saccharomyces cerevisiae*, con datos extraídos

de cinco bases de datos diferentes. Se reporta que las redes presentan distribuciones de grados tipo ley de potencias con exponentes que varían en $\gamma \in [2.1, 2.5]$. En [53] también es estudiada la red de interacción de proteínas de *Saccharomyces cerevisiae* a partir de 3 bases de datos. Cada base de dato comprende una caracterización de la red utilizando métodos distintos. Se reporta que $n \in \{1361, 2152, 4713\}$ y $M = \{3221, 2831, 14846\}$ respectivamente. La dos primeras de dichas redes presentan un distribución de grados ley de potencias con truncamiento exponencial en donde $\gamma = 2.5$ y $\kappa = \{30, 100\}$ respectivamente. La tercera de ellas es libre de escala con $\gamma = 2.1$. En [115] se reporta que para la red de interacción de proteínas de *C. elegans* con $n = 2898$ y 5460 conexiones dirigidas, las distribuciones de grados de las conexiones dirigidas corresponden a leyes de potencias. En [97] se reportan que redes metabólicas de 43 organismos presentan distribuciones de grados libres de escala para las conexiones entrantes y salientes con exponentes que varían en $\gamma_{in} \in [2.0, 2, 4]$ y $\gamma_{out} \in [2.1, 2.4]$. El tamaño y la cantidad total de conexiones de las redes varían entre $n \in \{178, \dots, 785\}$ y $M_{in} + M_{out} \in \{792, \dots, 5959\}$ respectivamente. Barabási *et. al.* [25] provee de una pequeña revisión del tema.

Observaciones: Perotti *et. al.* [164] incluyen bajo una misma perspectiva a las redes de interacción de proteínas, a las redes metabólicas, y a las redes ecológicas. Lo importante a resaltar aquí es que las redes de interacción de proteínas y las redes metabólicas exhiben una topología libre de escala, o al menos sus distribuciones de grados están caracterizadas en parte por leyes de potencias. El análisis correspondiente para las redes ecológicas quedará relegado a la sección 2.8.3.

2.1.3. El modelo de Barabási y Albert: El mecanismo de crecimiento y conexión preferencial

Hasta el advenimiento del modelo de Barabási–Albert [23], normalmente en los modelos de red considerados se partía de un número n de nodos que eran eventualmente conectados o reconectados siguiendo algún protocolo permaneciendo n fijo. En contraste, muchas de las redes reales son sistemas abiertos en donde el número de nodos varía en el tiempo. En particular en algunos casos como Internet o la WWW, claramente nacieron como pequeñas redes que eventualmente crecieron. Aunque seguramente en dichas redes cada tanto algunos nodos desaparecen, la principal tendencia en ellas ha sido un crecimiento sostenido*.

Por otro lado, en las redes aleatorias las conexiones se realizan al azar. Es esperable que en las redes reales éste no sea el caso, sino por el contrario que diferentes nodos presenten diferentes probabilidades de ser destinatarios de conexiones. Incluso es posible que dichas probabilidades

*Al menos han sufrido un evidente crecimiento por un relativamente largo período de tiempo.

estén determinadas por propiedades topológicas de la red, sin necesidad de invocar agentes externos a la red. Considérese por ejemplo una red de citas de publicaciones científicas en algún área dada. Es razonable esperar que una nueva publicación tenga más probabilidades de citar a publicaciones previamente muy citadas que a publicaciones desconocidas. En otras palabras, las publicaciones más citadas a su vez tienen más probabilidades de ser nuevamente citadas. Barabási y Albert denominaron a este tipo de dinámica de conexión, *preferential attachment* o *conexión preferencial*. Consideraciones análogas pueden aplicarse a una gran diversidad de redes tales como la WWW, Internet, red de colaboración de actores, etc, [23], razón por la cual es de esperarse que el conexión preferencial sea un mecanismo relativamente ubicuo que participa en la construcción de las redes reales.

Bajo estas consideraciones, Barabási y Albert* introducen un modelo de red en donde n crece sostenidamente en el tiempo, y los nuevos nodos se reconectan a los viejos utilizando una dinámica de conexión preferencial lineal con el grado de los mismos. Mas precisamente, se parte de una pequeña red semilla de n_0 nodos y m_0 links cuya estructura a la larga no importa, y luego a cada paso de tiempo:

- i) *Crecimiento*: Un nuevo nodo es introducido al sistema generando exactamente $m \leq n_0$ nuevos links con los nodos que ya estaban presentes en la red. Por lo que el tamaño de la red crece de n a $n + 1$.
- ii) *Conexión preferencial*: Cada uno de los nuevos m links se conecta con probabilidad:

$$\Pi_i = \frac{k_i}{\sum_t k_t}, \quad i = 1, \dots, n \quad (2.7)$$

al i -ésimo nodo de la red. Las dobles conexiones son evitadas.

Luego de t pasos de tiempo la red posee $n(t) = n_0 + t$ nodos y $M = m_0 + mt$ links. Simulaciones numéricas concluyen que las redes resultantes de éste mecanismo posee una distribución de grados ley de potencias $P(k) \sim k^{-\gamma}$ en donde $\gamma \simeq 3$ e independiente de m . Dado que la forma en que escala $P(k)$ es independiente del tiempo[†], se concluye que el sistema se auto-organiza en un estado libre de escala estacionario [23].

El cálculo de la distribución de grados ha sido resuelto analíticamente usando al menos tres enfoques diferentes. Originalmente Barabási y Albert[23, 10] utilizaron una teoría continua. En Dorogovtsev *et. al.* [60] se formula y resuelve el problema utilizando una ecuación maestra. En Krapivsky *et. al.* [110] se trata el problema en términos de *rate equations*,

*Haciendo simplificaciones tales como obviar las eventuales y más esporádicas desapariciones de nodos.

[†]Para tiempos suficientemente largos.

o en otras palabras, ecuaciones de tasas de crecimiento de los promedios $N_k(t)$ de las cantidades de nodos a grado k al tiempo t . Esencialmente, se encontró que cuando $t \rightarrow \infty$ y $k \gg m$:

$$P(k) = 2m^2 k^{-\gamma} \quad , \quad \gamma = 3 \quad (2.8)$$

lo cual coincide muy bien con los cálculos numéricos, incluyendo el prefactor $2m^2$ [23, 10].

Una pregunta válida es si los dos ingredientes, a saber : *crecimiento* y *conexión preferencial*, son necesarios para obtener redes libres de escala. Barabási y Albert [23, 10] estudian esta cuestión, encontrando numéricamente y analíticamente que en presencia de crecimiento, pero en ausencia de conexión preferencial, la distribución de grados de la red resultante decae exponencialmente con k :

$$P(k) = \frac{e}{m} \exp\left(\frac{-k}{m}\right) \quad (2.9)$$

perdiéndose el carácter libre de escala. Por otro lado, si se comienza con n nodos y luego a cada paso de tiempo se elige al azar un nodo, y se lo conecta a otro nodo i según la probabilidad de la ecuación (2.7), es decir, si se conserva el conexión preferencial pero no el crecimiento, tampoco se obtiene un estado estacionario con $P(k) \sim k^{-\gamma}$. Aunque numéricamente se encontró que para tiempos tempranos la red exhibe una distribución de grados libre de escala, eventualmente la distribución de grados se distorsiona. Ya para $t/n \simeq 40$ y $n = 10000$ la distribución ha perdido todo su carácter libre de escala [10]. En particular cuando $t \simeq \mathcal{O}(n^2)$ se obtiene una red totalmente conectada que claramente difiere de una red libre de escala.

Observaciones: El argumento de Barabási–Albert es que la particular estructura libre de escala es el resultado de un proceso dinámico, un *mecanismo* de construcción de la red. Esto contrasta con lo que ocurre para las redes aleatorias. En particular, se trata de un mecanismo fuera del equilibrio ya que presenta crecimiento, y en donde el conexión preferencial no es completamente aleatorio, sino que presenta un sesgo que introduce algún tipo de organización a la red. Aunque eventualmente se han encontrado contraejemplos en donde redes libres de escala se obtienen sin necesidad de recurrir ni al crecimiento ni al conexión preferencial (ver por ejemplo [37]), estos otros modelos utilizan información externa a la red para lograr reproducir redes libres de escala (en el caso del ejemplo antes mencionado, se asigna un “fitness” a cada nodo).

Variaciones del modelo de Barabási–Albert

Se han estudiado muchas variaciones al modelo de Barabási–Albert. La intención aquí no es repasarlas a todas, sino mencionar algunas de ellas y sus contribuciones para mostrar la flexibilidad de los conceptos introducidos por Barabási y Albert con su modelo.

Una primera modificación que ha sido estudiada consiste en generalizar el conexionado preferencial por medio de relajar la condición de linealidad en la ecuación (2.7). Mas precisamente, reemplazarla por:

$$\Pi_i \sim k_i^\alpha \quad (2.10)$$

En Krapivzky *et. al.* [110] se resuelve esta cuestión analíticamente utilizando el enfoque de las *rate equations*. Se encuentra que para el caso *sublineal*, es decir, cuando $\alpha < 1$, la distribución de grados deja de ser libre de escala siendo remplazada por un comportamiento exponencial estirado. En el caso *superlineal*, cuando $\alpha > 2$ ocurre un fenómeno tipo “ganador toma todo” en donde cuando $t \rightarrow \infty$ la fracción de nodos no conectada al nodo que acapara toda la red se va a cero. Cuando $2 > \alpha > 3/2$ la fracción de nodos que posee más de dos links no conectados al nodo acaparador, también se va a cero. En general, para $1 < (l+1)/l < \alpha < l/(l-1) < 2$ la fracción de nodos que posee l links de modo que ninguno de ellos está conectado al nodo acaparador se va a cero con $t \rightarrow \infty$. En otras palabras, en el régimen superlineal tampoco emergen redes libres de escala.

Una limitación del modelo de Barabási–Albert es que el exponente está fijo, es decir $\gamma = 3$. Tal como quedó en evidencia en el cuadro (2.1), las redes reales presentan distintos valores del exponente, preferentemente γ varía entre 1 y 3. Algunas variaciones al modelo original de Barabási–Albert han dado solución a esta cuestión. Por ejemplo, Krapivsky y Redner [109] introducen un modelo en donde a cada paso de tiempo un nodo ingresa a la red y un viejo nodo i es seleccionado uniformemente al azar como posible candidato a conectarse con el nuevo nodo. Luego con probabilidad $1 - r$ dicha conexión se efectiviza. Por el contrario con probabilidad r el nuevo nodo en vez de conectarse a i , se conecta a al ancestro j de i . Es decir, se conecta al nodo al que el nodo i se conectó al ingresar a la red. Estudiando el modelo desde el enfoque de las *rate equations* se puede ver que el modelo se comporta como un modelo de conexionado preferencial en donde la probabilidad de que un nuevo nodo se conecte a un nodo i de grado k_i viene dada por:

$$\Pi_i = k_i - 2 + \frac{1}{r} \quad (2.11)$$

lo que eventualmente conduce a una red cuya distribución de grados es una ley de potencias $P(k) = k^{-\gamma}$ en donde:

$$\gamma = 1 + \frac{1}{r} \in [2, \infty) \quad , \quad r \in (0, 1] \quad (2.12)$$

Observaciones: Dos conclusiones se pueden sacar de este ejemplo. La primera es que en un mecanismo de generación de redes puede subyacer el conexionado preferencial sin que éste haya sido explícitamente implementado. La segunda es que el conexionado preferencial más el crecimiento no necesariamente lleva a $\gamma = 3$. En [10] (ver tabla III de dicho

trabajo) hay un recuento de modelos de redes libres de escala en donde ambas cuestiones son corroboradas. Por otro lado, salvo en aquellos modelos que presentan *crecimiento acelerado**, en general los exponentes son tales que $\gamma \geq 2$. Lo cual es importante resaltar ya que en algunos casos de redes reales $\gamma < 2$ (ver tabla 2.1).

El conexasionado preferencial en redes reales

Determinar la participación de un mecanismo de *conexasionado preferencial* en redes empíricas que evidencian crecimiento requiere un registro de la evolución de las mismas. Dichos registros son relativamente fáciles de obtener para redes sociales o tecnológicas. En el caso de redes biológicas la obtención de registros de su evolución suele ser más difícil, pero hay trabajos en donde se ha hecho. En muchas de las redes tecnológicas y sociales hay evidencia a favor de que sobre ellas opera un mecanismo de conexasionado preferencial. En redes biológicas la cuestión no es del todo clara. En [96] se investiga el conexasionado preferencial en Internet a nivel sistema autónomo, en una red de citaciones de publicaciones en *Physical Review Letters*, en una red de colaboraciones en trabajos científicos en neuro-ciencias, y la red de colaboración de actores según la *Internet Movie Data Base*. Ellos encontraron que se satisface la ecuación (2.10) con $\alpha = 1.05, 0.95, 0.79$ y 0.81 respectivamente. Concluyen que las dos primeras, Internet y la red de citaciones, corresponden a un conexasionado preferencial lineal, mientras que las dos últimas, la de colaboraciones científicas y actorales, corresponden a un conexasionado preferencial sublineal. En Pastor-Satorrás *et. al.* [163, 197] es estudiada la probabilidad $\Pi(k)$ de que un nuevo nodo, o uno preexistente se conecte a otro de grado k utilizando registros de la evolución de *sampleos* de Internet entre los años 1997 y 2000. La probabilidad $\Pi(k)$ puede expresarse como:

$$\Pi(k) = \sum_{i:k_i=k}^{n_k} \Pi_i \quad (2.13)$$

donde Π_i es la probabilidad de que el nodo ingresante se conecte al nodo preexistente i , $n_k \simeq nP(k)$ es la cantidad de nodos de grado k , y la suma corre sobre todos los sitios i de grado $k_i = k$. Si la presión selectiva favorece algún tipo mecanismo de conexasionado preferencial, es decir, si los nodos ingresantes a la red tienen una probabilidad de conectarse a los ya preexistentes que sea proporcional al grado de los mismos elevado a alguna potencia α , entonces deberíamos esperar que:

$$\Pi_i = \frac{k_i^\alpha}{\sum_{j=1}^n k_j^\alpha} \quad (2.14)$$

Por otro lado $\sum_{j=1}^n k_j^\alpha \simeq n \langle k^\alpha \rangle$, de modo que:

*Redes en donde $\langle k \rangle$ crece con el tamaño de la red.

$$\begin{aligned}
\Pi(k) &= n_k \frac{k^\alpha}{\sum_{j=1}^n k_j^\alpha} \\
&\simeq nP(k) \frac{k^\alpha}{n \langle k^\alpha \rangle} \\
&\sim P(k)k^\alpha
\end{aligned} \tag{2.15}$$

Suponiendo que $P(k) \sim k^{-\gamma}$, se tiene que $\Pi(k) \sim k^{\alpha-\gamma}$. Pastor-Satorrás *et. al.* encuentran que $\gamma \simeq 2.2$ y $\Pi(k) \sim k^{-1.2}$ lo cual es consistente con $\alpha \simeq 1$, o lo que es lo mismo, con conexionado preferencial lineal. Newman [140] utilizando métodos similares estudia un par de redes de colaboraciones. Una en trabajos científicos en física (*Los Alamos e-print Archive*) y la otra en biología (*Medline*). Concluye que $\alpha \simeq 0.8$ y $\simeq 1.04$ respectivamente. En [131] se estudia el conexionado preferencial de las conexiones entrantes y salientes de la red “social” *Flickr**. En dicho trabajo se reporta que la cantidad de conexiones que recibe (genera)[†] cada usuario por día crece como una ley de potencias en función de la cantidad de la conexiones entrantes (salientes) que actualmente posee. Sorprendentemente se han hecho mediciones del conexionado preferencial en redes ecológicas [156]. El trabajo se centra en el estudio de la dinámica de conexionado en una red de polinización en el Ártico. En el período de florecimiento de las distintas especies de plantas diversas especies de polinizadores entran en contacto con las mismas. A lo largo de dicho período se registra la formación de la red de contactos entre los grupos de especies en función del tiempo. Más allá de la poca estadística y el entendible ruido del sistema, se encontró evidencia de conexionado preferencial. Para el conexionado de nuevas plantas floreciendo con polinizadores ya instalados $\alpha \simeq 0.6$. Para el conexionado de nuevos polinizadores con plantas ya florecidas $\alpha \simeq 0.8$. Las distribuciones de grados son consistentes con leyes de potencias con truncamiento exponencial. Se argumenta que la causa del truncamiento es una “capacidad de carga máxima” en la cantidad de conexiones de los nodos (en la siguiente sección se discutirá con más detalle el concepto de “capacidad de carga” en el grado de los nodos). Es importante resaltar que éste trabajo mencionado hace referencia a la existencia de conexionado preferencial pero no en la formación de una red ecológica a través de la evolución, sino en la dinámica de las conexiones al cambiar las estaciones del año. De hecho no está claro que la hipótesis de conexionado preferencial sea el principal mecanismo generador de topologías libres de escala en redes biológicas y en particular en redes ecológicas [100, 30]. Por ejemplo, en [116] se encuentra evidencia plausible de que el conexionado preferencial participa

*Flickr <http://www.flickr.com/> es una red social en donde se comparten imágenes.

[†]El estudio se realiza sobre una versión dirigida de la red. Por definición, una conexión que recibe un dado usuario es una conexión entrante a dicho usuario, mientras que una que genera es una conexión saliente desde dicho usuario.

en la formación de la red metabólica de la *E. Coli* a través del mecanismo de duplicación de nodos [161, 190]. Sin embargo también menciona que nuevas enzimas ingresan a la red a través del mecanismo de *transferencia genética horizontal**, y preferentemente lo hacen las enzimas de mayor conectividad. En [69] se reporta que en la red de interacción de proteínas de la *S. cerevisiae* se observa una correlación positiva entre la edad de las clases de proteínas y su conectividad, lo cuál constituye evidencias a favor del conexionado preferencial. Sin embargo en [112] se reporta que en la red de interacción de proteínas de la *S. cerevisiae*, la funcionalidad de una clase de proteínas es un factor tan o más importante que la antigüedad en la determinación de la conectividad de las proteínas. Paralelamente en [32] se reporta que en la red de interacción de proteínas de la *S. cerevisiae* un mecanismo de reconexión ocurre a una tasa significativamente mayor que el mecanismo de duplicación de nodos, sugiriendo que el mecanismo de reconexión constituye un factor tan o más relevante que el conexionado preferencial en la formación de la topología de la red.

Observación: En redes sociales o tecnológicas hay fuerte evidencia que el mecanismo de conexionado preferencial constituye uno de los principales factores que participan en la formación de la topología de dichas redes. Para el caso de las redes biológicas hay evidencia a favor de que otros mecanismos además del conexionado preferencial podrían estar participando en la formación de *hubs* en la red.

Modelos para topologías ni aleatorias ni libres de escala

Según se expuso en la sección 2.1.2 algunas redes empíricas no presentan la topología de las redes aleatorias, pero sin embargo tampoco son libres de escalas. Tal es el caso de distribuciones de grado como las dadas por la ecuación (2.6). En Amaral *et. al.* [17] se provee de un par de modelos que constituyen una modificación al modelo de Barabási–Albert que reproduce el comportamiento de la ecuación (2.6).

Básicamente, en estos modelos los nodos adquieren una nueva propiedad, la de estar activos ó inactivos. Mientras los nodos están activos pueden recibir nuevas conexiones de acuerdo al conexionado preferencial. Cuando un nodo se vuelve inactivo, se lo remueve del conjunto de nodos aptos para recibir conexiones. En otras palabras, a cada paso de tiempo t el nodo ingresante a la red desarrolla m nuevos links con los viejos nodos de acuerdo a la siguiente modificación de la ecuación (2.7):

$$\Pi_i = \begin{cases} \frac{k_i}{\sum_{j \in A(t)} k_j} & \text{si } i \in A(t) \\ 0 & \text{c.c.} \end{cases} \quad (2.16)$$

*La transferencia genética horizontal (o lateral), es el proceso por el cuál un organismo incorpora un gen procedente de otro organismo sin ser descendiente de éste último. Este tipo de transferencia genética es muy común y quizás predominante en los organismos unicelulares.

en donde $A(t)$ es el conjunto de nodos en estado activo al tiempo t .

En uno de los modelos el nuevo ingrediente es el *envejecimiento de los nodos*. Mas precisamente, cuando un nodo ingresa a la red lo hace en estado activo. Luego, a cada paso de tiempo cada nodo activo puede volverse inactivo con probabilidad $1/\tau$. Esto implica que la probabilidad de que un nodo pueda permanecer activo decae exponencialmente con el transcurso de su tiempo de vida activo, poseyendo un tiempo de vida característico $1/\ln \tau$.

En el otro modelo, el nuevo ingrediente es una *capacidad limitada* en el número de conexiones que un nodo puede soportar. Al igual que antes, cada nodo ingresa a la red en estado activo. Si eventualmente un nodo alcanza un grado máximo fijo y predeterminado k_{cc}^* , dicho nodo pasa a estado inactivo. Puede interpretarse a k_{cc} como la *capacidad de carga* que cada nodo de la red puede tener.

Obsérvese la figura (2.1). En el gráfico **a** puede apreciarse el comportamiento de la distribución de grados para el modelo que incorpora envejecimiento para diferentes valores de τ . Puede apreciarse como a medida que disminuye τ , es decir se acelera el envejecimiento, el truncamiento de la ley de potencias se hace más pronunciado. En el gráfico **b** se observa el comportamiento de la distribución de grados para el modelo que incorpora capacidad de carga. Claramente el apartamiento de la ley de potencias va incrementándose a medida que disminuye la capacidad de carga de los nodos. En ambos casos, para τ y k_{cc} relativamente largos, el comportamiento ley de potencias persiste en una región considerable de la curva, propiedad consistente con la ecuación (2.6).

Observaciones: Una situación típica interpretable en términos del modelo que incorpora envejecimiento es la red de colaboración de actores en donde cada nodo representa un actor, y un link existe entre un par de actores si han trabajado al menos una vez en una misma película. En este ejemplo es claro que incluso un nodo altamente conectado eventualmente dejará de recibir conexiones. El segundo modelo es interpretable en términos de la red mundial de aeropuertos. Las aerolíneas comerciales prefieren tener acceso a un pequeño número de hubs donde casi todas las rutas se conectan. Por otro lado por cuestiones de limitaciones físicas y temporales cada aeropuerto tiene que limitar el número de aterrizajes y despegues, lo que eventualmente impone un límite al número de conexiones que cada nodo posee [17].

Lo importante a resaltar en estos modelos es que incluso para casos en que las distribuciones de grados no son de cola larga, los mecanismos de crecimiento y conexionado preferencial son útiles para explicar ciertas diferencias encontradas al comparar redes reales con redes aleatorias.

Las redes ecológicas posiblemente correspondan a un caso en donde la capacidad de carga en el grado de los nodos de la red sea parte del mecanismo de ensamblado de la misma. Esta cuestión será analizada más

**cc* viene de *capacidad de carga*.

en profundidad en la sección 2.8.3.

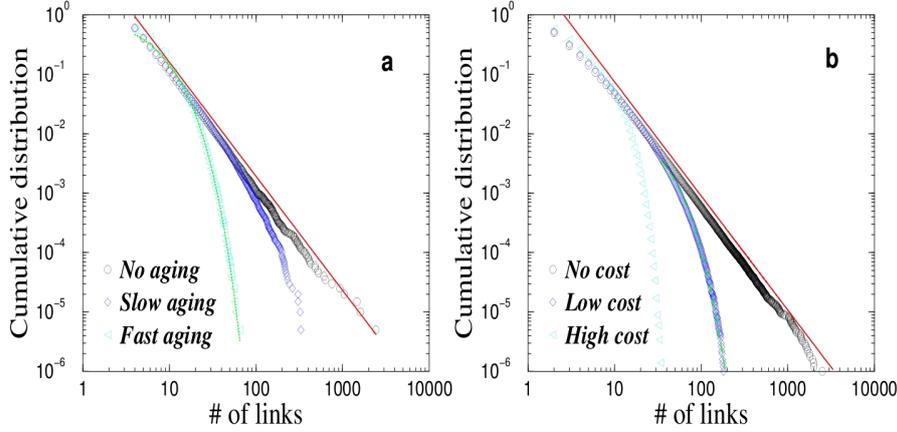


Figura 2.1: Distribuciones de grados para las versiones modificadas del modelo de Barabási–Albert que incorporan *envejecimiento* (gráfico **a**), ó *capacidad de carga* (gráfico **b**) [17]. Se observa como al aumentar la tasa de envejecimiento, o respectivamente disminuir la capacidad de carga, las distribuciones de grados se apartan de la ley de potencias predicha por el modelo de Barabási–Albert [23] (líneas continuas rojas correspondientes a un comportamiento $\sim k^{-3}$). Cuando ello ocurre el comportamiento queda bien descrito por una ecuación de la forma (2.6). Esta figura fué extraída de [10] (figura 27).

2.2. Las distancias topológicas entre nodos de la red

Una red *totalmente conectada* es tal que siempre es posible recorrer una secuencia de links que conecta a todo par de nodos en la misma. Dado un par de nodos (i, j) en una red conectada, la distancia topológica L_{ij} entre los mismos consiste en la menor cantidad de links que es necesario recorrer sobre la red para ir de uno de ellos al otro. Cuando $L_{i,j} = l \in \{1, 2, \dots, n-1\}$ se dice que i y j son vecinos topológicos de orden l^* . Una magnitud que suele estudiarse en las redes totalmente conectadas, es la distancia topológica promedio:

$$L = \left\langle \frac{1}{\frac{n(n-1)}{2}} \sum_{\langle i,j \rangle} L_{ij} \right\rangle \quad (2.17)$$

donde $\langle \cdot \rangle$ denota un promedio sobre el ensamble de redes, $\langle i, j \rangle$ denota el conjunto de todos los pares[†] de nodos posibles en cada red del ensamble,

*Es decir, $L_{i,j} = 1$ primeros vecinos (topológicos), $L_{i,j} = 2$ segundos vecinos, etc. A lo largo de la tesis sólo se hará referencia a éste concepto de vecinidad, y no al que se deduce a partir de (por ejemplo) una distancia euclídea.

[†]No ordenados.

siendo $n(n-1)/2$ el número de pares posibles para una red de n nodos.

Una magnitud que suele estudiarse en el mismo contexto, es el que se denomina el diámetro de la red:

$$D = \left\langle \max_{\langle i,j \rangle} L_{ij} \right\rangle \quad (2.18)$$

Aunque $D \geq L$, frecuentemente L y D son del mismo orden, siendo esa la razón por la cuál se estudia ambas magnitudes en el mismo contexto, o en muchos de los casos el análisis se reduce a sólo estudiar L .

2.2.1. La distancia topológica de un grafo aleatorio

Considérese un modelo de redes aleatorias en donde en vez de pre-determinarse los parámetros n y p que definen el ensamble de redes, se pre-determinan n y $P(k)$. Tal modelo se denomina *modelo configuracional*[153]* y a las redes en dicho modelo se las denomina *grafos aleatorios generalizados*† ó *redes aleatorias generalizadas*[10].

Dada una red, sea z_i es el promedio del número de vecinos de orden i que un nodo cualquiera posee (de modo que z_1 es el promedio del número de primeros vecinos $\langle k \rangle$, z_2 es el promedio de segundos vecinos, etc, en donde el promedio es tomado sobre todos los nodos de la red). En una red aleatoria generalizada pre-determinar $P(k)$ pre-determina la secuencia $\{z_1, z_2, \dots\}$. Se puede probar que para una red aleatoria generalizada se tiene que (en el límite $n \gg z_1$ y $z_2 \gg z_1$) [10]:

$$L \simeq \frac{\ln(n/z_1)}{\ln(z_2/z_1)} + 1 \ll n \quad (2.19)$$

Notar que dicha expresión sólo depende de z_1 y z_2 . En el caso de una red aleatoria (la común y silvestre en donde se pre-determina n y p) $z_2 \simeq z_1^2$ por lo que:

$$L \simeq \ln n / \ln \langle k \rangle \quad (2.20)$$

Cuando la distribución de grados es una ley de potencias truncada por una exponencial‡:

$$P(k) = Ck^{-\gamma}e^{-k/\kappa} \quad \text{para } k \geq 1 \quad (2.21)$$

se obtiene que la expresión de la ecuación (2.19) en el límite $n \gg 1$ se reduce a:

$$L = \frac{\ln n + \ln[Li_{\gamma}(e^{-1/\kappa})/Li_{\gamma-1}(e^{-1/\kappa})]}{\ln[Li_{\gamma-2}(e^{-1/\kappa})/Li_{\gamma-1}(e^{-1/\kappa}) - 1]} + 1 \quad (2.22)$$

*Del inglés *configuration model*.

†Del inglés *generalized random graph*.

‡Si no se introduce el truncamiento exponencial, el posterior cálculo de L puede divergir para todo valor de n debido a la presencia de probabilidades significativas de que existan nodos de grados arbitrarios.

donde $Li_n(x)$ es el n -ésimo polilogaritmo* de x .

Observación: El resultado a tener en cuenta es que las redes aleatorias en general (sin importar su distribución de grados) son tales que $L \sim \ln n \ll n$.

Nota: La cuestión de las distancias topológicas en las redes empíricas será abordada en la sección 2.4.2 después de introducirse el concepto de *mundo pequeño* en la sección 2.4.

2.3. El concepto de clusterización en redes

Una medida que resulta útil a la hora de estudiar las propiedades de las redes es el coeficiente de clusterización Cc . Para definir cuán clusterizada está una red, Watts y Strogatz [204] definen primero cuán clusterizada está la vecinidad de un nodo. Mas precisamente, dado un nodo i con k_i vecinos, la clusterización $Cc_i^{(WS)}$ de la vecinidad del nodo i es según Watts–Strogatz:

$$Cc_i^{(WS)} = \frac{E_i}{\frac{k_i(k_i-1)}{2}} \quad (2.23)$$

donde $E_i \in \{0, 1, \dots, k_i(k_i-1)/2\}$ es la cantidad de conexiones existentes entre los k_i primeros vecinos del nodo i . Notar que $Cc_i^{(WS)} \in [0, 1]$, de modo que el extremo $Cc_i^{(WS)} \simeq 0$ indica que la vecinidad del nodo i está poco clusterizada, mientras que $Cc_i^{(WS)} \simeq 1$ indica que la vecinidad del nodo i está altamente clusterizada. En general, la clusterización de un ensamble de redes según la formulación de Watts–Strogatz es:

$$Cc^{(WS)} = \left\langle \frac{1}{n} \sum_i Cc_i^{(WS)} \right\rangle \quad (2.24)$$

en donde n es la cantidad de nodos de cada red[†], y el promedio $\langle \cdot \rangle$ es sobre el ensamble en cuestión. Al igual que con los nodos, si $Cc^{(WS)} \simeq 1$ para un ensamble de redes, luego dicho ensamble está altamente clusterizado, mientras que si $Cc^{(WS)} \simeq 0$ el ensamble está compuesto en su mayoría por redes de baja clusterización.

Otra forma en que se suele definir el coeficiente de clusterización de un ensamble de redes viene dada por:

$$Cc^{(Tr)} = \left\langle \frac{3 \times (\text{número de triples cerrados})}{\text{número de triples conectados}} \right\rangle \quad (2.25)$$

la cuál se suele denominar también por *transitividad*[‡] [153, 66] en correspondencia con el concepto de “fracción de triples transitivos” utilizado en

* $Li_s(x) = \sum_{k=1}^{\infty} z^k / k^s$.

[†] En general se suelen considerar ensambles de redes de tamaños muy similares sino iguales.

[‡] Aquí el promedio $\langle \cdot \rangle$ también es sobre todo el ensamble de redes en cuestión.

sociología [203]. Ésta otra forma de calcular el coeficiente de clusterización no es completamente equivalente a la utilizada por Watts–Strogatz [185]. Básicamente la diferencia entre ambas formas es que se invierte el orden de los operadores suma y división. En el caso de $Cc^{(WS)}$ se divide primero y luego se suma, mientras que en $Cc^{(Tr)}$ ocurre al revés, primero se suman todos los “numeradores” y “denominadores” por separado, y luego se dividen los resultados. Esta otra forma de definir la clusterización también satisface $Cc^{(Tr)} \in [0, 1]$ y se interpreta de manera análoga a la de Watts–Strogatz.

Observación: El coeficiente de clusterización sirve para cuantificar el grado de organización de la red en términos de grupos de nodos que forman “comunidades” muy interconectadas entre si. Una interpretación intuitiva en términos de una red social es que Cc estima la probabilidad de que los “amigos” de alguien a su vez sean “amigos” entre si.

2.3.1. La clusterización de las redes aleatorias

Para una red aleatoria cualquiera sea el criterio de elección de un par de nodos no introduce tendencias sobre la probabilidad de que dicho par esté conectado entre si*. Por ésta razón la probabilidad de que los vecinos de un dado nodo estén conectados entre si es igual a la probabilidad de que dos nodos elegidos al azar estén conectados entre si. Ello lleva a que:

$$Cc = Cc^{(WS)} = Cc^{(Tr)} = p = \frac{\langle k \rangle}{n} \quad (2.26)$$

Es decir, en una red aleatoria las dos definiciones de clusterización aquí utilizadas resultan en el mismo valor. Por otro lado, fijado un grado medio $\langle k \rangle$, luego la clusterización de las redes aleatorias decrece linealmente con el tamaño $Cc \sim 1/n$. Tal como se verá en la sección siguiente, ésta particular propiedad resultó en un primer contraste entre las redes aleatorias y las empíricas.

2.3.2. La clusterización en redes reales

Partiendo de la ecuación (2.26), si el modelo de red aleatoria predice bién el comportamiento de Cc en redes reales, luego es de esperar que en dichas redes se satisfaga $Cc/\langle k \rangle \simeq 1/n$. Utilizando los datos de la tabla I del review de Alber–Barabási [10] se puede graficar el comportamiento de $Cc/\langle k \rangle$ en función de n para diversas redes empíricas (ver figura 2.2). Se observa que las redes empíricas presentan un comportamiento muy distinto al predicho por el modelo de red aleatoria, más precisamente:

$$\frac{Cc^{(WS)}}{\langle k \rangle}(n) \simeq \text{constante} \quad (2.27)$$

*Esto ocurre por la definición de red aleatoria, cada par tiene probabilidad p de conectarse independientemente si los otros pares están conectados o no.

Una forma más directa de resaltar la diferencia entre la clusterización de las redes reales y la de correspondientes redes aleatorias es combinando las ecuaciones (2.26) y (2.27):

$$\frac{C_c}{C_{c_{rand}}} \sim n \quad (2.28)$$

lo cual indica que la clusterización de las redes reales crece con el tamaño respecto de la clusterización de las correspondientes redes aleatorias*. Lo que éste resultado implica es que las redes empíricas poseen una clusterización mucho más alta que una red sin estructura.

En la figura (2.2) sólo unos pocos puntos corresponden a redes biológicas, conviene complementar la evidencia de la relativamente alta clusterización mencionando algunas referencias extras para el caso de las redes biológicas. En [176] se reporta que $C_c^{(WS)}$ es independiente de n para las redes metabólicas de 43 organismos diferentes (figura 2B en dicho trabajo). En [209, 53] 5 versiones[†] de la red de interacción de proteínas de la levadura *S. servisiae* poseen un relativamente alto clustering. En [115] el interactoma[‡] de *C. elegans* posee una alta clusterización respecto al de correspondientes redes aleatorias. En relación a las redes ecológicas la cuestión es brevemente discutida en la sección 2.8.2. La conclusión en dicha sección es que debido al pequeño tamaño y la ecuación (2.28) la relativamente alta clusterización no es tan evidente en las redes ecológicas.

2.3.3. La clusterización en el modelo de Barabási–Albert

Para el modelo de redes libres de escala de Barabási–Albert el coeficiente de clusterización $C_c^{(WS)}$ en función del tamaño n se estudió en un primer lugar numéricamente. Se encontró que $C_c^{(WS)}(n) \simeq n^{-0.75}$ [10]. Sin embargo, dicho comportamiento no es correcto, resultando ser un artefacto de la aproximación numérica. Klemm y Eguíluz demostraron analíticamente que en este modelo [105]:

$$C_c^{(WS)}(n) = \frac{m}{8} \frac{(\ln n)^2}{n} \quad (2.29)$$

Observación: Aunque éste comportamiento revela que las redes de Barabási–Albert poseen un mayor grado de clusterización que las redes aleatorias, no se satisface la ecuación (2.27), por lo que el modelo de Barabási–Albert no captura la alta clusterización encontrada en las redes reales.

*Las correspondientes redes aleatorias poseen el mismo n y M que las redes reales con las cuales se las compara.

[†]Cada una de las versiones es obtenida de una base de datos diferente. La red de cada base de datos se obtiene utilizando técnicas experimentales y/o de análisis de datos diferentes.

[‡]Interactoma es el nombre que se da a la red de interacciones proteína–proteína.

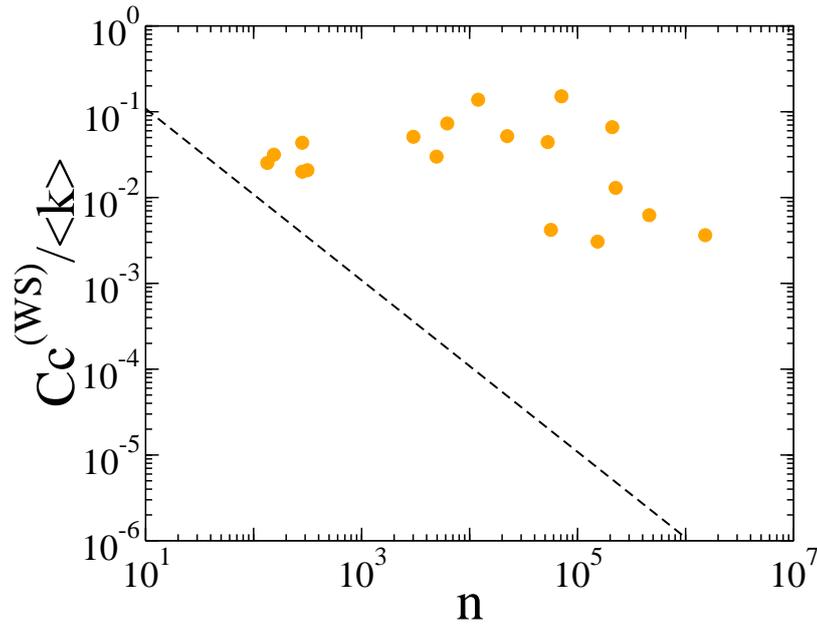


Figura 2.2: Comparación de $C_c^{(WS)}(n)/\langle k \rangle$ entre datos obtenidos de redes empíricas (círculos naranjas) y el comportamiento que se espera del modelo de redes aleatorias (línea discontinua). Se observa que el comportamiento de $C_c^{(WS)}(n)/\langle k \rangle$ para las redes empíricas es aproximadamente constante, a diferencia de lo que ocurre con una red aleatoria en donde $C_c^{(WS)} = \langle k \rangle/n$. El gráfico es una imitación del gráfico de la figura 9 del Review de Albert–Barabási [10], construido a partir de datos extraídos de la tabla I de dicho Review. Los datos empíricos corresponden a redes que cubren un amplio espectro de áreas de investigación. Desde redes tecnológicas de comunicación (como Internet, o la WWW) hasta redes ecológicas.

2.4. El concepto de Mundo Pequeño

El comportamiento de la distancia topológica L en una red aleatoria difiere significativamente si se lo compara con lo que ocurre en una red regular o *lattice**. En una *lattice* conformada por una red hipercúbica[†] en $d \ll n$ dimensiones se tiene que $L \sim \sqrt[d]{n}$. En otras palabras, la distancia topológica típica en una red aleatoria es mucho menor que en una *lattice*.

Considérese una red social. Los seres humanos vivimos en una superficie[‡] y por ende a priori sería posible hipotetizar que $L \sim \sqrt{n}$, es decir, escala como en una *lattice* en dimensión $d = 2$. Sin embargo la frase “6 grados de separación” es ya bien conocida[§]. La misma fue acuñada como

*En castellano tanto las redes regulares como las irregulares caen dentro de la categoría de la denominación *red*. En inglés sin embargo, *network* es el término que se suele aplicar a una red irregular, mientras que *lattice* es el que se suele aplicar a una red regular.

[†]Red hipercúbica denota la generalización del concepto de red cuadrada llevada a un número arbitrario de dimensiones.

[‡]La superficie de la tierra.

[§]Es suficiente con buscar en la red “six degrees of separation”.

resultado de un trabajo realizado por Milgram [184], quién estimó que la distancia topológica promedio en EEUU es 6. Cantidad mucho más parecida a $\ln(10^7) \simeq 16$ * que a $\sqrt{10^7} \simeq 3000$. A partir de éste resultado se acuñó a su vez la frase “mundo pequeño” que resume la idea de que a pesar del gran número de individuos que las sociedades poseen y de que la red de contactos sociales se correlaciona con una estructura bidimensional, se requiere en promedio recorrer pocos contactos para conectar dos personas elegidas al azar. Una explicación es proveída por el modelo de Watts–Strogatz presentado en la siguiente sección, en donde la presencia de estructura es cuantificada por una alta clusterización C_c .

2.4.1. El modelo de Watts–Strogatz

Inspirados en la cuestión de redes de mundo pequeño Watts y Strogatz propusieron un modelo que interpola entre una *lattice* y un grafo aleatorio [204]. Mas precisamente ellos construyen una red usando el siguiente algoritmo [10]:

- i) Comenzar con orden:* Crear un *ring lattice*[†] de n nodos, cada uno conectado con otros K vecinos ($K/2$ a cada lado). A fin de restringirse a redes conectadas pero a su vez diluidas[‡] se exige que $n \gg K \gg \ln n \gg 1$.
- ii) Aleatorizar:* Con probabilidad p reconectar cada uno de los links de la red a de nodo elegido al azar de forma tal de evitar las autoconexiones o las conexiones múltiples[§]. Este proceso introduce $pnK/2$ conexiones de largo alcance que conectan nodos que de otra manera estarían topológicamente distantes en la red.

Para $p = 0$ se conserva la estructura de *lattice*, mientras que para $p = 1$ se obtiene una red aleatoria. Variando $p \in [0, 1]$ se puede regular el grado de interpolación entre una *lattice* y una red aleatoria (ver figura (2.3)).

Las distancias topológicas en el modelo de Watts–Strogatz

Watts y Strogatz estudiaron $L(p)$, es decir, que ocurre con L al variar p . Encontraron que $L(0) \simeq n/2K \gg 1$ consistente con la forma en que

*Que resulte menor no está mal, recordar que para una red aleatoria vale la ecuación (2.19) que en general es menor a $\ln n$. Por otro lado estoy asumiendo que $n \simeq 10^7$ es una estimación del número de habitantes en EEUU. Pensar en toda la población de EEUU es a su vez otra aproximación.

[†]Un *ring lattice* es un caso particular de *lattice* en dimensión $d = 1$ y con condiciones de contorno periódicas. Cada nodo en la red puede tener mas de dos vecinos, lo cuál se logra por medio de conectarlos con los segundos vecinos, etc.

[‡]El término en inglés es *sparse*. La idea es que el número de links es mucho menor que el total que podría haber $M \ll n(n-1)/2$.

[§]Es decir, evitar que exista un par de nodos (o más) unidos por más de una conexión.

escala L para una lattice en dimensión $d = 1$. En el otro extremo cuando $p \rightarrow 1$ se encontró que $L \simeq \ln n / \ln K$ consistente con lo que ocurre en una red aleatoria. Una pregunta relevante corresponde a la naturaleza de la transición entre los casos límites $p = 0$ y $p = 1$. En la figura (2.4) se puede observar el comportamiento típico de $L(p)$. Posteriores trabajos determinan que hay una región de entrecruzamiento de valores de p que depende de n para la cuál la transición ocurre. Visto de otra manera, a un valor fijo de p existe un tamaño de transición n^* tal que si $n < n^*$ luego $L \sim n$, pero si $n > n^*$ luego $L \sim \ln n$ [10]. En general, hay evidencia analítica y empírica de que [29, 28, 26, 27, 154, 56, 10]:

$$L(n, p) \sim \frac{n}{K} f(pKn^d) \quad (2.30)$$

donde:

$$f(u) = \begin{cases} \text{constante} & \text{si } u \ll 1 \\ \frac{\ln u}{u} & \text{si } u \gg 1 \end{cases} \quad (2.31)$$

y donde la transición entre los distintos regímenes de comportamiento en la función ocurre aproximadamente para $n^* \sim p^{-1/d}$. Aquí d es la dimensión sobre la cuál se construye la red anillo. Entonces, para el modelo original en donde $d = 1$ la probabilidad a la cuál la transición ocurre escala como $p^* \sim 1/n$ [10].

Observación: Una de las cuestiones importantes a tener en cuenta es que aunque el resultado $L \sim \ln n$ es originalmente obtenido para redes aleatorias, luego se encuentra que el mismo comportamiento se observa sobre redes a las que se les ha aleatorizado una mínima fracción de las conexiones. En otras palabras, basta una pequeña fracción de aleatoriedad (del orden de $\sim n^{-d}$) para que la distancia topológica en una red de Watts–Strogatz se comporte como la de una red aleatoria.

La clusterización en el modelo de Watts–Strogatz

Una de las propiedades de las redes “semillas” que luego son reconectadas en el modelo de Watts–Strogatz, es que poseen una alta clusterización en comparación con redes aleatorias de igual tamaño y conectividad. En particular, para $p = 0$ el coeficiente de clusterización $C_c^{(WS)}$ es independiente de n . Mas precisamente, en una red anillo de dimensión $d = 1$ en donde cada nodo está conectado a los K vecinos más próximos, se tiene que [10]:

$$C_{c_i}^{(WS)} = \frac{3(K-2)}{4(K-1)} \quad , \text{ para todo nodo } i \quad (2.32)$$

y para todo $K \geq 4$ y múltiplo de 2. Luego trivialmente:

$$C_c^{(WS)} = 3(K-2)/4(K-1) \quad (2.33)$$

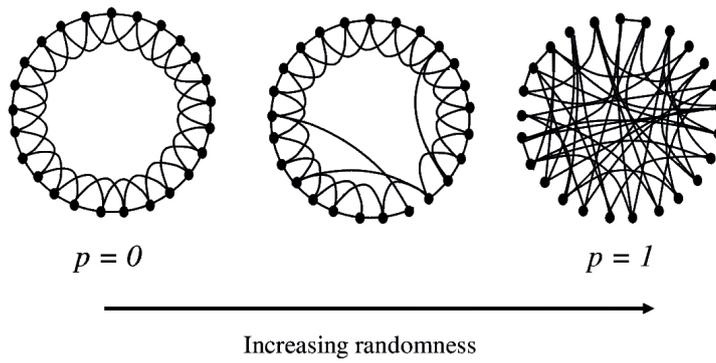


Figura 2.3: Redes generadas con el algoritmo de Watts–Strogatz para diferentes valores de p sobre una red anillo con $n = 20, K = 2$. Para $p = 0$ se ve claramente la forma de la red anillo sin ninguna alteración. La red en este estado posee una alta clusterización Cc , pero típicamente hay del orden de n/K links intermediarios entre los pares de nodos, i.e. la distancia topológica L es relativamente grande. Cuando $p = 1$ la red ha perdido toda la estructura topológica heredada de la red anillo. En esta otra situación la distancia topológica L se ha reducido drásticamente, sin embargo lo mismo ha ocurrido con la clusterización Cc . A un valor intermedio de p (red del medio) la red conserva buena parte de la estructura topológica de la red anillo original logrando que la clusterización se mantenga alta. Al mismo tiempo pueden apreciarse algunos links que han sido reconectados y sirven de atajos para acercar topológicamente nodos originalmente distantes reduciendo así el valor de L de una manera significativa. Esta figura fué extraída de [204] (figura 1 de dicho trabajo).

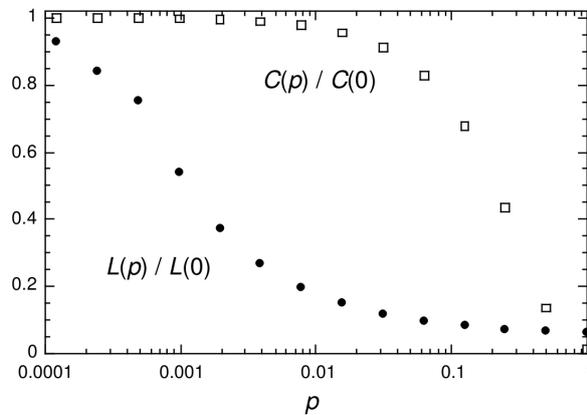


Figura 2.4: En esta figura se presenta el comportamiento de la distancia topológica típica L y el coeficiente de clusterización Cc para el modelo de Watts–Strogatz. Mas precisamente se grafican $L(p)$ y $Cc(p)$, es decir las mencionadas magnitudes en función del parámetro p , y renormalizados por $L(0)$ y $Cc(0)$ respectivamente. Cuando $p \rightarrow 0$ se aprecia que $Cc(p)/Cc(0) \rightarrow 1$ y $L(p)/L(0)$ indicando que la red resultante de aplicar el algoritmo de Watts–Strogatz tiende a generar redes similares a la red anillo original. Por el contrario cuando $p \rightarrow 1$ ambas cantidades disminuyen drásticamente correspondiendo a que la red resultante se asemeja más a una red aleatoria que a la red original. Lo importante es observar que aunque ambas cantidades decaen, no lo hacen simultáneamente. Existe una región de valores de p en donde $Cc(p)/Cc(0)$ permanece en valores altos mientras que $L(p)/L(0)$ ha decaído. En estos valores de p la red interpolante conserva buena parte de la estructura de la red original, sin embargo ha bastado reconectar unos pocos links para que la distancia topológica se asemejará a la de la red aleatoria. Esta figura fué extraída de [204] (figura 2 de dicho trabajo).

magnitud que es independiente de n y tiende a $3/4$ cuando $K \rightarrow \infty^*$. Esta propiedad de que Cc no dependa de n es característica de redes ordenadas tipo *lattice*, en donde Cc depende solamente en el número de coordinación K . Barrat y Weigt [27] encontraron numéricamente que la diferencia entre $Cc^{(WS)}(p)$ y $Cc^{(Tr)}(p)$ es relativamente pequeña para el modelo de Watts–Strogatz. Por otro lado, cuando $p > 0$ dos vecinos de un nodo i que estaban conectados para $p = 0$, siguen siendo vecinos de i y vecinos entre si con probabilidad $(1 - p)^3$. Esto es así ya que $(1 - p)^3 + \mathcal{O}(1/n)$ es la probabilidad de que dichos links no hayan sido reconectados hacia otros nodos. Luego:

$$Cc^{(Tr)}(p) \simeq Cc^{(WS)}(0)(1 - p)^3 \quad (2.34)$$

donde se usó el hecho de que $Cc_i^{(WS)}(0) = Cc^{(WS)}(0)$ para todo i . Posteriormente Newman [137] para un modelo similar[†] (el modelo de Newman–Watts [149] que es una versión de campo medio aunque no equivalente) encuentra la siguiente relación en el límite $n \rightarrow \infty$:

$$Cc^{(Tr)} = \frac{3(K - 2)}{4[(K - 1) + Kp(p + 2)]} \quad (2.35)$$

Es importante resaltar que las ecuaciones (2.34) y (2.35) presentan tendencias similares, sin embargo difieren significativamente. La primera de ellas tiende a cero con $p \rightarrow 1$, mientras que la segunda tiende al valor $3(K - 2)/4(4K - 1) \neq 0$.

Observación: En el modelo de Watts–Strogatz existe un rango de valores de p en donde coexisten una alta clusterización $Cc(p) \simeq Cc(0)$ y pequeñas distancias topológicas similares a las de las redes aleatorias $L(p) \simeq L(1)$ (ver figura (2.4)). La coexistencia de alta clusterización (respecto de las presentadas por redes aleatorias) y pequeñas distancias topológicas (similares a las de redes aleatorias), constituye la formalización proveída por Watts–Strogatz[204] de **las propiedades de mundo pequeño**. En otras palabras, en las redes de Watts–Strogatz una pequeña fracción de aleatoriedad en el conexionado de los links es suficiente para acortar significativamente las distancias topológicas sin destruir la presencia de estructura.

La distribución de grados del modelo de Watts–Strogatz

En el modelo de Watts–Strogatz para $p = 0$ todos los nodos tienen el mismo grado. Como resultado de ello, la distribución de grados es una función delta centrada en $k = K$. Un valor no nulo de p introduce desorden en la red generando variabilidad en el grado que cada nodo puede tener, manteniéndose la media $\langle k \rangle = K$. Dado que sólo un extremo

*Pero respetando que $K \ll n$

[†]En el Review de Barabási hay un pequeño error en la ecuación (73). Ocurre que Barrat–Weigt y Newman utilizan en las fórmulas $k = K/2$. En dicha ecuación del Review en ves de K lo correcto es k .

de cada link es reconectado, cada nodo conserva como mínimo $K/2$ links. Barrat y Weigt calcularon analíticamente la distribución de grados para el modelo de Watts–Strogatz encontrando que coincide muy bien con los datos calculados numéricamente [27]:

$$P(k) = \sum_{n=0}^{\min(k-K/2, K/2)} \binom{K/2}{n} (1-p)^n p^{K/2-n} \frac{(pK/2)^{k-K/2-n}}{(k-K/2-n)!} e^{-pK/2} \quad (2.36)$$

para $k \geq K/2$. La forma de la distribución de grados del modelo de Watts–Strogatz es similar a la de un grafo aleatorio (ver figura (2.5)). La misma posee un pico alrededor de $\langle k \rangle = K$ y decae exponencialmente para valores grandes de k . En otras palabras, todos los nodos poseen mas o menos el mismo grado.

Observación: El modelo de Watts–Strogatz provee de una explicación posible a la propiedad de mundo pequeño. Sin embargo la distribución de grados del modelo de Watts–Strogatz no explica el comportamiento ley de potencias de las distribuciones de grados de muchas redes empíricas.

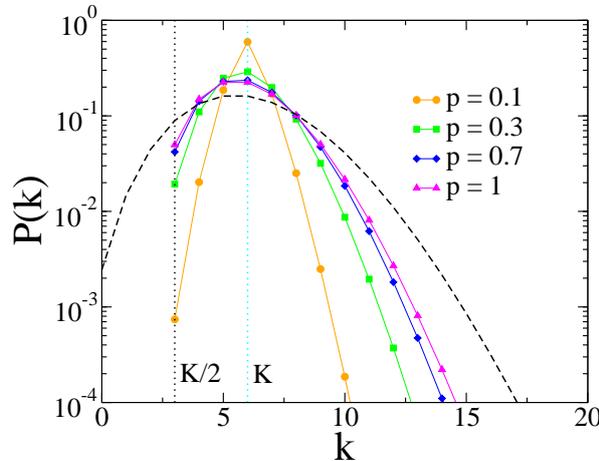


Figura 2.5: Distribución de grados del modelo de Watts–Strogatz para $K = 6$ y diferentes valores de p según la ecuación (2.36). Las líneas de puntos de colores negra y cian indican los valores de $k = K/2$ y $k = K$ respectivamente. El valor $k = K/2$ indica el mínimo grado que los nodos de las redes de Watts–Strogatz pueden tener, mientras que $\langle k \rangle = K$. La línea entrecortada corresponde a la distribución de grados de una red aleatoria de $n = 1000$ nodos, $\langle k \rangle = K$. Se aprecia que la distribución de grados del modelo de Watts–Strogatz es similar al de las redes aleatorias. Esta figura es una reproducción similar a la figura 2 de [27].

2.4.2. La distancia topológica en redes reales

Las consideraciones teóricas entorno al modelo de Watts–Strogatz ponen en evidencia el hecho de que aún cuando las redes presentan gran estructura – incluso estructura favorable a incrementar L – basta un poco de aleatoriedad para reducir drásticamente el valor de L asemejándolo al de una red aleatoria. Es de esperar entonces que en las redes de origen empírico dicha cualidad se conserve. En la tabla (2.2) se comparan los valores de L de las redes reales, con el correspondiente valor para una red aleatoria L_{rand} (ecuación (2.20)) y una red aleatoria generalizada de distribución de grados libre de escala L_{pow} (ecuación (2.22)). Se puede apreciar que los 3 valores resultan comparables confirmando que las redes empíricas presentan distancias topológicas similares al de redes aleatorias. Sin embargo, también es de esperar que las redes de origen empírico posean algún tipo de estructura, dando lugar a la posibilidad de manifestarse en alguna pequeña variación en la comparación de L con lo predicho por las ecuaciones para redes aleatorias generalizadas. Un análisis un poco más fino permite resaltar la pequeña diferencia. A partir de la ecuación (2.22) se obtiene que:

$$A(L_{pow} - 1) - B = \log_{10} n \quad (2.37)$$

donde:

$$A = \log_{10} \left[\frac{Li_{\gamma-2}(e^{-1/\kappa})}{Li_{\gamma-1}(e^{-1/\kappa})} - 1 \right] \quad \text{y} \quad B = \log_{10} \left[\frac{Li_{\gamma}(e^{-1/\kappa})}{Li_{\gamma-1}(e^{-1/\kappa})} \right] \quad (2.38)$$

En dichas expresiones se han utilizado logaritmos en base 10 en vez de logaritmos naturales con el fin de dar una mejor idea de los órdenes de magnitud a la hora de graficar los datos. Si los datos correspondientes a redes empíricas han de satisfacer las predicciones correspondientes a la ecuación (2.22), luego $A(L - 1) - B$ debería dar $\log_{10} n$ para los diferentes valores de n que cada red posee. La figura (2.6) muestra una comparación entre $A(L - 1) - B$ y $\log_{10} n$ en función de n . Se observa que los datos empíricos siguen la tendencia, sin embargo hay una discrepancia sistemática en donde los datos empíricos resultan levemente mayores.

Observación: La predicción de los modelos de redes aleatorias para las distancias topológicas L de las redes empíricas son muy buenas, aunque es apreciable una pequeña desviación sistemática consistente con lo observado en el modelo de Watts–Strogatz [204] en donde L crece levemente con la presencia de estructura. En resumen las redes reales exhiben la propiedad de *mundo pequeño*, es decir, distancias topológicas comparables a las de correspondientes redes aleatorias y un importante grado de estructura evidenciado por la alta clusterización (sección 2.3.2).

Red	n	L	L_{rand}	L_{pow}	referencia
WWW, site level, undir.	153127	3.1	3.35	-	[6]
WWW	325729	11.2	8.32	4.77	[11]
WWW	2×10^8	16	8.85	7.61	[36]
Internet, domain	6374	3.72	6.73	-	[163]
Internet, router	3888	12.5	8.75	7.67	[71]
Internet, router	150000	11	12.8	7.47	[83]
Movie actors	212250	4.54	3.65	4.01	[23]
Movie actors	225226	3.65	3.69	-	[204]
Coauthors, MEDLINE	1520251	4.6	4.91	-	[143, 142, 141]
Coauthors, SPIRES	56627	4	2.12	1.95	[143, 142, 141]
Coauthors, NCSTRL	11994	9.7	7.34	-	[143, 142, 141]
Coauthors, neuro.	209293	6	5.01	3.86	[24]
Coauthors, math.	70975	9.5	8.2	6.53	[24]
Metabolic, E. Coli	778	3.2	3.32	2.89	[97]
Substrate graph, E. Coli	282	2.9	3.04	-	[200]
Reaction graph, E. Coli	315	2.62	1.98	-	[200]
Food web, Ythan estuary	134	2.43	2.26	1.71	[133]
Food web, Silwood Park	154	3.4	3.23	2	[133]
Words, cooccurrence	460902	2.67	3.03	-	[92]
Power grid	4941	18.7	12.4	-	[204]
<i>C. Elegans</i>	282	2.65	2.25	-	[204]

Cuadro 2.2: Comparación para diferentes redes de origen empírico de, su valor de L , el correspondiente a una red aleatoria L_{rand} calculado según la ecuación (2.20), y el correspondiente a una red aleatoria generalizada con distribución de grados libre de escala L_{pow} en los casos en que es posible calcularlo según la ecuación (2.22). En la tabla también se presentan los tamaños n para evidenciar el carácter mundo pequeño. Se puede apreciar que las predicciones de los modelos de redes aleatorias L_{rand} y L_{pow} presentan valores muy similares a los empíricos. Esta tabla es un resumen de las tablas I y II de [10].

2.5. La clusterización en función del grado de los nodos

La mayoría de los datos de las tablas (2.1) y (2.2), y los puntos de la figura (2.2) corresponden a las mismas redes. En otras palabras, muchas de las redes reales presentan simultáneamente la propiedad de ser libres de escala y de mundo pequeño[10]. Según lo que se ha visto hasta ahora, por un lado las redes de Watts–Strogatz reproducen la propiedad de mundo pequeño, y por otro lado el modelo de Barabási–Albert reproduce la propiedad libre de escala. Sin embargo, las redes de Watts–Strogatz no explican las distribuciones de grados libres de escala (ecuación (2.5)), y las redes de Barabási–Albert no predicen el comportamiento de la clusterización de las redes de mundo pequeño (ecuaciones (2.27) y (2.28)).

La cuestión de la alta clusterización en redes libres de escala puede entenderse mejor al estudiar el coeficiente de clusterización promedio $Cc(k)$ como función del grado de los nodos k . Aquí interesa en particular el caso en que Cc se calcula según la definición de Watts–Strogatz, por lo que la dependencia de la clusterización con el grado de los nodos queda

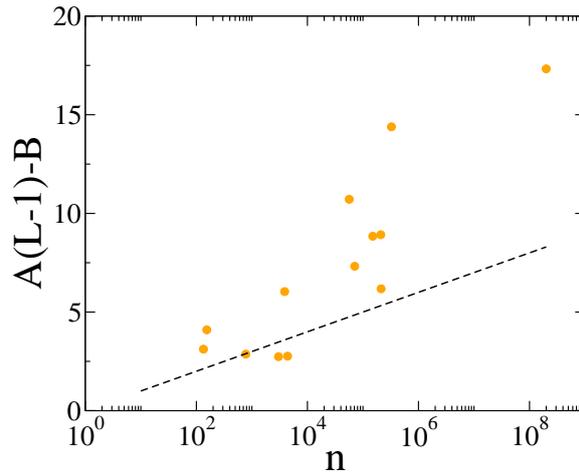


Figura 2.6: Comparación de L entre redes empíricas con distribución de grados libres de escala, y la predicción correspondiente a la ecuación (2.22). Mas precisamente, para cada red de la tabla (2.2) en que se ha calculado L y L_{pow} , se calculan A, B de acuerdo a las ecuaciones (2.38) utilizando datos extraídos de la tabla II de [10]. Luego, con dichos valores se grafica $A(L-1) - B$. La línea discontinua corresponde a graficar $\log_{10}(n)$ versus n . Dicho comportamiento es el que se esperaría cuando se satisface la ecuación (2.22). Los valores de los datos empíricos parecen seguir la tendencia de la línea discontinua para valores chicos de n pero se alcanza a observar que son sistemáticamente mayores para tamaños más grandes. Esta figura es una reproducción de la figura 13 de [10].

definida por la expresión:

$$C_c^{(WS)}(k) = \left\langle \frac{\sum_i C c_i^{(WS)} \delta_{k_i, k}}{\sum_i \delta_{k_i, k}} \right\rangle \quad (2.39)$$

donde las sumatorias son sobre todos los nodos, k_i es el grado del nodo i , y el promedio $\langle \cdot \rangle$ es sobre todo el ensamble de redes *.

2.5.1. La organización jerárquica de Ravasz *et. al.*

Estudiando la topología de redes metabólicas Ravasz *et. al.*[176] sugiere que la presencia simultánea de relativamente alta clusterización $C_c^{(WS)}$ y distribuciones de grados tipo ley de potencias, constituye fuerte evidencia de que las redes están compuestas por módulos altamente clusterizados organizados de manera jerárquica. La alta clusterización es consistente con que la red esté formada por módulos constituidos por nodos altamente conectados entre sí. Por otro lado la distribución de gra-

*Siendo matemáticamente mas precisos, habría que modificar la definición de $C_c^{(WS)}(k)$ de modo de excluir los términos en que $\sum_i \delta_{k_i, k} = 0$. En otras palabras evitar las divisiones por cero al considerar valores de k que no corresponden al grado de ningún nodo. Por cuestiones de simplicidad de notación se omite dicho detalle.

dos ley de potencias indica la presencia de *hubs** los cuales cumplirían la función de conectar estos módulos manteniendo la red unida. Para ilustrar lo que posiblemente podría estar ocurriendo introducen un modelo determinista de red†. En dicho modelo la red se construye partiendo de un módulo de cuatro nodos altamente interconectados. Luego se hacen tres copias de dicho modulo y se conectan los nodos externos de dichas copias al nodo central del módulo original. De esta manera se obtiene un módulo de 16 nodos menos densamente conectado que el original. A la red así obtenida se le pueden hacer 3 copias y conectarlas de manera análoga a la anterior, con la posibilidad de repetir el proceso la cantidad de veces que se desee (ver figura (2.7) para un ejemplo de 3 iteraciones). Este proceso recursivo da lugar a una red cuya distribución de grados es una ley de potencias en donde $\gamma = 1 + (\ln 4)/(\ln 3) \simeq 2.26$, y su coeficiente de clusterización es $C_c^{(WS)} \simeq 0.6$ independiente de n . Este modelo sugiere que es posible obtener una red de distribución de grados ley de potencias y alta clusterización, cuando la misma presenta una estructura modular organizada jerárquicamente. En particular, si en dicho modelo se estudia la dependencia de la clusterización con el grado de los nodos se encuentra la siguiente ley:

$$C_c^{(WS)}(k) \sim k^{-1} \quad (2.40)$$

indicando que los clusters que integran los nodos son menos “apretados” cuando mayor es el grado de los mismos. En particular encuentran el comportamiento de la ecuación (2.40) en las 43 redes metabólicas que estudiaron (gráfico de la derecha en la figura (2.8)), las cuales evidencian una organización modular jerárquica según otros medios. En otro trabajo, Ravasz y Barabási[176] estudian $C_c^{(WS)}(k)$ para las siguientes redes: una red de colaboración de actores extraída de www.imdb.com, en donde dos actores están conectados si actuaron juntos en alguna película; una red sinónimos en donde dos palabras están conectada si figuran como sinónimos en el *Merrian Webster Dictionary*; en un muestreo de la WWW obtenida mapeando el dominio www.nd.edu; y en una subred a nivel de interdominios o sistema autónomo de Internet. Encuentran que la clusterización en función del grado de los nodos es bien representada por la ecuación (2.40), y argumentan en favor de presencia de modularidad jerárquica en las mismas. Por otro lado, para Internet a nivel routers que presenta una distribución de grados ley de potencias, y para una red de alimentación eléctrica‡ que posee una distribución de grados exponencial, se tienen que $C_c^{(WS)}(k)$ es constante. Lo mismo ocurre en el modelo de Barabási–Albert[23, 176], el cuál particularmente presenta una relativamente baja clusterización. Esto pone de manifiesto que no necesariamente la clusterización decae con el grado de los nodos, y por

*Nodos con $k \gg \langle k \rangle$.

†Es decir las redes resultantes a partir de dicho modelo no son el producto de un proceso estocástico.

‡Del oeste de los EE.UU.

el contrario suele hacerlo en redes que evidencian algún tipo de organización modular jerárquica. Más ejemplos empíricos pueden encontrarse en [199] donde se puede observar que diferentes redes presentan diversos comportamientos de la función $Cc^{(WS)}(k)$, algunos constantes y otros asemejan leyes de potencia.

Diferentes modelos de redes libres de escala predicen el comportamiento de la ecuación (2.40) [105, 199, 135, 81], sin embargo el exponente -1 no es necesariamente universal. En [53] se observa que para una red de interacción de proteínas ocurre que:

$$Cc^{(WS)}(k) \sim k^{-\beta} \quad (2.41)$$

con $\beta \simeq 0.48$ (gráfico de la izquierda en la figura (2.8)). Por otro lado en [209] redes de proteínas con distribuciones de grado libres de escala presentan el comportamiento de la ecuación (2.41) con $\beta \simeq 2$. Alternativamente, Ravasz y Barabási[176] proveen de un simple modelito de red libre de escala en donde β puede ser regulado entre 1 y 2, mientras que en otro modelo de red basado en dinámica Hamiltoniana $\beta \in [0, 1]$ [20].

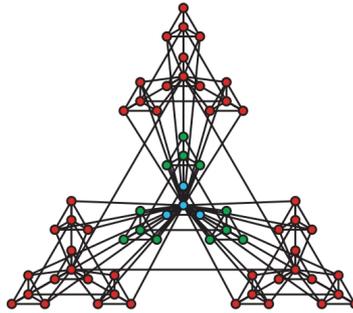


Figura 2.7: Red determinística que presenta organización jerárquica construída a partir de la receta proveída en [176]. Se observan pequeños módulos densamente interconectados por lo que sus nodos presentan alta clusterización (módulos más hacia la periferia). Y al mismo tiempo módulos que contienen nodos altamente conectados que poseen baja clusterización (módulos más hacia el centro). La figura fué extraída de [176](figura 1C).

2.6. La correlación entre los grados de nodos vecinos: Asortatividad

En general los grados de nodos vecinos pueden presentar algún tipo de correlación. Una posibilidad es que los nodos muy conectados tiendan a conectarse con otros nodos muy conectados (y *vice versa*). Cuando ello ocurre se dice que la red presenta una “mezcla” selectiva o *asortativa* del grado de los nodos vecinos*. Si ocurre lo opuesto, es decir nodos muy co-

*Del inglés: *assortative mixing by degree*.

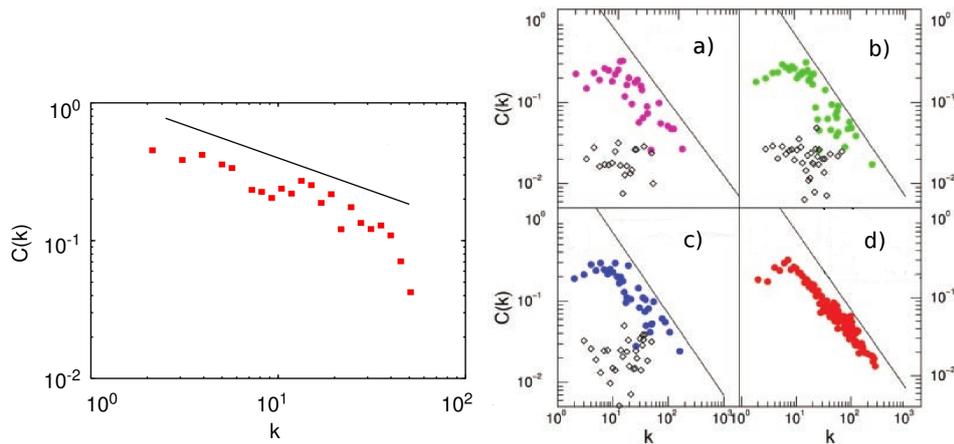


Figura 2.8: **Izquierda:** Se observa el coeficiente de clusterización en función del grado de los nodos $C_c(k)$ para una versión de la red de interacciones proteína–proteína de la levadura *S. cerevisiae* (cuadrados rojos). La línea continua satisface $C_c(k) \sim k^{-0.48}$ y corresponde al ajuste de los datos empíricos. La figura fué extraída de [53] (gráfico II izquierdo de la figura 3). **Derecha:** En **a)**, **b)** y **c)** se observa $C_c(k)$ para las redes metabólicas de distintos organismos (círculos llenos de diferentes colores) y una comparación con $C_c(k)$ de correspondientes redes libres de escala obtenidas a partir del modelo de Barabási–Albert[23] (rombos vacíos) en donde $C_c(k) \sim \text{constante}$. Más precisamente: **a)** (color magenta) corresponde a la célula procariota *Aquifex aeolicus*; **b)** (color verde) corresponde a *E. Coli*; **c)** (color azul) corresponde a *S. cerevisiae*. En **d)** (color rojo) se presenta al promedio de $C_c(k)$ de las 43 redes metabólicas de los diferentes organismos estudiados en [176]. En todos los casos la línea continua corresponde a un comportamiento $C_c(k) \sim k^{-1}$. La figura es una adaptación de la figura 2 en [176].

nectados tienden a ser vecinos de nodos poco conectados entonces se dice que la red presenta una “mezcla” *disasortativa* de los grados de los nodos vecinos. Con el fin de simplificar el vocabulario simplemente se dirá que correlaciones positivas entre los grados de los nodos vecinos corresponde a redes *asortativas*, mientras que correlaciones negativas corresponde a redes *disasortativas*. Una interpretación intuitiva en términos de una red social es que en una red asortativa las personas sociables tienden a relacionarse con personas sociables (y *vice versa*).

A continuación es abordada la cuestión de como caracterizar el grado de asortatividad (ó disasortatividad) de una red (ó ensamble de redes). Se discutirán dos formas de hacerlo. La primera de ellas consiste en una versión del *coeficiente de correlación de Pearson* entre el grado de los nodos vecinos de una red. Más precisamente, considérese una única red. Si uno elige al azar un nodo en dicha red, luego la probabilidad de que este nodo posea k links viene dada por la distribución de grados $P(k)$ de la red. Por otro lado supóngase que se selecciona un link al azar, y luego se elige también al azar un nodo en uno de sus dos extremos. La probabilidad de que éste nodo posea k links no es $P(k)$, sino que es proporcional a $kP(k)$ [144]. Esto es esperable ya que nodos de grados

altos acaparan más links que nodos con pocas conexiones. Si el nodo al extremo de un link elegido al azar tiene grado k su *grado restante* es $k - 1^*$, por lo que la probabilidad asociada al *grado restante* k de un nodo en el extremo de un link elegido al azar es:

$$q_k = \frac{(k+1)P(k)}{\sum_j jP(j)} \quad (2.42)$$

La probabilidad conjunta de los grados restantes j y k de los dos nodos en los extremos del link elegido al azar es denotada por $e_{j,k} = e_{k,j}$ y satisface:

$$\sum_{j,k} e_{j,k} = 1 \quad \sum_j e_{j,k} = q_k \quad (2.43)$$

En una red en la que no hay correlación entre los grados restantes de los nodos al extremo de los links, ocurre que $e_{j,k} = q_j q_k$. Una tendencia en la correlación entre j y k , puede ser determinada por la correlación conectada entre los grados restantes de los nodos vecinos de una red, es decir:

$$\langle jk \rangle - \langle j \rangle \langle k \rangle = \sum_{j,k} jk(e_{j,k} - q_j q_k) \quad (2.44)$$

donde el promedio $\langle \cdot \rangle$ se calcula sobre todo los links de la red. Con el fin de eventualmente comparar el grado de asortatividad de diversas redes, es conveniente normalizar la expresión de la ecuación (2.44) dividiéndola por el máximo valor que puede alcanzar. Este máximo es igual a la varianza de la distribución q_k [144]:

$$\sigma_q^2 = \sum_k k^2 q_k - (k q_k)^2 \quad (2.45)$$

que corresponde a la expresión (2.44) para una red máximamente asortativa (i.e. $e_{jk} \propto \delta_{jk}$). Luego, la correlación conectada y normalizada entre el grado remanente de los nodos vecinos en una red es [144]:

$$r = \frac{1}{\sigma_q^2} \sum_{jk} kj(e_{jk} - q_j q_k) \quad (2.46)$$

Ésta expresión corresponde al coeficiente de correlación de Pearson del grado de los nodos vecinos en una red. Tal magnitud está confinada a los valores $-1 \leq r \leq 1$.

Observación: Valores positivos de r corresponden a redes asortativas, y negativos a redes disasortativas.

En la práctica, si uno posee una red de n nodos y M links, r puede obtenerse a partir de la expresión [144]:

*Definición.

$$r = \frac{M^{-1} \sum_l j_l k_l - [M^{-1} \sum_l \frac{1}{2}(j_l + k_l)]^2}{M^{-1} \sum_l \frac{1}{2}(j_l + k_l) - [M^{-1} \sum_l \frac{1}{2}(j_l + k_l)]^2} \quad (2.47)$$

donde las sumatorias son sobre todos los links de la red, y el par (j_l, k_l) son los grados de los nodos en los extremos del link l^* .

Otra forma que ha sido utilizada para determinar la correlación entre los grados de nodos vecinos consiste en analizar el comportamiento de el grado promedio $k_{nn}(k)$ de los nodos primeros vecinos de los nodos de grado k [†] [163, 197]. Más precisamente, la función $k_{nn}(k)$ viene definida por:

$$k_{nn}(k) = \left\langle \sum_i \delta_{k_i, k} \frac{\sum_{j \in \vartheta_i} k_j}{\sum_{j \in \vartheta_i} 1} \right\rangle \quad (2.48)$$

donde $\langle \cdot \rangle$ es sobre un ensamble de redes, k_i es el grado del nodo i , y ϑ_i es el conjunto de primeros vecinos del nodo i [‡].

Observación: Si el comportamiento de la función $k_{nn}(k)$ es ascendente, entonces en promedio los nodos más conectados tienden a conectarse con otros muy conectados lo cuál corresponde a redes asortativas. Por el contrario, un comportamiento descendente corresponde a redes disasortativas. Cuan acentuado es el comportamiento creciente (decreciente) determina cuál es el grado de asortatividad (disasortatividad) de la red en cuestión.

2.6.1. La asortatividad/disasortatividad en redes reales

M. E. J. Newman en un review [146] presenta un resumen del estudio del coeficiente de correlación de Pearson r en redes de diversos orígenes. En la tabla (2.3) se muestran los valores de r para las diferentes redes. Newman hace una distinción entre las redes sociales y el resto resaltando que para las primeras r resulta positivo, mientras que para las segundas resulta negativo. En particular, los resultados de Newman implican que las redes biológicas presentan disasortatividad.

Al intentar realizarse la clasificación asortatividad/disasortatividad utilizando $k_{nn}(k)$ frecuentemente se obtienen resultados equivalentes a los obtenidos por Newman. Por ejemplo, en [53] se encuentra que para la red de interacciones proteína-proteína de la levadura *S. cerevisiae*, $k_{nn}(k)$ presenta un comportamiento descendente (ver figura 2.9) en consistencia con la clasificación disasortativa que Newman atribuye a las redes

*Notar que se han aproximado los grados remanentes de los nodos por sus grados. Según Newman dicha aproximación no afecta los resultados [144].

[†]La expresión *nn* proviene del inglés *nearest-neighbors*.

[‡]Siendo matemáticamente más precisos, habría que modificar la expresión que define $k_{nn}(k)$ para excluir los términos en donde $\sum_{j \in \vartheta_i} 1 = 0$, evitándose así la división por cero. Dicho detalle se lo omite por cuestiones de simplicidad de notación.

biológicas. Sin embargo también pueden observarse escenarios más complicados en donde los nodos pueden presentar una tendencia asortativa ó disasortativa dependiendo del grado de los mismos. Mas precisamente, Piraveenan *et. al.* [168, 171] encuentra que el coeficiente de correlación de Pearson del grado de los nodos (ecuación (2.46)) puede escribirse como una suma de contribuciones r_i de cada nodo i de la red:

$$r = \sum_i r_i \quad (2.49)$$

donde:

$$r_i = \frac{j_i(j_i + 1)(\bar{k}_i - \mu_q)}{2M\sigma_q^2} \quad (2.50)$$

donde j_i es el grado del nodo i , \bar{k}_i el el grado promedio de los primeros vecinos del nodo i , $\mu_q = \langle kq_k \rangle = \sum_k kq_k$ es una magnitud independiente de i , M es el número de links de la red, y σ_q fué definido en la ecuación (2.45). Según Piraveenan *et. al.* si el valor de r_i es positivo (negativo) luego el nodo i es asortativo (disasortativo). A r_i se la denomina *asortatividad local* del nodo i y a partir de ella se puede definir la asortatividad local (promedio*) en función del grado de los nodos[†]:

$$r(k) = \frac{\sum_i \delta_{k_i,k} r_i}{\sum_i \delta_{k_i,k}} \quad (2.51)$$

En algunos casos se pueden observar comportamientos mixtos en donde un grupo de nodos presenta una asortatividad contraria a la global. Por ejemplo, en [170] Piraveenan *et. al.* mencionan que la red de interacciones proteína–proteína del *H. Sapiens* es una red asortativa con *hubs* disasortativos ya que $r = 0.075$, mientras que $r(k)$ es negativo para valores grandes de k (ver figura (2.10))[‡].

Como detalles adicionales, las redes aleatorias y el modelo de Barabási–Albert[23] poseen $r = 0$ (al menos en el límite $n \rightarrow \infty$)[144]. Modelos inspirados en redes sociales reproducen valores de r positivos[38, 151]. Y argumentos del tipo entrópico (principio de parsimonia) han sido utilizados para explicar porqué la mayoría de las redes son disasortativas[99]. En otras palabras, las redes asortativas ($r > 0$) no son probabilísticamente favorecidas en el ensamble de redes libres de escala.

*Menciono promedio ya que $r(k)$ es el valor promedio de las asortatividades locales r_i de los nodos de grado k .

[†]Al igual que con anteriores definiciones, en la definición de la ecuación (2.51) se está asumiendo que existe al menos un nodo de grado k para cada valor de k . Si ese no fuera el caso ocurrirían divisiones por cero en la ecuación. Para cubrir la posibilidad en que ello ocurriese, se podría apelar a una definición más sofisticada pero no se lo hace por cuestiones de simplicidad.

[‡]Es importante resaltar que la expresión de la ecuación (2.50) fué inicialmente propuesta de otra forma [168]. Luego dicha expresión es corregida [171] y es la que aquí se presenta. En [169] hay un gráfico análogo (figura 4) al de la figura (2.10) pero realizado con la expresión antes de corregirse, por lo que la conclusión es diferente.

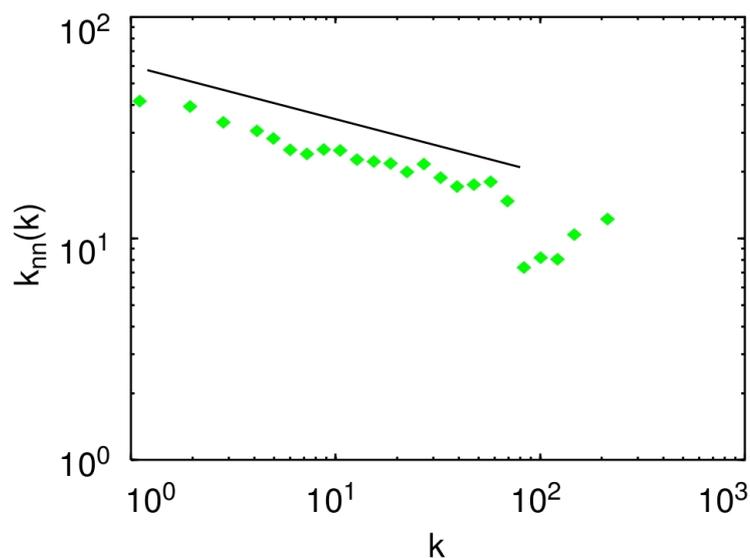


Figura 2.9: Caracterización de la asortatividad/disasortatividad para una versión de la red de interacciones proteína–proteína de la levadura *S. cerevisiae* utilizando la función $k_{nn}(k)$ (ver ecuación (2.48) y texto relacionado). Se observa que $k_{nn}(k)$ decae como una ley de potencias $\sim k^{-\delta}$ en donde $\delta \simeq 0.24$. La tendencia decreciente corresponde a un carácter disasortativo de la red. Ésta figura fué extraída de [53] (gráfico III derecho de la figura 3 en dicho trabajo).

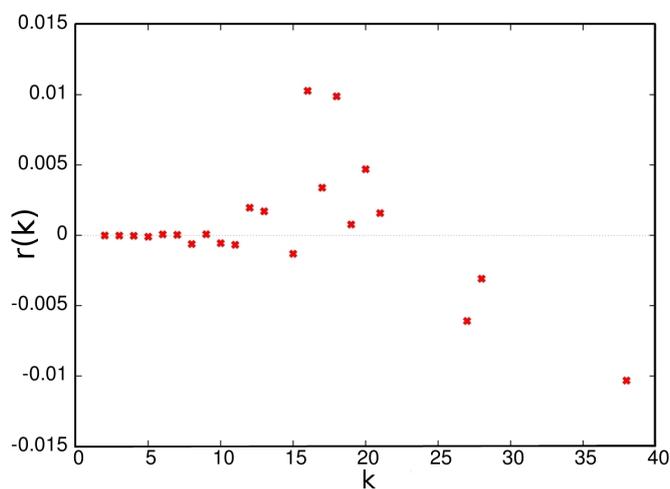


Figura 2.10: Coeficiente de asortatividad local en función del grado de los nodos $r(k)$ (ecuación (2.51)) en función de k para la red de interacción de proteínas del *H. Sapiens*. El coeficiente de correlación de Pearson a nivel global (ecuación (2.46)) indica que la red es globalmente asortativa ($r = 0.075$). Por otro lado puede observarse que $r(k)$ es negativo para valores grandes de k indicando un carácter disasortativo en los *hubs* de la red. Esta figura fué extraída de [170] (figura 3).

2.6.2. Algunas consideraciones teóricas sobre la asortatividad

Newman y Park [151] mencionan que en redes pequeñas se genera una tendencia disasortativa si se restringe a dichas redes a poseer no más de un link entre cada par de nodos. Esta cuestión es confirmada numéricamente en [123] y analíticamente en [160]. La idea es que en la definición del coeficiente de correlación de Pearson r (ecuación (2.46)), el grado de asortatividad se mide en relación al modelo nulo utilizado que es el *modelo configuracional* (ver sección 2.2.1). En el modelo configuracional se permiten redes con múltiples conexiones, es decir, se permite más de un link por cada par de nodos. Un modelo de red que evita las conexiones múltiples inevitablemente introduce correlaciones que se reflejan disminuyendo r y por ende provocando tendencias disasortativas. La probabilidad de que aparezcan conexiones múltiples en el modelo configuracional está asociada a la relación entre los grados de los nodos más conectados y el tamaño de la red. Para observar conexiones múltiples en el modelo configuracional, los nodos con los grados más altos tiene que tener grados alrededor de \sqrt{n} para una red de tamaño n [151]. Para una red libre de escala en donde $P(k)$ posee una cola larga, la anterior condición se puede aproximar por $k_{max} \gtrsim \sqrt{n}$ donde k_{max} es el grado máximo de entre todos los nodos de la red. Teniendo en cuenta que para las redes libres de escala k_{max} es en general significativamente mayor que para las correspondientes redes aleatorias, luego la tendencia disasortativa es particularmente acentuada para las redes libres de escala. Por ejemplo, en el modelo configuracional con distribución ley de potencias $P(k) \sim k^{-\gamma}$ se tiene que $k_{max} \sim n^{1/(\gamma-1)}$ [151] indicando que para $\gamma < 3$ tendencias disasortativas se observarían a todos los tamaños si se evitasen las conexiones múltiples.

Por otro lado, Newman y Park [151] usando estimaciones analíticas sobre el mencionado modelo configuracional concluyen que el coeficiente de clusterización $Cc^{(Tr)}$ de las redes disasortativas puede ser bien explicado (al menos cualitativamente). Sin embargo, dichas estimaciones de $Cc^{(Tr)}$ resultan mucho menores que los valores reales para redes asortativas. Newman y Park [151] resaltan que las redes asortativas en general son redes sociales que presentan un significativo grado de *estructura de comunidades*^{*†}. Ellos argumentan utilizando un modelo que, dicha acentuada organización en comunidades es el nexo entre una alta clusterización no explicada por el modelo configuracional y la asortatividad del grado de los nodos. En otras palabras, el grado de asortatividad y clusterización pueden presentar vínculos. Por ejemplo, en [210, 99] muestran que acentuar la disasortatividad eventualmente provoca que $Cc(k)$ decaiga con

*En la sección 2.5.1 se habla de *módulos* para hacer referencia al menos a un concepto similar.

†Una teoría alternativa es que los individuos suelen “presentar” sus conocidos a sus otros conocidos dando lugar a la formación de triángulos dentro de la red[151].

Clasificación	Red	r	referencia
Social	Movie actors	0.208	[204]
	Compay directors	0.276	[55]
	Math coauthorship	0.120	[42]
	Physics coautorship	0.363	[142]
	Biology coautorship	0.127	[142]
	Email address book	0.092	[147]
	Student relationships	-0.029	[31]
Information	WWW nd.edu	-0.067	[11]
	Roget's Thesaurus	0.157	[106]
Technological	Internet	-0.189	[71]
	Power grid	-0.003	[204]
	Train routes	-0.033	[181]
	Software packages	-0.016	[145]
	Software classes	-0.119	[195]
	Electronic circuits	-0.154	[91]
	Peer to peer network	-0.366	[8]
Biological	Metabolic network	-0.240	[97]
	Protein interactions	-0.156	[95]
	Marine food web	-0.263	[90]
	Fresh water food web	-0.326	[119]
	Neural network	-0.226	[204]

Cuadro 2.3: El coeficiente de correlación de Pearson r para el grado remanente de los nodos vecinos, para diferentes tipos de redes reales. Se observa una clara distinción entre redes sociales (que resultan en su mayoría asortativas) y el resto (que resultan en su mayoría disasortativas). La tabla es una réplica de la tabla 3.1 en [146].

k . Por otro lado, en [75] se concluye que exigir una alta clusterización provoca asortativización en la red. Recíprocamente asortativizar la red afecta la clusterización, aunque en un grado mucho menor, haciéndose más acentuado el efecto cuanto “más ancha” es la distribución de grados. También pone de manifiesto que aumentos en la clusterización provoca aumentos en la modularidad, mientras que aumento en la asortatividad tiene una mínima influencia.

2.7. Resistencia ante fallas aleatorias o ataques dirigidos en una red

Eventualmente es posible pensar un conjunto de redes cada una totalmente conectada*, como una sola red fragmentada. Cada elemento del conjunto de redes totalmente conectadas es llamada una *componente* de la red fragmentada. La componente con la mayor cantidad de nodos se llama la *componente gigante*†.

Dada una red de n nodos, sea n_g el número de nodos en la componente gigante. Luego sea:

*La definición de totalmente conectada fué dada en la sección 2.2.

†Del inglés *giant component*.

$$S = \frac{n_g}{n} \in (0, 1] \quad (2.52)$$

el tamaño relativo de la componente gigante respecto del número de nodos. Por otro lado sea:

$$\langle s \rangle \in [1, c] \quad , \text{donde } c = \begin{cases} n/2 & \text{si } n \text{ es par} \\ (n-1)/2 & \text{c.c.} \end{cases} \quad (2.53)$$

el promedio de los tamaños de las componentes distintas a la componente gigante. Considerense procesos consistentes en una secuencia de pasos en cada uno de los cuales se aísla un nodo por medio de removerle sus links. Ante una secuencia de *ataques dirigidos* se van removiendo sucesivamente los links a los nodos más conectados en orden descendiente del grado de los mismos. Ante una secuencia de *fallas aleatorias*, se van removiendo sucesivamente los links de una secuencia de nodos ordenada al azar. Por una cuestión de simplicidad, cuando se diga *un proceso de aislación de nodos* se estará haciendo referencia o bien a una secuencia de ataques dirigidos, o bien a una secuencia de fallas aleatorias. Si a un dado instante de un proceso de aislación de nodos la cantidad de nodos a los que se les han removido sus links es n_r , luego $n_r \leq n_a$, donde n_a es la cantidad de nodos que han quedado aislados en dicho instante del proceso. Mas precisamente, n_r es la cantidad de nodos que son forzados a aislarse (de entre los que aún no están aislados) a lo largo del proceso, mientras que n_a es la cantidad de nodos que quedan aislados como un efecto secundario de la aislación de los primeros. La fracción f de nodos a los que se les han removido sus links viene dada por:

$$f = \frac{n_r}{n} \quad (2.54)$$

Las funciones $S(f)$ y $\langle s \rangle (f)$ permiten caracterizar la respuesta de una red ante un proceso de remoción de nodos. En otras palabras, permiten caracterizar como y cuando se fragmenta una red ante la aislación sucesiva de algunos de sus nodos. Un nodo aislado representa un constituyente* del sistema que no está funcionando.

Las funciones $S(f)$ y $\langle s \rangle (f)$ fueron utilizadas por Albert *et. al.* [12],[13],[10] con el fin de caracterizar la tolerancia de redes de distintas topologías a ataques dirigidos o a fallas aleatorias. En dichos trabajos $S(f)$ y $\langle s \rangle (f)$ fueron promediados sobre un ensamble de redes. Las motivaciones eran por ejemplo, tratar de entender porqué organismos relativamente simples crecen, persisten y se reproducen a pesar de que drásticas perturbaciones ambientales o farmacéuticas los afectan. Una explicación plausible reside en la tolerancia o robustez topológica de las redes metabólicas o genéticas subyacentes a dichos organismos[10]. Una motivación equivalente es

*Normalmente usaría la palabra componente, pero aquí sería susceptible a confundirlo con el término “componente de un grafo”.

la aparente robustez que presentan diferentes redes de comunicación, en donde las fallas de algunos de sus constituyentes raramente conllevan a un colapso del funcionamiento de toda la red [10]. La robustez de dichas redes puede ser atribuible a un conexionado redundante, pero surge la pregunta de cómo afecta la topología de dicho conexionado a los distintos tipos de fallas. Por ésta razón, en Albert *et. al.* [12],[13],[10] se concentraron en caracterizar la diferencia en respuesta de redes aleatorias y redes de topología libre de escala ante ataques dirigidos y fallas aleatorias.

Si en un proceso de aislación de nodos la red nunca se fragmentase, luego ocurriría que $S = 1 - f$ y $\langle s \rangle = 1$ para toda instancia del proceso. En la práctica, para una ensamble de redes de tamaño finito $S \simeq 1 - f$ cuando $f \ll 1$, pero a medida que f crece el decaimiento de S se acentúa indicando que componentes de tamaños mayores a 1 se desprenden de la componente gigante. Esto a su vez se refleja en un crecimiento de $\langle s \rangle$ por encima de 1 (ver figura (2.11)). En general existe una fracción crítica f_c en donde la componente gigante adquiere el mismo tamaño que $\langle s \rangle$, y las redes del ensamble se fragmentan quedando compuestas por muchas componentes aisladas de tamaños similares y pequeños. En dicho punto $\langle s \rangle$ alcanza un pico (máximo) y luego decae para $f > f_c$. En el caso de calcular las funciones $S(f)$ y $\langle s \rangle(f)$ para una única red (en vez de promediarlas sobre un ensamble) el comportamiento es análogo aunque no es suave sino que ruidoso. Lo importante a resaltar aquí es que f_c depende del tipo de proceso de aislación de nodos – ataques dirigidos, ó fallas aleatorias – y de la topología de la red. Las redes aleatorias poseen un f_c mayor que el de las redes libres de escala ante ataques dirigidos. Por el contrario, poseen un f_c menor ante ataques aleatorios.

Observación: El tema ha sido estudiado con mucha más profundidad que con la que es expuesta aquí. Por ejemplo, existen resultados analíticos a tamaño infinito en donde se demuestra que las redes libres de escala no fragmentan nunca ante fallas aleatorias si $\gamma < 3$ [51, 39]. La intención aquí no es resaltar las propiedades del proceso de fragmentación de redes a tamaño infinito. Sino poner en evidencia que f_c sirve para distinguir entre redes con distintas distribuciones de grados. En la sección 2.8.4 y el apéndice A se discute un poco más la cuestión pero en relación a la robustez en redes ecológicas.

2.8. La densidad de conexiones en una red: “Conectividad”

La *conectividad* o *densidad de conexiones* en una red viene definida por:

$$C^{(un)} = \frac{M}{\frac{n(n-1)}{2}} \quad (2.55)$$

donde M es el número de links presentes en la red y n es su número de

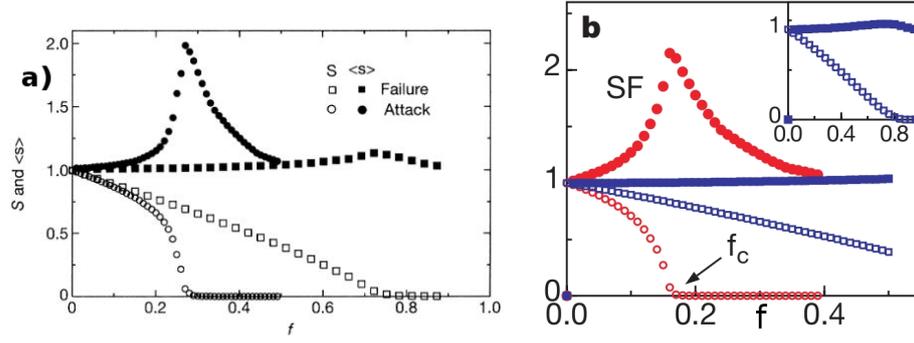


Figura 2.11: En la figura **a)** se pueden apreciar el proceso de fragmentación en redes aleatorias ante ataques dirigidos (cuadrados) o fallas aleatorias (círculos) a través del comportamiento de $S(f)$ (símbolos vacíos) y $\langle s \rangle(f)$ (símbolos llenos) para un ensamble de redes aleatorias de $n = 10000$ y $\langle k \rangle = 4$. Se observa que para $f \ll 1$, S decae como $\simeq 1 - f$, mientras que $\langle s \rangle \simeq 1$. Pero a medida que f crece la caída en $S(f)$ se acentúa colapsando rápidamente a $S \simeq 0$ en f_c . Para el ataque dirigido $f_c \simeq 0.28$, mientras que para las fallas aleatorias $f_c \simeq 0.66$. En ambos casos simultáneamente se observa un pico en $\langle s \rangle$ indicando que los fragmentos que se desprenden de la componente gigante no son nodos aislados. En **b)** se observa el mismo análisis pero para redes libre de escala correspondientes al modelo de Barabási–Albert [23]. El comportamiento de las curvas es similar al observado para las redes aleatorias, pero con valores distintos de las fracciones críticas. Ante ataques dirigidos $f_c \simeq 0.18$, mientras que ante un fallas aleatorias $f_c \simeq 0.8$ (inset). En resumen, las redes libres de escala son más resistentes a fallas aleatorias. Por otro lado, las redes aleatorias son más resistentes a ataques dirigidos. Las figuras fueron extraídas de [12],[13].

nodos. $n(n - 1)/2$ es el máximo valor que M puede alcanzar, de ahí el término densidad de conexiones. El supraíndice (*un*) denota *undirected*^{*} y corresponde a que esta definición de la conectividad es válida para redes no dirigidas. Dado que el grado medio de una red viene dado por $\langle k \rangle = 2M/n$, luego existe una simple relación entre la conectividad y el grado medio:

$$C^{(un)} = \frac{\langle k \rangle}{n - 1} \simeq \frac{\langle k \rangle}{n} \quad \text{cuando } n \gg 1 \quad (2.56)$$

Debido a esta simple relación, si la conectividad decae con el tamaño más rápido que $\sim 1/n$, luego la red a medida que crece se va diluyendo en conexiones. Eventualmente ocurrirá que $\langle k \rangle < 1$ con lo cuál la red inevitablemente se fragmentará.

En el caso de una red dirigida la definición de la conectividad es:

$$C^{(d)} = \frac{Q}{n(n - 1)} \quad (2.57)$$

donde (*d*) denota *directed*[†] y Q es la cantidad de conexiones dirigidas que la red posee de entre el máximo posible $n(n - 1)$ [‡]. Sea $k_i^{(t)} = k_i^{(in)} + k_i^{(out)}$,

*No dirigida

†Dirigida.

‡Se están omitiendo las autoconexiones.

es decir la suma de las conexiones entrantes y salientes al nodo i . Por otro lado sea $c_{i,j}$ la matriz de adyacencia a la red dirigida. Es decir, $c_{i,j} = 1$ si hay una conexión dirigida desde j hacia i , y vale cero en caso contrario. Entonces es posible derivar una expresión análoga a la ecuación (2.56) para $C^{(d)}$, a saber:

$$\begin{aligned}
C^{(d)} &= \frac{1}{n(n-1)} \sum_{i,j} c_{i,j} \\
&= \frac{1}{2n(n-1)} \sum_{i,j} c_{i,j} + c_{j,i} \\
&= \frac{1}{2n(n-1)} \sum_i \left[\sum_j c_{i,j} + c_{j,i} \right] \\
&= \frac{1}{2n(n-1)} \sum_i k_i^{(t)} \\
&= \frac{1}{2(n-1)} \langle k^{(t)} \rangle \\
&\simeq \frac{1}{2n} \langle k^{(t)} \rangle \quad \text{cuando } n \gg 1
\end{aligned} \tag{2.58}$$

Siempre es posible convertir una red dirigida a una no dirigida. Entre dos nodos cualesquiera hay dos casos posibles. En un caso hay una sola conexión dirigida que los une. En el otro los nodos están conectados por dos conexiones dirigidas, una para cada sentido. Cuando ello ocurre diremos que la *conexión es doble*. Supóngase que una red dirigida de Q conexiones dirigidas es convertida en una no dirigida resultando en M conexiones no dirigidas. Si $\zeta \in [0, 1]$ es la fracción de las M conexiones no dirigidas que originalmente correspondían a una conexión doble, luego se tiene:

$$\begin{aligned}
Q &= [2\zeta + (1 - \zeta)]M \\
&= (\zeta + 1)M \\
\frac{Q - M}{M} &= \zeta
\end{aligned} \tag{2.59}$$

o lo que es equivalente:

$$C^{(d)} = \frac{1}{2}(\zeta + 1)C^{(un)} \tag{2.60}$$

En otras palabras, para especificar completamente la relación entre $C^{(d)}$ y $C^{(un)}$ al variar n es necesario conocer también como varía ζ con n .

Además de ζ , una magnitud relacionada que puede ser de interés es la fracción de conexiones dobles η_i del nodo i :

$$\eta_i = \frac{k_i^{(t)} - k_i}{k_i} \tag{2.61}$$

donde k_i es el grado del nodo i -ésimo en la versión no dirigida de la red*. Luego se puede definir el promedio de las conexiones dobles por nodo de la red como:

$$\langle \eta \rangle = \frac{1}{n} \sum_i \eta_i \quad (2.62)$$

donde la suma es sobre todos los nodos de la red. Notar que en general:

$$\sum_i \eta_i \neq \frac{\sum_i k_i^{(t)} - k_i}{\sum_i k_i} = \frac{2Q - 2M}{2M} = \zeta \quad (2.63)$$

Por lo que, en general $\langle \eta \rangle \neq \zeta/n$, aunque para una red aleatoria vale la igualdad y ζ/n es reinterpretable como la proporción de conexiones dobles por nodo. Al igual que ocurre con $Cc^{(WS)}$ y $Cc^{(Tr)}$ (ver sección 2.3), la diferencia entre $\sum_i \eta_i$ y ζ corresponde a invertir el orden de las operaciones suma y división, y sólo la aparición de correlaciones genera la desigualdad. La igualdad $\langle \eta \rangle = \zeta/n$ para redes aleatorias es probada numéricamente en la figura (4.37) del Capítulo 4.

Observaciones: Por un lado, como se verá mas adelante, la conectividad ha sido un factor de estudio ya que se ha relacionado la misma a cuestiones de estabilidad en sistemas dinámicos a los que se le puede asociar una matriz de interacciones. Por otro lado, la proporción de conexiones dobles de una red aleatoria se puede calcular analíticamente[†], con lo cuál se dispone de otra magnitud que puede estudiarse en un dado modelo de red (o alguna dada red empírica) y comparársela fácilmente con lo que se obtiene en el modelo de redes aleatorias (modelo nulo). En otras palabras, la proporción de conexiones dobles provee de otra herramienta para detectar en algún dado modelo de red (o red empírica) algún tipo de correlación que las distinga de las redes aleatorias.

Detalle de notación: A lo largo del trabajo se utilizará la definición de conectividad dada por $C^{(d)}$. Por ésta razón la notación se simplificará, y se denotará por C a $C^{(d)}$.

2.8.1. La conectividad y otras propiedades en redes ecológicas

Una interacción trófica es una relación presa–predador existente entre un par de especies en un dado ecosistema. En un ecosistema, el conjunto de especies y las relaciones presa–predador entre las mismas constituyen una red denominada *red trófica*[‡]. En general en los ecosistemas también pueden darse otros tipos de interacciones diferentes a las interacciones tróficas dando lugar al concepto más amplio de *redes ecológicas*. En las

*Es decir, $k_i = \sum_{j \neq i} 1 - \delta_{0, (c_{i,j} + c_{j,i})}$, donde $c_{i,j}$ es la matriz de adyacencia de la red dirigida y $\delta_{i,j}$ es la delta de Kronecker.

[†]Este cálculo se presenta en los resultados.

[‡]Del inglés *food-web*.

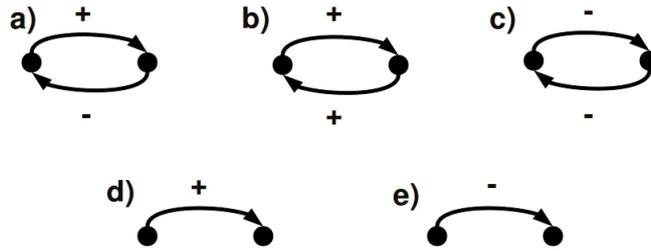


Figura 2.12: Los distintos tipos de interacción posibles en redes dirigidas pesadas. En el contexto de las redes ecológicas se las denomina por: **a** presa-predador y huesped-parásito, **b** cooperativa, **c** competitiva, **d** comensalismo, **e** antibiosis o amensalismo.

redes ecológicas las distintas formas en que dos especies pueden interactuar se resumen en la figura (2.12). Por ejemplo, dos especies pueden cooperar (ejemplos: plantas y polinizadores; plantas y frugívoros [93]) beneficiándose mutuamente sin necesidad de que una se alimente de la otra, por lo que no corresponde a una interacción trófica. Por otro lado, las interacciones tróficas generalmente se las suele representar por un link dirigido indicando quién es la presa y quién el predador. Por ello una red ecológica corresponde a un nivel más detallado de descripción que una red trófica. Una interacción trófica corresponde a una interacción ecológica en donde la presa se ve perjudicada y el predador beneficiado, por ende, una interacción trófica corresponde a una interacción ecológica del tipo **a**) (ver figura 2.12). Por ésta razón en este trabajo para el cálculo de la conectividad C asociaremos a cada red trófica una red ecológica en donde todas las conexiones son dobles, es decir, $\zeta = 1$ en la ecuación (2.60). Por otro lado, en las redes tróficas suelen existir autoconexiones (links que conectan a un nodo consigo mismo) que representan canibalismo. La cantidad de autoconexiones en una red trófica en general constituye una fracción muy pequeña* del total de conexiones, por lo que en el presente trabajo serán despreciadas.

Las redes tróficas son redes pequeñas si se las compara con redes empíricas incluidas dentro del contexto de las redes complejas. Esto conlleva a que sea al menos difícil caracterizar sus propiedades topológicas de manera análoga a como se lo hace con los otros tipos de redes. Dada ésta problemática, la conectividad parece ser un buen candidato a estudiar en el contexto de las redes tróficas ya que es una “propiedad global” de la red. Otras herramientas de caracterización tales como la distribución de grados, capturan muchos detalles locales que introducen serias fluctuaciones. En redes más grandes las fluctuaciones suelen ser absorbidas dando lugar a caracterizaciones más claras.

Sin embargo este razonamiento – el de pensar que la conectividad puede ser un buen candidato a estudiar en las redes tróficas – resulta

*La fracción de conexiones que son autoconexiones es del orden de $1/n$.

engañoso. Es un número que no dice mucho si no se lo pone en un contexto. Por ejemplo, la comparación de la conectividad con un modelo de hipótesis nula no tiene sentido ya que la conectividad es un parámetro del modelo de hipótesis nula (modelo de redes aleatorias). Entonces, para poner en contexto la conectividad en las redes tróficas se estudia el comportamiento de la misma para redes de diferentes tamaños con la esperanza de encontrar alguna regularidad.

J.A. Dunne hace un repaso de la cuestión [63, 64]. En dichos trabajos se menciona que a lo largo de varias décadas previas a los 1970s, el consenso general era que comunidades más complejas eran más estables. En dichas consideraciones el significado de la palabra complejidad era impreciso pero, en particular, se asociaba mayor complejidad a mayor conectividad y mayor número de especies. Este particular punto de vista fue puesto a prueba por un trabajo de Robert May [125] (sobre el que se detallará más adelante) en donde se presentan argumentos matemáticos en la dirección contraria. Aunque el modelo que May propone en dicho trabajo es una gran simplificación de un sistema ecológico, da lugar a una controversia sobre la que se ha trabajado y discutido ampliamente con el fin de intentar resolverla. La conclusión del trabajo de May es que un sistema es estable si y sólo si $Cn < \alpha$ (donde por el momento α es una constante), volviéndose inestable en el caso contrario. Como consecuencia de ello, el borde de estabilidad corresponde a $C \sim 1/n$, o lo que es lo mismo a $\langle k \rangle$ constante respecto de n . Bajo la suposición de que α es la misma para todos los sistemas, la contradicción entre el modelo de May y el consenso previo salta a la vista. Según May, un sistema no puede poseer valores de C y n arbitrariamente grandes sin que eventualmente inestabilice. Según Dunne [63, 64] esta contradicción dió lugar a una primera etapa de investigación en donde se concluyó que los ecosistemas aparentemente existen en el borde de estabilidad de May, mas precisamente que $\langle k \rangle = 2M/n \simeq 4$, o en otras palabras $\langle k \rangle$ es independiente de n [49]*. Sin embargo, posteriores trabajos realizados sobre redes tróficas caracterizadas con más detalle experimental revelaron ciertas anomalías. G.A. Polis [173] utilizando datos más confiables de la red trófica de *Coachella Valley* encuentra que $\langle k \rangle \simeq 19.2$, casi un orden de magnitud superior al valor previamente consensuado. Otro estudio realizado por Martinez [119] sobre la red trófica de *Little Rock Lake* resulta en $\langle k \rangle \simeq 22$ y da argumentos que pone de manifiesto que las anteriores catalogaciones de redes tróficas eran defectuosas – pecando de baja “resolución experimental”. El mejoramiento de los datos y estudios enfocados en el efecto de agregación de especies† entre otras cosas, indicaron que ciertos patrones presentes en los estudios realizados sobre redes tróficas, como por ejemplo el de $\langle k \rangle$ constante, pueden ser el resultado de datos incompletos o metodológica-

*Normalmente en los textos de ecología M es denotado por L (de *link*), y n por S (de *species*).

†Donde diferentes especies son confundidas como una misma y representadas por un único nodo – llevando a subestimaciones de n .

mente mal determinados [63, 64]. Incluso hay trabajos que reportan una conectividad constante (es decir, $\langle k \rangle \sim n$) [202], o al menos que el grado medio crece $\langle k \rangle \sim n^\mu$ con μ entre 0.54 y 1.04 dependiendo de que redes sean consideradas en el ajuste [120]. A partir de allí una serie de distintos estudios soportaron la hipótesis de que $\langle k \rangle \sim n^\mu$ con $\mu \sim 0.6 \pm 0.3^*$. En un Review muy reciente en donde colaboran muchos autores [93], se concluye que $\langle k \rangle \simeq 2.08n^\mu$ con $\mu = 0.49$. También muestra como otro tipo de redes ecológicas: las de interacciones *huesped-parasito* y las de interacciones *mutualistas* están comenzando a ser estudiadas, y el paulatino incremento de la resolución empírica lleva por caminos similares. Como detalle agregado, es interesante resaltar que el comportamiento $\langle k \rangle \sim n^\mu$ con $\mu \simeq 0.5$ es también observado en redes metabólicas [97] (ver figuras 3c y 3d de dicho trabajo).

Observacion: En resumen, la cuestión de como escala la conectividad C con el tamaño n en redes ecológicas no está completamente resuelta debido a la dispersión de los datos. Asumiendo un comportamiento tipo ley de potencias de la forma $C \sim n^{-(1+\epsilon)}$, algunos datos sugieren que $\epsilon = -\mu \lesssim 0$. Otros resultados (que serán discutidos en el Capítulo 4) sugieren que si se extrapola la curva de $C(n)$ incluyendo otras redes biológicas que presentan tamaños mayores a los que alcanzan las redes ecológicas, entonces $\epsilon \gtrsim 0$ [164].

2.8.2. Mundo pequeño y clusterización en redes ecológicas

Montoya y Solé [133] analizan la distancia topológica típica L y el coeficiente de clusterización $C_c^{(WS)}$ para cuatro redes ecológicas. Para cada red, encuentran que L es del orden de correspondientes redes aleatorias resultando levemente mayor. Por otro lado el coeficiente de clusterización $C_c^{(WS)}$ resulta sensiblemente más grande en las redes ecológicas. Estas apreciaciones son consistentes con la interpretación de que las redes ecológicas son de mundo pequeño. Por otro lado en Dunne *et. al.* [65, 64] se encuentra que aunque la mayoría presenta un coeficiente de clusterización $C_c^{(Tr)}$ mayor al de correspondientes redes aleatorias, sólo 5 de las 16 redes tróficas estudiadas presentan un coeficiente de clusterización significativamente superior al de correspondientes redes aleatorias ($C_c^{(Tr)}/C_{c_{rand}} > 2$), e incluso en algunos casos las fluctuaciones pueden llegar a invertir la relación (4 de los 16 casos). En particular, las 5 redes que poseen una relativa alta clusterización son las que presentan las conectividades C más bajas. En otras palabras, la diferencia entre el coeficiente de clusterización para las redes tróficas y sus pares aleatorias disminuye si la conectividad es alta. Luego aclaran que en el caso de las redes tróficas, la ausencia de un coeficiente de clusterización signifi-

*Promedio y desviación estandard sobre los 8 valores reportados en [63] que soportan la hipótesis $\langle k \rangle \sim n^\mu$. Sección “Constant Connectance”.

cativamente mucho mayor al de correspondientes redes aleatorias es de esperarse como consecuencia de la ecuación (2.28) y del pequeño tamaño n que las redes tróficas presentan.

Observación: En resumen, las redes tróficas presentan distancias topológicas pequeñas (valores de L similares al de correspondientes redes aleatorias), y un coeficiente de clusterización en en la mayoría de los casos mayor al de redes aleatorias pero la diferencia no es muy acentuada. Se puede concluir que las propiedades de mundo pequeño se cumplen en las redes tróficas pero no son tan marcadas como en redes de tamaños mayores.

2.8.3. La distribución de grados en las redes ecológicas

En algunos trabajos se ha reportado que las redes tróficas no poseen una distribución de grados libre de escala [40, 65]. Particularmente en [40] se propone que el comportamiento de las distribuciones de grados de las distintas redes tróficas allí consideradas pueden hacerse colapsar en una única curva indicando una forma universal tipo exponencial. Mientras que en [65] se reportan una diversidad de comportamientos: distribuciones de grado libres de escala, decaimiento exponencial, y uniformes. Otros sin embargo, reportan distribuciones de grado tipo ley de potencias con $\gamma \in \{0.94, 1.03, 1.13\}$ [133]. También se han reportado sobre redes ecológicas leyes de potencias con $\gamma = 1.51$ y leyes de potencias truncadas por colas exponenciales*[132]. Montoya y Solé[134] mencionan que la forma de la distribución de grados de una red trófica varía con tamaño de la red. Más precisamente, en redes pequeñas $P(k)$ se comporta como una distribución de cola corta, y a medida que el tamaño de la red aumenta $P(k)$ va adoptando la forma de una distribución más ancha hasta alcanzar el comportamiento de una distribución de cola larga. Ellos corroboran el ensanchamiento de $P(k)$ con n de la siguiente manera. Calculan el *coeficiente de variación CV* dado por:

$$CV = \frac{\sigma_k}{\langle k \rangle} = \frac{\sqrt{\langle (k - \langle k \rangle)^2 \rangle}}{\langle k \rangle} \quad (2.64)$$

para cada una de las redes tróficas estudiadas por ellos (el promedio se calcula según $\langle f(k) \rangle = \sum_k P(k)f(k)$). Luego comprueban que *CV* correlaciona positivamente con el tamaño n de las redes en cuestión.

En general caracterizar la distribución de grados en las redes ecológicas, y en particular en las redes tróficas conlleva un problema. La razón es que son redes muy pequeñas y no poseen una buena representabilidad del rango de conectividades k de los nodos a la hora de calcular $P(k)$. Es por ello que determinar la forma funcional de $P(k)$ en redes ecológicas es una cuestión delicada sobre la que hay que mantener ciertas reservas.

*Los valores de γ en dichos ajustes son 0.01 y 0.28 indicando que la cola exponencial predomina el ajuste de la curva.

Observación: Es importante resaltar que las distribuciones de grados no son Poissonianas [133, 65], o en otras palabras, no corresponden a las de redes aleatorias. Por otro lado, aún cuando las redes ecológicas no presentan distribuciones de grados libres de escala se pueden reconciliar con el enfoque de Barabási–Albert[23]. Recuérdese el modelo de Amaral *et. al.*[17] discutido en la sección 2.1.3, en donde se agrega *capacidad de carga máxima* del grado de los nodos al mecanismo de generación de redes del modelo de Barabási–Albert. En dicho modelo aún en presencia de crecimiento y conexión preferencial las redes resultantes pueden exhibir distribuciones de cola corta y diferenciarse de las redes aleatorias.

2.8.4. El ataque dirigido y las fallas aleatorias en redes ecológicas

Solé y Montoya estudiaron el ataque dirigido y las fallas aleatorias (ver sección 2.7) en redes ecológicas[189]. Las redes tróficas que ellos estudian en dicho trabajo poseen – al menos – distribuciones de grado de cola larga. Encontraron que ante fallas aleatorias las redes se fragmentan a valores irrealísticamente grandes de f , aunque no reportan el valor obtenido. Dicho valor puede estimarse utilizando el cálculo analítico presentado en [51] obteniéndose $f_c \simeq 0.9$. Por otro lado para ataques dirigidos las redes fragmentan alrededor de $f_c \simeq 0.25$ en dos de los tres casos (*Ytahn Estuary* y *Silwood Park*), y $f_c \simeq 0.35$ en el otro caso (*Little Rock Lake*). Ellos concluyen que la gran tolerancia a fallas aleatorias y relativamente alta fragilidad ante ataques dirigidos es consistente con sus otras observaciones de que dichas redes presentan distribuciones de grados de cola larga. Sin embargo la tolerancia o fragilidad ante un proceso de aislación de nodos, originalmente se plantea como una medida relativa a lo que ocurre en una red aleatoria [12],[13]. En otras palabras, Solé y Montoya no comparan el comportamiento de $S(f)$ y $\langle s \rangle(f)$ que ellos obtienen para las redes ecológicas estudiadas, con el comportamiento que obtendrían en versiones aleatorizadas de dichas redes (redes aleatorias de mismo tamaño y número de links que las originales). Esta cuestión será rediscutida en el capítulo 4.1.6 y el apéndice A en donde se presentan nuevos resultados.

2.9. La estabilidad en las redes de interacción: El modelo de R. May

En los 70's Gardner y Ashby [78] sugieren la idea de que sistemas complejos de muchas componentes que interactúan a través de muchas conexiones distribuidas aleatoriamente, sólo pueden ser estables por debajo de un nivel crítico de conectividad. En otras palabras, sistemas demasiado conectados resultan inestables. Poco después R. May complementa dicho trabajo con argumentos analíticos [125]. El trabajo de May

constituye un importante antecedente para la elaboración de las ideas y modelos propuestos en éste trabajo (Capítulo 3), por lo que conviene presentarlo con cierto detalle. May considera un sistema de n variables $\vec{x} = (x_1, \dots, x_n)$, que en general pueden obedecer algún sistema de ecuaciones diferenciales no lineales de primer orden:

$$\frac{d\vec{x}}{dt} = \vec{f}(\vec{x}) \quad (2.65)$$

La estabilidad de las posibles trayectorias de equilibrio o independientes del tiempo del sistema definido por f pueden ser estudiadas a través del primer término no nulo de la expansión de Taylor de f en las cercanías del punto de equilibrio. En otras palabras, la estabilidad de los puntos de equilibrio es caracterizada por la ecuación:

$$\frac{d\vec{x}}{dt} = \mathbb{A}\vec{x} \quad (2.66)$$

La estabilidad del sistema queda determinada por los autovalores de \mathbb{A} . El sistema es estable si y sólo si todos ellos tienen parte real negativa. Siguiendo ideas de Gardner y Ashby, R. May asume que si las n componentes del sistema no interactúan, cada una respondería a perturbaciones retornando a un punto de equilibrio en un tiempo característico dado. Además asume la aproximación en que todos los tiempos característicos son iguales*, ello equivale a exigir que $a_{i,i} = -1$ para toda componente i . May encuentra que en el caso en que sólo una fracción C de las $n(n-1)$ interacciones posibles representadas por \mathbb{A} son no nulas y sorteadas a partir de una distribución de media cero y varianza finita α^\dagger , la probabilidad de que el sistema sea estable es tal que:

$$P(n, C, \alpha) \xrightarrow{n \rightarrow \infty} \begin{cases} 1 & \text{si } C < \frac{1}{\alpha n} \\ 0 & \text{si } C > \frac{1}{\alpha n} \end{cases} \quad (2.67)$$

Notar que C es la conectividad de la red que representa el sistema, y α cuantifica la “fuerza” de las interacciones. En general en correspondencia con la ecuación (2.67), dados α y $n \gg 1$ hay una región de transición centrada en $C_{trans}(n, \alpha)$ por encima de la cual el sistema es inestable, y por debajo estable. C_{trans} disminuye al aumentar α o n . A tamaños finitos el ancho de la región de transición decrece rápidamente con el tamaño según $\sim n^{-2/3}$ [125].

*Se puede probar que los resultados no cambian cualitativamente si los tiempos característicos son sorteados de una distribución de varianza finita.

†En el paper de May[125] dice que α es el promedio del cuadrado de las entradas no nulas de la matriz. Ello corresponde a que α es la varianza. Pero luego define la curva que separa las regiones de estabilidad/inestabilidad por $C = 1/(\alpha^2 n)$. Ello está mal, lo correcto es $C = 1/(\sigma^2 n)$ donde $\sigma^2 = Var(\{a_{i,j} \neq 0\})$ denota la varianza de las entradas no nulas del ensamble de matrices[87].

Observaciones: R. May pensando en una red ecológica sugiere una interpretación de su modelo. Suponiendo que mayores C y α corresponden a mayor complejidad, luego los sistemas poseen un límite a la complejidad que pueden alcanzar sin antes volverse inestables. En particular el modelo de May indica que los sistemas “viven” en la región $C < 1/(\alpha n)$, lo cual entra en conflicto con lo observado en redes tróficas (ver sección 2.8.1).

2.9.1. Analizando el concepto de estabilidad utilizado por May: Extensiones y alternativas

El trabajo de May dió lugar a un extenso debate, el dilema de la *diversidad–estabilidad* [127, 94]. En la sección 2.8.1 ya se ha discutido en parte la cuestión. En general el debate se centra en el contexto de las redes ecológicas. El concepto de estabilidad utilizado por May* es claramente una gran simplificación ya que asume que un sistema es estable si y sólo si el sistema es gobernado por una dinámica que converge a un *equilibrio* estable determinado por una aproximación lineal. Esta simplificación deja afuera escenarios más complejos en donde por ejemplo la dinámica converge a un ciclo límite. Sin embargo hay trabajos que indican que aún cuando se consideran criterios de estabilidad más generales como el de persistencia [45], o dinámicas no lineales como en [45, 98, 183], no se observan cambios cualitativos respecto de lo encontrado por May.

Son numerosos los trabajos que buscan resolver el dilema de la *diversidad–estabilidad*. Algunos de ellos se enfocan en estudiar sistemas en donde las entradas no nulas de las matrices $a_{i,j}$ en cuestión no son elegidas completamente al azar. Por ejemplo, en [107, 108] se propone agregar a la dinámica de poblaciones una dinámica adaptativa de las relaciones tróficas. En [136] la estabilidad incrementa cuando largos *loops* de conexiones poseen suficientes interacciones débiles. En [126] se concluye que en un modelo no lineal de dinámica de poblaciones en redes tróficas, la presencia de interacciones débiles ayuda a estabilizar la dinámica, por ejemplo a través de amortiguar las oscilaciones. Allesina y Pascual [15] ponen de manifiesto que la estabilidad aumenta con la fracción de interacciones tipo presa–predador (interacciones *anticorrelacionadas* o del tipo **a** en la figura (2.12)). Más precisamente, encuentran que a las matrices de May que se les exige que $a_{i,j}a_{j,i} \leq 0$ (al menos en la mayoría de los casos), presentan una mayor probabilidad de resultar estables respecto de las matrices originales de May en donde $a_{i,j}$ y $a_{j,i}$ eran estadísticamente independientes.

Otros enfoques buscan analizar la relación del concepto de estabilidad de May respecto de la topología de la red de interacciones. En [182] se

*Conviene resaltar que se hace referencia al concepto de estabilidad de May y no al modelo de May. Por ejemplo se podría modificar el modelo estudiando redes de topologías libres de escala en vez de redes aleatorias pero manteniéndose el concepto de estabilidad.

estudia la estabilidad de May en redes mundo-pequeño*. La conclusión en dicho trabajo es que la propiedad de mundo pequeño no afecta la ubicación del “punto” de transición C_{trans} , pero sí afecta el ancho de la región haciéndose más ancha para redes con alta clusterización que para redes aleatorias. Esto es importante ya que la mayor variabilidad permitida por la alta clusterización puede ser eventualmente aprovechada por un proceso de selección de matrices estables. En [35] se concluye que tanto para redes aleatorias como para redes libre de escala, la estabilidad tiende a decrecer con una mayor asortatividad. En [196] se estudia el concepto de estabilidad de May en redes de interacción generadas por medio de un proceso de selección a través de un algoritmo genético aplicado sobre un ensamble de sistemas originalmente constituidos por matrices aleatorias. La selección opera sobre el *fitness*[†] de cada matriz definido como el número de autovalores con parte real negativa. El proceso de selección conduce a que las matrices evolucionen hacia sistemas cada vez más estables[‡]. En una red dirigida, dos nodos i y j se dicen estar en *componentes conectadas* (CC) separadas si no existe un camino dirigido que los conecte en ninguna de las dos direcciones posibles. Por otro lado, dos nodos i y j están en una misma *componente fuertemente conectada* (CFC) si y sólo si existe un camino dirigido de i hacia j y otro de j hacia i . En el artículo, encuentran que el proceso de selección da lugar a redes que se organizan jerárquicamente en módulos. Más precisamente, la red se fragmenta en varias (CC)’s las cuales a su vez se fragmentan en varias (CFC)’s. Ésto ocurre así debido a que los autovalores de una matriz, son los autovalores de sus (CFC)’s. Estabilizar pequeñas (CFC)’s resulta eventualmente más probable que estabilizar una sola gran (CFC), lo que lleva a la modularización del sistema. En [159] el estudio de la cuestión modularidad-estabilidad es más detallado. Esencialmente se estudian redes con n y M fijos que resultan de minimizar la siguiente “energía” utilizando un proceso de *simulated annealing* sobre las conexiones de la red:

$$E(\omega) = \omega L + (1 - \omega)\lambda_{max} \quad (2.68)$$

donde λ_{max} es el máximo de las partes reales de todos los autovalores de las matrices en cuestión, y L es la distancia típica entre nodos. Para $\omega \simeq 1$ la optimización principalmente lleva a minimizar L favoreciendo topologías tipo estrellas[§]. Para valores intermedios de ω se obtienen estructuras modulares en donde cada módulo posee esencialmente una forma de estrella más algunas conexiones que interconectan los diferentes módulos. La estabilidad de los módulos domina la estabilidad del sistema, y la estabilidad de cada módulo está principalmente determinada por el

*Redes de Watts–Strogatz [204].

[†]No encontré una traducción de *fitness* al castellano que me pareciera adecuada.

[‡]Según el concepto de estabilidad de May.

[§]Por ejemplo, para una red de n nodos y $M = n - 1$ links, el mínimo $L = 2$ corresponde a una estrella.

grado del *hub* que constituye el centro del mismo. Como consecuencia de esto último, la estabilidad del sistema aproximadamente viene determinada por $\lambda_{max} \sim \sqrt{n/m}$ donde m es el número de módulos. En el caso en que los *hubs* no dominan la estabilidad de sus correspondientes módulos, luego la modularización tiene un efecto negativo sobre la estabilidad, incluso si los módulos son organizados jerárquicamente[158].

En el presente trabajo se propone un mecanismo evolutivo de generación de redes de interacción, en donde las redes crecen sujetas a una presión selectiva basada en el criterio de estabilidad de May [165, 164]. Se verá como las diversas propiedades observadas tanto en los modelos como en redes reales emergen de manera auto-organizada a partir de éste mecanismo. Estos resultados contribuyen a afianzar la idea de que el concepto de estabilidad de May participa en la formación de ciertas redes de interacción, particularmente en las biológicas.

2.9.2. Espectro de autovalores en redes complejas

Dado que los autovalores juegan un rol central en el concepto de estabilidad de May, es provechoso estudiar el espectro de autovalores asociado a dichas redes sobre diferentes topologías. Cuando $C = 1$, las matrices de May son de la forma:

$$\mathbb{A} = \mathbb{B} - \mathbf{1} \quad (2.69)$$

donde \mathbb{B} es una matriz aleatoria. Más precisamente, los elementos $b_{i,j}$ de \mathbb{B} son estadísticamente independientes entre sí y están distribuidos a partir de una misma distribución $\tilde{P}(b)$ de media nula y varianza finita $\tilde{\sigma}^2$. Los autovalores ω' de \mathbb{A} son soluciones de la ecuación:

$$\begin{aligned} \det(\mathbb{A} - \omega' \mathbf{1}) &= 0 \\ \det(\mathbb{B} - \mathbf{1} - \omega' \mathbf{1}) &= 0 \\ \det(\mathbb{B} - \omega \mathbf{1}) &= 0 \end{aligned} \quad (2.70)$$

donde $\omega = \omega' + 1$ son los autovalores de \mathbb{B} . En otras palabras, los autovalores de \mathbb{A} se obtienen restando 1 a los de \mathbb{B} . Tener ello en mente ya que en los capítulos subsiguientes son los autovalores de \mathbb{A} los que interesarán, pero la discusión a continuación es mas sencilla si se hace en términos de los autovalores de \mathbb{B} , por lo que así será.

Cuando $C < 1$, se puede pensar que $1 - C$ es la probabilidad de que los elementos de matriz de \mathbb{B} resulten nulos. En particular, se puede pensar que dichos elementos son distribuidos a partir de una distribución $P(b)$ definida en términos de $\tilde{P}(b)$:

$$P(b) = C\tilde{P}(b) + (1 - C)\delta(b) \quad (2.71)$$

por lo que simplemente a la matriz diluída* le corresponde la varianza:

*Matriz con entradas nulas, o lo que es lo mismo, matriz en donde $C < 1$.

$$\sigma^2 = C\tilde{\sigma}^2 \quad (2.72)$$

A matrices aleatorias como \mathbb{B} , en el límite $n \rightarrow \infty$ le corresponde una densidad $P(\omega)$ de autovalores complejos $\omega = (\lambda + i\varphi) \in \mathbf{C}$ dada por[191]:

$$P(\omega) = \begin{cases} \frac{1}{\pi n \sigma^2} & , |\omega| \leq \sqrt{n\sigma^2} \\ 0 & \text{c.c.} \end{cases} \quad (2.73)$$

Esencialmente, los autovalores están dentro de un círculo en el plano complejo. Este resultado vale particularmente para matrices en donde $\langle b_{i,j} b_{j,i} \rangle = 0$. Para matrices correlacionadas, se obtienen resultados un tanto diferentes en donde el círculo es remplazado por una elipse. En el caso particular de correlación(anticorrelación) total, o lo que es lo mismo, matrices simétricas(antisimétricas), uno de los ejes de la elipse se va a cero. Por ejemplo, para el caso de matrices simétricas se obtiene la famosa ley del semi-círculo de Wigner[206]. Aquí, nos interesa estudiar la estabilidad del sistema, en otras palabras, particularmente interesa la densidad de probabilidades $P(\lambda)$ de las partes reales λ de los autovalores ω :

$$P(\lambda) = \int d\varphi P(\omega) = \begin{cases} \frac{2}{\pi n \sigma^2} \sqrt{n\sigma^2 - \lambda} & , |\lambda| \leq n\sigma^2 \\ 0 & \text{c.c.} \end{cases} \quad (2.74)$$

donde es importante resaltar que la integración es sobre la parte imaginaria φ de ω . Lo importante, es que en general (salvo el caso totalmente antisimétrico) se tiene que la máxima parte real de entre todos los autovalores satisface:

$$\lambda_{max} \sim \sqrt{n\sigma^2} = \sqrt{nC\tilde{\sigma}^2} \equiv \sqrt{nC\alpha} \quad (2.75)$$

para alguna constante de proporcionalidad, y la equivalencia se pone para unificar notación con el modelo de May (ver sección 2.9). El caso de la ecuación (2.74) corresponde a una topología completamente aleatoria. Cuando la topología presenta algún tipo de estructura ocurren cosas diferentes. En [35] se concluye que las redes tipo May (pero con diagonales nulas) sobre topologías libres de escala del modelo de Barabási–Albert[23] son tales que $P(\lambda)$ decae de forma exponencial al crecer $|\lambda|$ salvo cuando $|\lambda| \ll 1$ en donde $P(\lambda) \sim \lambda^{-1/2}$ (figura 2.13). Las redes de Barabási–Albert presentan algún tipo de correlación introducida por el mecanismo de crecimiento que las genera. Si se aleatoriza dichas redes por medio de aplicar *reconectado cruzado*[122] de modo que la distribución de grados conserva su forma pero se pierden las mencionadas correlaciones, luego la caída exponencial se conserva mientras que se pierde el comportamiento $\sim \lambda^{-1/2}$ cuando $|\lambda| \ll 1$ (inset derecho en la figura 2.13). Dada una red, la matriz de adyacencia \mathbb{G} es tal que $g_{i,j} = g_{j,i} = 1$ cuando existe un link entre i y j , y vale cero en el caso contrario. Comúnmente se estudia el espectro de autovalores de las matrices de adyacencia de una red. Sin embargo,

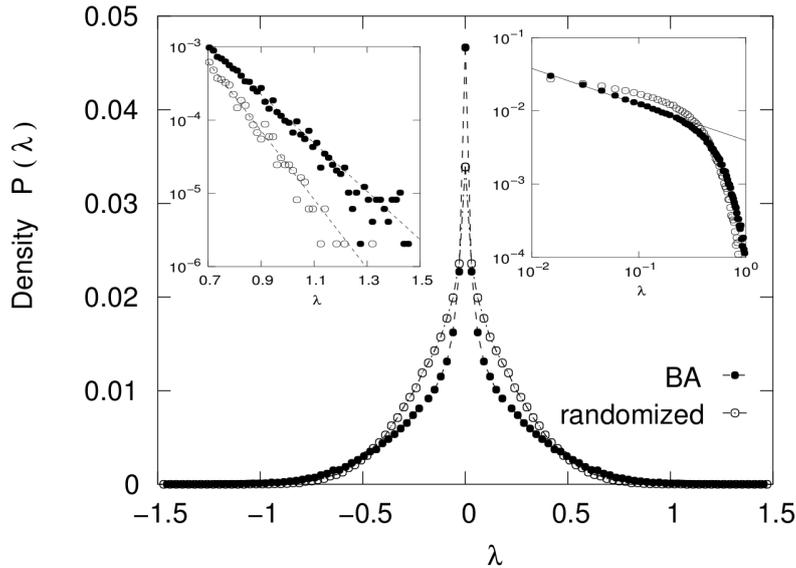


Figura 2.13: **Gráfico principal:** Densidad de probabilidades $P(\lambda)$ de las partes reales de los autovalores asociados a un ensamble de matrices tipo May construídas sobre redes de topología libre de escala generadas por el modelo de Barabási–Albert[23] (círculos llenos), y versiones aleatorizadas de dichas redes que preservan la distribución de grados $P(k)$ (círculos vacíos). Las redes de Barabási–Albert alcanzaron el tamaño $n = 5000$, y fueron construídas para $m = 10$. Los elementos no nulos de las correspondientes matrices fueron elegidos de una distribución Gaussiana de media cero y varianza $\sigma^2 = 0.2$. En el **inset de la izquierda** ilustra la caída exponencial de las colas de $P(\lambda)$ para ambos casos. En el **inset de la derecha** se observa que cuando $\lambda \ll 1$ entonces $P(\lambda) \sim \lambda^{-1/2}$ para el caso en que las matrices corresponden a redes de Barabási–Albert (círculos llenos), cosa que no ocurre para las versiones aleatorizadas de dichas redes (círculos vacíos). La comparación pone de manifiesto que las correlaciones introducidas por el proceso de crecimiento en el modelo de Barabási–Albert afectan el espectro de autovalores. Esta figura fué extraída de [35] (figura 2).

notar que el comportamiento de las distribuciones $P(\lambda)$ correspondientes a las redes tipo May es significativamente diferente al de las matrices de adyacencia para las mismas topologías[72, 79]. Por ejemplo, para matrices tipo May sobre topologías libres de escala del tipo Barabási–Albert ($\gamma = 3$)[23] ó Bianconi–Barabási ($\gamma \simeq 2.1$)[33], se tiene que $\lambda_{max} \sim n^{-1/2}$ lo cuál contrasta con el comportamiento $\lambda_1 \sim n^{-1/4}$ reportado para las matrices de adyacencia de las redes de Barabási–Albert[79].

Observación: Es importante resaltar que la distribución de autovalores de las matrices de May es diferente a la de matrices de adyacencia de igual topología. La diferencia es consecuencia de que en las matrices de adyacencia todas las entradas no nulas son 1, mientras que en las matrices tipo May son elegidas de una distribución de media nula y varianza finita*.

*En esta consideración estoy omitiendo la diagonal de las matrices.

Capítulo 3

Un mecanismo de presión selectiva basado en la estabilidad

Los mecanismos propuestos hasta el momento exitosamente han logrado explicar el origen de la complejidad en algunos tipos de redes. Sin embargo, es reconocido que otro gran número de casos son de carácter esencialmente distintos, y por ende no pueden ser explicados por este tipo de modelos. Básicamente, los mecanismos referidos van dando forma a la topología de la red procesando información contenida en la topología misma. El ejemplo típico es el conexionado preferencial del modelo de Barabási–Albert [23]. Sin embargo existe la posibilidad de que algún fenómeno no caracterizado por la topología de la red participe en el mecanismo que le va dando forma a la misma. Una posibilidad de importancia, al menos en el contexto de redes biológicas, es la de algún criterio de aptitud fundamentado en la dinámica del sistema que subyace a la red que lo representa.

Las redes biológicas comúnmente exhiben el acoplamiento de al menos dos procesos dinámicos. El primero de estos procesos concierne a la incorporación de agentes a la red, los cuales se incorporan a través de un lento proceso evolutivo. El segundo de estos procesos corresponde a la dinámica de las magnitudes asociadas a los agentes de la red, la cuál en principio afecta y es afectada por el primero de estos procesos. Es razonable esperar que la estructura que presentan estas redes emerja (al menos en parte) como resultado del acoplamiento de estos dos procesos.

Este trabajo consiste en identificar un mecanismo de evolución de redes de interacción que presumiblemente es implementado por algunos sistemas reales. En particular nos interesa un mecanismo en el que se acoplan los dos procesos antes mencionados. El acoplamiento que proponemos consiste en definir un criterio de aptitud de los agentes que se incorporan al sistema fundamentado en la estabilidad de la dinámica subyacente a la red de interacciones. Comprobaremos que redes de topologías complejas y de propiedades observables en los sistemas reales,

particularmente en las redes biológicas, surgen como resultado de un mecanismo de evolución que incorpora dicho acoplamiento. Es así como se dispone de una forma alternativa, y en ciertos casos presumiblemente más apropiada, para explicar las propiedades observadas en muchas de las redes complejas de origen natural, a saber, en aquellos tipos de sistemas en los cuales la dinámica subyacente a la red juega un rol en el mecanismo de conformación de la topología de la misma.

Para fijar ideas, veamos algunos ejemplos concretos de sistemas en los cuales el proceso de incorporación y remoción de agentes está entrelazado a la dinámica del sistema de agentes interactuantes. Primero consideremos una red ecológica, en la cuál un conjunto de especies interactúan a través de relaciones tipo presa–predador, parasitarias, comensalistas, etc., de forma tal que da lugar a una dinámica de poblaciones. Este tipo de sistemas se auto–construyen a través de reglas de ensamblado propias de las comunidades, las cuales están fuertemente influenciadas por la dinámica subyacente de las especies y las interacciones específicas entre ellas [205, 167]. Por ejemplo, en las reglas de ensamblado de las redes tróficas participa un flujo de especies (que se incorporan o disgregan del sistema a través de migraciones, especiaciones, extinciones, etc.) acoplado con la dinámica poblacional. Otro ejemplo podría ser las redes neuronales, en donde la incorporación de cientos de miles de nuevas neuronas es regulada por un proceso en el cuál la dinámica de las neuronas y la conectividad están interrelacionadas de una manera no entendida completamente. Un tercer ejemplo en una escala diferente, puede ser imaginado en el contexto de redes sociales, en donde la incorporación de un miembro novato puede ser aceptada o rechazada basándose en su contribución individual al interés global de la comunidad, su aptitud, su desempeño, o las ganancias económicas que éste proporciona. Otro ejemplo de gran importancia son las redes de interacción a nivel celular (redes de interacción de proteínas, metabólicas, de transcripción genética, etc.) en las cuales la incorporación de nuevos agentes es producida por procesos de variabilidad genética, sujetas a selección natural a través del fenotipo asociado. En estos ejemplos es relativamente fácil visualizar los dos procesos mencionados anteriormente, es decir, en primer lugar la incorporación o remoción de agentes al sistema, y en segundo lugar la dinámica entre los agentes del sistema la cuál afecta al (y se ve afectada por el) primero de estos procesos. La consecuencia de afectar la estabilidad (inestabilidad) global del sistema provocada por la incorporación de un nuevo miembro de una dada conectividad, está representada en estos ejemplos por la abundancia (carencia) de comida, el bienestar (o la muerte) neuronal, el incremento (decremento) de las ganancias, o el *fitness** asociado a los genes. Nótese que cada nuevo miembro no sólo puede resultar en su propia incorporación o rechazo del sistema, sino que además puede promover una avalancha de extinciones/remociones

*Según Wikipedia *fitness* se traduce al español como *aptitud*. En ésta circunstancia me pareció mejor utilizar el anglicismo.

entre los miembros previamente existentes, un efecto que mostraremos que influencia significativamente la topología de las redes resultantes.

Todos estos ejemplos pueden ser pensados como un sistema de n agentes regido por una dinámica del tipo:

$$\frac{d\vec{x}}{dt} = \vec{F}(\vec{x}) \quad (3.1)$$

donde \vec{x} es un vector de n componentes que describe las variables de estado relevantes para cada agente, y \vec{F} es una función no lineal arbitraria. Uno puede imaginar que \vec{x} en diferentes sistemas puede representar diferentes cosas, por ejemplo, la concentración de hormonas en una red de regulación de hormonas, o el promedio de las densidades de población en una red trófica, o las concentraciones de químicos en una red bioquímica, o la actividad de los diferentes genes en una red de regulación genética, etc. Asumiremos que un dado agente i interactúa con un conjunto limitado de $k_i < n$ otros agentes, de modo que F_i depende únicamente de las variables pertenecientes a dicho conjunto. Mas precisamente, la interacción entre las variables del sistema viene dada por el conjunto de funciones F_i de modo que la variable i -ésima interacciona con la j -ésima si es que F_i depende explícitamente de x_j y/o F_j depende explícitamente de x_i . De este modo, puede asociarse esquemáticamente una red al sistema en donde cada nodo de la red está asociado a un agente, mientras que cada conexión de la red está asociada a una interacción entre las variables asociadas a los correspondientes agentes.

En relación a los dos procesos antes mencionados en el planteamiento del problema, asumiremos que en los sistemas en consideración rigen dos escalas temporales en la dinámica. Una escala a tiempos cortos en donde las ecuaciones de la dinámica, y por ende la red de interacciones, permanecen fijas. La otra, una escala de tiempos largos (mucho mas larga que los tiempos de observación), en donde el sistema está sujeto a un flujo externo de nuevos agentes (por ejemplo a través de migraciones en las redes tróficas, o mutaciones en las redes de regulación genética, etc), los cuales al ingresar al sistema interaccionan con algunos de los agentes que ya estaban presentes en la red. Estos nuevos agentes eventualmente pueden ser efectivamente incorporados al sistema o ser rechazados. En el caso de ser efectivamente incorporados n y todo el conjunto de ecuaciones diferenciales (3.1) cambia.

El criterio de si la incorporación de un nuevo agente al sistema se efectiviza o no, consiste en *evaluar la estabilidad* de la solución del nuevo conjunto de ecuaciones diferenciales que resulta tras la incorporación del nuevo agente. Es importante resaltar dos cosas. La primera es que la elección del criterio a utilizar para decidir si el sistema presenta un comportamiento estable o no, queda abierta. En este trabajo se utilizará un criterio que, aunque de carácter bastante general, comprende una elección particular que deja afuera otras posibles formas de caracterizar la estabilidad que podrían ser estudiadas en futuros trabajos (en [166] se

puede ver un sumario de distintos conceptos relacionados al de estabilidad). Lo segundo es hacer notar que este criterio constituye el vínculo entre el proceso de incorporación/remoción de agentes (que se relaciona a los cambios topológicos de la red de interacciones) y la dinámica subyacente del sistema, es decir, el vínculo explícito entre los dos procesos mencionados en el planteamiento del problema. Concretamente, un nuevo agente se incorpora efectivamente al sistema si su inclusión resulta en una transformación tal en la dinámica de manera que la misma lleva de un comportamiento estable a otro nuevo comportamiento estable. En caso contrario, la incorporación del nuevo agente es rechazada, o eventualmente ocurre una avalancha de extinciones/remociones de los agentes que ya estaban presentes en la red hasta que la estabilidad de la dinámica es restablecida.

Para aclarar ideas, en la figura (3.1) se observa una representación esquemática del proceso. La transición de la dinámica entre un comportamiento estable y el subsecuente ocurre en una escala de tiempos cortos τ_C mucho menor a la escala de tiempos τ_L que representa el largo característico de los períodos en que el sistema permanece sin sufrir desestabilizaciones. Los tiempos t_r, t_{r+1}, \dots representan las instancias en las cuales el sistema sufre una desestabilización como resultado de la incorporación de un nuevo agente al sistema, posiblemente a expensas de la remoción de otros. Cada uno de estos cambios, por definición del proceso en cuestión, llevan la dinámica de un comportamiento estable a otro, representados por la secuencia x_r^*, x_{r+1}^*, \dots . De esta forma nos encontramos ante un proceso basado en un mecanismo en el cuál la estabilidad actúa como un agente de selección natural. Este proceso da lugar a un ilimitado ensamble de sistemas caracterizados por una historia de crecimiento a través de la estabilidad de los mismos, a cada uno de los cuales se les puede asociar una red de interacciones. La aleatoriedad de las redes en el ensamble es auto-generada a través del proceso de incorporación de nuevos agentes. Esto es así, ya que asumiremos que la naturaleza de los agentes que intentan ingresar al sistema no satisface ningún patrón determinista en particular. Por otro lado, cada conjunto específico de interacciones define un sistema dinámico particular, por ende el análisis subsecuente para las escalas de tiempo entre sucesivas incorporaciones y/o remociones de agentes es puramente determinista, lo que incluye la determinación de la estabilidad del sistema. Una posible generalización de estas ideas podría considerar la posibilidad de introducir aleatoriedad a la determinación de la estabilidad del sistema.

Estas consideraciones fueron plasmadas en un modelo. Descubrimos a partir del mismo que este mecanismo de selección darwiniana sobre los cambios topológicos en la red de interacciones, conduce a la emergencia de características topológicas que son comúnmente consideradas en la literatura como propias de las redes complejas, tales como la propiedad mundo pequeño, una topología libre de escala, la presencia de jerarquías, y la robustez de la estructura de la red frente a la eliminación aleatoria

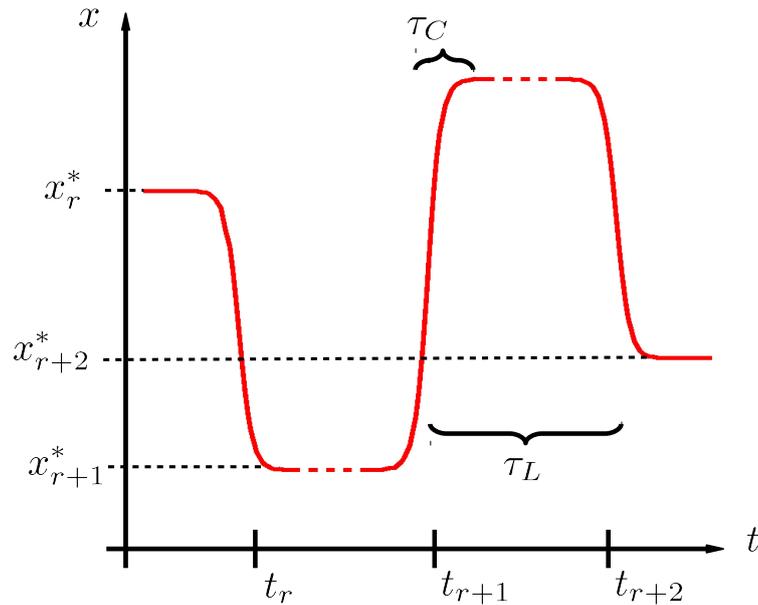


Figura 3.1: Representación esquemática de la dinámica de los sistemas en consideración. Se observa períodos de tiempos característicos relativamente largos τ_L (indicados por la región entrecortada de la curva roja) en donde el comportamiento del sistema es estable, intercalados por períodos de tiempo relativamente mas cortos τ_C en donde el sistema sufre una desestabilización temporal de su dinámica, relacionada a cambios topológicos en la red de interacciones.

de sus nodos [165, 164]. En particular, muchos de los resultados se ajustan a propiedades encontradas en redes de origen biológico. Entre ellas, una distribución de grados de cola larga, la disasortatividad del grado de los nodos de las redes, y la emergencia de interacciones anticorrelacionadas. Esto último ha sido previamente relacionado a un incremento en la estabilidad del sistema [15]. Además, encontramos evidencia de que los sistemas que crecen bajo este mecanismo, no pueden crecer mas allá de cierto tamaño, o al menos resulta cada vez menos probable. Quizás sea esta una posible explicación de porqué las redes biológicas son relativamente mucho mas pequeñas que por ejemplo las redes tecnológicas o sociales. Particularmente, algunos aspectos presentes en nuestros resultados están relacionados a una cuestión aún en debate, a saber, el dilema presente en la relación entre la estabilidad y la complejidad de los sistemas modelados por redes ecológicas [125, 15, 107, 136].

3.1. El modelo

El modelo consiste en una abstracción simplificada de lo hasta ahora considerado como ingredientes principales que constituyen el mecanismo de selección basado en la estabilidad. La mayor simplificación la com-

prende el análisis de estabilidad de la dinámica del sistema. Básicamente, siguiendo las ideas de May [125], asumiremos que el comportamiento estable de la dinámica corresponde a un punto fijo de la ecuación (3.1), es decir, la dinámica converge siempre a un punto de equilibrio estable \vec{x}^* tal que $\vec{F}(\vec{x}^*) = 0$. La estabilidad de dicha solución \vec{x}^* es determinada por el autovalor de máxima parte real λ_{max} de la matriz Jacobiana:

$$a_{i,j} \equiv \left(\frac{\partial F_i}{\partial x_j} \right) \Big|_{\vec{x}^*} \quad (3.2)$$

de manera que \vec{x}^* es un punto fijo estable si y solo si $\lambda_{max} < 0$. Dicha matriz representa a su vez las interacciones en el estado de equilibrio, por lo que en dicha circunstancia de estabilidad la red asociada al sistema presenta una conexión entre los nodos i y j si y solo si $a_{i,j}$ y/o $a_{j,i}$ son no nulos. La dinámica a tiempos cortos no necesita ser considerada al detalle, es decir, no es necesario considerar soluciones de las ecuaciones de movimiento, en principio es suficiente considerar que a tiempos suficientemente largos, se presenta una secuencia de puntos estables, $x_1^*, x_2^*, \dots, x_r^*, \dots$, y que a cada uno de ellos le corresponde una red de interacciones representadas por una matriz. De modo que se dispone de una secuencia de matrices $a_{ij}^{(1)}, a_{ij}^{(2)}, \dots, a_{ij}^{(r)}, \dots$ que determinan la evolución topológica de la red de interacciones del ensamble de sistemas en consideración.

Se podría considerar que la simplificación del criterio de estabilidad reduciendolo a la existencia o no de un punto fijo estable de la dinámica es demasiado restrictiva. Sin embargo, se verá a lo largo del trabajo que dicha simplificación admite cierto tipo de relajación de la condición. Por ejemplo, se verá que los resultados que se obtienen bajo este criterio de estabilidad son cualitativamente los mismos aún cuando se permite que λ_{max} sea levemente positivo (en relación al tamaño típico de las entradas de la matriz de interacciones), o incluso cuando, se evalúa la estabilidad de una fracción del sistema representado por una submatriz de $a_{i,j}$. Además, es posible pensar que los x_i no representan las variables directamente observadas del sistema, sino que consisten en un promedio de la evolución de las mismas sobre ciertos intervalos de tiempo lo suficientemente largos como para suavizar el comportamiento de las mismas filtrando así un posible comportamiento no lineal, pero lo suficientemente cortos como para no arruinar el análisis de la estabilidad.

Suponiendo que los agentes considerados aisladamente alcanzarán estados estables por cuenta propia luego de cierto tiempo característico, los elementos diagonales $a_{i,i}$ se asumen negativos, y de valor unidad en orden de simplificar el análisis [125]. Los elementos no diagonales $a_{i,j}$ con $i \neq j$, adoptarán valores aleatorios, tanto positivos como negativos elegidos a partir de alguna distribución de probabilidades. El resto de la definición del modelo queda determinada por el siguiente algoritmo. Se parte de una red de un solo nodo, la cual se la representa con una matriz de rango 1 dada por $a_{1,1} = -1$, y luego se intenta crecer el sistema por medio de

la incorporación de nodos, cada una de las cuales corresponde a agregar una fila y una columna a la matriz. Dichas incorporaciones se pueden efectivizar o no basándose en el criterio de estabilidad ya mencionado.

Más específicamente, supóngase que se dispone de una red de tamaño n , a la cuál le corresponde una matriz de rango n que representa un sistema estable, es decir, todos los autovalores de dicha matriz poseen parte real negativa. Luego para la incorporación del $(n + 1)$ -ésimo nodo se elige aleatoriamente su grado k_{n+1} sorteando el valor del mismo a partir de una distribución de probabilidad uniforme entre 1 y n . Luego las interacciones de este nuevo agente con cada miembro previamente existente i se eligen de manera tal que los elementos no diagonales de matriz $a_{i,n+1}$ y $a_{n+1,i}$ son nulos con probabilidad $1 - k_{n+1}/n$ y diferentes de cero con probabilidad k_{n+1}/n . Los valores particulares que adoptan los elementos no nulos de la matriz son elegidos al azar de una distribución uniforme en $[-b, b]$. b determina el rango de variabilidad de las interacciones del sistema y es uno de los dos parámetros del modelo. Luego se calcula numéricamente λ_{max} para la matriz $(n + 1) \times (n + 1)$ resultante. Si $\lambda_{max} < 0$ el nuevo nodo es aceptado, caso contrario significa que la introducción del nuevo nodo ha desestabilizado el sistema, y por ende se impondrá que, o bien el nuevo agente es removido volviéndose a la matriz original, o bien éste permanece pero a costa de provocar la extinción de un cierto número de agentes previamente existentes en la red. Con el fin de simplificar aún más los cálculos numéricos, se permite la extinción de a lo sumo q agentes del sistema ($q \leq k_{n+1}$). Los mismos se eligen aleatoriamente de entre los k_{n+1} primeros vecinos del nuevo nodo incorporado. q constituye el segundo parámetro del modelo. Mas precisamente, para elegir cuales nodos serán eliminados, primero se elige uno al azar de entre los k_{n+1} disponibles y se lo remueve. Si la matriz $n \times n$ resultante corresponde a un sistema estable ($\lambda_{max} < 0$), entonces el nuevo sistema es aceptado y se recomienza con la incorporación de nodos. Si por el contrario la remoción del nodo no estabilizó el sistema, se remueve otro más de entre los $k_{n+1} - 1$ remanentes repitiéndose el anterior procedimiento. Si tras la remoción de q nodos el sistema no estabiliza, entonces se retorna al sistema original de la matriz $n \times n$ antes de haber introducido el nuevo nodo, para volver a comenzar con nuevos intentos de incorporar nodos. El proceso se repite de modo que eventualmente la red va incorporando nodo tras nodo y así va creciendo.

Capítulo 4

Resultados

A continuación se presentan los resultados obtenidos a partir de estudiar estadísticamente las redes generadas con el modelo. Comúnmente cada proceso de crecimiento de una red se lo deja evolucionar hasta que la misma alcanza un tamaño máximo n_{max} (típicamente hasta valores de $n_{max} = 200$). Luego se repite del orden de 3×10^5 veces el proceso de crecimiento completo hasta el tamaño máximo n_{max} con el fin de obtener un muestreo lo suficientemente representativo como para que el análisis estadístico de las redes obtenidas a lo largo del proceso sea adecuado. En general, la disminución en las fluctuaciones de la cola en la distribución de grados $P(k)$ es un buen indicador de si se han calculado suficientes muestras.

4.0.1. Detalles técnicos

El cálculo de los autovalores se realiza utilizando el método numérico de factorización o descomposición QR que incorpora el uso de reflexiones de Householder lo cuál mejora su rendimiento tanto en velocidad como en precisión numérica. Al calcular numéricamente los autovalores se presentan inevitablemente ciertas dificultades, además de la precisión numérica, ya que en algunos casos el algoritmo no converge. Cuando ello ocurre, simplemente descartamos el proceso de crecimiento y comenzamos nuevamente con otro. Este inconveniente ocurre muy esporádicamente, por lo que la proporción de procesos de crecimientos descartados es poco representativa dentro del muestreo del ensamble.

Otra dificultad técnica que se ha presentado corresponde a la instancia del proceso de crecimiento en que se intenta incorporar un nodo a la red. Para cada paso de crecimiento, denominemos por ν al número de intentos que se han de realizar con la intención de incorporar un nuevo nodo que lleve a un sistema estable. No muy frecuentemente, algunas veces la red alcanza un estado en el que se necesitan demasiados intentos ν para lograr que la red crezca, haciendo que el algoritmo se vuelva numéricamente demasiado costoso. Para eludir esta dificultad nos hemos visto ante la necesidad de introducir un parámetro artificial ν_{max} , de

modo que se intenta incorporar un nuevo nodo a la red como máximo un número ν_{max} de veces. Si tras $\nu_{max} \simeq \mathcal{O}(n_{max})$ intentos no se ha logrado incorporar un nuevo nodo, entonces el proceso de crecimiento se descarta y se comienza uno nuevo. Se han realizado simulaciones para diferentes valores del parámetro ν_{max} encontrándose que a partir de cierto valor los resultados obtenidos resultan indistinguibles. Desde el punto de vista práctico ν_{max} es finito, pero desde el punto de vista teórico ν_{max} es infinito por lo que no se lo considera un parámetro del modelo, sino un artificio numérico para poder lidiar con el problema de reducir el tiempo de ejecución de los programas.

A partir de la siguiente sección, primero se discutirán los resultados concernientes a las propiedades topológicas de las redes generadas con el modelo y como se relacionan estas con otros resultados previamente presentados en la bibliografía. Luego se discutirán resultados que muestran que el proceso de selección no sólo afecta su topología, sino que también afectan la estructura del tipo de interacciones influenciando así la dinámica asociada.

4.0.2. Las distintas variaciones del modelo.

A lo largo del resto de la presentación de los resultados, eventualmente se introducirán variaciones al modelo con el fin de estudiar alguna propiedad de interés. Las definiciones precisas de las distintas variaciones del modelo, junto a las razones de *porqué* son introducidas, serán detalladas a lo largo del texto. En esta sección se provee de una lista de las distintas variaciones del modelo, junto con una breve descripción de cada una de ellas.

La lista es la siguiente:

- i) *el modelo básico*: Este modelo es el que fué introducido en el Capítulo 3.
- ii) *el modelo de estabilidad local*: En el *modelo básico* a la hora de evaluar la estabilidad del sistema se analizan los autovalores de la matriz $a_{i,j}$ asociada a todo el sistema. En cambio, en esta variación del modelo se modifica la forma en que se evalúa la estabilidad. Básicamente, en vez de analizar los autovalores de la matriz que representa todo el sistema, se analizan los autovalores de una *submatriz* que representa una porción del sistema en cuestión. Mas precisamente, se evalúa la estabilidad de un *sub-sistema* del sistema original. Este *sub-sistema* está constituido por el último nodo incorporado al sistema más los vecinos de orden $R \in \{1, 2, \dots, n-1\}$ del mismo. Es así que R constituye un parámetro propio de esta variación del *modelo básico*. Esta variación del modelo es introducida y discutida en la sección 4.1.2.

iii) *modelo de interacciones anticorrelacionadas*: En esta variación del *modelo básico* las interacciones representadas por las entradas $a_{i,j}$ de la matriz asociada al sistema, no son sorteadas a partir de una distribución uniforme en $[-b, b]$ de manera tal que no hayan a priori correlación entre ellas *. Por el contrario, en esta variación se exigirá a priori que las interacciones $a_{i,j}$ y $a_{j,i}$ presenten una dada probabilidad de estar anticorrelacionadas (esto es, que tengan signo contrario entre si cuando ambas son no nulas). El grado de anticorrelación que presentarán a priori vendrá regulado por un parámetro $h \in [0, 1]$ propio de esta variación del *modelo básico*. Esta variación del modelo es introducida y discutida en la sección 4.2.2.

4.1. Propiedades topológicas

4.1.1. Conectividad

En primer lugar, se discute el análisis de el promedio sobre el ensamble de procesos de crecimiento de la conectividad de las redes C en función del tamaño n (ver sección 2.8). Conviene hacerlo de esta manera, ya que provee de una conveniente primera visualización de que es lo que ocurre a lo largo del proceso de crecimiento. En la figura (4.1) se observa el comportamiento típico de $C(n)$ para diferentes valores de parámetro b . Las curvas no dependen del parámetro q † (ver figura (4.2)). En general se observa que C decrece con n , pero hay dos comportamientos distintos separados por una región de transición. Para valores chicos de n (típicamente hasta $n \simeq 10$) la conectividad va disminuyendo lentamente con n manteniéndose alrededor de $C \simeq 0.5$. Luego de la región de transición entre $n \simeq 10$ y $n \simeq 25$, la relación entre C y n viene dada por un comportamiento tipo ley de potencias $C(n) = Bn^{-(1+\epsilon)}$, donde $\epsilon > 0$ pero bastante menor que 1. El ensamble de matrices producidas por nuestro modelo es un subensamble de las matrices estudiadas por R. May (ver sección (2.9)). A partir de lo obtenido por May, es posible dar una explicación del comportamiento de la curva $C(n)$ en nuestro modelo. Dado que el grado de los nodos que intentan ingresar a la red son sorteados entre el mínimo y el máximo posibles (es decir en $\{1, n-1\}$ para una red de tamaño n), en ausencia de selección o bajo una fuerza selectiva muy débil, la conectividad tiende al promedio $C \simeq 0.5$ que es lo que ocurre para valores chicos de n . A medida que el tamaño del sistema crece, la curva $C(n)$ penetra demasiado en la zona de inestabilidad por lo que a partir de ahí la presión selectiva impuesta por el criterio de estabilidad se impone ya que la conectividad supera el límite de May para el tamaño

*A priori en el sentido de que si aparecen correlaciones, entonces son productos del proceso de selección inducido a través de exigir estabilidad en el sistema.

†Sólo se observa una pequeña variación para valores pequeños de n que se discutirá más adelante en la sección 4.1.2.

dato, desestabilizando el sistema. El resultado final surge así como un equilibrio entre dos tendencias opuestas que compiten: el flujo externo constante que tiende a aumentar la conectividad y la desestabilización que tiende a disminuirla. Allí es donde el comportamiento de la curva va cambiando hasta alcanzar el régimen $C \sim n^{-(1+\epsilon)}$. En los insets de la figura (4.1) se observa que al variar b , ϵ se mantiene mas o menos constante dentro del rango $[0.1, 0.3]$, mientras que B cambia al menos un orden de magnitud. El comportamiento de $C(n)$ ante la variación de b es razonablemente bien representado por la siguiente relación [164]:

$$C \sim \alpha^{-\omega} n^{-(1+\epsilon)} \quad (4.1)$$

en donde $\omega = 0.7 \pm 0.1$, y $\alpha = b^2/3$ dado que las entradas de la matrices son sorteadas a partir de una distribución uniforme en $[-b, b]$.

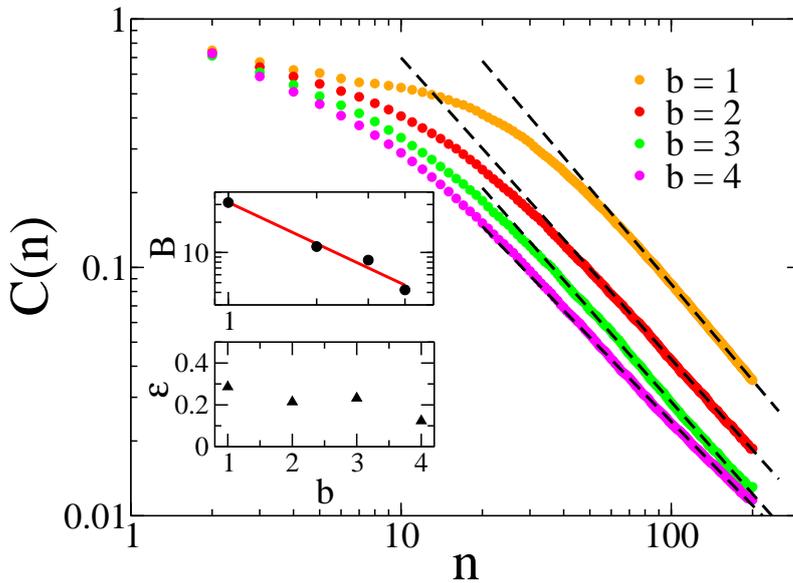


Figura 4.1: Conectividad en función del tamaño de las redes para $q = 3$, $n_{max} = 200$ y diferentes valores de b [164]. Los símbolos corresponden a simulaciones numéricas, y las líneas punteadas a ajustes tipo ley de potencias de las colas de las curvas $C = Bn^{-(1+\epsilon)}$. Los insets muestran los valores ajustados de B y ϵ en función de b . Los valores de de los parámetros en dichos ajustes están en la tabla (4.2) (sección 4.1.7).

En la figura (4.3) se observa las curvas de $C(n)$ obtenidas con nuestro modelo comparadas con las correspondientes curvas de transición del modelo de May, las cuales en este caso vienen dadas por $C_{trans} = 1/(b^2/3)n$. Es notable como las matrices generadas por nuestro modelo crecen penetrando significativamente en la zona de inestabilidad. Esto es un indicativo de que las redes que crecen bajo la presión selectiva del criterio de estabilidad desarrollan algún tipo de estructura muy particular cuya probabilidad de ocurrir en un ensamble completamente aleatorio es

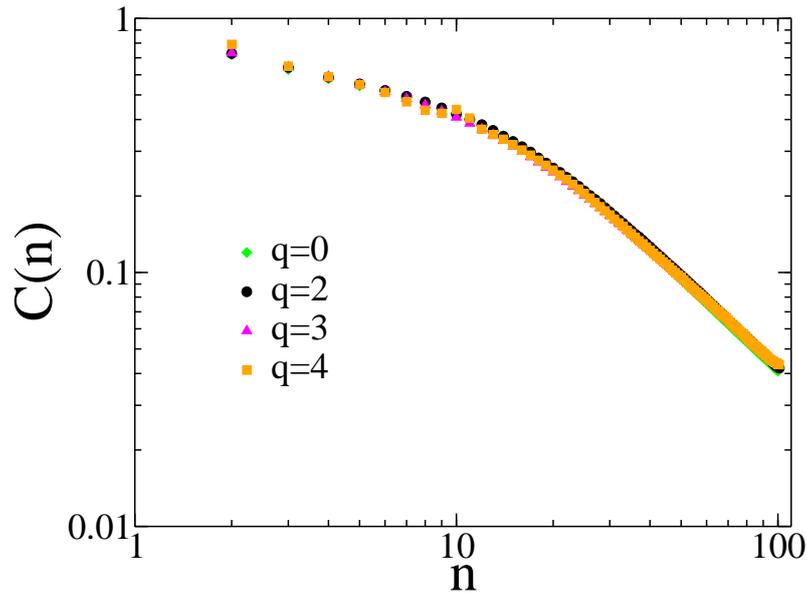


Figura 4.2: Conectividad en función del tamaño de las redes para $b = 2$, $n_{max} = 100$ y diferentes valores de q . Las curvas solapan indicando que el comportamiento de $C(n)$ es independiente de q . En otras palabras, el proceso de extinciones no se ve reflejado en la cantidad de conexiones que adopta el sistema. Sin embargo se observará al analizar la distribución de grados $P(k)$ que sí afecta la forma en que se distribuyen dichas conexiones (sección 4.1.2).

prácticamente cero. En otras palabras las matrices en cuestión pertenecen a un subensamble muy poco probable dentro del ensamble aleatorio, y por ende son obtenibles únicamente a partir de un proceso de selección. En las secciones venideras, se estudiará la estructura de las redes correspondientes a dicho subensamble.

En la figura (4.4) se grafica la conectividad para diferentes redes biológicas a lo largo de tres ordenes de magnitud del tamaño de las redes a partir de información obtenida en la literatura. Se observa que los datos son bien ajustados por una única ley de potencias $C(n) \sim n^{-1.2}$ en buena concordancia con el valor promedio de $\epsilon = 0.2$ predicho por el presente modelo. El comportamiento de escala de $C(n)$ para las redes tróficas es el centro de un viejo debate en ecología (ver sección 2.8.1). Mientras que es esperado un comportamiento tipo ley de potencias, el valor del exponente (y de las asociadas interpretaciones) es controversial, en parte debido a la gran dispersión de los datos disponibles, el relativamente pequeño rango de tamaños correspondientes, y en algunos casos la baja resolución de los datos. Por ejemplo, un ajuste tipo ley de potencias de los datos únicamente correspondientes a las redes tróficas de la figura (4.4) resulta en un exponente menor a 1 con un error relativo de alrededor del 50% y un coeficiente de correlación $R^2 = 0.5$. La consistencia del escaleo ob-

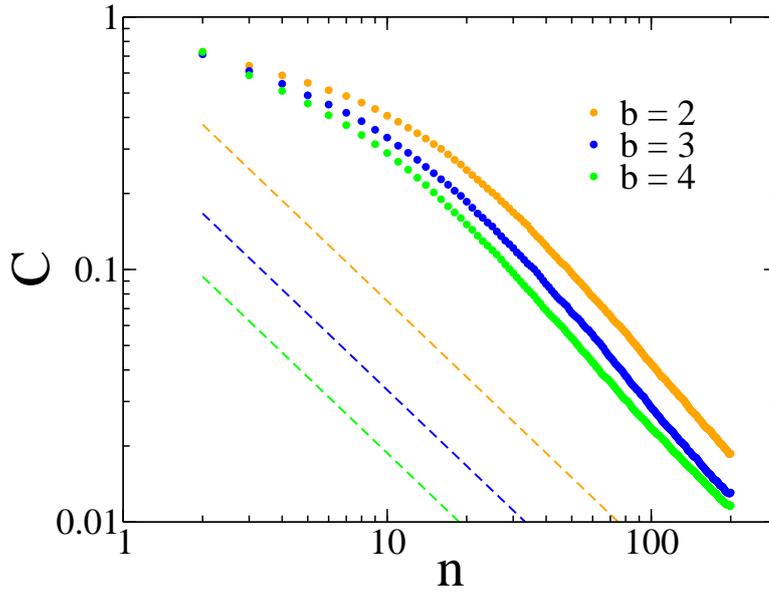


Figura 4.3: Comparación de las curvas de conectividad versus el tamaño obtenidas con el modelo aquí propuesto (círculos), respecto de las curvas de transición entre las zonas de estabilidad e inestabilidad encontradas por May [125] (líneas discontinuas). Mismo color corresponde al mismo valor de b . Se observa que las matrices generadas por nuestro modelo se encuentran bien adentro de la zona de inestabilidad, lo cuál es indicativo de que han adquirido algún tipo de estructura particular que les permite ser estables aún en dicha región.

servado en la figura (4.4) para un rango de tamaños mucho más amplio sugiere que el debate en ecología debe ser reconsiderado en un contexto evolutivo más amplio. Es importante mencionar que el comportamiento $C(n) \sim n^{-(1+\epsilon)}$ también ha sido obtenido en un modelo de criticalidad auto-organizada en redes tróficas [186].

Observación: En la sección dedicada al comportamiento de $C(n)$ para las redes ecológicas (sección 2.8.1) se vió que aún está inconcluso si $\epsilon < 0$ ó $\epsilon > 0$. El resultado presentado en ésta sección es evidencia teórica en favor del segundo caso. Sin embargo, una pequeña modificación al *modelo básico* (la variación *iii*) ó *modelo de interacciones anticorrelacionadas* mencionado en la sección 4.0.2 que será estudiada en la sección 4.2.2) también permite reproducir el caso $\epsilon < 0$. En otras palabras, el enfoque aquí propuesto también provee de un explicación posible en caso de que las redes ecológicas presenten el comportamiento $\epsilon < 0$ diferente al encontrado en el marco más amplio presentado en la figura (4.4) en donde $\epsilon > 0$.

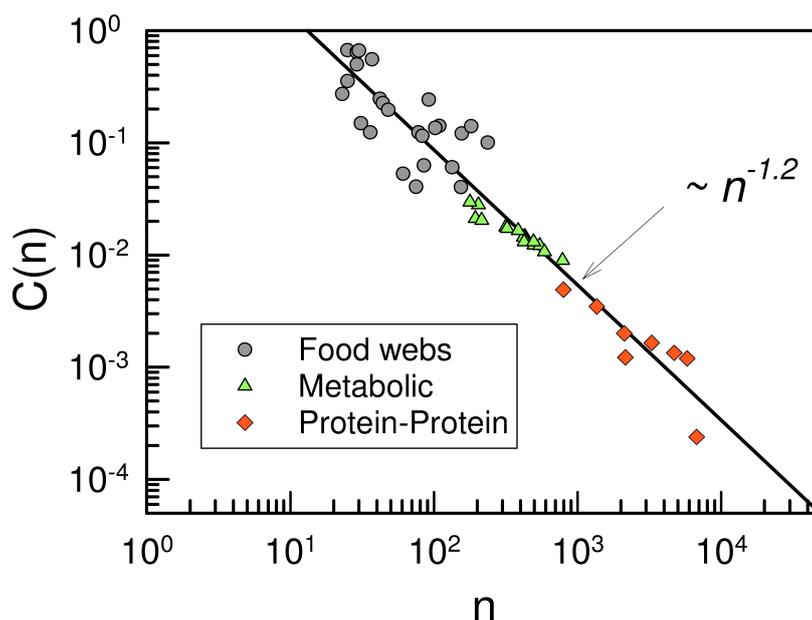


Figura 4.4: Conectividad en función del tamaño de las redes para diferentes redes biológicas. La línea recta es un ajuste tipo ley de potencias $C = Bn^{-(1+\epsilon)}$ resultando en un exponente $\epsilon = 0.2 \pm 0.1$ ($R=0.92$) [164]. Los datos fueron extraídos de: [209, 53] (redes de interacciones proteína-proteína), [97] (redes metabólicas), [129, 65, 64] (redes tróficas).

4.1.2. La distribución de grados

En la figura (4.5) se observa el promedio de las distribuciones de grado de las redes resultantes a lo largo de varios procesos de crecimiento. La misma presenta para cierto rango de valores de k pertenecientes a la cola de la curva un comportamiento tipo ley de potencias:

$$P(k) \sim k^{-\gamma} \quad (4.2)$$

cuyo exponente γ depende de la elección de los valores de los parámetros b y q . En otras palabras, la presión selectiva basada en un criterio de estabilidad induce una topología libre de escala a las redes producidas por nuestro modelo. La desviación más notable del comportamiento tipo ley de potencias se observa a valores pequeños de k . Lo que ocurre es que, tal como ya se ha discutido en la sección 4.1.1, al comienzo del proceso de crecimiento la red aún no ha llegado al borde de la región en donde no es necesario una restricción topológica para lograr estabilidad. En dicha región la red aún posee las características topológicas de las redes aleatorias, y conformará un *núcleo* que sobrevivirá en gran medida al resto del proceso de crecimiento. El lomo de la curva corresponde a este núcleo de características topológicas aleatorias. Es fácil ver que la distribución de grados de redes que no han superado un tamaño de alrededor de $n \simeq 25$ poseen esta forma de lomo que corresponde a una distribución de grados

de cola corta. Según lo observado en la sección 2.8.3, éste resultado es interesante ya que provee de una explicación a la variación encontrada en el comportamiento de $P(k)$ con n en redes ecológicas según es reportado en [134]. Más precisamente, provee una explicación de porqué redes tróficas de tamaños pequeños presentan una distribución de grados de cola corta, mientras que redes tróficas de tamaños grandes una distribución de grados de cola larga. Otra desviación del comportamiento ley de potencias en la distribución de grados del presente modelo puede notarse en el extremo de la cola, la cuál decae abruptamente. Esto es un efecto de tamaño finito como puede verse al comparar las curvas tomadas a diferentes tamaños n_{max} del proceso de crecimiento (triángulos, cuadrados y círculos de la figura (4.5)) en donde se comprueba que γ es independiente de n_{max} . Se consideró una posible explicación a la caída de la cola de la curva en la $P(k)$. La curva $P(k)$ se calcula promediando sobre sistemas que alcanza el tamaño máximo predeterminado n_{max} . Sistemas a tamaños n inferiores a n_{max} son frecuentemente el producto del encogimiento de sistemas a tamaños mayores como resultado de la extinción de algunos de sus nodos. Ello no ocurre en sistemas que alcanzan el tamaño n_{max}^* , lo cuál podría ser la causa de que se observe la caída de la cola al medir la distribución de grados $P(k)$. Sin embargo, no es el caso ya que la caída de la cola también se ha observado al medirse $P(k)$ en tamaños n inferiores a n_{max}^\dagger , e incluso en ausencia de extinciones ($q = 0$).

En el inset de la figura (4.5) se observa como varía el exponente γ del ajuste tipo ley de potencias en función del parámetro b para diferentes elecciones del parámetro que regula el tamaño máximo de las extinciones q . Se observa que el exponente varía aproximadamente entre $\gamma \simeq 2.4$ y $\gamma \simeq 5$, obteniéndose un mínimo en aproximadamente $b \simeq 2$ y $q = 3$ o 4. Es importante resaltar que existe un rango de tamaño considerable de valores de b para los cuales $\gamma < 3$. Más aún, en dicho rango de valores los exponentes tienden a ser aparentemente independientes de q . En la siguiente sección se discutirá la importancia de éste resultado.

La importancia de las extinciones

Cuando se discutió el mecanismo de presión selectiva basado en la estabilidad, se mencionó que la incorporación de un agente al sistema podría eventualmente provocar una avalancha de extinciones de agentes previamente residentes en el sistema. Dicho ingrediente puede parecer natural si uno piensa por ejemplo en redes tróficas en donde la llegada de una nueva especie al ecosistema puede provocar un desequilibrio que eventualmente conduzca a la extinción de otras especies. También es viable si uno piensa en una red de expresión genética, en donde podría ocurrir que la expresión simultánea de ciertos genes podría ser incom-

*por la evidente razón de que no hay sistemas a tamaños mayores a n_{max} .

†Se ha utilizado para la prueba $n = 100$ y $n_{max} = 120$. Las figuras no son mostradas por resultar triviales.

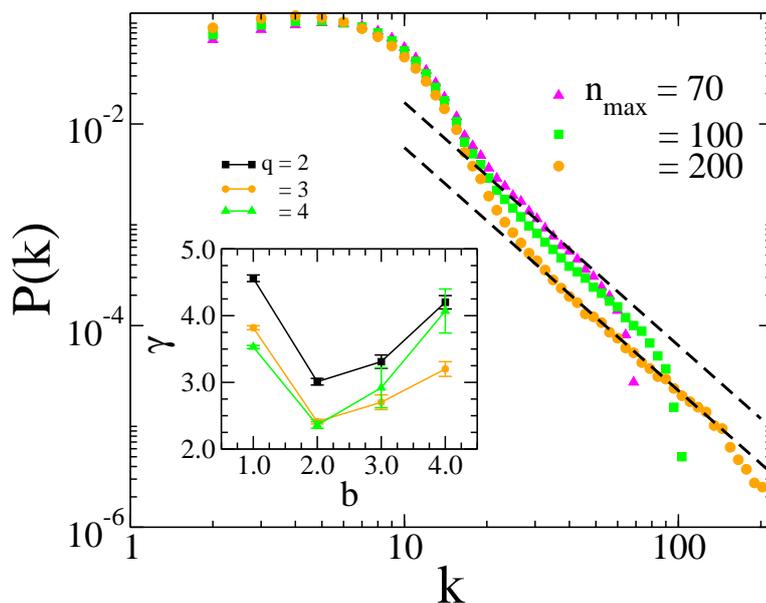


Figura 4.5: Distribución de grados $P(k)$ para diferentes valores de n_{max} cuando $b = 2, q = 3$ (adaptación de la figura 1 en [165]). Se observa que $P(k)$ presenta una cola que decae como una ley de potencias truncada. El comienzo del comportamiento ley de potencias siempre ocurre alrededor de $k \simeq 25$, valor que no depende de n_{max} . El truncamiento corresponde a un decaimiento más pronunciado que una ley de potencias y es un efecto a tamaño finito ocurriendo a valores de $k \simeq n_{max}$. En conclusión, extrapolando a tamaño infinito los sistemas corresponden a redes que presentan una distribución de grados cuya cola se comporta de acuerdo a $P(k) \sim k^{-\gamma}$. El exponente γ resulta independiente de n_{max} tal como sugieren las líneas entrecortadas. A las curvas se le aplicó un *binning logarítmico* para mejorar la visualización de las mismas. *Inset*: Dependencia del exponente γ con los parámetros b y q para $n_{max} = 100$ (reproducción del inset de la figura 1 en [165]). Se puede apreciar que existe un rango de valores de b para los cuales $\gamma < 3$ resultando en un comportamiento no trivial de la distribución de grados.

patible para la vida, por lo que la incorporación de ciertos genes podría eventualmente inducir la desaparición de otros. Sin embargo a priori no queda claro que la posibilidad de permitir extinciones sea crucial para que el mecanismo de presión selectiva conduzca a resultados contrastables, al menos cualitativamente, con la realidad. El incorporar la posibilidad de que ocurran avalanchas de extinciones al modelo claramente es una complicación del mismo. Sin embargo la ausencia de este ingrediente, lo que corresponde a elegir $q = 0$, conduce a que las distribuciones de grado resultantes presenten un comportamiento tipo ley de potencias, pero con exponentes en donde $\gamma > 3$. Que el exponente sea $\gamma > 3$ implica que la varianza de la distribución de grados es finita, por lo que a la red le corresponde un grado característico tal como ocurre para las redes aleatorias. En otras palabras $\gamma > 3$ es consistente con una distribución de grados de cola corta. Mas precisamente, se ha establecido que cuando

$\gamma > 3$ se pierden muchas de las propiedades que diferencian a las redes de cola larga de las de cola corta [25]. Por ejemplo, en redes de *cola larga* en la distribución de grados, ciertas estrategias de búsqueda se ven favorecidas [8], hay ausencia de un mínimo umbral en la tasa de infección para la prosperación de una epidemia [162], hay resistencia a la fragmentación de la red frente a fallas [12, 51, 10], y responden de manera diferente ante ataques dirigidos a como lo hacen redes aleatorias [12, 52, 10]. En definitiva, la dinámica de extinciones – en este caso controlada por el parámetro q – juega un papel importante sobre los resultados cualitativos que se desprenden del mecanismo de presión selectiva.

Cuando se incorporó el mecanismo de extinciones al modelo, una de las intenciones era estudiar la distribución de tamaños de avalanchas de las mismas. Sin embargo se encontró que para $q \geq 5$ las redes no crecen más allá del tamaño $n \simeq 25$. No queda claro si este resultado es un artificio introducido por la particular forma de implementar los procesos de extinciones en el modelo, o si el resultado posee una validez más general. En particular, existe la posibilidad de que lo que ocurre sea un artificio numérico y no se descarta a priori que las redes no puedan crecer más allá de $n \simeq 25$ en tiempos de simulación prohibitivamente largos. Por otro lado, también es posible que ocurra algún cambio cualitativo en la dinámica de crecimiento cuando $q \geq 5$, de manera tal que efectivamente impone una cota máxima al tamaño al cuál las redes pueden crecer. Si ese es el caso, sería interesante estudiar la naturaleza de la transición. Hay evidencia de que al menos en el contexto de las redes tróficas, la distribución de tamaños de extinciones corresponde a una ley de potencias [102], algo que es consistente con los registros de extinciones fósiles [138], y que también ha sido modelado [188, 152]. Dicha fenomenología no ha podido ser corroborada por el presente modelo. Nuevos modelos o variantes de este mismo, eventualmente puedan arrojar luz sobre esta cuestión y su relación con la presión selectiva basada en un criterio de estabilidad.

La robustez de la topología compleja

Considerando que los sistemas biológicos nunca se encuentran en una situación completamente estable, relajamos el criterio de estabilidad $\lambda_{max} < 0$, permitiendo a λ_{max} tomar pequeños valores positivos de modo que el tiempo característico para que el sistema abandone un punto fijo inestable sea $\tau = \lambda_{max}^{-1} \gg 1$. Aceptando nuevos nodos siempre y cuando $\lambda_{max} < \Delta$, el cálculo de la distribución de grados $P(k)$ para distintos valores de Δ (positivos y negativos) muestra resultados cualitativamente similares a los ya obtenidos para el caso $\Delta = 0$ (ver figura (4.6)). Sólo se encuentran pequeñas variaciones en la curva que afectan levemente el exponente γ . Puede concluirse así que la emergencia de una topología compleja es robusta ante variaciones del criterio de estabilidad. No es necesario que el proceso de selección permita sobrevivir a aquellos

sistemas que se encuentren efectivamente en una situación de estabilidad, sino que basta con que el comportamiento resulte lo suficientemente estable en la escala de tiempo en consideración.

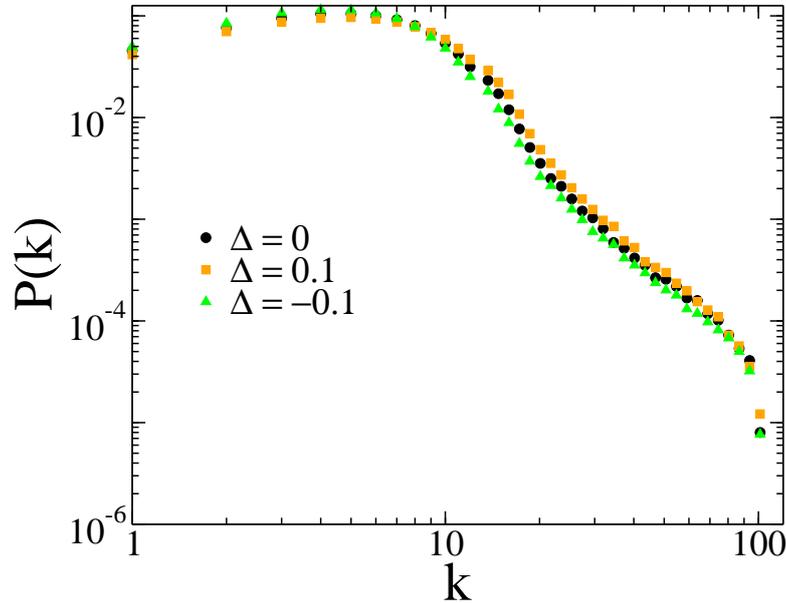


Figura 4.6: Distribución de grados $P(k)$ para $b = 2, q = 3, n_{max} = 100$ y diferentes valores de Δ [165]. A las curvas se le aplicó un *binning logarítmico*. No hay un cambio significativo en la $P(k)$ ante variaciones de Δ , indicando que la emergencia de una topología compleja es robusta ante variaciones del criterio de estabilidad.

Hasta el momento el criterio de estabilidad ha sido aplicado a nivel global en el sistema. Sin embargo podría ocurrir que eventualmente la mayor parte del sistema presente un comportamiento estable, mientras que quizás una pequeña fracción del mismo aún no haya alcanzado el equilibrio o quizás se haya desestabilizado sin que la desestabilización se haya propagado al resto del sistema. Para ello se estudia y analiza una variación del modelo. Dicha variación del modelo es: denominada *el modelo de estabilidad local* (ver sección 4.0.2, ítem *ii*). Mas precisamente, considérese ahora el efecto de la presión selectiva, pero basado en un criterio de estabilidad local en vez de global. La idea es, en vez de evaluar la estabilidad de la matriz que representa a todo el sistema, evaluar la estabilidad de una submatriz que representa una porción del sistema, una subred. El algoritmo de generación de matrices se modifica como sigue. Cuando un nuevo nodo ingresa al sistema, la incorporación del mismo se efectiviza o no evaluando la estabilidad de la subred compuesta de los vecinos de hasta orden R de este nuevo nodo ($R = 0$ el nuevo nodo, $R = 1$ primeros vecinos, $R = 2$ segundos vecinos, etc). En la figura (4.7) se muestra la distribución de grados $P(k)$ a tamaño $n_{max} = 100$ para diferentes valores de R . Se observa que la cola larga tipo ley de potencias

aparece a partir de $R \geq 3$. Como se verá más adelante al estudiar la propiedad de mundo pequeño de la topología de las redes del presente modelo para redes de tamaño $n = 100$ la distancia topológica típica entre nodos elegidos al azar es aproximadamente $L \simeq 3$ (ver figura (4.14) y texto relacionado). Esto sugiere una correlación entre la condición de estabilidad y la emergencia auto-organizada de la topología de mundo pequeño. Nótese además que considerar la estabilidad local da lugar a una mayor variación del valor del exponente γ ($\gamma = 0.9$ para $R = 3$), aunque γ rápidamente converge al resultado obtenido cuando se considera la estabilidad global (para $R \gtrsim 6$ ambos resultados son prácticamente indistinguibles).

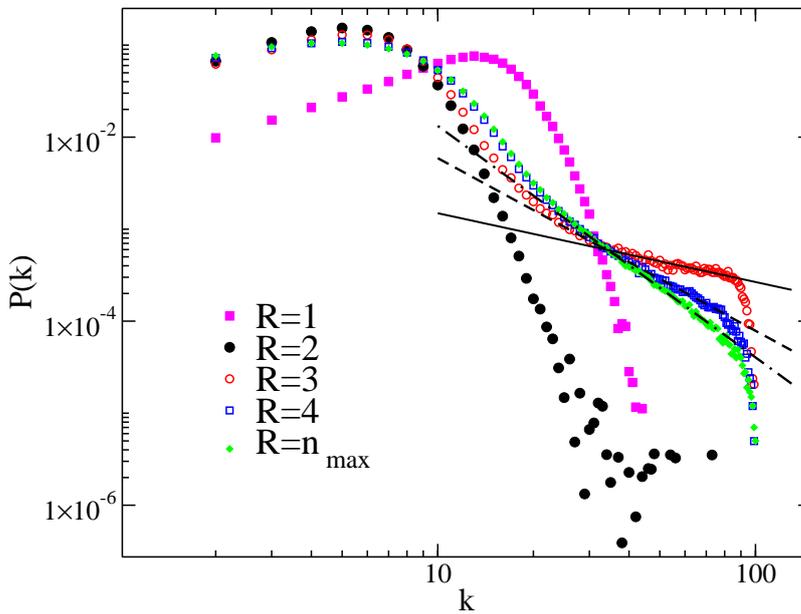


Figura 4.7: Distribución de grados $P(k)$ para $n_{max} = 100, b = 2, q = 3$ y distintos valores de R [165]. Se observa que la cola larga aparece cuando $R \geq 3$ que coincide con el valor de L para el correspondiente tamaño de la red. Nótese que la consideración de la estabilidad local permite una mayor variabilidad en el exponente γ ($\gamma \simeq 0.9$ para $R = 3$), aunque γ converge rápidamente al valor del resultado con estabilidad global (para $R \gtrsim 6$ ambos resultados se hacen prácticamente indistinguibles).

En la figura (4.8) se observa el tamaño promedio $n_s(n) \leq n$ de las subredes a las que se le evalúa la estabilidad al considerar la estabilidad local, en función del tamaño n . Se observa que para $R \geq 6$ el tamaño de las subredes es igual al tamaño del sistema, es decir $n_s(n) \simeq n$. Para $R \leq 5$ el tamaño promedio de las subredes va por debajo del tamaño del sistema, siendo más acentuada la diferencia cuanto más chico es R . En particular entre $R = 3$ y $R = 2$ se observa una relativamente drástica disminución del tamaño de las subredes frente al tamaño del sistema. Para $R \geq 3$ la subred a la que se le evaluará la estabilidad posee – para

n 's grandes – aproximadamente el 75 % de los nodos (como mínimo, mas grande es R mayor es el porcentaje). En cambio para $R \leq 2$ dicha subred representa no más del 25 % del sistema. Este cambio abrupto coincide con la pérdida del carácter cola tipo ley de potencias de la distribución de grados $P(k)$. Este resultado refuerza la hipótesis de que existe una correlación entre el criterio de estabilidad y la emergencia auto-organizada de la propiedad mundo pequeño. Finalmente, en las figuras (4.9) y (4.10) se grafica $n_s(n)$ para diferentes valores de b y q respectivamente. Vemos que no se observan variaciones significativas de $n_s(n)$ con b y q .

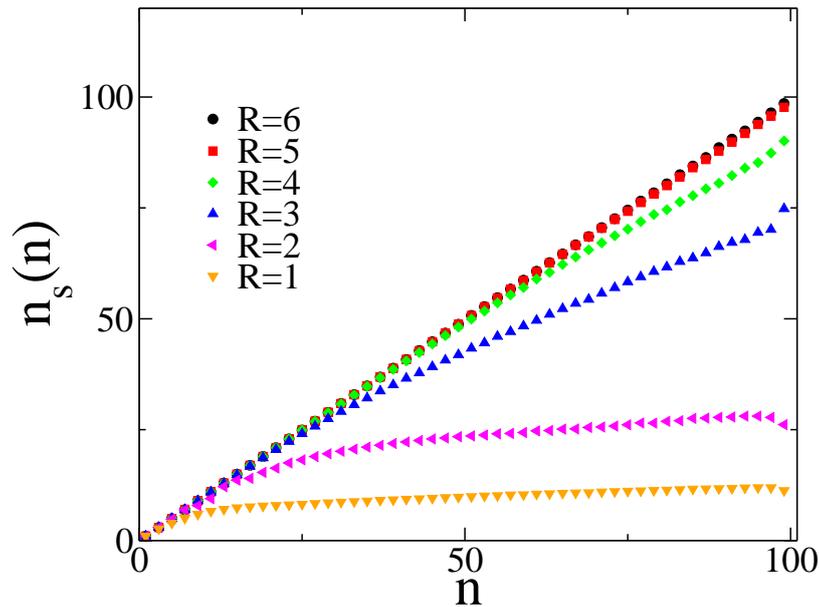


Figura 4.8: Tamaño promedio de las subredes $n_s(n)$ en función del tamaño n para distintos valores de R . Se observa que a $R = 6$ el tamaño promedio escala de la forma $n_s(n) \simeq n$, lo cual es un indicador de que las subredes poseen el mismo tamaño que la red que corresponde al sistema completo. Lo mismo ocurre para $R > 6$. A partir de $R \leq 5$ el tamaño promedio de las subredes comienza a decrecer respecto del tamaño del sistema. En particular entre $R = 3$ y $R = 2$ hay un cambio brusco del tamaño promedio de las subredes, lo cual corresponde a la transición entre el comportamiento con cola tipo ley de potencias de la distribución de grados $P(k)$, y el comportamiento en donde se pierde el carácter ley de potencias (ver figura (4.7)).

Los resultados obtenidos son consistentes con la hipótesis inicialmente barajada de que lo relevante es que sea la mayor parte del sistema lo que presenta un comportamiento estable. Si el criterio de selección ignora la estabilidad de una pequeña fracción del mismo, el resultado cualitativo se conserva, redes de topologías complejas emergen. Por el contrario si la fracción del sistema que presenta estabilidad no es la más representativa, luego el carácter complejo de la topología de la red se pierde en gran medida resultando en una distribución de grados de cola corta. En este sentido, se puede aseverar que la topología emergente posee otra cualidad

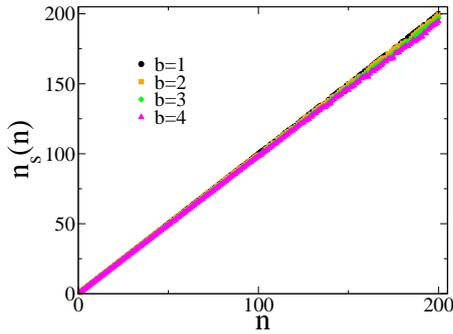


Figura 4.9: Tamaño promedio de las subredes $n_s(n)$ en función del tamaño n versus distintos valores de b para $q = 3$, $n_{max} = 100$, $R = 6$. Se observa que el comportamiento de $n_s(n)$ es prácticamente independiente de b en los rangos estudiados. Apenas se distingue una leve caída de las curvas a medida que aumenta b para valores grandes de n .

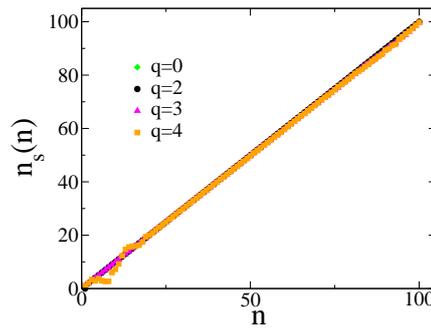


Figura 4.10: Tamaño promedio de las subredes $n_s(n)$ en función del tamaño n versus distintos valores de q para $b = 2$, $n_{max} = 100$, $R = 6$. No se aprecia una dependencia significativa de $n_s(n)$ con q .

de robustez ante variaciones del criterio de estabilidad.

La red de conexiones dirigidas

Algunas de las redes reales estudiadas son de carácter dirigido, en donde es claro que la relación que matemáticamente se representa por una conexión es de carácter asimétrico. Tal es el caso de redes como la World–Wide–Web, la red de citaciones en trabajos científicos, redes de e–mails, redes metabólicas, y redes ecológicas tales como las redes tróficas entre otras [10, 34, 41, 173]. Sin embargo, muchas veces se estudia la versión simetrizada de dichas redes dirigidas [133, 199, 200]. En este trabajo se estudiará ampliamente la versión simetrizada de las redes que emergen del proceso de selección basado en el criterio de estabilidad. Sin embargo, dado que uno de los resultados más importantes es que la distribución de grados es libre de escala, es relevante verificar si dicha propiedad se mantiene al no simetrizar las conexiones de la red. Por esta razón mostramos la distribución de grados de las conexiones entrantes/salientes (ver figura 4.11). Se observa que estas distribuciones presentan una cola tipo ley de potencias, al igual que la distribución de grados de la versión simetrizada de la red. Este resultado es de esperarse ya que en principio el criterio de estabilidad considerado en este modelo no distingue entre conexiones entrantes o salientes, por lo que por una cuestión de simetría el resultado correspondiente a las conexiones en un sentido tiene que ser idéntico al de las conexiones en el sentido opuesto. Es importante resaltar que esta simetría no necesariamente ocurre en las redes reales, sin embargo la presión selectiva basada en la estabilidad, según este modelo, no es la causa de dichas asimetrías. Esto es de esperarse, ya que a priori

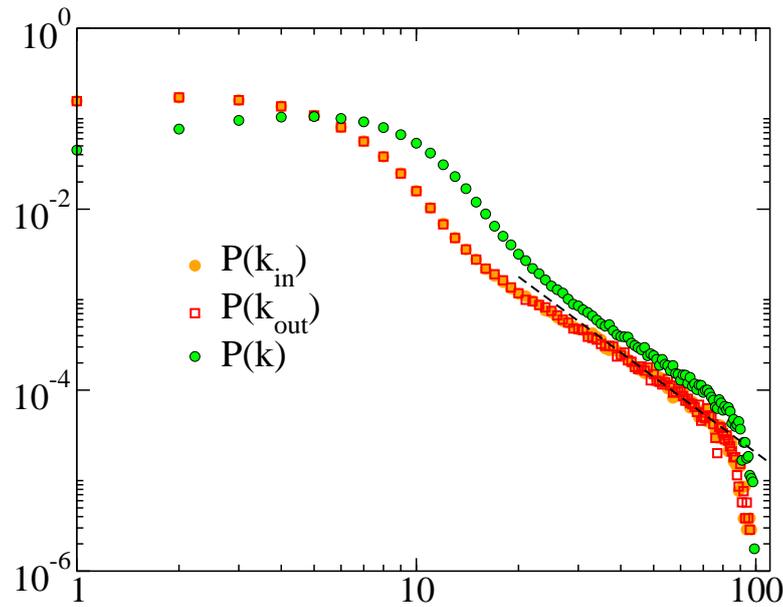


Figura 4.11: Comparación de las distribuciones de grados de las conexiones dirigidas $P(k_{in})$ y $P(k_{out})$, con la distribución de grados de las conexiones simetrizadas $P(k)$ para $n_{max} = 100$, $b = 2$ y $q = 3$. Vemos que $P(k_{in})$ y $P(k_{out})$ poseen el mismo comportamiento, lo cuál es de esperarse puesto que la presión selectiva no distingue entre conexiones entrantes y salientes. También puede verse que, aunque estas curvas son diferentes a la de la $P(k)$, presentan colas tipo ley de potencias con un exponente $\gamma_{in} = \gamma_{out} \simeq 2.8$ (dado por el ajuste que corresponde a la línea discontinua) levemente mayor al de la $P(k)$ que es aproximadamente $\simeq 2.5$.

el mecanismo pretende explicar un amplio espectro de tipos de sistemas, algunos de los cuales presumiblemente si presentan dicha simetría.

Sistemas en que $\langle a_{i,j} \rangle \neq 0$, tendencias cooperativas ó competitivas

Una cuestión interesante de estudiar es ver que ocurre si la simetría correspondiente a que la cantidad de interacciones positivas (estadísticamente hablando) sea la misma que la cantidad de interacciones negativas se rompe. Para ello estudiamos la distribución de grados de redes a tamaño máximo que fueron crecidas por medio del mecanismo de presión selectiva pero con una modificación al algoritmo. En vez de sortear las entradas no nulas $a_{i,j}$ de las matrices a partir de una distribución uniforme $[-b, b]$, se las sorteó a partir de una distribución uniforme en $[-b+x, b+x]$. De esta manera las matrices resultantes poseen una media $\langle a_{i,j} \rangle \neq 0$. Cuando $\langle a_{i,j} \rangle > 0$ ($x > 0$), corresponde a un sistema que presenta una tendencia cooperativa entre las componentes del mismo (ver figura (2.12)). Por el contrario, cuando $\langle a_{i,j} \rangle < 0$ ($x < 0$), el sistema presenta una tendencia competitiva entre sus componentes. En la figura

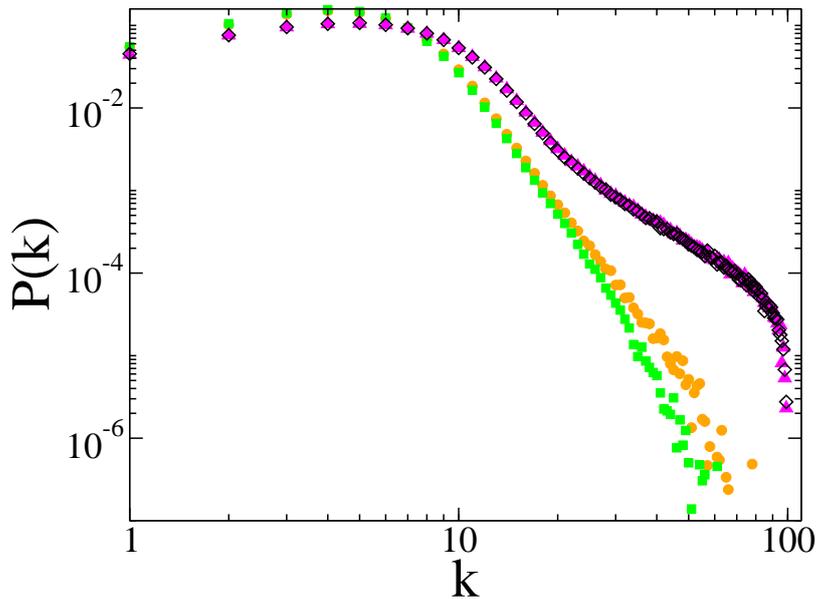


Figura 4.12: Gráfico de $P(k)$ para distintos valores de $\langle a_{ij} \rangle$ y para $n_{max} = 100, b = 2, q = 3$. Las curvas se comparan con el caso en que $\langle a_{ij} \rangle = 0$ (triángulos magenta). Cuando la ruptura de la simetría es pequeña ($\langle a_{ij} \rangle = 0.1$, rombos negros vacíos) no hay cambios apreciables en el comportamiento de $P(k)$. Cuando la ruptura es más acentuada (tendencia cooperativa $\langle a_{ij} \rangle = 1.0$ cuadrados verdes, tendencia competitiva $\langle a_{ij} \rangle = -1.0$ círculos anaranjados) se observa que la cola de la curva cae drásticamente aumentando el exponente γ por encima de 3. El cambio es más abrupto para el caso cooperativo que para el competitivo.

(4.12) se compara la distribución de grados del caso simétrico $\langle a_{i,j} \rangle = 0$ con casos en que se rompe la simetría tanto hacia una tendencia cooperativa como hacia una tendencia competitiva. Una pequeña ruptura de la simetría no genera cambios apreciables en la $P(k)$, por lo cuál la topología libre de escala es robusta frente a desviaciones de la situación simétrica. Sin embargo, cambios considerables en $\langle a_{i,j} \rangle$ resulta en cambios apreciables en la distribución de grados. Aunque el comportamiento tipo ley de potencias persiste, el exponente γ puede eventualmente crecer más allá de 3 y por lo tanto, las redes pierden muchas de las propiedades de las redes libres de escala (ver sección 4.1.2). Curiosamente la ruptura de la simetría no es simétrica, una tendencia cooperativa conduce más rápidamente a que $\gamma > 3$ frente a una tendencia competitiva. En otras palabras, si un valor bajo del exponente γ es deseable por un sistema en donde la estabilidad como presión selectiva juega un rol en la conformación de la red, luego las interacciones competitivas se ven levemente favorecidas.

El conexionado preferencial

Para excluir la posibilidad de que la propiedad libre de escala de la topología observada en las redes producidas por nuestro modelo, esté trivialmente asociada con la oculta ocurrencia del mecanismo de conexionado preferencial, hemos calculado la probabilidad $\Pi(k)$ de que un nodo ingresante se conecte a un nodo de grado k (ver ecuación (2.13) y texto relacionado en la sección (2.1.3)). En la figura (4.13) se muestra la probabilidad relativa de conectarse $\Pi(k)/P(k)$ de nuestro modelo para b fijo, y diferentes tamaños n , la misma es comparada con el resultado correspondiente para una red del mismo tamaño y conectividad $C(n)$ obtenida con el algoritmo de Barabási–Albert. Esta cantidad presenta el comportamiento esperado $\Pi(k)/P(k) \sim k$ en las redes de Barabási–Albert en consistencia con la ecuación (2.15). En el presente modelo $\Pi(k)/P(k)$ permanece prácticamente constante para un amplio rango de valores de k , incluyendo un rango de valores en donde la distribución de grados presenta comportamiento tipo ley de potencias. Sin embargo se observa que para valores suficientemente grandes de k , la probabilidad relativa $\Pi(k)/P(k)$ presenta una tendencia a crecer. En otras palabras, en el presente modelo (a diferencia de lo que ocurre con el modelo de Barabási–Albert), a medida que la red crece, el mecanismo de conexionado provocado por la presión selectiva muestra un entrecruzamiento (*crossover*) entre dos regímenes: uno dominado por el conexionado preferencial, y el otro no.

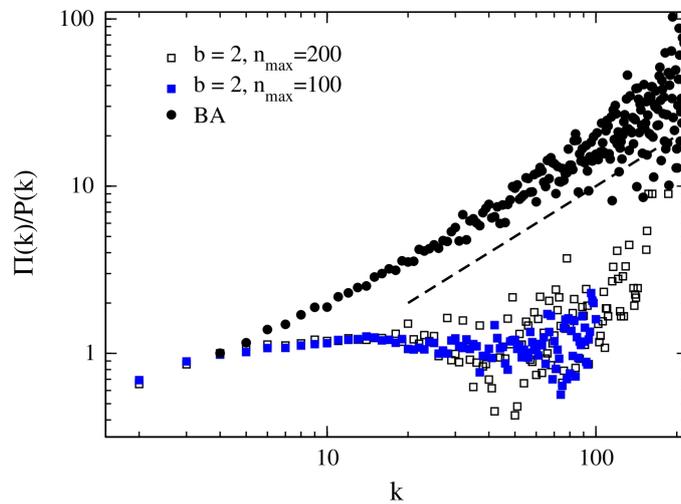


Figura 4.13: Comparación de $\Pi(k)/P(k)$ para distintos valores de n_{max} (cuadrados azules llenos y negros vacíos) con el de una correspondiente red de Barabási–Albert (círculos negros llenos) para $b = 2$ y $q = 3$ [165]. Los resultados son independientes de q y levemente dependientes de b . La tendencia a crecer de la cola se incrementa levemente con b . La línea punteada corresponde a un comportamiento lineal $\Pi(k)/P(k) \sim k$. Se observa de este modo que el conexionado preferencial no es al menos el principal mecanismo que rige el crecimiento de las redes de nuestro modelo.

4.1.3. Propiedades de mundo pequeño

Hemos visto que la topología de las redes que resultan del proceso de selección no corresponde a la de redes puramente aleatorias. Sin embargo, la construcción de las mismas claramente es no determinista, sino que el proceso posee una gran componente aleatoria. Se ha establecido que en dichas situaciones las redes presentan comúnmente la propiedad de mundo pequeño, a saber, que la distancia topológica L típicamente encontrada entre nodos elegidos al azar es mucho menor al tamaño del sistema, más precisamente que $L \sim \ln n$, mientras que el coeficiente de clusterización C_c es mucho mayor que en una red aleatoria del mismo tamaño n y conectividad C (ver sección 2.4). En la figura (4.14) se compara la dependencia de L con el tamaño n de las redes generadas con el presente modelo, con el de redes aleatorias de misma conectividad C y tamaño, y distintos datos empíricos. Se confirma el comportamiento logarítmico para la dependencia de L con n , encontrándose que $L(n) \simeq A \ln(B + n) + C$. La distancia típica entre nodos es levemente mayor para las redes generadas por el mecanismo de presión selectiva que para redes aleatorias, diferencia que disminuye con el tamaño. Los datos empíricos de L son del orden de los predichos por ambos modelos (redes que crecen bajo presión selectiva y redes aleatorias), y (aunque dispersos) tienden a crecer con n .

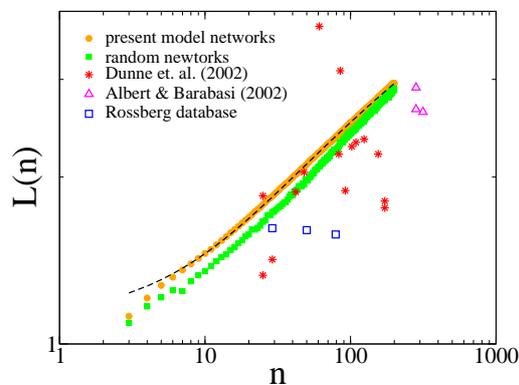


Figura 4.14: Comparación de L como función de n entre redes generadas por el presente modelo (círculos naranjas para $b = 2, q = 3$), redes aleatorias (cuadrados verdes) y datos empíricos obtenidos de redes que representan sistemas reales. Las redes aleatorias se construyen a partir de las conectividades $C(n)$ y tamaños n correspondientes a las redes generadas por el presente modelo. La línea discontinua corresponde al ajuste $L(n) \simeq 0.71 \ln(16.24 + n) - 0.86$ de la curva generada por el presente modelo. Los datos empíricos (aunque dispersos) son del orden y siguen la tendencia creciente. Triángulos magenta corresponden a redes tróficas, red de sustratos y metabólica de *E. Coli* y red neuronal de *C. Elegans* (Review de Albert–Barabási [10]). Asteriscos rojos (Dunne *et. al.* [65, 64]) y cuadrados azules (calculados a partir de las redes en la base de datos de A.G. Rossberg, Apéndice B) corresponden a redes tróficas.

En esta sección el coeficiente de clusterización $Cc(n)$ se obtiene promediando sobre el ensamble de matrices de tamaño n generadas por el presente modelo. Existen al menos dos definiciones distintas del coeficiente de clusterización (ver sección 2.3). Una es la definida por Watts–Strogatz [204] (ver ecuación (2.24)) aquí denotada por $Cc^{(WS)}$, la otra comúnmente denominada *transitividad* [153, 66] (ver ecuación (2.25)) aquí denotada por $Cc^{(Tr)}$. Las definiciones no son equivalentes. A lo largo del texto Cc hará referencia al coeficiente de clusterización en general, mientras que $Cc^{(WS)}$ y $Cc^{(Tr)}$ detallan específicamente que tipo de definición ha sido utilizado en un dado cálculo. En primer lugar se estudia como varía la clusterización Cc con el tamaño n de las redes generadas por el presente modelo. Mas específicamente se estudia la función $Cc(n)$ para diferentes elecciones de los parámetros b y q . En la figura (4.15) se puede observar el comportamiento de $Cc^{(WS)}(n)$, y como cambia el mismo al variar b . Se observa que $Cc^{(WS)}(n)$ en general decae. En particular para valores grandes de n decae mas lento que una ley de potencias. La curva decrece mas pronunciadamente cuando menor es b . De hecho las curvas se cruzan en $n \simeq 70$. Para valores menores de $n \simeq 70$ ocurre que $Cc^{(WS)}$ es mayor cuanto menor es b lo cuál es consistente con que cuanto menor es b mayor es la conectividad C (ver figura (4.1))*. Sin embargo para $n \gtrsim 70$ ocurre lo opuesto indicando que a pesar de disminuir la conectividad C se forma una estructura mas clusterizada dando lugar a un aumento de $Cc^{(WS)}$. En la figura (4.16) se puede observar el comportamiento de la otra definición de la clusterización con el tamaño $Cc^{(Tr)}(n)$ al variar b . El comportamiento cambia significativamente respecto de la definición anterior. Se encuentra que para $b < 2$ la curva $Cc^{(Tr)}(n)$ decae mas lento que una ley de potencias, mientras que si $b > 2$ la curva decae más rápido que una ley de potencias. El caso particular de $b = 2$ decae como una ley de potencias, mas precisamente el *scaling* es $Cc^{(Tr)}(n) \sim n^{-0.75}$ [165]†. Barabási y Albert reportaron dicho comportamiento para el modelo de conectado preferencial [10], aunque después se encontró que dicho resultado es incorrecto siendo consecuencia de una aproximación numérica. El verdadero *scaling* para el modelo de conectado preferencial es $Cc \sim (\ln^2 n)/n$ lo cuál para un rango pequeño de valores de n se puede confundir fácilmente por el comportamiento $\sim n^{-0.75}$. Dicha incertidumbre está presente en los resultados aquí presentados. En cuanto a la variación de $Cc(n)$ respecto de q , en la figura (4.17) se puede observar el caso particular para $Cc^{(WS)}(n)$. Para valores grandes de n ($n \gtrsim 15$) la dependencia con q es pequeña según se puede observar en el inset de la figura, encontrándose que la curva tiende a decaer mas lento cuanto mas grande es q . Para valores pequeños de

*Una mayor conectividad aumenta las chances de formar clusters y por ende de incrementar $Cc^{(WS)}$.

†En el paper [165] hay un error. En el texto del paper se define el coeficiente de clusterización a lo Watts–Strogatz, pero luego para los cálculos se utiliza la definición de la transitividad.

n ($n \lesssim 15$) también la variación es pequeña (o despreciable) salvo para $q = 4$ en donde se puede observar una abrupta disminución de $Cc^{(WS)}$ si se utiliza el *modelo de estabilidad local* (ver ítem *ii*) en la sección 4.0.2). Lo importante aquí es resaltar que no hay una dependencia significativa de $Cc^{(WS)}$ con q para valores grandes de n . La dependencia de $Cc^{(Tr)}(n)$ con q es presentada en la figura (4.18). No se aprecian cambios cualitativos en las curvas al variar q . Sólo se distingue que la curva presenta valores más pequeños cuando q aumenta, es decir, al revés de como lo hace la curva $Cc^{(WS)}(n)$ para valores grandes de n .

Es interesante comparar los valores de Cc obtenidos con el modelo, con el de redes aleatorias y redes de origen empírico. Normalmente los valores empíricos presentan una amplia dispersión, sin embargo se puede apreciar si el presente modelo predice mejor o no los valores de Cc que el modelo de redes aleatorias. En la figura (4.19) se compara el coeficiente de clusterización $Cc^{(WS)}$ entre redes generadas por el presente modelo con el de redes aleatorias y con datos empíricos. Los datos empíricos concuerdan mejor con el presente modelo que con las redes aleatorias. Si se utiliza la otra definición del coeficiente de clusterización $Cc^{(Tr)}$ los datos extraídos de Dunne *et. al.* también corresponden mejor con lo predicho por el presente modelo que por las redes aleatorias (ver figura (4.20)). Lamentablemente, los autores en Dunne *et. al.* no especifican con cuál de las 2 definiciones han calculado Cc en las redes que estudiaron. Sin embargo, para las redes tróficas *Ythan Estuary (1,2)* y *Little Rock Lake*, se puede comparar los valores de Cc calculados en Dunne *et. al.* con los calculados por Montoya y Solé [133] ya que ambos utilizan la misma fuente de datos * [119, 90, 86]. Al comparar se ve que los datos de Dunne *et. al.* son significativamente menores a los calculados por Montoya y Solé. Esta aparente inconsistencia se debe a que utilizan diferentes formas de calcular Cc . Montoya y Solé utilizan la definición dada por la ecuación (2.24), mientras que en Dunne *et. al.* se utiliza la definición dada por la ecuación (2.25) †. La comparación entre los datos de Montoya–Solé y Dunne *et. al.* es consistente con lo que aquí se encuentra (y con lo reportado en [146]), a saber, que la definición de la ecuación (2.25) arroja valores inferiores a la definición de la ecuación (2.24).

En la sección 2.3.2 se establece que $Cc^{(WS)}(n)/\langle k \rangle(n)$ es aproximadamente independiente del tamaño para un amplio conjunto de redes empíricas de gran tamaño. En este trabajo se compara el comportamiento de $Cc(n)/\langle k \rangle(n)$ como función de n para redes generadas por el presente modelo con el de redes aleatorias de misma conectividad y tamaño,

*Siendo más específicos, a pesar que parten de la misma fuente de datos en Dunne *et. al.* los datos son estandarizados de manera diferente a como lo hacen Montoya y Solé. En el apéndice B se describen las distintas estandarizaciones utilizadas. En Dunne *et. al.* se utiliza el nivel de estandarización L . Montoya y Solé utilizan el C .

†Según pude corroborar con mis cálculos utilizando las redes empíricas sacadas de la base de datos proveída por A.G.Rossberg (ver apéndice B para más detalles sobre la base de datos).

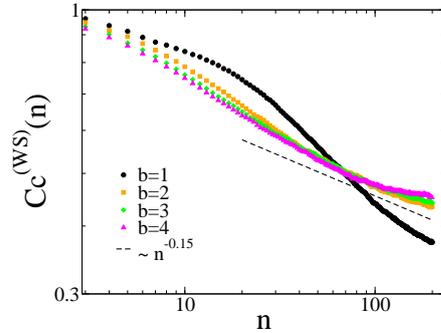


Figura 4.15: Dependencia con el parámetro b del coeficiente de clusterización $Cc^{(WS)}$ en función del tamaño n para $q = 3$. La línea entrecortada corresponde a una ley de potencias ($\sim n^{-0.15}$). Sirve de guía al ojo para ver que las curvas de $Cc(n)$ decaen mas lento que una ley de potencias. En particular, las curvas decaen mas lento cuanto mas grande es b .

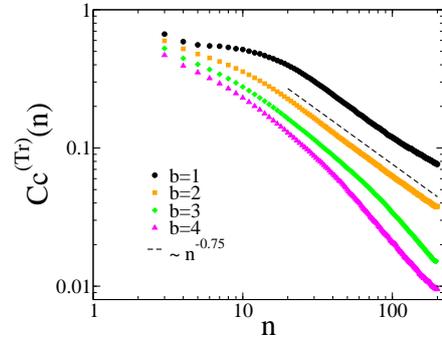


Figura 4.16: Dependencia con el parámetro b del coeficiente de clusterización $Cc^{(Tr)}$ en función del tamaño n para $q = 3$. $Cc^{(Tr)}(n)$ decae como una ley de potencias $\sim n^{-0.75}$ para el caso $b = 2$ cuando $n \gtrsim 20$. Para $b < 2$ la curva decae más lento que una ley de potencias, y lo opuesto ocurre para $b > 2$. En general $Cc^{(Tr)} \ll Cc^{(WS)}$ para n grande.

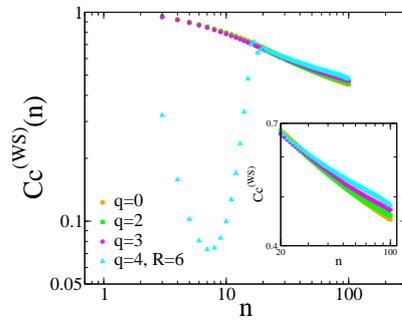


Figura 4.17: Dependencia con el parámetro q del coeficiente de clusterización $Cc^{(WS)}$ en función del tamaño n para $b = 2$. Cuando $n \gtrsim 15$ la curva posee una dependencia leve respecto de q (detalles en el inset). A tamaños menores para el caso $q = 4$ y $R = 6$ aparece una diferencia que no afecta los resultados a tamaños grandes.

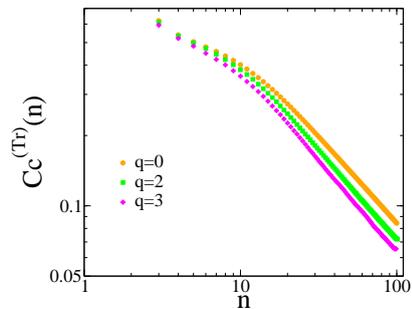


Figura 4.18: Se observa una leve dependencia del comportamiento de $Cc^{(Tr)}(n)$ respecto del parámetro q . Las curvas fueron calculadas para $b = 2$.

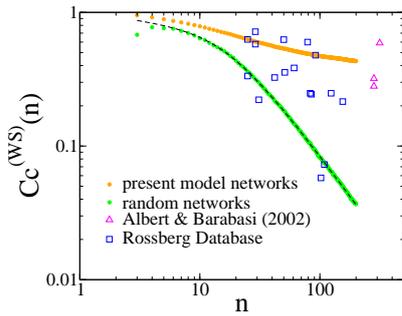


Figura 4.19: Comparación de $C_c^{(WS)}(n)$ en función de n entre redes del presente modelo (círculos naranjas $b = 2, q = 3$), redes aleatorias (círculos verdes) de misma conectividad C y tamaño n , y algunas de las redes empíricas consideradas en la figura (4.14). Los datos empíricos corresponden mejor con la alta clusterización exhibida en el presente modelo que con lo observado en el modelo de red aleatoria. La línea discontinua corresponde a la expresión analítica para redes aleatorias $C_c(n) = p = C(n)(2 - C(n))$ (ver sección 2.3.1 y ecuación (4.11)).

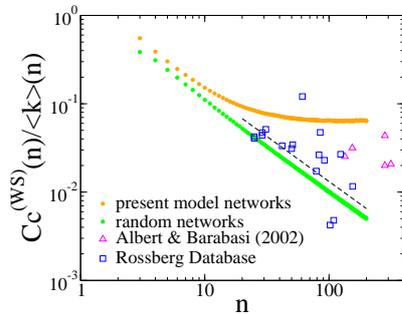


Figura 4.21: Comparación de $C_c^{(WS)}(n)/\langle k \rangle(n)$ en función de n entre redes del presente modelo (círculos naranjas), redes aleatorias (círculos verdes) de misma conectividad C y tamaño n , y algunas de las redes empíricas consideradas en la figura (4.14). La curva correspondiente al presente modelo tiende a una constante para tamaños grandes, resultado previamente reportado para redes reales (ver figura (2.2) y texto relacionado). Para las redes aleatorias $C_c/\langle k \rangle \sim 1/n$ tal como lo indica la línea discontinua. Los datos empíricos se dispersan en la región comprendida entre la curva correspondiente a las redes generadas con el presente modelo, y la curva correspondiente a las redes aleatorias.

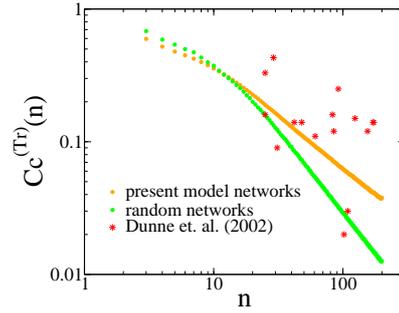


Figura 4.20: Comparación de $C_c^{(Tr)}(n)$ en función de n entre redes del presente modelo (círculos naranjas $b = 2, q = 3$), redes aleatorias de misma conectividad C y tamaño n (círculos verdes), y redes tróficas (asteriscos rojos, fuente Dunne *et. al.* [65, 64]). Los datos empíricos de $C_c^{(Tr)}$ tienden a ser sensiblemente mayores a los generados por el presente modelo aunque se corresponden mejor que con el modelo de red aleatoria que presenta baja clusterización.

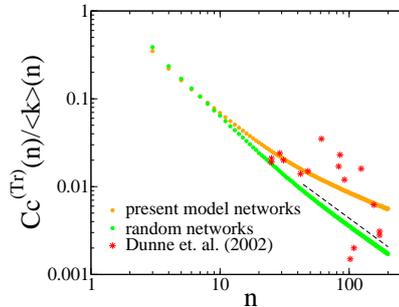


Figura 4.22: Comparación de $C_c^{(Tr)}(n)/\langle k \rangle(n)$ en función del tamaño entre redes del presente modelo (círculos naranjas), redes aleatorias (círculos verdes) de misma conectividad C y tamaño n , y las redes tróficas en Dunne *et. al.* [65, 64] (asteriscos rojos). El presente modelo predice mejor el comportamiento de las redes empíricas que el modelo de redes aleatorias. Para las redes aleatorias $C_c^{(Tr)}/\langle k \rangle \sim 1/n$ tal como lo indica la línea discontinua.

y también con los datos empíricos correspondientes a las mismas fuentes bibliográficas con las que se construyó la figura (4.14). La comparación se ha realizado utilizando ambas definiciones del coeficiente de clusterización Cc (ecuaciones (2.24) y (2.25)). En la figura (4.21) puede observarse dicha comparación para el caso en que se utilizó la definición de Watts–Strogatz (ecuación (2.24)). En este caso se encuentra que el presente modelo confirma la condición $Cc^{(WS)}(n)/\langle k \rangle(n) \simeq \text{constante}$ para tamaños suficientemente grandes. Los datos empíricos están dispersos en la región comprendida entre la curva correspondiente a las redes generadas por el presente modelo y la curva correspondiente a las redes aleatorias. En la figura (4.22) la comparación se realiza pero utilizando la otra definición de Cc correspondiente a la transitividad (ecuación (2.25)). En este caso se pierde la propiedad de que la curva $Cc^{(Tr)}(n)/\langle k \rangle(n)$ tienda a una constante para valores grandes de n . Los datos de Dunne *et. al.* permiten corroborar que con esta otra definición de Cc los datos empíricos se corresponden mejor con la relativamente alta clusterización que exhibe el presente modelo respecto de lo observado en las redes aleatorias. Estos resultados indican que algunas propiedades son independientes de la definición de Cc pero otras no. Al comparar con los datos empíricos la distinción entre las predicciones proveídas por el modelo aquí presentado frente al modelo de redes aleatorias, es más acentuada al analizar $Cc(n)$ que al analizar $Cc(n)/\langle k \rangle(n)$. Esto sugiere que el comportamiento de $\langle k \rangle(n)$ de las redes del presente modelo puede discrepar al compararlo con el de las redes empíricas. Esta discrepancia es confirmada en la figura (4.23) en donde se observa que en el presente modelo $\langle k \rangle(n)$ decrece como una ley de potencias, mientras que los datos empíricos extraídos de la bibliografía antes mencionada muestran que $\langle k \rangle(n)$ tiende a crecer con n (aunque la tendencia no es clara).

De acuerdo a las figuras (4.14) y (4.19), redes mas grandes se vuelven menos clusterizadas y aumentan la distancia típica entre nodos. Esto puede observarse en la figura (4.25)[165], en donde los datos fueron calculados cada 50 intentos de incorporar un nodo independientemente de si el intento resulta en la incorporación efectiva o no del nodo. En algún sentido, así es como una red natural luciría para un observador si este tomara fotos instantáneas a lo largo del tiempo. Claramente ambas cantidades (Cc y L) fluctúan inversamente en el tiempo, hecho que se corrobora en el inset de la figura (círculos negros), donde los datos son comparados con las mismas redes aleatorizadas (círculos rojos). El comportamiento del coeficiente de clustering Cc y la distancia topológica L está conectado a la dinámica de selección que determina que nodos son agregados, y que nodos son descartados. La presión selectiva favorece la incorporación de nodos con pocas conexiones, dado que modifican menos la matriz de interacciones $a_{i,j}$ que nodos con muchas conexiones (obviamente esta preferencia de incorporar nodos de pocas conexiones se ve reflejada en la distribución de grados $P(k)$). Entonces, en la mayoría de los casos la red crece a expensas de incorporar nodos con una o pocas conexiones,

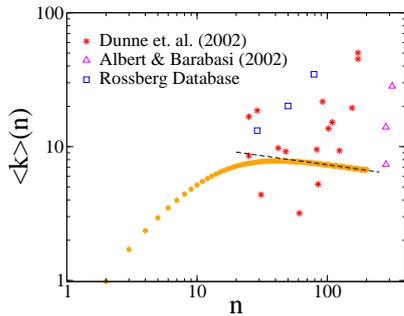


Figura 4.23: Comparación de $\langle k \rangle(n)$ en función de n calculado entre redes del presente modelo (círculos naranjas $b = 2, q = 3$) y las redes empíricas consideradas en la figura (4.14). La línea discontinua es un ajuste de la forma $\langle k \rangle(n) \sim n^{-1.13}$ indicando que $\langle k \rangle$ disminuye con n para tamaños grandes. Los datos empíricos de $\langle k \rangle$ tienden crecer con n aunque la tendencia no es clara.

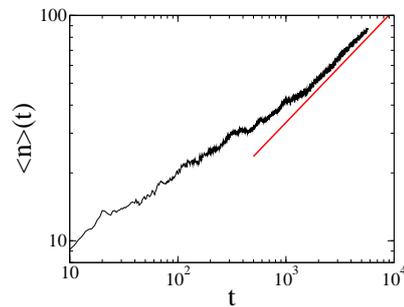


Figura 4.24: Promedio del tamaño de las redes en función del tiempo medido en número de intentos para $b = 2, q = 3$ [165, 164]. La línea continua roja corresponde a la ley de potencias $\sim t^{1/2}$.

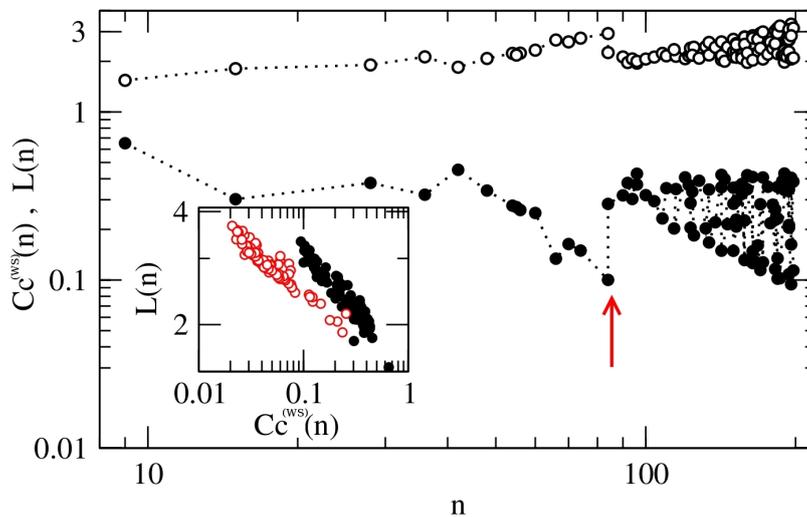


Figura 4.25: Gráficos de $C_c^{(ws)}(n)$ (círculos negros) y $L(n)$ (círculos vacíos) no promediados a lo largo de una corrida para el caso $b = 2, q = 3$ [165]. Se aprecia el momento en que se incorpora un hub a la red (flecha roja). El inset muestra los mismos datos graficados uno contra otro (círculos negros), también se muestran datos calculados para redes aleatorias del mismo tamaño n y conectividad C (círculos vacíos). Los datos son obtenidos cada 50 pasos de tiempo del mecanismo de crecimiento independientemente de la estabilidad resultante.

provocando un incremento de L y un decremento de Cc . La mayor parte del tiempo nodos con muchas conexiones desestabilizan la red y son rechazados, pero cuando uno es finalmente aceptado, un significativo decremento en L junto con un incremento en Cc es observado (flecha roja en el ejemplo de la figura (4.25)). Este repentino cambio es característico de la incorporación de un nuevo *hub* a la red. También se ha comprobado que dichas fluctuaciones llevan a un crecimiento lento del tamaño n , más precisamente un crecimiento tipo difusivo $\langle n \rangle \sim t^{1/2}$ donde el tiempo es medido en número de intentos (ver figura (4.24)) [165, 164].

4.1.4. Organización jerárquica

Otra cantidad de interés es el coeficiente de clusterización promedio $Cc(k)$ como función del grado de los nodos k . En esta sección el coeficiente de clusterización en función del grado de los nodos se calcula promediando la expresión dada por la ecuación (2.39) sobre el ensamble de matrices generadas por el modelo a tamaño n_{max} . Un ejemplo típico del comportamiento de $Cc^{(WS)}(k)$ en las redes crecidas a través del mecanismo de presión selectiva es el que se presenta en la figura (4.26). Se observa que $Cc^{(WS)}(k)$ decrece monótonamente con k y posee una cola tipo ley de potencias $Cc^{(WS)}(k) \sim k^{-\beta}$ con un exponente cercano a -1, $\beta \simeq 0.9$. Notar que el comportamiento tipo ley de potencias aparece aproximadamente para $k \gtrsim 25$, precisamente el mismo rango de valores para el cuál la distribución de grados $P(k)$ presenta la cola tipo ley de potencias. Este comportamiento resulta prácticamente independiente de b y q , el exponente β varía levemente al variar dichos parámetros (ver figuras (4.27) y (4.28)). Como se vió en la sección 2.5.1 este comportamiento en $Cc(k)$ sugiere de una estructura modular con organización jerárquica, lo cuál ha sido observada en redes metabólicas y de interacción de proteínas (ver figura (2.8) en la sección 2.5.1). El pequeño tamaño de las redes ecológicas resulta en general en un comportamiento irregular de la función $Cc(k)$, aunque en este trabajo se ha encontrado una excepción (ver sección A.1 del apéndice A). Es así que el modelo aquí analizado presenta otra característica más que también es observada en el contexto de las redes biológicas.

4.1.5. Correlación en el grado de los nodos vecinos

Una primera forma en que se ha analizado la correlación entre el grado de los nodos de las redes elegidas por el proceso de selección basado en la estabilidad, consistió en calcular el grado promedio $k_{nn}(k)$ de los nodos primeros vecinos de los nodos de grado k (ver sección 2.6). En la figura (4.29) se ve que $k_{nn}(k)$ presenta una cola que decae como una ley de potencias $k_{nn} \sim k^{-\delta}$ a partir de aproximadamente $k \simeq 20$, con un coeficiente del exponente cercano a $\delta \simeq 0.25$. Si se considera que la mayor porción de la curva de $k_{nn}(k)$ es representada por la cola (que a su vez

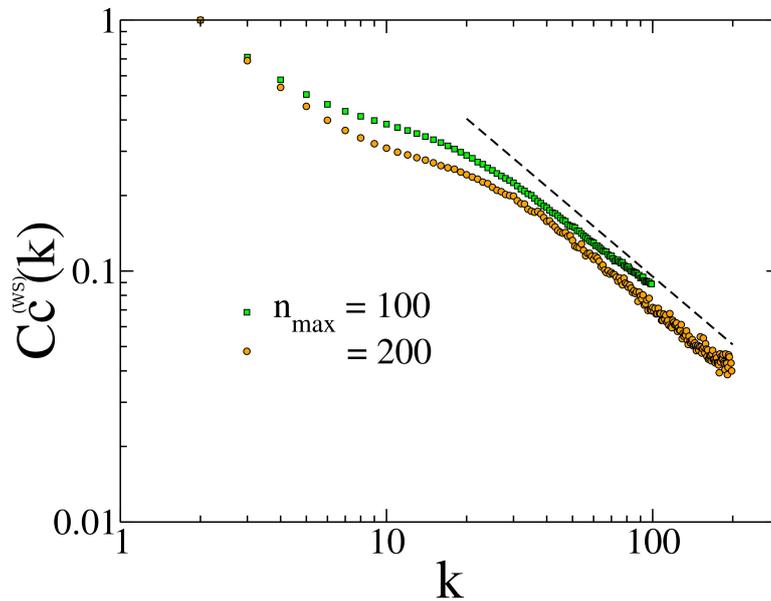


Figura 4.26: Coeficiente de clusterización promedio $C_c^{(WS)}$ en función del grado de los nodos k para $b = 2$, $q = 3$ y diferentes valores de n_{max} [164]. La línea discontinua corresponde a la ley de potencias $\sim k^{-0.9}$. Este tipo de comportamiento se ha interpretado en términos de que las redes presentan modularidad jerárquica en el sentido de Ravasz *et. al.* [175, 176].

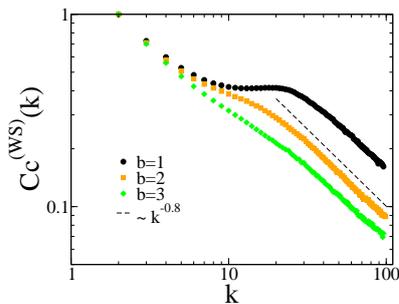


Figura 4.27: Coeficiente de clusterización promedio $C_c^{(WS)}$ en función del grado de los nodos k para $q = 3$, $n_{max} = 100$ y diferentes valores de b . En todos los casos se observa que $C_c^{(WS)}(k) \sim k^{-\beta}$ (con $\beta \simeq 1$) para valores grandes de k tal como lo indica la línea discontinua correspondiente a $\sim k^{-0.8}$. Un apartamiento de la ley de potencias aparece a valores chicos de k , el cuál se acentúa al disminuir b .

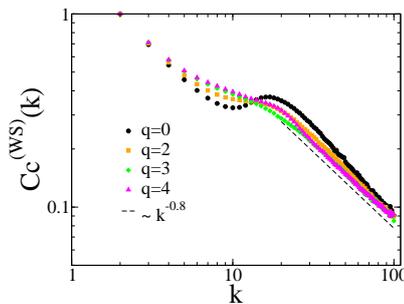


Figura 4.28: Coeficiente de clusterización promedio $C_c^{(WS)}$ en función del grado de los nodos k para $b = 2$, $n_{max} = 100$ y diferentes valores de q . En todos los casos se observa que $C_c^{(WS)}(k) \sim k^{-\beta}$ (con $\beta \simeq 1$) para valores grandes de k , tal como lo indica la línea discontinua correspondiente a $\sim k^{-0.8}$. Un apartamiento de la ley de potencias aparece a valores chicos de k , el cuál se acentúa al disminuir q .

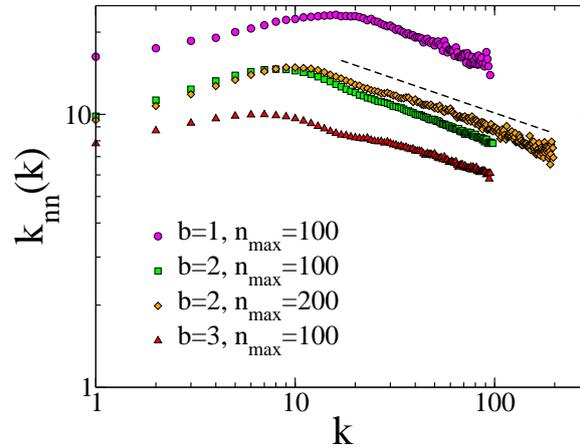


Figura 4.29: Promedio del grado de los primeros vecinos de los nodos de grado k para $q = 3$, y diferentes valores de b y n_{max} [164]. La línea discontinua corresponde a la ley de potencias $\sim k^{-0.25}$. Este resultado corresponde a que las redes generadas por el modelo son principalmente disasortativas. El comportamiento de $Cc(k)$ no se ve afectado cualitativamente ante variaciones del tamaño n o el parámetro b . En particular, se puede apreciar que el grado típico de los vecinos de los nodos disminuye al aumentar b lo cual es consistente con que la conectividad C disminuya bajo la misma condición (ver figura (4.1)).

corresponde a la parte libre de escala de la distribución de grados $P(k)$, luego, el comportamiento dominante (en la región de interés) es disasortativo [164]. Un comportamiento similar llevado al mismo exponente δ ha sido observado para redes de interacción proteína–proteína [53] (ver figura (2.9)). El comportamiento disasortativo es además consistente con otros resultados teóricos en donde se concluye que en redes de topología con tendencia asortativa la estabilidad se ve disminuída, es decir, la máxima parte real λ_{max} de entre todos los autovalores de matrices aleatorias como las aquí consideradas crece más rápidamente para redes asortativas que para redes disasortativas (ver sección 2.9.1). Para valores chicos de k , la curva crece indicando un comportamiento asortativo, por lo que $k_{nn}(k)$ presenta un comportamiento mixto al considerarse todo el rango de valores de k . El comportamiento cualitativo de $k_{nn}(k)$ se preserva al variar b, q y n_{max} (figuras (4.29) y (4.30)).

Es importante resaltar que la función $k_{nn}(k)$ aquí calculada es un promedio sobre el ensamble de matrices a tamaño n_{max} . Una sola de las matrices a tamaño n_{max} correspondería a una curva de $k_{nn}(k)$ en donde muchos de los grados no se verían representados por ningún valor de k_{nn} debido a que las redes poseen tamaños muy pequeños y por ende los nodos existentes no cubren todos los valores de k posibles. Ello puede entenderse mejor al notar que la distribución de grados en la región de la curva que corresponde al comportamiento libre de escala tiene valores en el rango $P(k) \in [10^{-6}, 10^{-3}]$ (ver figura (4.5)), mientras que las redes

poseen del orden de $n \simeq 10^2$ nodos lo cuál quiere decir que cada red presenta muy pocos nodos correspondientes a la región ley de potencias de la $P(k)$. Es por ésta razón que a menudo se utiliza el coeficiente de correlación de Pearson del grado de los nodos vecinos (ver sección 2.6) para caracterizar la asortatividad de una red. En la figura (4.31) se presenta una comparación del promedio $\langle r \rangle (n)$ (sobre el ensamble de matrices a tamaño n) del coeficiente de correlación de Pearson entre las redes generadas por el presente modelo y redes aleatorias de misma conectividad C y tamaño n . Para las redes aleatorias $\langle r \rangle$ es nulo, salvo para tamaños pequeños en donde pueden aparecer tendencias disasortativas (ver sección 2.6.2). Lo mismo ocurre para las redes del presente modelo, sin embargo se obtienen valores de $\langle r \rangle$ significativamente más negativos que los que se obtienen para las correspondientes redes aleatorias. Ésto indica una tendencia disasortativa en las redes del presente modelo. Junto con los valores de $\langle r \rangle (n)$ se han graficado barras que indican la desviación estandar de la dispersión de los valores de $r(n)$ *. Nótese que los valores de $r(n)$ para las redes aleatorias están mucho menos dispersos que para las redes del presente modelo. En otras palabras, $r(n)$ fluctúa relativamente mucho en el ensamble de redes del presente modelo. A modo de comparación se presentan también $r(n)$ para redes reales – redes tróficas para ser más precisos ya que los tamaños de las mismas permiten una comparación directa con los resultados obtenidos con el presente modelo (figura (4.31) y cuadro (4.1)). Para valores pequeños de n el coeficiente de correlación de Pearson de las redes reales se entremezcla entre la curva del correspondiente modelo y la de las redes aleatorias. Para valores mayores de n las redes reales presentan valores de $r(n)$ significativamente más negativos que los valores de $\langle r \rangle (n)$ de las redes generadas por el modelo en cuestión. A tamaños $n \gtrsim 175 \pm 5$ la situación parece revertirse para el presente modelo y los valores de $\langle r \rangle (n)$ se vuelven positivos aunque muy pequeños† (del orden de 10^{-2}). Este resultado parece entrar en contradicción con lo obtenido a priori a partir de la curva $k_{nn}(k)$. Sin embargo, recordar que en el presente modelo la curva de $k_{nn}(k)$ presenta un comportamiento mixto y que en dicha situación el signo de r no está unívocamente determinado. Ésta cuestión fue discutida en la sección (2.6.1), en donde en particular se hace referencia al caso de una red de interacción de proteínas que presenta *hubs* disasortativos y al mismo tiempo un valor de r levemente positivo tal como ocurre en el presente modelo.

Observación: El comportamiento mixto de $k_{nn}(k)$ (nodos de baja conectividad con tendencia asortativa, y nodos de alta conectividad con tendencia disasortativa) lleva a cierta indeterminación respecto de si la red es asortativa o disasortativa a nivel global (lo cuál se ve reflejado en el cambio de signo de $\langle r \rangle (n)$). Sin embargo, el comportamiento tipo

*Notar que he quitado el promedio de la expresión.

†Incluso los valores de $\langle r \rangle$ del presente modelo se vuelven más grandes que los de las redes aleatorias a partir de $n \gtrsim 150$.

ley de potencias de la distribución de grados $P(k)$ ocurre a partir de $k \gtrsim 25$, por lo que es el comportamiento disasortativo que corresponde a la cola de $k_{nn}(k)$ el que resulta relevante. En otras palabras, en las redes del presente modelo se observa un comportamiento disasortativo en la región de valores de k que resultan de interés.

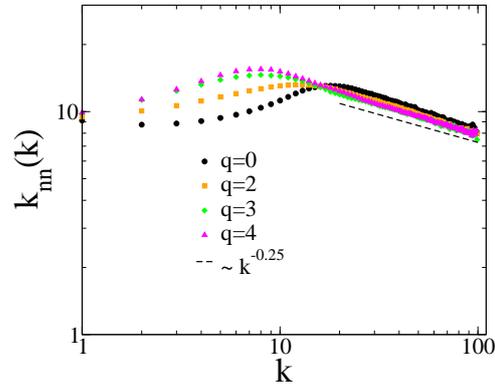


Figura 4.30: Promedio del grado de los primeros vecinos de los nodos de grado k para $b = 2$, $n_{max} = 100$, y diferentes valores de q . A valores grandes de k , el comportamiento de $k_{nn}(k)$ no se ve significativamente afectado por variaciones en q . El comportamiento asortativo a valores pequeños de k se acentúa al incrementarse q .

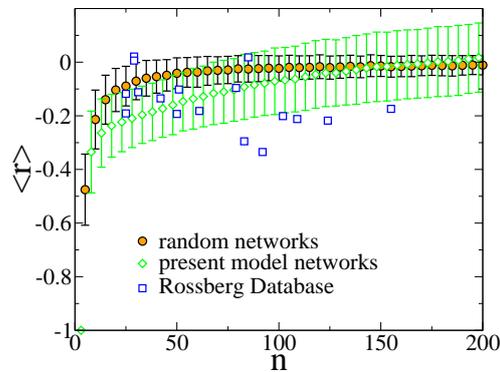


Figura 4.31: Promedio sobre todo el ensamble del coeficiente de Pearson de la correlación del grado remanente de los nodos vecinos $\langle r \rangle (n)$ en función del tamaño de las redes n . Se realiza una comparación entre redes que crecen bajo el efecto de la presión selectiva (círculos naranjas); redes aleatorias de igual conectividad $C(n)$ y tamaño n (rombos verdes); y redes tróficas (cuadrados azules) extraídas de la base de datos de A.G. Rossberg (ver Apéndice B). Se observa que por cuestiones de tamaño finito, incluso las redes aleatorias presentan correlación entre el grado de sus nodos vecinos. Las redes generadas con el presente modelo presentan sistemáticamente valores más negativos de $\langle r \rangle$ para valores de $n \lesssim 175 \pm 5$. Por encima de dichos valores de n los valores de $\langle r \rangle$ se vuelven positivos pero muy pequeños (del orden de 10^{-2}). Salvo para las redes tróficas, se grafican las desviaciones estandar. Se observa que para el presente modelo las fluctuaciones son mucho mayores que para las redes aleatorias en correspondencia con la existencia de *hubs*. Las curvas fueron obtenidas para $b = 2$ y $q = 3$.

Red	n	M	r
Benguela Current	29	121	0.021
Bridge Brook Lake	25	103	-0.191
British Grassland	61	97	-0.182
Canton Creek	102	696	-0.201
Caribbean Reef	50	503	-0.193
Chesapeake	31	67	-0.112
Coachella Valley	29	221	0.006
El Verde Rainforest	155	1438	-0.174
Little Rock Lake	92	960	-0.335
Northeast US Shelf	79	1371	-0.096
Scotch Broom	85	219	0.018
Skipwith Pond	25	185	-0.118
St. Marks Seagrass	51	264	-0.102
St. Martin Island	42	205	-0.135
Stony Stream	109	827	-0.212
Ythan Stuary (91)	83	390	-0.295
Ythan Stuary (96)	124	574	-0.218

Cuadro 4.1: Coeficiente de correlación de Pearson entre los grados remanentes de los nodos vecinos para las distintas redes tróficas extraídas de la base de datos de A.G. Rossberg (ver apéndice B). Las redes tróficas entran en la categoría de las redes biológicas y exhiben una tendencia disasortativa (valores negativos de r).

4.1.6. Tolerancia a las fallas topológicas

En la sección 2.7 se introdujeron los conceptos de resistencia ante *ataques dirigidos* o *fallas aleatorias*. En otras palabras como fragmenta o colapsa una red ante diferentes estrategias de aislación de nodos. Se calculan $S(f)$ y $\langle s \rangle(f)$ para ataques dirigidos y fallas aleatorias a una red típica del presente modelo a tamaño $n = 100$ (ver figura (4.32)), y se compara los resultados con lo que ocurre en una correspondiente red aleatoria. La forma particular de $S(f)$ y $\langle s \rangle(f)$ varía con cada red, pero las conclusiones cualitativas permanecen. Ante ataques dirigidos la red del presente modelo colapsa en $f_c \simeq 0.3$, mientras que la red aleatoria lo hace en $f_c \simeq 0.55$. Es importante tener en cuenta que dichos valores sólo son sugestivos, ya que no corresponden a promedios. Por otro lado, ante fallas aleatorias ambas redes responden de manera similar colapsando a valores muy grandes de f . Graficar $S(f)$ y $\langle s \rangle(f)$ para una sola red, tiene la ventaja de poder comparar los resultados obtenidos en el presente modelo con los de redes reales.

En la sección y 2.8.4 se discutió brevemente un trabajo de Solé y Montoya[189] donde se calculan $S(f)$ y $\langle s \rangle(f)$ ante ataques dirigidos y fallas aleatorias a un conjunto de redes ecológicas. Sin embargo (tal como es mencionado en la sección 2.8.4) ellos no comparan los resultados que obtienen con lo que obtendrían para versiones aleatorizadas de las redes estudiadas, lo cuál permitiría eventualmente distinguir alguna particularidad que las redes ecológicas presentan y las redes aleatorias no. Ello motivó realizar tal comparación utilizando las redes ecológicas ob-

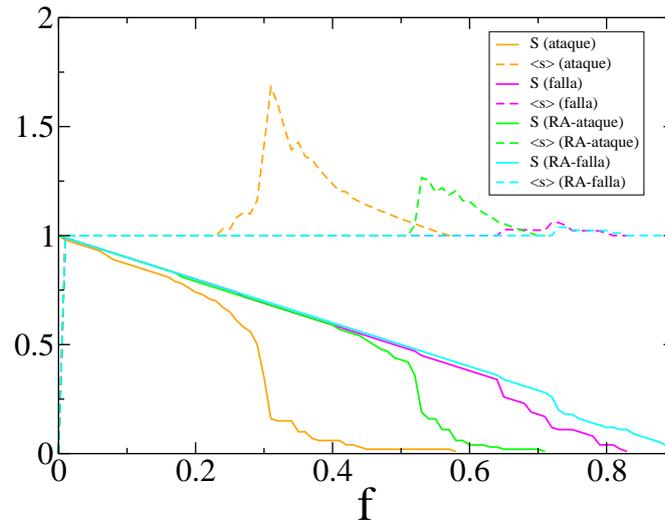


Figura 4.32: Fragmentación de las redes ante una secuencia de fallas aleatorias (colores magenta y cian) o una secuencia de ataques dirigidos (colores naranja y verde). Se grafica el tamaño relativo del cluster más grande S (líneas continuas) y el tamaño promedio de los clusters aislados $\langle s \rangle$ (líneas discontinuas) como función de la fracción de nodos f a los que se les han removido sus links (ecuación (2.54)). Los colores naranja y magenta corresponden a una red típica generada por el presente modelo para $b = 2, q = 3$ a $n = 100$. Los colores verde y cian corresponden a una red aleatoria generada a partir del mismo tamaño y número de conexiones. Las etiquetas en la leyenda ayudan a identificar las curvas: **ataque** y **RA-ataque** indican ataque dirigido; **falla** y **RA-falla** indican fallas aleatorias; **ataque** y **falla** corresponden a la red del presente modelo; **RA-ataque** y **RA-falla** corresponden a la red aleatoria (RA son las siglas de Red Aleatoria). Ante ataques dirigidos la red generada por el presente modelo colapsa ($S(f)$ se va a cero y $\langle s \rangle(f)$ tiene un pico) a un valor de f significativamente menor al de la red aleatoria. Frente a fallas aleatorias las redes responden de manera similar colapsando a valores muy grandes de f .

tenidas de la base de datos de A.G. Rossberg*. Los resultados de dicha comparación son presentados en el apéndice A, y además de confirmar lo apreciado por Solé y Montoya permiten concluir que la respuesta de $S(f)$ y $\langle s \rangle(f)$ de las redes ecológicas y las redes del presente modelo contrastan de manera similar con la respuesta de las redes aleatorias. Más precisamente, las redes ecológicas y las redes del presente modelo son menos resistentes a ataques dirigidos que las correspondientes redes aleatorias, mientras que ante fallas aleatorias las redes ecológicas, las redes del presente modelo y las redes aleatorias responden de manera similar.

4.1.7. Límites al tamaño máximo

En la figura (4.4) se puede ver que en las redes ecológicas el tamaño n cubre (aproximadamente) el rango entre 10 y 10^2 , en las metabólicas entre

*Ver apéndice B para más información sobre la base de datos.

10^2 y 10^3 , y en las de proteínas entre 10^3 y 10^4 . En otras palabras, las redes biológicas presentan una cota (aproximada) N al tamaño máximo que pueden alcanzar. N es del orden de 10^4 . Otro tipo de redes como las sociales o tecnológicas pueden llegar a presentar tamaños muy superiores. Por ejemplo, el sitio [1] estima que en la actualidad la World-Wide-Web posee del orden de 10^{10} nodos. En [146] se reportan redes sociales con tamaños de hasta del orden de 10^7 nodos*. Es natural preguntarse si hay alguna razón por la cuál las redes biológicas no crecen a tamaños tan grandes como (por ejemplo) las redes sociales o tecnológicas.

Como consecuencia de que en el presente modelo se cumple la relación:

$$C(n) \simeq Bn^{-(1+\epsilon)} \quad , \text{con } 1 > \epsilon > 0 \text{ cuando } n \gg 1 \quad (4.3)$$

(ver sección 4.1.1), puede concluirse la existencia de un tamaño máximo N que las redes generadas por el mismo no pueden superar. La existencia de dicho tamaño (y eventualmente los valores del mismo) constituye una explicación posible al tamaño relativamente pequeño que las redes biológicas presentan. El tamaño máximo N se obtiene para el caso en que no hay extinciones ($q = 0$) partiendo de la siguiente expresión:

$$\begin{aligned} C(n+1) &= \frac{1}{2n(n+1)} \sum_{i,j=1}^{n+1} c_{i,j} + c_{j,i} \\ &= \frac{1}{2n(n+1)} \left\{ \left[\sum_{i,j=1}^n c_{i,j} + c_{j,i} \right] + \left[\sum_{i=1}^n c_{i,n+1} + c_{n+1,i} \right] + \left[\sum_{j=1}^n c_{n+1,j} + c_{j,n+1} \right] \right\} \\ &= \frac{1}{2n(n+1)} 2n(n-1)C(n) + \frac{1}{2n(n+1)} 2k_{n+1}^{(t)} \\ &= \frac{n-1}{n+1} C(n) + \frac{1}{n(n+1)} k_{n+1}^{(t)} \end{aligned} \quad (4.4)$$

donde $k_i^{(t)} = k_i^{(in)} + k_i^{(out)}$ es la suma de la cantidad de conexiones entrantes y salientes al nodo i (ver sección 4.1.1). La hipótesis de que no hay extinciones se utilizó para deducir la tercera línea en (4.4). Utilizando las ecuaciones (4.4) y (4.3) se tiene:

$$\begin{aligned} k_{n+1}^{(t)} &\simeq Bn(n+1) \left[(n+1)^{-(1+\epsilon)} - \frac{n-1}{n+1} n^{-(1+\epsilon)} \right] \\ &\simeq Bn \left[(n+1)^{-\epsilon} - n^{-\epsilon} + n^{-(1+\epsilon)} \right] \\ &\simeq Bn^{1-\epsilon} \left[(1+1/n)^{-\epsilon} - 1 + 1/n \right] \\ &\simeq Bn^{1-\epsilon} \left[1 - \epsilon/n + \dots - 1 + 1/n \right] \\ &\simeq B(1-\epsilon)n^{-\epsilon} \\ k_n^{(t)} &\simeq B(1-\epsilon)n^{-\epsilon} \quad , \text{ya que } n+1 \simeq n \end{aligned} \quad (4.5)$$

*La red de llamadas telefónicas.

Es decir, la cantidad de links (dirigidos) $k_n^{(t)}$ que posee el último nodo ingresado a la red decrece con el tamaño de la misma. Dado que en el proceso de crecimiento del modelo no se permiten ingresar nodos aislados, un tamaño máximo N surge de la condición $k_N^{(t)} \geq 1$, o lo que es lo mismo:

$$\begin{aligned} 1 &= k_N^{(t)} \simeq B(1 - \epsilon)N^{-\epsilon} \\ [B(1 - \epsilon)]^{1/\epsilon} &\simeq N \end{aligned} \quad (4.6)$$

Recordando que $C(n)$ no depende de q (figura (4.2) y texto relacionado en la sección 4.1.1), pueden obtenerse valores de N en función de b utilizando los valores de B y ϵ resultantes de los ajustes de la figura (4.1). Dichos valores se resumen en la tabla (4.2) encontrándose que en la mayoría de los casos N es del orden de 10^4 en buena concordancia con lo observado para las redes biológicas. A partir del ajuste de la figura (4.4) se puede realizar una estimación de N según los datos empíricos y de acuerdo a la ecuación (4.6) encontrándose que $N \simeq 10^6$. Aunque es una estimación un tanto mayor que las proveídas por el presente modelo, es apreciablemente menor a la encontrada en algunas redes tecnológicas y sociales.

En el caso en que se permiten extinciones en el algoritmo de crecimiento del modelo ($q > 0$), es posible determinar una estimación del tamaño máximo directamente a partir del cálculo numérico de $k_n^{(t)}$ en función de n . En la figura (4.33) se presenta dicho cálculo para el caso $q = 3$ y diferentes valores de b . Se observa que $k_n^{(t)}$ crece para valores chicos de n , pero a partir de cierto tamaño comienza a decrecer siguiendo una ley de potencias. Más precisamente:

$$k_n^{(t)} \simeq k_0 n^{-\xi}, \quad n \gg 1 \quad (4.7)$$

donde la constante k_0 y el exponente ξ se obtienen a partir de ajustes y dependen del parámetro b . El comportamiento indicado por la ecuación (4.7) es cualitativamente consistente con el de la ecuación (4.5), sin embargo los valores de ξ resultan sensiblemente mayores a los de ϵ (ver tabla 4.2), por lo que el proceso de extinciones afecta el tamaño máximo que el sistema puede alcanzar. A partir de la ecuación (4.7) se puede obtener una condición análoga a la de la ecuación (4.6) resultando en otra estimación N_1 al tamaño máximo:

$$\begin{aligned} 1 &= k_{N_1}^{(t)} \\ &\simeq k_0 N_1^{-\xi} \\ N_1 &\simeq k_0^{1/\xi} \end{aligned} \quad (4.8)$$

En la tabla 4.2 se presentan los valores de N_1 que se obtienen a partir de la ecuación (4.8) y los valores ajustados para k_0 y ξ . Los valores de N_1 resultantes son del orden de 10^3 , es decir, un orden de magnitud inferior a los valores de N .

b	B	ϵ	N	k_0	ξ	N_1
1	31.9	0.28	7.3×10^4	333	0.77	1.9×10^3
2	11.4	0.21	3.5×10^4	47.2	0.47	3.6×10^3
3	8.37	0.23	3.3×10^3	21.6	0.34	8.4×10^3
4	4.23	0.12	5.7×10^4	–	–	–
★	21	0.2	1.3×10^6	–	–	–

Cuadro 4.2: En este cuadro se presentan para diferentes valores de b los valores de los parámetros B y ϵ que resultan de los ajustes en la figura (4.1). A partir de dichos valores de B y ϵ se calculan y presentan en la tabla las estimaciones N al tamaño máximo que las redes pueden alcanzar según la ecuación (4.6). También son presentados los valores de k_0 y ξ que resultan de los ajustes en la figura (4.33) a partir de los cuales se calculan y presentan en la tabla las estimaciones N_1 al tamaño máximo según la ecuación (4.8). En la última fila de la tabla (indicada por $b = \star$) se calcula N según la ecuación (4.6) para el valor de B y ϵ que resulta del ajuste a los datos empíricos en la figura (4.4).

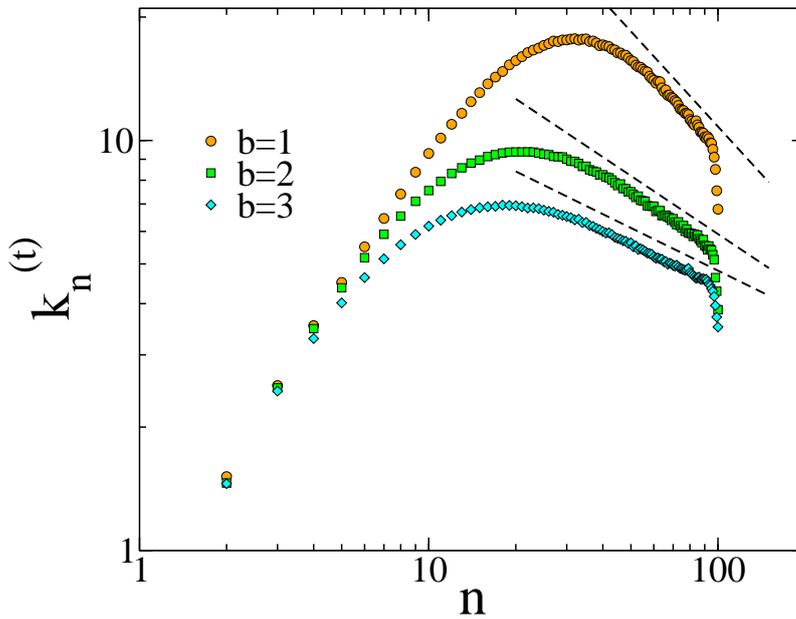


Figura 4.33: Grado total $k_n^{(t)}$ de los nodos ingresantes a sistemas de tamaño $n-1$ como función de n para distintos valores de b . Se observa que $k_n^{(t)}$ crece para valores chicos de n . Eventualmente a partir de cierto valor de n (el cuál depende de b) comienza a disminuir de forma tal que la cola presenta un comportamiento tipo ley de potencias. En estas simulaciones se utilizaron los valores de los parámetros $n_{max} = 100$, $q = 3$.

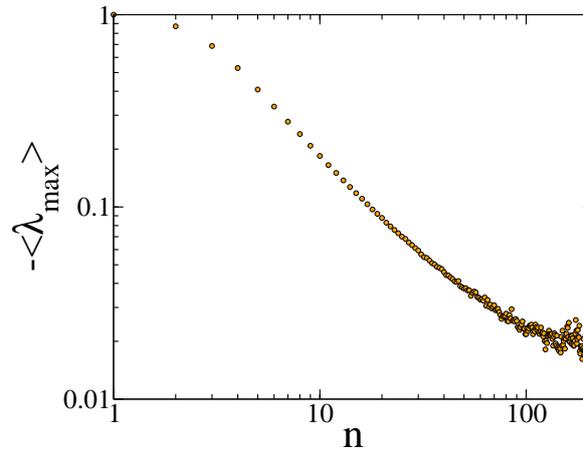


Figura 4.34: Gráfico log-log de $-\langle\lambda_{max}\rangle(n)$ donde λ_{max} es la máxima parte real de entre todos los autovalores de una matriz (el promedio es sobre un ensamble de matrices generadas para $n_{max} = 200$, $b = 2$ y $q = 3$). $-\langle\lambda_{max}\rangle(n)$ cae a cero más lento que una ley de potencias indicando que a medida que el sistema va creciendo le cuesta más estabilizar.

4.2. Propiedades dinámicas

En las secciones previas se han analizado diferentes propiedades topológicas que son seleccionadas por la condición de estabilidad, es decir, propiedades asociadas a la matriz de adyacencia $c_{i,j}$ definida por:

$$c_{i,j} = \begin{cases} 1 & a_{i,j} \neq 0 \\ 0 & \text{c.c.} \end{cases} \quad (4.9)$$

la cuál sólo posee información que es independiente de la intensidad de las interacciones $a_{i,j}$. En lo que sigue, se analiza las características de la dinámica asociada a las redes que emergen del mencionado criterio de selección. En otras palabras, se investiga las propiedades estadísticas de los elementos no nulos $a_{i,j} \neq 0$ de las matrices generadas por el mecanismo de presión selectiva, y las propiedades estadísticas de los autovalores asociados a dichas matrices.

En primer lugar se estudia la dependencia del promedio de la máxima parte real de entre todos los autovalores λ_{max} con el tamaño de las redes n (ver sección 2.9.2). En las redes construidas sin presión selectiva basada en la estabilidad, eventualmente ocurre que $\lambda_{max} > 0$ (por ejemplo aumentando n o α en el modelo de May). Las matrices de nuestro modelo satisfacen $\lambda_{max} < 0$ para todo valor de n por construcción*. En la figura (4.34) se grafica $-\langle\lambda_{max}\rangle(n)$ donde el promedio es sobre los valores de λ_{max} correspondientes a un ensamble de matrices a tamaño n

*Salvo en el caso en que se exija la versión relajada de la condición de estabilidad $\lambda_{max} < \Delta$ en el caso $\Delta \gtrsim 0$.

generado por el presente modelo. La función $-\langle \lambda_{max} \rangle (n)$ decae a cero (desde valores positivos) más lento que una ley de potencias, por lo que extrapolando el comportamiento de dicha función se concluye que λ_{max} nunca alcanzará el cero. El hecho de que $-\langle \lambda_{max} \rangle (n)$ sea una función decreciente que tiende a cero, puede interpretarse como que a medida que el sistema crece le cuesta más lograr estabilizarse, y cuando lo logra, dicha estabilidad es más endeble.

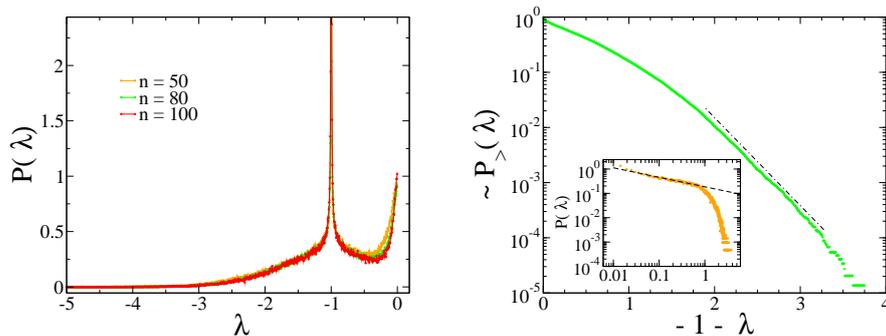


Figura 4.35: **Izquierda:** Densidad de probabilidad $P(\lambda)$ de las partes reales de los autovalores λ correspondientes al ensamble de matrices estables generadas por el presente modelo para $b = 2, q = 3$ y a diferentes tamaños n . La distribución varía poco con el tamaño, encontrándose la mayor discrepancia en la región de valores de λ cercanos a cero. Todas las distribuciones presentan un pico en $\lambda = -1$. **Derecha:** En círculos verdes se grafica la distribución acumulativa complementaria $P_{>}(\lambda)$ en función de $-1 - \lambda$ de la distribución $P(\lambda)$ para amortiguar el ruido de ésta última (ver ecuación (A.1) y texto relacionado). La cola de la curva decae exponencialmente tal como es indicado por la línea entrecortada de puntos y rayas. *Inset:* En círculos naranjas se grafica $P(\lambda)$ en función de $-1 - \lambda$. Se observa que $P(\lambda)$ decae como una ley de potencias $(-1 - \lambda)^{-0.4}$ para valores intermedios de $-1 - \lambda$ tal como es indicado por la línea entrecortada. El comportamiento es similar al observado en otras redes tipo May libres de escala (figura 2.13 y texto relacionado).

La densidad de probabilidades $P(\lambda)$ de las partes reales λ de los autovalores de las redes a distintos tamaños n generadas por el presente modelo, pueden apreciarse en la figura (4.35). Lo primero que se observa es que $P(\lambda)$ no cambia significativamente con el tamaño n . Además $P(\lambda)$ es no nula sólo para valores negativos de λ en concordancia con que $\lambda_{max} < 0$. Más precisamente, el comportamiento de $P(\lambda)$ en el presente modelo se parece mucho al obtenido en matrices de May sobre topologías libre de escala correspondientes a redes de Barabási–Albert[23] (ver figura 2.13 y texto relacionado). Sin embargo hay importantes diferencias. En la figura (2.13) las densidades son simétricas y centradas en $\lambda = 0$ en donde presentan un pico. En el presente modelo el correspondiente pico se ha corrido hacia $\lambda = -1$. Ello tiene que ver con que en las matrices de la figura (2.13) los elementos diagonales son nulos*, mientras que en las matrices del presente modelo los elementos diagonales valen -1 (ver ecuación (2.70) y texto relacionado). Otra importante diferencia es que

*O tienen media nula sobre el ensamble de matrices.

en el presente modelo $P(\lambda)$ no es simétrica en torno al pico en $\lambda = -1$. Hacia la derecha del pico (hacia valores crecientes de λ) $P(\lambda)$ decae para luego volver a subir conduciendo a un segundo pico en $\lambda = 0$. Este otro pico corresponde a una acumulación de autovalores en el entorno negativo cercano a $\lambda = 0$ provocando que la probabilidad de que el sistema crezca disminuya con el tamaño. Más precisamente, dado que $P(\lambda)$ no cambia significativamente con n y el número de autovalores crece con n , luego la probabilidad de encontrar autovalores acumulados en el entorno cercano a $\lambda = 0$ aumenta con n . Este incremento en la dificultad de que el sistema crezca es consistente con algo que se ha observado en las simulaciones, a saber, el número de intentos típicamente necesarios para que el sistema incorpore un nodo crece con el tamaño.

4.2.1. Interacciones anticorrelacionadas

Lo primero que se calcula sobre la estadística de los elementos no nulos $a_{i,j} \neq 0$ de las matrices generadas por el presente modelo, es la distribución $P(a_{i,j})$ de valores que cualquiera de dichos elementos puede adoptar para matrices a tamaño n_{max} . El comportamiento típico es el observado en la figura (4.36). La distribución es una función par, y prácticamente uniforme en el intervalo $[-b, b]$ mostrando una pequeña elevación para valores chicos de $a_{i,j}$. La paridad de $P(a_{i,j})$ implica que la estabilidad de las redes no es acentuada por un particular signo de los coeficientes de interacción. En relación a trabajos previos, el comportamiento prácticamente uniforme de $P(a_{i,j})$ sugiere que la presencia significativa de interacciones débiles no son el principal factor estabilizante de los sistemas (ver sección 2.9.1 y referencias [136],[126]).

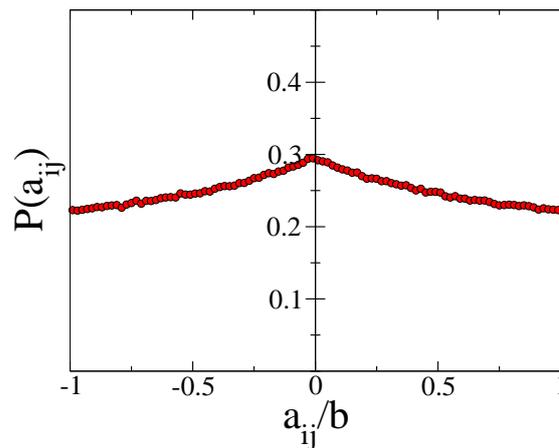


Figura 4.36: Densidad de probabilidades de los elementos de matriz $a_{i,j}$ no nulos, para $b = 2$, $q = 3$ y $n_{max} = 100$ [164]. La función $P(a_{i,j})$ resulta simétrica y prácticamente uniforme.

El siguiente paso en el estudio de las propiedades estadísticas de los elementos $a_{i,j} \neq 0$ de las matrices generadas por el presente modelo es el estudio de las correlaciones de a pares, la cuál viene definida por la expresión:

$$\langle a_{i,j}a_{j,i} \rangle (n) = \left\langle \frac{1}{n_{cd}} \sum_{\langle i,j \rangle} a_{i,j}a_{j,i} \right\rangle \quad (4.10)$$

donde n_{cd} es el número *conexiones dobles* (pares i, j tales que $a_{i,j} \neq 0$ y $a_{j,i} \neq 0$), la sumatoria es sobre el conjunto de todos los pares no repetidos $\langle i, j \rangle^*$, y el promedio $\langle \cdot \rangle$ en la expresión de la derecha es sobre todas las matrices de tamaño n en el ensamble de matrices generadas por el presente modelo. Como se vió en la sección 2.9.1[15], la presencia de interacciones anticorrelacionadas entre pares de nodos ($\langle a_{i,j}a_{j,i} \rangle < 0$) probó ser un factor estabilizante de las matrices en cuestión. Más precisamente, interacciones anticorrelacionadas permiten obtener matrices estables (matrices en donde $\lambda_{max} < 0$) para valores mayores de $C\alpha n$ que para el caso $C\alpha n = 1$ (ecuación (2.67)) correspondiente al modelo estándar de May en donde $\langle a_{i,j}a_{j,i} \rangle = 0$. En la figura (4.37) se puede observar el comportamiento de $\langle a_{i,j}a_{j,i} \rangle$ como función del tamaño n de las redes generadas por el presente modelo (rombos magenta). La correlación resultante es negativa para todo valor de n y satura en $\langle a_{i,j}a_{j,i} \rangle \simeq -0.65$ para valores grandes de n . Éste resultado indica que la presencia de interacciones anticorrelacionadas es una propiedad emergente del mecanismo de selección al cuál están sujetas las matrices en el presente modelo.

Dado que $\langle a_{i,j}a_{j,i} \rangle$ se calcula sobre todos los pares de nodos unidos por conexiones dobles, es útil estudiar la proporción de conexiones dobles en la red. En el inset de la figura (4.37) se compara la fracción de conexiones dobles por nodo $\langle \eta \rangle$ (ver sección 2.8 y ecuación (2.62)) de las redes generadas por el presente modelo con la de redes aleatorias construidas a partir de las mismas conectividades $C(n)$ y tamaños n . Para las redes aleatorias una expresión analítica de $\langle \eta \rangle (n)$ puede ser calculada. La probabilidad de poseer una conexión entre un par arbitrario (i, j) de nodos es:

$$\begin{aligned} p &= P(c_{i,j} = 1, c_{j,i} = 1) + P(c_{i,j} = 1, c_{j,i} = 0) + P(c_{i,j} = 0, c_{j,i} = 1) \\ &= P(c_{i,j} = 1)P(c_{j,i} = 1) + P(c_{i,j} = 1)P(c_{j,i} = 0) + P(c_{i,j} = 0)P(c_{j,i} = 1) \\ &= C^2(n) + 2C(n)(1 - C(n)) \\ &= C(n)(2 - C(n)) \end{aligned} \quad (4.11)$$

lo cuál está relacionado al grado promedio por nodo, $\langle k \rangle = p(n-1) \simeq pn$. Para redes aleatorias tenemos que:

*La notación $\langle i, j \rangle$ corresponde al conjunto de todos los pares (i, j) no ordenados, y por ende no hay repetición sumando sobre (i, j) y luego sobre (j, i) .

$$\langle \eta \rangle_{rand} = \frac{\zeta}{n} = \frac{\langle k^{(t)} - k \rangle}{\langle k \rangle} \quad (4.12)$$

donde:

$$\begin{aligned} k_i^{(t)} - k_i &= \sum_j c_{i,j} + c_{j,i} - [1 - (1 - c_{i,j})(1 - c_{j,i})] \\ &= \sum_j c_{i,j}c_{j,i} \end{aligned} \quad (4.13)$$

por lo que (usando las mismas probabilidades que en (4.11)):

$$\begin{aligned} \langle k^{(t)} - k \rangle &= \left\langle \sum_j c_{i,j}c_{j,i} \right\rangle \\ &= C^2(n)n \end{aligned} \quad (4.14)$$

Así, remplazando en la ecuación (4.12), se obtiene:

$$\langle \eta \rangle_{rand} = \frac{C^2(n)n}{\langle k \rangle} = \frac{C(n)}{2 - C(n)} \quad (4.15)$$

por lo que para $n \gg 1$ se tiene que $\langle \eta \rangle_{rand} \sim n^{-(1+\epsilon)}$ lo cuál coincide perfectamente con los cálculos numéricos (línea continua y triángulos naranjas del inset). Por otro lado para las redes de nuestro modelo se observa que $\langle \eta \rangle \sim n^{-0.68}$ cuando $n \gg 1$ (círculos verdes del inset). La comparación de $\langle \eta \rangle$ y $\langle \eta \rangle_{rand}$ entre las redes del presente modelo y correspondientes redes aleatorias presentados en la figura (4.37) indica que las redes generadas con el presente modelo exhiben una relativamente alta proporción de conexiones dobles. Aún así, $\langle \eta \rangle(n)$ es una función decreciente y presenta valores pequeños para valores grandes de n lo cuál indica que la proporción de conexiones dobles es significativamente más pequeña que la proporción de conexiones simples. Ésto se puede observar cualitativamente al comparar el gráfico de la izquierda y el de la derecha en la figura (4.39).

Para tener una visión más completa de que es lo que ocurre, se calculó la correlación $\langle a_{i,j}a_{j,i} \rangle / \langle |a_{i,j}| \rangle^2$ entre los elementos de matriz que conectan los nodos i y j como función del grado k_i del nodo i . El promedio se toma sólo sobre conexiones dobles. En la figura (4.38) se observa que el valor absoluto de la correlación presenta un mínimo alrededor de $k_i \simeq 25$ y luego tiende a cero a medida que el grado k_i crece. El inset de dicha figura muestra que la fracción de conexiones dobles que están anticorrelacionadas $\langle \kappa \rangle$ tiende a $1/2$ para k_i grandes. Se puede concluir a partir de estos resultados que las interacciones entre los *hubs* y sus vecinos están esencialmente descorrelacionados. Esto sugiere que la influencia de

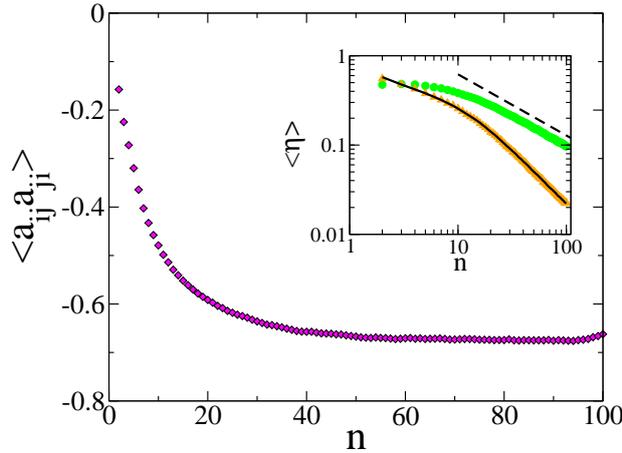


Figura 4.37: Función de correlación $\langle a_{i,j} a_{j,i} \rangle$ como función del tamaño de la red n , para $b = 2$, $q = 3$ [164]. El promedio fué calculado sobre todos los pares de nodos unidos por una conexión doble para las redes de un mismo tamaño. El inset muestra una comparación entre la fracción de conexiones dobles $\langle \eta \rangle$ en las presentes redes (círculos verdes) y redes aleatorias de misma conectividad C y tamaño n : $\langle \eta \rangle_{rand} = C/(2 - C)$ (línea continua correspondiente a ecuación (4.15), y triángulos anaranjados correspondientes a cálculos numéricos). La línea discontinua corresponde a un ajuste tipo ley de potencias $\sim n^{-0.68}$ del comportamiento de $\langle \eta \rangle (n)$ en el presente modelo.

los *hubs* en la estabilización de la dinámica está principalmente asociada a su rol topológico, como por ejemplo el de mantener conectada la red.

Por otro lado, al considerar las sub-redes constituidas únicamente por conexiones dobles se observa una estructura modular, en donde cada módulo tiene una forma de estrella tal como se puede observar en el gráfico de la derecha en la figura (4.39). En un trabajo previo se observó la aparición de éste tipo de estructuras en un modelo en donde se optimiza la estabilidad en redes de May sujetas a restricciones topológicas. Más precisamente, se aplica un proceso de selección sobre redes de May que optimiza simultáneamente la estabilidad disminuyendo λ_{max} , y la eficiencia en la comunicabilidad disminuyendo la distancia topológica L (ver discusión en la sección 2.9.1 en relación a la referencia [159]). Notar que una estructura de estrellas interconectadas por unas pocas conexiones implica una baja clusterización, por lo cual es la estructura de las conexiones simples la que da lugar a la alta clusterización en las redes del presente modelo.

4.2.2. Forzado de la anticorrelación

Es bien sabido que existen tanto interacciones excitatorias como inhibitorias entre las neuronas que componen un cerebro. Sin embargo ello no necesariamente conduce a la existencia de interacciones anticorrelacionadas entre diferentes regiones del cerebro. En [76] se menciona que

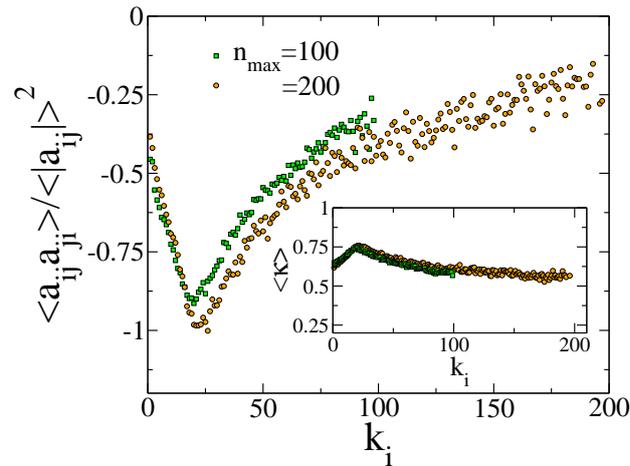


Figura 4.38: Función de correlación entre los elementos de matriz que conecta a un nodo i y su vecino j como función de su grado k_i , para $b = 2$, $q = 3$ y diferentes valores de n_{max} [164]. El inset muestra la fracción (promedio) de conexiones dobles que anticorrelacionan $\langle \kappa \rangle$ como función del grado k_i .

algunas regiones del cerebro intensifican su actividad mientras que otras la disminuyen (respecto de la actividad en reposo) cuando una persona realiza tareas cognitivas que demandan concentración. Esto define una red funcional en donde la actividad de las diferentes regiones pueden estar (positivamente ó negativamente) correlacionadas entre sí. También se agrega que el patrón de correlaciones que las respectivas regiones exhiben ante la realización de tareas cognitivas, es similar al que se observa al estudiar las correlaciones de las fluctuaciones en la actividad espontánea de las distintas regiones del cerebro en reposo (las RSN o *Resting State Networks*[46]). Incluso la similitud persiste durante el sueño[77] o bajo el efecto de anestesia[198]. Además la intensidad de las correlaciones (positivas o negativas) que la red funcional presenta es proporcional a la dificultad de la tarea en cuestión[76]. Éstos resultados son consistentes con la existencia de interacciones anticorrelacionadas entre diferentes regiones del cerebro, de manera que el incremento de actividad en algunas suprime el nivel de actividad en otras mejorando la relación señal-ruido de las primeras. En otras palabras, la existencia de interacciones anticorrelacionadas entre las distintas regiones del cerebro podría estar asociada a la estabilidad funcional del mismo. Por otro lado, en la sección 2.9.1 se menciona un trabajo de Allesina y Pascual[15] en donde se argumenta que un aumento en la proporción de interacciones tipo presa-predador lleva a un incremento de la estabilidad en redes ecológicas. La intencionalidad en dicho trabajo es arrojar luz sobre el dilema *complejidad-estabilidad* que nació con el trabajo de May (sección 2.9.1). Ocurre que no necesariamente la mayoría de las interacciones ecológicas corresponden a interacciones presa-predador (ver por ejemplo [155]), lo cual dificulta

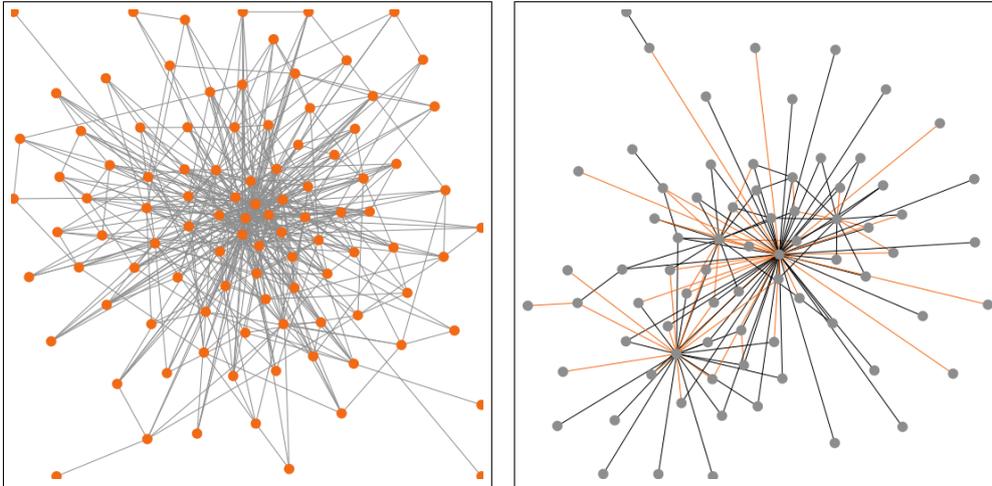


Figura 4.39: **Izquierda:** Estructura de una red típica generada por el presente modelo. En este caso la red tiene $n = 100$ nodos, $M = 488$ links no dirigidos, y fue generada para $b = 2, q = 3$. **Derecha:** Sub-red constituida sólo por conexiones dobles y sus nodos asociados extraída de la red en el gráfico de la izquierda. La sub-red conserva los $n = 100$ nodos pero la cantidad de conexiones no dirigidas se redujo a 133. Notar que está constituida por módulos con forma de estrella interconectados entre sí por unos pocos links. Los links pintados de negro corresponden a interacciones anticorrelacionadas, los naranjas a interacciones positivamente correlacionadas.

argumentar en favor de una alta proporción de interacciones de este tipo. Las consideraciones teóricas de Allesina y Pascual sólo hacen uso de que las interacciones presa–predador son interacciones anticorrelacionadas. Las interacciones presa–predador no son las únicas con ésta propiedad, las interacciones huesped–parásito también son interacciones ecológicas anticorrelacionadas (ver figura (2.12)), son frecuentemente omitidas en la compilación de las redes tróficas *[113, 93], y al mismo tiempo representan una proporción significativa de componentes de la red[114]. Más precisamente, en Lafferty *et. al.*[113] se menciona que los parásitos pueden jugar un rol importante en las propiedades estructurales de las redes ecológicas, e incluso en la estabilidad de los sistemas ecológicos. Por ejemplo, se menciona que posiblemente las redes tróficas de *British Grassland*[121] y *Scotch Broom*[130] (aquí estudiadas) presentan distribuciones de grados de cola larga (y no de cola corta como generalmente ocurre con el resto de las redes) como resultado de la incorporación de parásitos a las mismas[†].

Como se vió anteriormente, una tendencia hacia interacciones anticorrelacionadas es una propiedad emergente del presente modelo. Sin embargo, en vista de las consideraciones previas sería de utilidad poder

*La razón de dicha omisión reside en que originalmente muchas de las interacciones parásitarias son considerados dentro del contexto de los los procesos infecciosos y no de los procesos tróficos.

[†]Se menciona *posiblemente* ya que no es posible afirmar dicha cuestión con certeza debido a que la inclusión de los parásitos a dichas redes tróficas aún no es una cuestión resuelta.

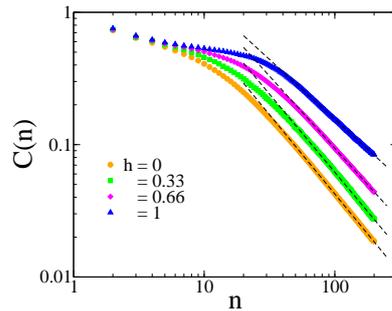


Figura 4.40: Conectividad C en función del tamaño para $b = 2, q = 3$ y diferentes valores de la probabilidad de forzar la anticorrelación h (símbolos). Las líneas discontinuas corresponden a ajustes de la forma $C \sim n^{-(1+\epsilon)}$ para tamaños n suficientemente grandes. ϵ es decreciente con h y eventualmente cambia de signo volviéndose negativo (ver tabla 4.3).

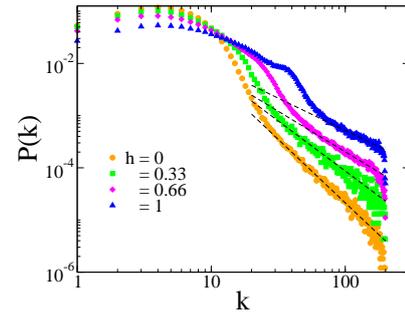


Figura 4.41: Distribución de grados $P(k)$ para distintos grados del forzado de la anticorrelación de las interacciones regulado por el parámetro h (símbolos). Las líneas discontinuas corresponden a ajustes tipo ley de potencias $P(k) \sim k^{-\gamma}$ para valores grandes de k . γ es decreciente con h alcanzando el mínimo valor $\gamma \simeq 1.3$ en $h = 1$ (ver tabla 4.3). Estas curvas fueron obtenidas para $b = 2, q = 3$ y $n_{max} = 200$.

controlar la proporción de interacciones anticorrelacionadas y estudiar el efecto que ello produce en las propiedades emergentes de las redes generadas con el presente modelo. Para estudiar esta posibilidad se modificó levemente el algoritmo de crecimiento de redes introduciendo la posibilidad de forzar la anticorrelación en las interacciones. Mas precisamente, cada vez que se sortean un par de interacciones $a_{i,j}$ y $a_{j,i}$, luego con probabilidad h se le modifica el signo a cualquiera de ellas de manera que queden anticorrelacionadas ($a_{i,j}a_{j,i} \leq 0$). Es importante resaltar que cuando $h = 0$ se recupera el algoritmo de crecimiento sin el forzado de la anticorrelación de las interacciones. Esta variación del modelo es denominada: *modelo de interacciones anticorrelacionadas* (ver sección 4.0.2, item *iii*). En la figura (4.40) se observa la conectividad C en función del tamaño n para distintos valores del parámetro h . Claramente las curvas se desplazan hacia valores mayores de n cuanto más grande es h . Esto es indicativo de que forzar la anticorrelación tiende a estabilizar las matrices en concordancia con lo observado por Allesina y Pascual[15]. Interesantemente, cuando h es lo suficientemente grande la relación $C \sim n^{-(1+\epsilon)}$ cambia de modo que ϵ se vuelve negativo (ver tabla 4.3). Este resultado automáticamente implica la desaparición de la cota al tamaño máximo mencionada en la sección (4.1.7). Por otro lado, recordar que el ajuste de datos empíricos en muchas redes tróficas arroja valores $\epsilon < 0$.

Para ver como el forzado de la anticorrelación afecta la topología de la red de interacciones en la figura (4.41) se muestra el comportamiento de la distribución de grados para distintos valores del parámetro h . Se observa que la emergencia de una cola tipo ley de potencias $\sim k^{-\gamma}$ es preservada, ocurriendo a todos los valores del parámetro h estudiados. Sin embargo,

h	ϵ	γ
0	0.21	2.43
0.33	0.18	1.92
0.66	0.09	1.55
1	-0.10	1.30

Cuadro 4.3: Valores de ϵ y γ que resultan de los ajustes de las figuras (4.40) y (4.41) respectivamente.

el exponente asociado a la ley de potencias cambia, encontrándose que un aumento de h corresponde a una disminución de γ de forma tal que el mínimo correspondiente a $h = 1$ resulta en $\gamma \simeq 1.3$ (ver tabla 4.3). Este valor en torno de 1 ha sido previamente reportado en la literatura en el contexto de las redes tróficas [133, 132], y en algunos casos en redes metabólicas [200] (si bien en la mayoría de los casos se encuentran exponentes de orden 2 o mayores [97]). Los exponentes asociadas a las redes de proteínas son del orden de 2 o mayores [122, 10, 25, 53], aunque también hay reportes de valores menores [192]. En definitiva, jugando con los parámetros b, q, R y h se puede cubrir un amplio espectro de valores de γ . Ello en particular pone de manifiesto que las razones de que una red tenga γ dentro de cierto rango, no necesariamente posea una única explicación.

Capítulo 5

Conclusiones

En este trabajo se propone que un mecanismo de presión selectiva basada en la estabilidad puede constituir una explicación a la emergencia de topologías no triviales en redes naturales. Para verificar la validez de esta hipótesis se introdujo un modelo simplificado, pero de gran generalidad, que implementa las características básicas de dicho mecanismo. Mediante el estudio de las propiedades topológicas de las redes resultantes y las propiedades dinámicas asociadas, se encontró que muchas de las características observadas en redes biológicas pueden explicarse como propiedades emergentes de dicho mecanismo. Se mostró así mismo que dichas propiedades emergentes resultan robustas ante variaciones del modelo, es decir, ante variaciones del criterio de estabilidad. Por otro lado, dichas variaciones permiten a su vez dar evidencia de la robustez del criterio de estabilidad. Por ejemplo, la flexibilidad introducida a través del parámetro Δ (sección 4.1.2) muestra que un criterio de estabilidad lineal tiene un amplio alcance. Esto es así ya que las propiedades emergentes de las redes generadas con éste mecanismo se conservan incluso para $\Delta > 0$, lo cuál es suficiente para apantallar efectos no lineales que no sean muy fuertes. La introducción del parámetro R (sección 4.1.2) pone en evidencia que las propiedades emergentes persisten incluso cuando la estabilidad no es total pero si predominante. El análisis en torno a la conectividad permite englobar dentro de un mismo marco a diferentes tipos de sistemas biológicos que cubren un amplio espectro de escalas, desde las microscópicas como las redes de interacción de proteínas o metabólicas, hasta las macroscópicas de las redes ecológicas. Incluso provee de una explicación, al menos cualitativa, del caso particular de las redes tróficas en donde eventualmente podría ocurrir que $\epsilon < 0$ cuando predominan las interacciones anticorrelacionadas (sección 4.2.2). Muchas redes biológicas presentan topologías libres de escala, propiedad frecuentemente asociada a redes complejas. El modelo aquí presentado es capaz de reproducir ésta topología no trivial, sugiriendo que es una consecuencia muy probable de un mecanismo de presión selectiva basado en la estabilidad. Incluso provee de una explicación de porqué a tamaños pequeños las redes tróficas no exhiben una distribución de grados de cola larga y

a tamaños grandes si (secciones 2.8.3[134] y 4.1.2). Notablemente en el régimen de $\epsilon < 0$, que corresponde a *crecimiento acelerado** se obtiene $\gamma \simeq 1.3 < 2$ en concordancia con los modelos de crecimiento acelerado (sección 2.1.3) y lo reportado en las redes tróficas libres de escala (sección 2.8.3). En particular, se observa que el predominio de las interacciones tipo *presa-predador* o *huesped-parásito* (correspondientes a interacciones anticorrelacionadas) resulta ser un recurso estabilizante mucho más fácil de obtener que el debilitamiento de las interacciones (sección 4.2.1). Esto refuerza la idea de que la topología de las redes presenta una relación íntima con la estabilidad del sistema. El modelo también provee evidencias de que un balance entre competitividad y cooperación favorece la complejidad (sección 4.1.2). Al mismo tiempo, pone en evidencia que (al menos para sistemas pequeños) existen explicaciones alternativas al *conexionado preferencial* para dar lugar a distribuciones de grados de cola ancha (sección 4.1.2). En relación a esta cuestión, el modelo provee de una posible explicación a porqué las redes biológicas presentan tamaños mucho más chicos que, por ejemplo, las redes sociales (sección 4.1.7). Probablemente, la particular estructura de las redes biológicas sea el resultado de que operen mecanismos evolutivos (en el sentido de selección natural) en los sistemas biológicos subyacentes. En contraste, en sistemas sociales parecen operar mecanismos de naturalezas diferentes como (por ejemplo) el *conexionado preferencial*. Más precisamente, nadie que ingresa a una sociedad establece contacto simultáneamente con un montón de personas, es decir, nadie se convierte en un *hub* de la noche a la mañana. Por el contrario en el caso de una red de proteínas, una mutación puede dar lugar a la aparición inmediata de un *hub*. La relación del presente modelo con las redes biológicas se ve reforzada ya que reproduce el carácter principalmente disasortativo de las mismas, e incluso presenta un comportamiento mixto similar al observado en algunas redes de proteínas (sección 4.1.5). Incluso en el análisis en torno a las propiedades de mundo pequeño se encuentra buena concordancia con lo observado en redes biológicas. Las distancias topológicas de las redes biológicas son similares a la del modelo. Por otro lado, tanto las redes biológicas como las del presente modelo, exhiben una relativamente alta clusterización en comparación con el exhibido por redes aleatorias. El análisis de la clusterización sugiere además una posible organización jerárquica frecuentemente observada en redes biológicas (sección 4.1.4). Entre otras cosas, se confirma lo encontrado en un modelo previo en donde el sistema estabiliza modularizándose, y en donde cada módulo posee esencialmente una forma de estrella tal como las observadas en las sub-redes de conexiones dobles del presente modelo (sección 4.2.1). En el presente modelo, la propiedad libre de escala en la distribución de grados se determina promediando sobre un ensamble de redes generadas por el mismo. No es posible determinar tal propiedad a partir de una sola de dichas redes dado

*Crecimiento acelerado es cuando $\langle k \rangle$ crece con el tamaño.

el tamaño relativamente pequeño que las mismas poseen. Sin embargo, cada una de las redes generadas por el modelo exhibe un respuesta ante daños topológicos muy similar al que normalmente se le atribuye a las redes con distribuciones de grados de cola ancha (sección 4.1.6). Además la respuesta ante los daños es muy similar a la observada en las redes tróficas (apéndice A). El análisis estadístico de los autovalores asociados a las matrices generadas en el presente modelo resultan similares a otros obtenidos para redes de May libres de escala (sección 4.2). Notar que el modelo y sus diferentes variaciones permiten cubrir un amplio rango de valores de γ . Los rangos de γ cubiertos por las distintas variaciones muchas veces solapan, indicando que no se puede identificar inmediatamente (observando a γ) qué particular/es ingrediente/s está/n en juego en un dado sistema. Finalmente, los resultados aquí obtenidos sugieren que los mecanismos evolutivos en donde los sistemas crecen sujetos a una presión selectiva basada en la estabilidad explican mejor lo observado en redes biológicas que muchos otros enfoques mejor adaptados a las redes sociales y/o tecnológicas.

5.1. Posibles extensiones al presente trabajo

En lo que sigue se hará repetidamente referencia al mecanismo propuesto en este trabajo que dió origen al modelo que ha sido ampliamente estudiado, más precisamente al mecanismo evolutivo de generación de redes de interacción en donde actúa una presión selectiva basada en la estabilidad de la dinámica del sistema. Conviene introducir una forma abreviada por lo que se hablará del *mecanismo STS* (*STS*: Selection Through Stability mechanism) cuando se haga referencia al mecanismo en cuestión.

A continuación voy a introducir una lista de posibles extensiones al presente trabajo:

- i) Avalanchas de extinciones:* En la sección 4.1.2 quedó de manifiesto que la presencia de extinciones juega un rol importante en el presente modelo ya que la presencia de las mismas permite generar redes con distribuciones de grados con exponentes $\gamma < 3$ (distribuciones de cola larga o varianza infinita), mientras que la ausencia de extinciones inevitablemente conduce a $\gamma > 3$ (distribuciones de cola corta ó varianza finita) a diferencia de lo que se observa en redes empíricas en donde $\gamma < 3$ (sección 2.1.2). Sin embargo el mecanismo de extinciones implementado en el presente modelo posee un aspecto arbitrario e irrealista: cada vez que el sistema inestabiliza, tan sólo un máximo q de especies pueden resultar extintas. La introducción del parámetro q ha sido necesaria para evitar que los tiempos de simulación se disparen conduciendo a un crecimiento

demasiado lento del sistema. Más precisamente, si se permiten procesos de extinciones de tamaño arbitrario, el sistema no alcanza a crecer (en tiempos de simulación razonables) a tamaños adecuados para estudiar las propiedades topológicas de las correspondientes redes.

¿Como es la estadística de extinciones en sistemas ecológicos? En algunos *Reviews* se discute la relación entre la estadística de extinciones en ecología de acuerdo a lo observado en registros fósiles con lo obtenido a partir de diferentes modelos (Newman y Palmer[150], Drossel[61]). En ellos se menciona que el problema está sujeto a debate principalmente por dos cuestiones. Por un lado la baja resolución de los datos empíricos no permite determinar de una manera suficientemente confiable el comportamiento de las distintas distribuciones estadísticas. Por otro lado no hay consenso respecto de si los eventos de extinción masiva tienen como causa perturbaciones externas al ecosistema (como pueden ser la caída de meteoritos, cambios en el nivel del mar, etc.) o causas internas debidas a efectos no lineales en la dinámica del ecosistema. Respecto a la primera de estas cuestiones, a pesar de la baja resolución de los datos en los registros fósiles, el consenso es que existen eventos en donde muchas especies* se extinguen “simultáneamente” (en otras palabras, una *avalancha de extinciones*), y que la distribución $P(s)$ del número de especies s extintas en un mismo evento decae como una ley de potencias $P(s) \sim s^{-\tau}$ de exponente $\tau \simeq 2$. La segunda de estas cuestiones está íntimamente relacionada al trabajo teórico en el área ya que diversos modelos que adoptan una u otra posición reproducen la mencionada estadística de extinciones. Más precisamente, no hay consenso respecto de cuál es el mecanismo que da lugar a la estadística observada. Por un lado están los modelos basados en coevolución en donde la extinción de una especie afecta el *fitness* de especies relacionadas pudiendo causar una avalancha de extinciones. En particular, Bak y Sneppen propusieron un simple modelo de coevolución que exhibe *criticalidad auto organizada* (*SOC: Self Organized Criticality*) y desencadenó gran actividad académica en el área[22]. Dicho modelo no reproduce el exponente $\tau = 2$ pero si el comportamiento tipo ley de potencias en $P(s)$, de todas maneras existen otros modelos que exhiben *SOC* y reproducen el exponente $\tau = 2$ [188]. Por otro lado, desde un enfoque diametralmente opuesto Newman [139] propone un modelo en donde las extinciones son causadas por las presiones que el entorno ejerce sobre las especies. Más precisamente, en dicho modelo no hay interacción entre las especies y también reproduce el exponente $\tau = 2$. La cuestión de si

*Para ser más precisos, los registros de especies en general tratan de familias de especies constituidas por grupos de especies con similitudes taxonómicas. La razón es que es más fácil determinar el momento de origen y extinción de una familia de especies que de una única especie.

la dinámica macroevolutiva es un proceso a gran escala que hereda (o no) una fenomenología crítica inherente en procesos ecológicos a menor escala es una cuestión aún en debate[187]. Evidencias de la existencia de grandes fluctuaciones en la estadística de extinciones en ecosistemas actualmente en actividad[102], sugieren que la actividad a nivel ecosistemas y la actividad a nivel macroevolutivo están íntimamente relacionados.

Claramente el modelo aquí estudiado *no* reproduce la estadística de extinciones $P(s) \sim s^{-\tau}$ debido a las limitaciones introducidas por el parámetro q (ya que $s \leq q$ y q es pequeño). Quizás sea posible implementar un modelo basado en el *mecanismo STS* (por ejemplo una variación del modelo aquí estudiado) que eventualmente reproduzca la estadística de extinciones. Esto contribuiría a esclarecer la cuestión de si la macroevolución está principalmente dominada por agentes exógenos ó endógenos a los ecosistemas. Un enfoque basado en el *mecanismo STS* se asemeja más al enfoque normalmente planteado en los modelos coevolutivos que exhiben *SOC* ya que los mismos se centran en las interacciones entre las componentes del sistema. Ésta observación sugiere que un modelo que implemente el *mecanismo STS* y reproduzca la estadística de extinciones $P(s) \sim s^{-\tau}$ plausiblemente exhiba *SOC*.

En el presente trabajo, se enfatizó el estudio de las propiedades topológicas de las redes que emergen del proceso de selección, encontrándose que el *mecanismo STS* reproduce muchas propiedades topológicas observadas en redes biológicas. Comúnmente en los trabajos que se proponen y estudian modelos que reproducen la estadística de extinciones, no se estudian las propiedades topológicas emergentes en la red de interacciones [188], o se restringen a topologías que reproducen (en buena medida) las observadas experimentalmente [85]. Surge entonces la opción de estudiar la relación entre la (presumiblemente adecuada) estadística de extinciones y las propiedades topológicas emergentes de las redes de interacción en un modelo que implementa el *mecanismo STS*. En particular ello permitiría estudiar la relación entre el patrón de conexiones que presenta una especie, su susceptibilidad a extinguirse, y su capacidad de provocar avalanchas de extinciones. Por otro lado, se podría investigar la correlación entre el tamaño de una avalancha y el grado de estabilidad global que el sistema presenta previo a la misma (por ejemplo, una correlación entre λ_{max} y s).

- ii) *Dinámica no lineal*: El presente modelo constituye una fuerte evidencia de que en los sistemas generados por un modelo que implementa el *mecanismo STS*, emergen propiedades comúnmente observadas en sistemas biológicos o al menos merecen ser estudiados con mayor profundidad. Sin embargo, el modelo está basado en una aproximación lineal de la dinámica, y aunque variaciones del mismo

(correspondientes a los parámetros Δ y R introducidos en la sección 4.1.2) sugieren que la aproximación lineal es buena, estudiar modelos que implementen explícitamente una dinámica no lineal reforzaría la generalidad del *mecanismo STS*. La introducción de una dinámica no lineal tiene a su vez otras ventajas, permite introducir explícitamente la evolución de $\vec{x}(t)$, es decir, explícitamente calcular la dinámica de poblaciones, o el nivel de expresión de diferentes metabolitos, etc. (ecuación (3.1)). En el presente modelo el criterio de estabilidad utilizado corresponde a la existencia (o ausencia) de una dinámica *localmente estable* en donde el sistema retorna al punto de equilibrio inicial tras aplicarle una pequeña perturbación. Sin embargo, el cálculo explícito de $\vec{x}(t)$ permite considerar formas alternativas de determinar la “estabilidad”* del sistema. En general, éstas formas alternativas son propuestas dentro del contexto de los sistemas ecológicos[166],[127],[94].

Introducir una dinámica no lineal casi inevitablemente implica considerar la cuestión de extinciones explícitamente, y de hecho es un recurso utilizado frecuentemente para dar realismo a los modelos [62],[128],[64]. Por ejemplo, se podría plantear un modelo evolutivo en donde la especie i -ésima se extingue al tiempo t cuando $x_i(t) < \epsilon$ (es decir, cuando la densidad de población de la especie i -ésima decae por debajo de cierto umbral arbitrario ϵ), y dejar que la red de interacciones se auto-organice seleccionando las especies de acuerdo a ésta dinámica de extinciones. En otras palabras, considerar un modelo en donde la selección ocurre a través de un mecanismo explícito de extinciones, más brevemente un *mecanismo STEE* (*STEE*: Selection Through Explicit Extinctions mechanism). No necesariamente emerge un *mecanismo STS* de un modelo que explícitamente implementa un *mecanismo STEE*, sin embargo sería interesante estudiar la relación entre ambos (tal como es sugerido al final del ítem i). Un criterio de estabilidad dentro de las dinámicas no lineales en ecología denominado *variabilidad* (la varianza de cada $x_i(t)$ a lo largo del tiempo[127]) podría ser un buen candidato para estudiar la relación entre un *mecanismo STS* y un *mecanismo STEE*, ya que especies con dinámicas poblacionales que presentan mayores fluctuaciones podrían estar más propensas a extinguirse. Es importante resaltar que hay dos grandes grupos de sistemas dinámicos, los de tiempo continuo, y los de tiempo discreto. En particular el análisis lineal de la estabilidad propuesto en el presente modelo se adapta bien a ambos casos (debido a que el tiempo no es considerado explícitamente) pero por otro lado presenta una desventaja. En un análisis lineal de la estabilidad las densidades

*Pongo estabilidad entre comillas ya que como consecuencia del debate estabilidad-complejidad, el término se ha usado frecuentemente para indicar conceptos muy distintos tanto desde el punto de vista semántico como del punto de vista matemático.

de poblaciones $\vec{x}(t)$ no son explícitamente calculadas, y por ende la mecánica de extinciones ha de introducirse *ad hoc*. Más precisamente, la decisión de cuales especies se extinguen en caso de inestabilidad adolece de demasiada arbitrariedad. En un modelo en donde $\vec{x}(t)$ es calculado explícitamente, criterios más realistas de extinciones de especies pueden ser implementados, por lo cuál ésta es otra ventaja de introducir una dinámica no lineal en un modelo que implemente el *mecanismo STS*.

El cálculo explícito de $\vec{x}(t)$ también ha sido utilizado para estudiar un mecanismo de selección de la red de interacciones basado en la sincronización de sus componentes[81], o más brevemente, un *mecanismo STSync* (*STSync*: Selection Through Synchronization). En dicho modelo, las componentes tienden a generar interacciones entre si cuando sincronizan y a perderlas en caso contrario. Este modelo genera redes con propiedades topológicas muy similares a la del modelo aquí estudiado. Quizás exista un factor común detrás de todos estos mecanismos, un mecanismo de selección que busca maximizar la correlación entre las componentes de $\vec{x}(t)$, o más brevemente, un *mecanismo STC* (*STC*: Selection Through Correlations mechanism). En un modelo que implementa un *mecanismo STC*, la idea sería generar interacciones entre componentes que tienden a correlacionar su actividad, eliminarlas en caso contrario, y eventualmente extinguir componentes que no correlacionan (significativamente) con ninguna otra. Éste particular enfoque podría ser de utilidad para modelar lo observado en redes obtenidas a partir de estudios en neuroimágenes ya que las mismas se definen a partir de la correlación dinámica entre sus componentes[68],[5],[73]. En otras palabras un *mecanismo STS* podría ser parte de una familia de mecanismos relacionados (por ejemplo los mecanismos *STEE*, *STSync* y *STC*) lo cual lo convierte en un interesante tópico de estudio. La consideración de dinámicas no lineales es necesaria para estudiar la relación entre el *mecanismo STS* y el resto.

- iii) *Análisis detallado de la presencia de organización modular jerárquica*: En el presente modelo, el comportamiento de $Cc(k)$ (sección 4.1.4) sugiere la existencia de una organización modular jerárquica en las redes generadas por el mismo en correspondencia con lo propuesto por Ravasz *et. al.*[176] (sección 2.5.1). El estudio de redes con organización en forma de estructuras de comunidades o modulares es un área muy activa en la actualidad tal como se puede apreciar en un muy reciente *Review* de Fortunato[74]. Fortunato menciona que frecuentemente las redes empíricas presentan una organización modular en donde los módulos se componen por módulos más pequeños dando lugar a una organización jerárquica. Además agrega que la generación y evolución de un sistema organizado en subsistemas estables e interrelacionados es mucho más rápida que

si el sistema fuera desestructurado, debido a que es más fácil primero ensamblar las subpartes y luego utilizarlas como unidades en el siguiente nivel de organización. Este escenario considerado por Fortunato, en donde la estabilidad y la organización modular jerárquica juegan un rol central, está íntimamente relacionado con lo observado en el presente modelo. Más precisamente, una estructura modular jerárquica en la topología de las redes generadas por el presente modelo induce una organización modular jerárquica de la estabilidad ya que a cada módulo le corresponde a una submatriz (de la matriz que representa toda la red) que posee asociada una estabilidad. En términos más generales, existe la posibilidad que la dinámica de un sistema presente algún tipo de organización modular jerárquica que de alguna manera esté en correspondencia con una organización jerárquica en la topología de la red de interacciones entre las componentes del mismo. Trabajos en ésta dirección fueron previamente publicados. Por ejemplo en Arenas *et. al.*[18] se encuentra una relación entre la organización jerárquica en la sincronización de un sistema de osciladores acoplados y la organización jerárquica en la estructura de la red de interacciones. En Arenas *et. al.*[18] la red de interacciones es estática, pero en [81],[82] Portillo y Gleiser encuentran que un acoplamiento entre la evolución de la red de interacciones y la sincronización de sus componentes (un *mecanismo STSync*) dan lugar a la emergencia de organización jerárquica en la topología de la red. En particular mencionan similitudes en lo observado en dichos modelos y lo observado en redes derivadas a partir de estudios de resonancia magnética en la actividad del cerebro (también llamada resonancia magnética funcional ó *FMRI: functional magnetic resonance imaging*) [68],[5],[73]. Otros trabajos sobre resonancia magnética funcional encuentran que la actividad del cerebro exhibe características propias de sistemas con una dinámica crítica[104],[172], lo cual sugiere que una dinámica crítica es muy plausible en éste tipo de sistemas. En otras palabras, éstas consideraciones están relacionadas a las discutidas en los items *i*) y *ii*) (extinciones y dinámica no lineal), lo cuál afianza la posibilidad de que un *mecanismo STS* esté asociado a la emergencia de organización jerárquica tanto en la topología de las interacciones como en la dinámica, y que haya una íntima relación entre ambas.

Investigar estas cuestiones requiere inevitablemente de métodos para determinar la existencia y forma de organización modular jerárquica en redes. Distintos métodos han sido propuestos, muchos de los cuales son analizados y comparados en el review de Fortunato[74]. Fortunato clasifica los métodos propuestos en dos grandes clases. Por un lado están los *métodos de resolución múltiple* en donde la existencia de un parámetro libre permite predecir el tamaño de los módulos a detectar (ejemplos:[174],[19]).

Por otro lado están los *métodos jerárquicos* (ejemplos:[180],[48]) los cuales se basan en definir una similaridad entre los distintos pares de vértices (sin importar si están conectados o no), y luego utilizar una técnica de *clusterización jerárquica* la cuál a su vez puede ser implementada a través de *métodos aglomerativos* (en donde los clusters son iterativamente aglomerados si la similaridad entre ellos es alta) ó *métodos divisivos* (en donde los clusters son iterativamente fraccionados por medio de remover links que conectan vértices con baja similaridad entre si). Los *métodos de resolución múltiple* suelen ser más confiables pero en general más lentos que los *métodos jerárquicos*, por lo que es una cuestión a explorar cual de ellos es el más adecuado. Finalmente, cabe resaltar que algunos métodos permiten estudiar la organización modular jerárquica en redes con interacciones pesadas (redes en donde los links tienen asociado un “factor de relevancia”). Ésto podría ser particularmente útil para estudiar la organización jerárquica en la matriz de correlaciones de la actividad de las componentes del sistema, lo cuál permitiría comparar la estructura topológica y la funcional en sistemas con dinámicas no lineales.

- iv) *Modelos más específicos*: Estudiar modelos que implementan el *mecanismo STS* sobre redes restringidas a poseer propiedades particulares exhibidas por otros modelos o sistemas empíricos, es un enfoque que resultaría fructífero, tanto en el entendimiento del *mecanismo STS*, como en el entendimiento de dichas topologías específicas y su relación con la naturaleza. Por ejemplo, podría establecerse que determinadas restricciones topológicas no son consistentes con sistemas estables.

Allesina *et. al.* [14] introduce un enfoque en el cuál propone comparar la probabilidad de reproducir con diferentes modelos la estructura observada en un conjunto de redes tróficas obtenidas empíricamente. Más precisamente, la idea es comparar las *verosimilitudes* con que cada modelo genera las redes empíricas. En particular hacen un estudio comparativo utilizando cuatro sencillos modelos (ordenados de menor a mayor verosimilitud asociada), el *cascade model*[50], el *niche model*[207], el *nested hierarchy model*[44], y el *minumum potential niche model* introducido por Allesina *et. al.* en el mismo trabajo. Ellos mencionan que realizar dicho estudio en modelos que explícitamente incorporan procesos evolutivos es impracticable (al menos actualmente), lo cuál afecta directamente al presente modelo (y a modelos basados en un *mecanismo STS*). Sin embargo, es posible modificar el presente modelo para que genere redes de topologías restringidas de acuerdo a lo permitido por los modelos anteriores. Ello brindaría la posibilidad de estudiar el *mecanismo STS* sobre las topologías más específicas de las redes tróficas. Trabajos en direcciones similares han sido reciente-

mente publicados lo cuál indica que es un área activa de investigación[85],[101],[84]. Otras posibilidades consisten en combinar el *mecanismo STS* con mecanismos previamente introducidos en el contexto de las redes complejas. Por ejemplo, se podría combinar con el concepto de capacidad de carga del grado de los nodos introducido en[17]. Este particular enfoque permitiría estudiar como afecta la ausencia de *hubs* las propiedades observadas en las redes generadas por el presente modelo. Por otro lado, versiones más específicas del modelo orientadas al estudio de redes celulares pueden obtenerse al combinarlo con un mecanismo de duplicación de nodos[103],[53] comúnmente asociado a un mecanismo de duplicación genética[193]. Finalmente otra posibilidad consiste en combinarlo con un mecanismo que favorezca la asortatividad como el introducido en [43], y así estudiar el efecto que tiene la asortatividad sobre un proceso de selección basado en la estabilidad. Ello permitiría eventualmente iniciar un estudio sobre la estabilidad en sistemas sociales.

Apéndice A

Tolerancia a ataques dirigidos y fallas aleatorias en redes tróficas

Según se discutió en la sección 2.7 la respuesta de una red ante un proceso de aislación de nodos puede caracterizarse a través de las funciones $S(f)$ y $\langle s \rangle(f)$. En las secciones 2.8.4 y 4.1.6 se mencionó que en la literatura no existe una comparación entre la respuesta que presentan las redes ecológicas ante un proceso de aislación de nodos, y la respuesta que presentan versiones aleatorizadas de dichas redes. Dicha comparación es importante, ya que sólo si existe una diferencia apreciable entre los valores de f_c obtenidos para las redes ecológicas y los obtenidos para las redes aleatorias, puede concluirse que las redes ecológicas presentan alguna particularidad en relación a la robustez topológica. Ello motivó a realizar tal comparación utilizando las redes ecológicas obtenidas de la base de datos de A.G. Rossberg (ver apéndice B para más información sobre la base de datos).

En la serie de gráficos exhibidos en las figuras (A.1),(A.2),(A.3),(A.4) y (A.5) se presenta el comportamiento de $S(f)$ y $\langle s \rangle(f)$ ante ataques dirigidos y fallas aleatorias para cada una de las redes en la base de datos, y de sus correspondientes versiones aleatorizadas (redes aleatorias de mismo tamaño n y número de links M). Lo concluido a partir de lo observado en dichas figuras se resume en la tabla A.1. En dicha tabla se hace referencia al tipo de distribución de grados correspondiente a cada red trófica según la clasificación echa por Dunne *et. al.* [65]. Tres de las redes, a saber: *Benguela Current*, *Caribbean Reef* y *Northeast US Shelf* no fueron clasificadas en Dunne *et. al.*, por lo que la correspondiente clasificación es realizada en los gráficos **a**),**b**) y **c**) de la figura (A.6). El método de clasificación consiste en calcular la *distribución cumulativa complementaria* $P_{\geq}(k)$ de la distribución de grados $P(k)$:

$$P_{\geq}(k) = \sum_{r \geq k}^{\infty} P(r) \quad (\text{A.1})$$

y luego determinar a través de un ajuste el correspondiente comportamiento. Trabajar con la distribución acumulativa complementaria $P_{\geq}(k)$ (en vez de trabajar con la distribución de grados $P(k)$) es conveniente para obtener una clasificación correcta a través del ajuste. Ésto es así ya que suaviza el ruido existente en la $P(k)$ el cuál es muy acentuado en las redes ecológicas por ser redes pequeñas.

$P_{\geq}(k)$ es una función no creciente, y en la práctica suele decaer. En un gráfico normal-log* $P_{\geq}(k)$ decae como una recta cuando $P(k)$ decae exponencialmente. Si $P(k)$ es ley de potencias, entonces $P_{\geq}(k)$ decae más lento que una recta y presenta una curvatura hacia arriba. Si $P(k)$ es uniforme, entonces $P_{\geq}(k)$ posee una curvatura hacia abajo. Lo mismo ocurre si $P(k)$ es una distribución de Poisson pero con diferencias respecto del caso uniforme. La distribución de Poisson no decae al principio (la uniforme si), y presenta un truncamiento para valores más chicos de k . En el gráfico **a**) de la figura (A.6) se presenta una comparación entre la $P_{\geq}(k)$ correspondiente a una adecuada distribución $P(k)$ de Poisson [†] y el comportamiento que presenta la $P_{\geq}(k)$ de la red ecológica de *Northeast US Shelf* que corresponde a una $P(k)$ uniforme. Ejemplos de como son las curvas para los casos en que $P_{\geq}(k)$ corresponde a una ley de potencias (*British Grassland*) y a una exponencial (*El Verde Rainforest*) son presentados en el gráfico **d**) de la figura (A.6) y la figura (A.7) respectivamente.

Las conclusiones que se pueden extraer de la tabla (A.1) – teniendo en cuenta algunas excepciones – son:

- i*) No sólo en las redes tróficas libres de escala (o de cola larga) se aprecia una respuesta diferente a los ataques dirigidos respecto de lo que ocurre con las redes aleatorias. Dicha diferencia también es apreciable en las redes en donde la distribución de grados es de tipo exponencial.
- ii*) Si la conectividad C de una red ecológica es alta y/o su distribución $P(k)$ es uniforme, la diferencia en la respuesta ante ataques dirigidos es baja.
- iii*) La diferencia ante fallas aleatorias no es tan significativa como la diferencia ante ataques dirigidos.

El ítem *i*) es particularmente llamativo ya que en los trabajos que originalmente dieron lugar a la idea de estudiar $S(f)$ y $\langle s \rangle(f)$ [12],[13], el énfasis se puso en resaltar la diferencia encontrada entre redes libres de escala y redes aleatorias[‡]. Aquí se encuentra una diferencia significativa

*Es decir, el eje de las x con espaciado normal, y el eje de las y con espaciado logarítmico.

[†]Adecuada significa red aleatoria al mismo n y M .

[‡]Notar que en Albert *et. al.* [12] (el paper original) se habla de redes exponenciales. Sin embargo se menciona que dichas redes exponenciales poseen una distribución de

Red	Diferenciabilidad	Dist. $P(k)$	Conectividad C
British Grassland	alta	ley de potencias	0.053
El Verde Rainforest	alta	exponencial	0.11
Little Rock Lake	alta	exponencial	0.23
Canton Creek	alta	exponencial	0.14
Chesapeake	alta	exponencial	0.08
Scotch Broom	alta	ley de potencias	0.061
Stony Stream	alta	uniforme	0.14
Ythan Estuary (91)	alta	exponencial	0.12
Ytahn Estuary (96)	alta	exponencial	0.076
Bridge Brook Lake	media	uniforme	0.34
Coachella Valley	media	uniforme	0.54
Northeast US Shelf	media – baja	uniforme	0.45
St. Marks Seagrass	media – baja	uniforme	0.21
St. Martin Island	media – baja	exponencial	0.24
Benguela Current	baja	uniforme	0.47
Caribbean Reef	baja	uniforme	0.41
Skipwith Pond	baja	uniforme	0.61

Cuadro A.1: Clasificación de las redes tróficas extraídas de la base de datos de A.G. Rossberg según sus respuestas ante ataques dirigidos o fallas aleatorias. La primera columna corresponde al nombre de la red trófica. La segunda columna sugiere el grado en el que se diferencia la respuesta de la red ante los distintos tipos de daño al comparar con la respuesta que presenta una correspondiente red aleatoria. La tercera columna es la clasificación de la distribución de grados según Dunne *et. al.* [65] (y según los gráficos **a**), **b**) y **c**) de la figura (A.6)). La cuarta columna corresponde a la conectividad de la red.

entre redes aleatorias que poseen distribuciones de grados de Poisson y redes ecológicas que poseen distribuciones de grados exponenciales, ambas consideradas distribuciones de cola corta. Posiblemente, ésta cuestión tiene su origen en el hecho de que las redes consideradas en éstos cálculos son pequeñas.

A.1. Modularidad jerárquica en redes tróficas

Según se vió en la sección 2.5.1 en redes de origen biológico y libres de escala que presentan tamaños mayores al de las redes ecológicas, Cc decae como una ley de potencias ($Cc(k) \sim k^{-\beta}$, ecuación (2.40)) sugiriendo una posible organización jerárquica en las mismas. Aunque tan sólo dos de las redes en la base de datos de A.G. Rossberg fueron catalogadas en Dunne *et. al.* [65] como redes libres de escala, es interesante al menos para estos dos casos estudiar el comportamiento de $Cc(k)$. En la figura (A.8) se presenta el comportamiento de $Cc(k)$ para *British Grassland* y *Scotch*

grados de Poisson. Ello lleva a una confusión, convendría llamar *red exponencial* a una red cuya distribución de grados es exponencial, y no a una red aleatoria cuya distribución de grados es de Poisson.

Broom que son las redes catalogadas como libres de escala. Se observa que $Cc(k)$ decae como una ley de potencias para la primera de ellas, mientras que para la segunda se observa un comportamiento irregular. En el caso del resto de las redes de la base de datos (que no son libres de escala), el comportamiento de $Cc(k)$ también es irregular como en el caso de *Scotch Broom*. En resumen, según el comportamiento observado en las $Cc(k)$ de las diferentes redes, hay evidencia que la presencia de modularidad jerárquica es plausible en al menos una de ellas. La cuestión de que en otro tipo de redes biológicas la ley $Cc(k) \sim k^{-\beta}$ es más ubícua que en las redes ecológicas quizás tenga su fundamento en el relativamente pequeño tamaño de estas últimas.

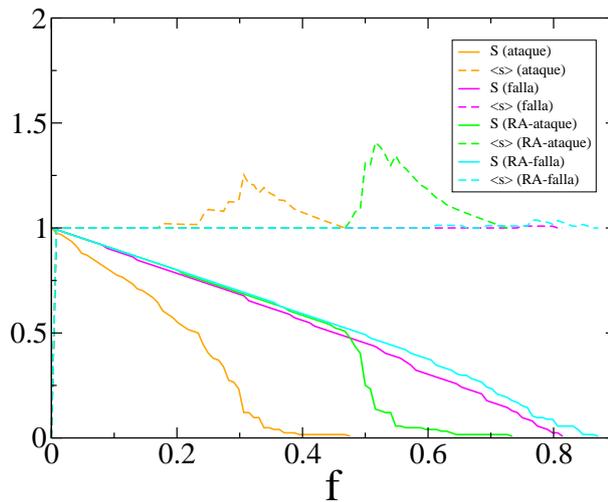


Figura A.1: Comportamiento del tamaño relativo de la componente más grande $S(f)$, y el tamaño promedio de las componentes secundarias $\langle s \rangle(f)$ en función de la fracción de nodos a los cuales se le removieron sus links f , para la red trófica *Ythan Estuary* (96). Dicha red posee $n = 124$ nodos y $M = 574$ links. La descripción de las curvas es la misma que la de la figura (4.32) en la sección 4.1.6. Las líneas continuas denotan $S(f)$, las discontinuas denotan $\langle s \rangle(f)$. Las líneas naranjas y magenta corresponden a las redes tróficas (indicadas por las etiquetas: **ataque** y **falla** en las leyendas). Las verdes y cian corresponden a las redes aleatorias (indicadas por las etiquetas: **RA-ataque** y **RA-falla** en las leyendas, donde RA son las siglas de Red Aleatoria). Las líneas naranjas y verdes corresponden a ataques dirigidos (indicadas por las etiquetas: **ataque** y **RA-ataque** en las leyendas), las magenta y cian a fallas aleatorias (indicadas por las etiquetas: **falla** y **RA-falla** en las leyendas). Se observa que ante ataques dirigidos, la red trófica colapsa (se fragmenta) a una fracción $f_c \simeq 0.3$, valor significativamente menor al cual fragmenta la versión aleatorizada de dicha red ($f_c \simeq 0.5$). Ante fallas aleatorias ambas redes responden de manera similar colapsando a valores muy grandes de f .

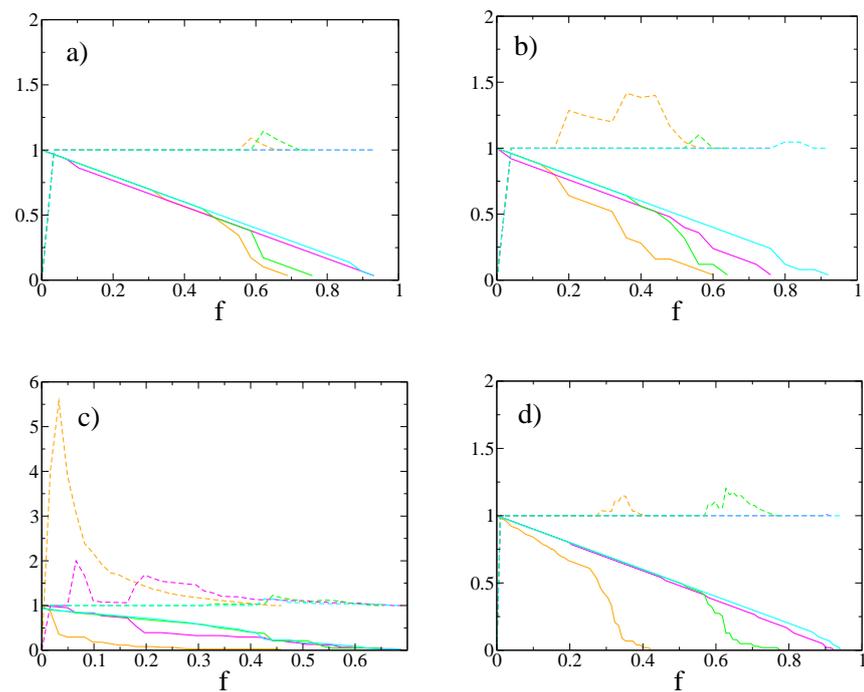


Figura A.2: La descripción de las curvas en esta figura es la misma que en la figura (A.1). A continuación se describe el nombre, el tamaño n y el número de links M de la red trófica asociada a cada gráfico en la figura: **a)** *Benguela Current*, $n = 29$, $M = 191$; **b)** *Bridge Brook Lake*, $n = 25$, $M = 103$; **c)** *British Grassland*, $n = 61$, $M = 97$; **d)** *Canton Creek*, $n = 102$, $M = 696$.

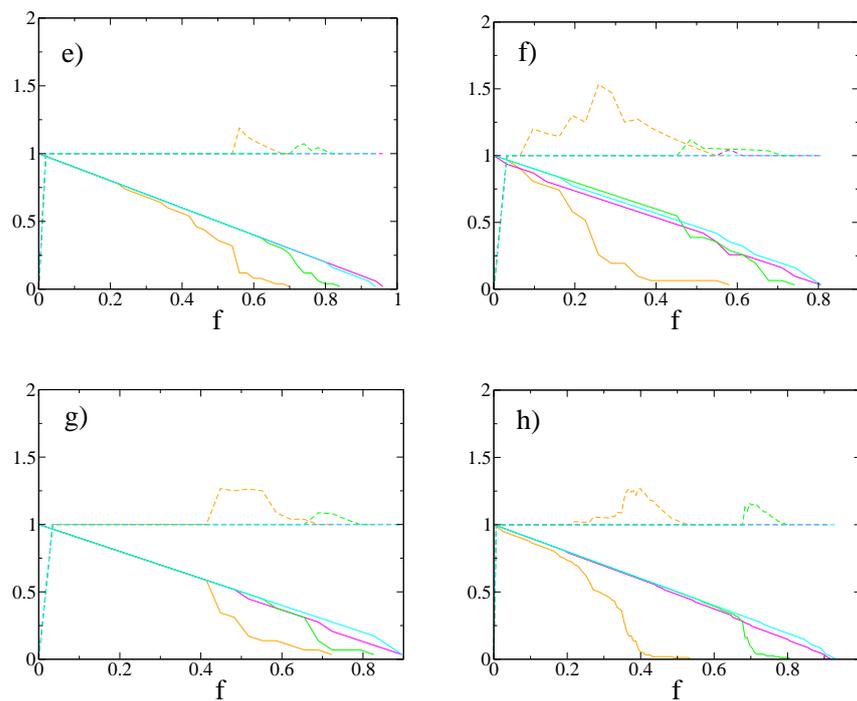


Figura A.3: La descripción de las curvas en esta figura es la misma que en la figura (A.1). A continuación se describe el nombre, el tamaño n y el número de links M de la red trófica asociada a cada gráfico en la figura: **e)** *Caribbean Reef*, $n = 50$, $M = 503$; **f)** *Chesapeake*, $n = 31$, $M = 67$; **g)** *Coachella Valley*, $n = 29$, $M = 221$; **h)** *El Verde Rainforest*, $n = 155$, $M = 1438$.

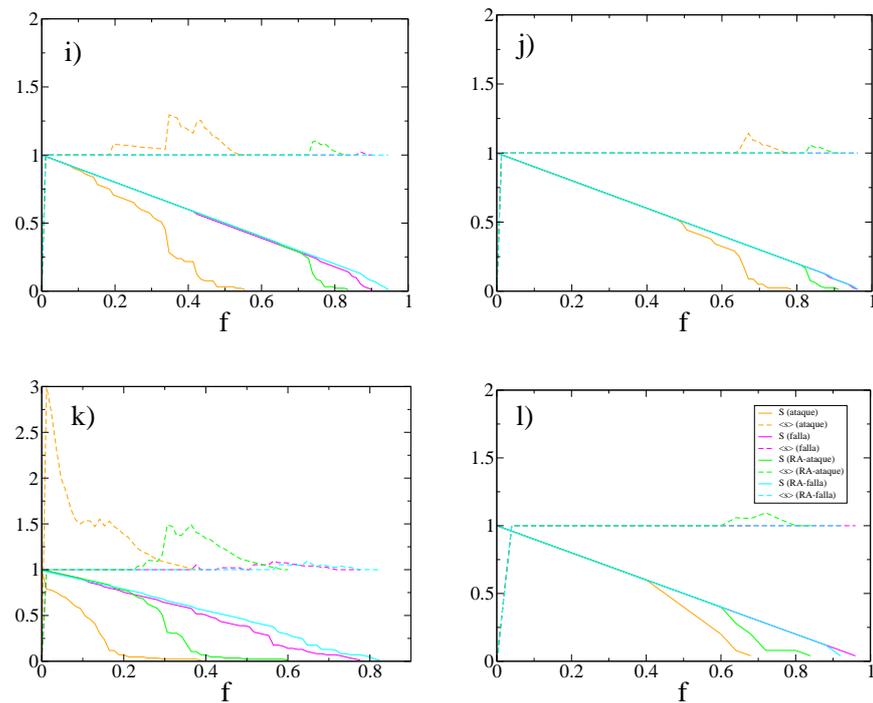


Figura A.4: La descripción de las curvas en esta figura es la misma que en la figura (A.1). A continuación se describe el nombre, el tamaño n y el número de links M de la red trófica asociada a cada gráfico en la figura: **i)** *Little Rock Lake*, $n = 92$, $M = 960$; **j)** *Northeast US Shelf*, $n = 79$, $M = 1371$; **k)** *Scotch Broom*, $n = 85$, $M = 219$; **l)** *Skipwith Pond*, $n = 25$, $M = 185$.

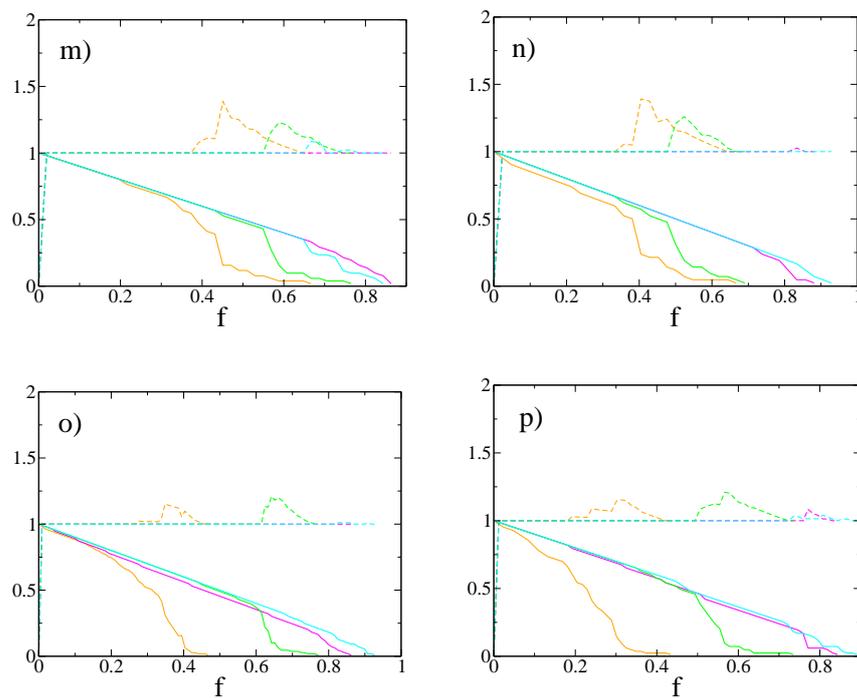


Figura A.5: La descripción de las curvas en esta figura es la misma que en la figura (A.1). A continuación se describe el nombre, el tamaño n y el número de links M de la red trófica asociada a cada gráfico en la figura: **m)** *St. Marks Seagrass*, $n = 51$, $M = 264$; **n)** *St. Martin Island*, $n = 42$, $M = 205$; **o)** *Stony Stream*, $n = 109$, $M = 827$; **p)** *Ythan Estuary (91)*, $n = 83$, $M = 390$.

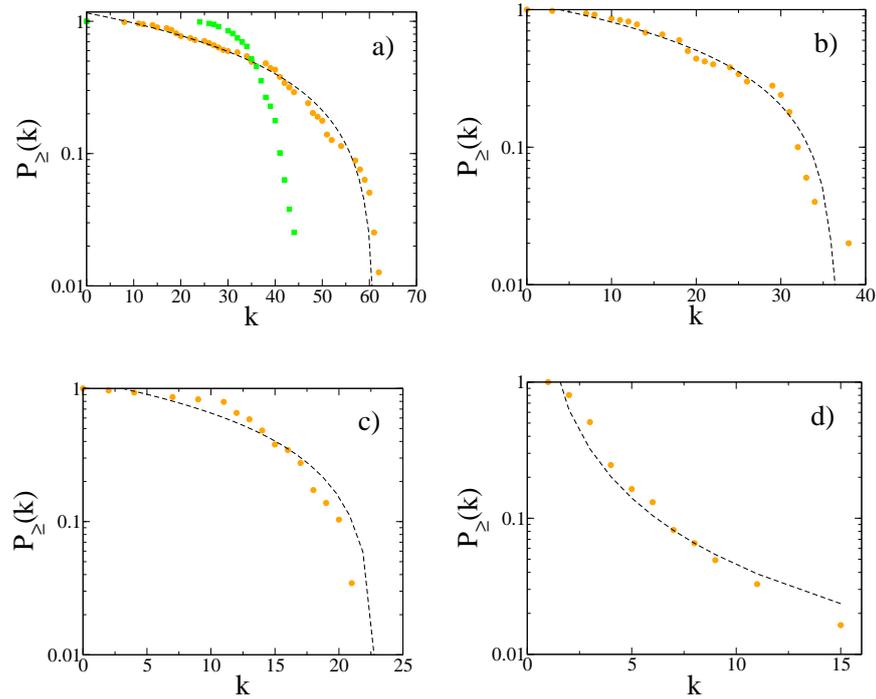


Figura A.6: Distribución de grados cumulativa complementaria $P_{\ge}(k)$ para diversas redes. **a)** Los círculos anaranjados corresponde a la red trófica *Northeast US Shelf*, mientras que los cuadrados verdes corresponden a una red aleatoria (distribución $P(k)$ tipo Poisson) de mismo tamaño y conectividad que la anterior. La línea discontinua es un ajuste que corresponde a una distribución $P(k)$ uniforme, es decir, un ajuste de la forma $P(k) = A + Bk$ con B pequeño. **b)** Red trófica *Caribbean Reef* (círculos anaranjados), ajuste correspondiente a una distribución uniforme (línea discontinua). **c)** Red trófica *Benguela Current* (círculos anaranjados), ajuste correspondiente a una distribución uniforme (línea discontinua). **d)** $P_{\ge}(k)$ para la red trófica *British Grassland* (círculos anaranjados). La línea discontinua corresponde a un ajuste tipo ley de potencias de la distribución de grados ($P(k) = Ak^{-\gamma}$). Este último gráfico es una reproducción del gráfico **a.** de la figura 2 en [65].

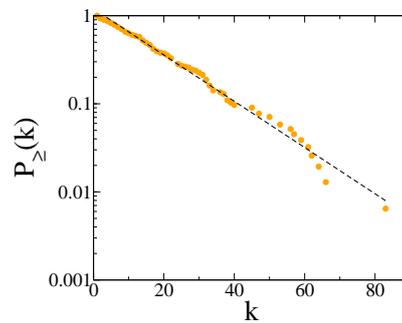


Figura A.7: Distribución de grados cumulativa complementaria $P_{\ge}(k)$ para la red trófica *El Verde Rainforest* (círculos anaranjados). La línea discontinua corresponde a un ajuste exponencial para la distribución de grados $P(k) = A \exp(-k/\kappa)$. Este gráfico es una reproducción del gráfico **e.** de la figura 2 en [65].

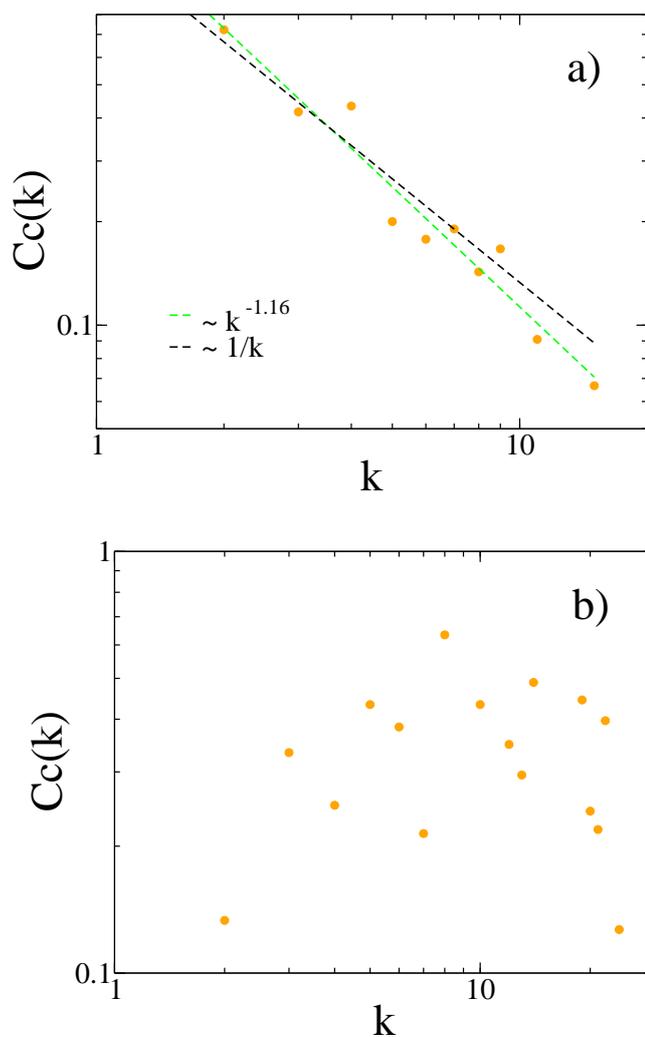


Figura A.8: Gráficos de $C_c(k)$ para las dos redes que Dunne *et. al.*[65] cataloga como redes con distribuciones de grados leyes de potencias, a saber las redes correspondientes a *British Grassland* y *Scotch Broom*. En **a)** se aprecia que *British Grassland* presenta un comportamiento $C_c(k) \sim 1/k$ indicativo de modularidad jerárquica, una propiedad asociada a algunas redes con distribuciones de grados libres de escala. Por otro lado, en **b)** se aprecia que *Scotch Broom* presenta un comportamiento irregular en $C_c(k)$.

Apéndice B

Sobre los datos obtenidos de la *home page* de A.G. Rossberg

Las redes generadas por el presente modelo sólo alcanzan tamaños no mayores a $n \simeq 200$ por una cuestión de costo computacional. Las redes tróficas resultan particularmente convenientes para comparar algunas propiedades de las redes generadas por el presente modelo con las de redes biológicas ya que presentan tamaños similares. Sin embargo, el uso de redes tróficas trae acarreado algunas inconveniencias. Según comenta J.A. Dunne [63],[64] (ver sección 2.8.1 para más detalles) hay una gran cantidad de datos recopilados sobre redes tróficas que no son adecuados para una correcta caracterización topológica de dichas redes tróficas por diversas razones. Una de las principales razones es que constituyen representaciones sobresimplificadas de los ecosistemas reales, por ejemplo, describiendo sólo una fracción de ellos. Otro problema principal reside en el nivel y tipo de resolución aplicado en el *agregado** de los diferentes *taxones*[†] dentro de las distintas *especies tróficas*[‡]. Por ejemplo, en la base de datos de redes tróficas *ECOWeB* (Ecologist Co-Operative Web Bank)[2] hay mas de 200 redes tróficas, y la tarea de decidir si una dada red está lo suficientemente resuelta como para ser estudiada topológicamente necesariamente requiere de la intervención de gente experta en el área. Dunne es una investigadora de amplia experiencia en el área y menciona que hay un pequeño conjunto de redes tróficas que poseen un aceptable nivel de resolución y cuyas redes están siendo utilizadas en análisis topológicos [64]. De hecho, ella y sus colaboradores lo utilizan en sus publicaciones [207],[65],[63],[64].

A.G. Rossberg posee un espacio en su *home page* [3] de donde puede

*Del ingles *aggregate*.

[†]Del ingles *taxon* (plural *taxa*). Un taxón es un grupo de organismos que es considerado una unidad por un taxonomista.

[‡]Una especie trófica es un grupo de taxones que comparten las mismas relaciones tróficas con el resto de las especies tróficas.

bajarse una base de datos con redes tróficas. Rossberg comenta que dicha base de datos fué cortesía de Dunne y N.D. Martínez*, y las redes que en ella aparecen son (en su gran mayoría) las mismas que aparecen en los trabajos de Dunne y Martinez. En particular utilizando los datos reportados en [65], pude comprobar que las redes corresponden a las originales (por ejemplo a través de comparar los valores de Cc).

En dicha base de datos las redes se encuentran disponibles en 4 grados de estandarización.

- R : Los datos crudos de los modelos empíricos de las redes tróficas sin ningún tipo de estandarización aplicada.
- C : Solamente la mayor componente conectada de la red en el nivel R .
- L : Se obtiene a partir de C por medio de agrupar especies trópicamente idénticas.
- S : Se obtiene a partir de R por medio de primero agrupar el nivel trófico más bajo (nivel basal[†]), y luego agrupar el resto las especies tróficamente idénticas.

Dunne *et. al.* en [65, 64] utiliza estas redes en el nivel de estandarización L . En este trabajo también se utiliza el nivel de estandarización L .

Las redes tróficas en cuestión y las correspondientes fuentes están citadas en un trabajo de Rossberg [179] y son las siguientes: *Benguela Current* [208], *Bridge Brook Lake* [88], *British Grassland* [121], *Canton Creek* [194], *Caribbean Reef* [157], *Chesapeake Bay* [21], *Coachella Valley* [173], *El Verde Rainforest* [177], *Little Rock Lake* [119], *Northeast US Shelf* [118], *Scotch Broom* [130], *Skipwith Pond* [201], *St. Marks Seagrass* [47], *St. Martin Island* [80], *Stony Stream* [194], *Ythan Estuary 1* (1991) [86], *Ythan Estuary 2* (1996) [90].

Sólo por aclaración, las siguientes redes tróficas aparecen en los anteriormente mencionados trabajos de Dunne *et. al.* [65, 64] y no aparecen en la base de datos Rossberg: *Lake Tahoe*, *Mirror Lake*. Y la visceversa: *Benguela Current*, *Caribbean Reef*, *Northeast US Shelf* aparecen en la base de datos de Rossberg, pero no en los mencionados trabajos de Dunne *et. al.*. Esta aclaración la realizo por la siguiente razón. En algunos casos [‡] los datos proveídos en Dunne *et. al.* me son útiles y lo cito sin necesidad de calcularlos a partir de las redes disponibles en la base de datos de Rossberg. En otros casos [§] utilizo las redes disponibles en la base de datos de Rossberg para calcular magnitudes de interés.

*N.D. Martínez es también un investigador de amplia experiencia en el área.

[†]Constituido por especies que no necesitan alimentarse de otras especies.

[‡]Como en el caso de $L(n)$, o de $Cc(n)$ calculado según la ecuación (2.25), o de $\langle k \rangle(n)$.

[§]Como en el caso de $Cc(n)$ calculado según la ecuación (2.24) o el coeficiente de correlación de Pearson para el grado de los nodos $r(n)$.

Como nota al margen, he hecho algunos cálculos utilizando las redes tróficas que se pueden obtener de la página del Pajek [4] pero he obtenido resultados sospechosos que me hicieron dudar de la confiabilidad de los datos.

Bibliografía

- [1] <http://www.worldwidewebsite.com/>.
- [2] <http://dspace1.rockefeller.edu/handle/10209/306>.
- [3] <http://axel.rossberg.net/datatable/datatable.html>.
- [4] <http://vlado.fmf.uni-lj.si/pub/networks/data/bio/foodweb/foodweb.htm>.
- [5] S. Achard, R. Salvador, B. Whitcher, J. Suckling, E. Bullmore, A Resilient, Low-Frequency, Small-World Human Brain Functional Network with Highly Connected Association Cortical Hubs, *The Journal of Neuroscience*, **26**, 1, págs. 63–72, 2006.
- [6] L. Adamic, The Small World Web, *Research and Advanced Technology for Digital Libraries* (eds. S. Abiteboul, A.-M. Vercoustre), tomo 1696 de *Lecture Notes in Computer Science*, pág. 852, Springer, 1999.
- [7] L. A. Adamic, B. A. Huberman, A.-L. Barabási, R. Albert, H. Jeong, G. Bianconi, Power-Law Distribution of the World Wide Web, *Science*, **287**, 5461, pág. 2115a, 2000.
- [8] L. A. Adamic, R. M. Lukose, A. R. Puniyani, B. A. Huberman, Search in power-law networks, *Physical Review E*, **64**, 4, pág. 046135, 2001.
- [9] W. Aiello, F. Chung, L. Lu, A random graph model for massive graphs, *Proceedings of the thirty-second annual ACM symposium on Theory of computing*, págs. 171–180, Portland, Oregon, USA, 2000.
- [10] R. Albert, A.-L. Barabási, Statistical mechanics of complex networks, *Reviews of Modern Physics*, **74**, 1, pág. 47, 2002.
- [11] R. Albert, H. Jeong, A.-L. Barabási, Internet: Diameter of the World-Wide Web, *Nature*, **401**, 6749, págs. 130–131, 1999.
- [12] R. Albert, H. Jeong, A.-L. Barabási, Error and attack tolerance of complex networks, *Nature*, **406**, 6794, págs. 378–382, 2000.

-
- [13] R. Albert, H. Jeong, A.-L. Barabási, correction: Error and attack tolerance of complex networks, *Nature*, **409**, 6819, pág. 542, 2001.
- [14] S. Allesina, D. Alonso, M. Pascual, A General Model for Food Web Structure, *Science*, **320**, 5876, págs. 658–661, 2008.
- [15] S. Allesina, M. Pascual, Network structure, predator–prey modules, and stability in large food webs, *Theoretical Ecology*, **1**, 1, págs. 55–64, 2008.
- [16] N. A. Alves, A. S. Martinez, Inferring topological features of proteins from amino acid residue networks, *Physica A*, **375**, 1, págs. 336–344, 2007.
- [17] L. A. N. Amaral, A. Scala, M. Barthélémy, H. E. Stanley, Classes of small-world networks, *Proceedings of the National Academy of Sciences of the United States of America*, **97**, 21, págs. 11149–11152, 2000.
- [18] A. Arenas, A. Díaz-Guilera, C. J. Pérez-Vicente, Synchronization Reveals Topological Scales in Complex Networks, *Physical Review Letters*, **96**, 11, pág. 114102, 2006.
- [19] A. Arenas, A. Fernández, S. Gómez, Analysis of the structure of complex networks at different resolution levels, *New Journal of Physics*, **10**, 5, pág. 053039, 2008.
- [20] M. Baiesi, S. S. Manna, Scale-free networks from a Hamiltonian dynamics, *Physical Review E*, **68**, 4, pág. 047103, 2003.
- [21] D. Baird, R. E. Ulanowicz, The Seasonal Dynamics of The Chesapeake Bay Ecosystem, *Ecological Monographs*, **59**, 4, pág. 329, 1989.
- [22] P. Bak, K. Sneppen, Punctuated equilibrium and criticality in a simple model of evolution, *Physical Review Letters*, **71**, 24, pág. 4083, 1993.
- [23] A.-L. Barabási, R. Albert, Emergence of Scaling in Random Networks, *Science*, **286**, 5439, págs. 509–512, 1999.
- [24] A.-L. Barabási, H. Jeong, Z. Néda, E. Ravasz, A. Schubert, T. Vicsek, Evolution of the social network of scientific collaborations, *Physica A*, **311**, 3-4, págs. 590–614, 2002.
- [25] A.-L. Barabási, Z. N. Oltvai, Network biology: understanding the cell’s functional organization, *Nature Reviews Genetics*, **5**, 2, págs. 101–113, 2004.
- [26] A. Barrat, Comment on “Small-world networks: Evidence for a crossover picture”, *arXiv:cond-mat/9903323*, 1999.

- [27] A. Barrat, M. Weigt, On the properties of small-world network models, *The European Physical Journal B*, **13**, 3, págs. 547–560, 2000.
- [28] M. Barthélémy, L. A. N. Amaral, Erratum: Small-World Networks: Evidence for a Crossover Picture [Phys. Rev. Lett. 82, 3180 (1999)], *Physical Review Letters*, **82**, 25, pág. 5180, 1999.
- [29] M. Barthélémy, L. A. N. Amaral, Small-World Networks: Evidence for a Crossover Picture, *Physical Review Letters*, **82**, 15, pág. 3180, 1999.
- [30] J. Bascompte, D. B. Stouffer, The assembly and disassembly of ecological networks, *Philosophical Transactions of the Royal Society B: Biological Sciences*, **364**, 1524, págs. 1781–1787, 2009.
- [31] P. S. Bearman, J. Moody, K. Stovel, Chains of affection: The structure of adolescent romantic and sexual networks, *American Journal of Sociology*, **110**, 1, pág. 44–91, 2004.
- [32] J. Berg, M. Lassig, A. Wagner, Structure and evolution of protein interaction networks: a statistical model for link dynamics and gene duplications, *BMC Evolutionary Biology*, **4**, 1, pág. 51, 2004.
- [33] G. Bianconi, A.-L. Barabási, Competition and multiscaling in evolving networks, *Europhysics Letters*, **54**, 4, págs. 436–442, 2001.
- [34] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D. Hwang, Complex networks: structure and dynamics, *Physical Reports*, **424**, pág. 175, 2006.
- [35] M. Brede, S. Sinha, Assortative mixing by degree makes a network more unstable, *arXiv:cond-mat/0507710*, 2005.
- [36] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, S. Rajagopalan, R. Stata, A. Tomkins, J. Wiener, Graph structure in the web, *Computer Networks*, **33**, 1-6, págs. 309–320, 2000.
- [37] G. Caldarelli, A. Capocci, P. D. L. Rios, M. A. Muñoz, Scale-Free Networks from Varying Vertex Intrinsic Fitness, *Physical Review Letters*, **89**, 25, pág. 258702, 2002.
- [38] D. S. Callaway, J. E. Hopcroft, J. M. Kleinberg, M. E. J. Newman, S. H. Strogatz, Are randomly grown graphs really random?, *Physical Review E*, **64**, 4, pág. 041902, 2001.
- [39] D. S. Callaway, M. E. J. Newman, S. H. Strogatz, D. J. Watts, Network Robustness and Fragility: Percolation on Random Graphs, *Physical Review Letters*, **85**, 25, pág. 5468, 2000.

- [40] J. Camacho, R. Guimerà, L. N. Amaral, Robust Patterns in Food Web Structure, *Physical Review Letters*, **88**, 22, 2002.
- [41] J. Camacho, D. Stouffer, L. Amaral, Quantitative analysis of the local structure of food webs, *Journal of Theoretical Biology*, **246**, 2, págs. 260–268, 2007.
- [42] R. Castro, J. W. Grossman, Famous trails to Paul Erdős, *The Mathematical Intelligencer*, **21**, 3, págs. 51–53, 1999.
- [43] M. Catanzaro, G. Caldarelli, L. Pietronero, Social network growth with assortative mixing, *Physica A*, **338**, 1-2, págs. 119–124, 2004.
- [44] M. Cattin, L. Bersier, C. Banasek-Richter, R. Baltensperger, J. Gabriel, Phylogenetic constraints and adaptation explain food-web structure, *Nature*, **427**, 6977, págs. 835–839, 2004.
- [45] X. Chen, J. E. Cohen, Global Stability, Local Stability and Permanence in Model Food Webs, *Journal of Theoretical Biology*, **212**, 2, págs. 223–235, 2001.
- [46] D. R. Chialvo, D. Fraiman, Optimal information-sharing in brain resting state networks, *arXiv:1011.1192*, 2010.
- [47] R. Christian, J. Luczkovich, Organizing and understanding a winter's seagrass foodweb network through effective trophic levels, *Ecological Modelling*, **117**, 1, págs. 99–124, 1999.
- [48] A. Clauset, C. Moore, M. E. J. Newman, Hierarchical structure and the prediction of missing links in networks, *Nature*, **453**, 7191, págs. 98–101, 2008.
- [49] J. E. Cohen, F. Briand, C. M. Newman, A Stochastic Theory of Community Food Webs III. Predicted and Observed Lengths of Food Chains, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **228**, 1252, págs. 317–353, 1986.
- [50] J. E. Cohen, C. M. Newman, A Stochastic Theory of Community Food Webs: I. Models and Aggregated Data, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **224**, 1237, págs. 421–448, 1985.
- [51] R. Cohen, K. Erez, D. ben-Avraham, S. Havlin, Resilience of the Internet to Random Breakdowns, *Physical Review Letters*, **85**, 21, pág. 4626, 2000.
- [52] R. Cohen, K. Erez, D. ben-Avraham, S. Havlin, Breakdown of the Internet under Intentional Attack, *Physical Review Letters*, **86**, 16, pág. 3682, 2001.

- [53] V. Colizza, A. Flammini, A. Maritan, A. Vespignani, Characterization and modeling of protein-protein interaction networks, *Physica A*, **352**, 1, págs. 1–27, 2005.
- [54] L. da F. Costa, F. A. Rodrigues, G. Travieso, P. R. V. Boas, Characterization of complex networks: A survey of measurements, *Advances in Physics*, **56**, 1, pág. 167, 2007.
- [55] G. F. Davis, M. Yoo, W. E. Baker, The small world of the American corporate elite, 1982-2001, *Strategic organization*, **1**, 3, pág. 301, 2003.
- [56] M. A. de Menezes, C. F. Moukarzel, T. J. P. Penna, First-order transition in small-world networks, *Europhysics Letters*, **50**, 5, págs. 574–579, 2000.
- [57] S. N. Dorogovtsev, A. V. Goltsev, J. F. F. Mendes, Critical phenomena in complex networks, *Reviews of Modern Physics*, **80**, 4, pág. 1275, 2008.
- [58] S. N. Dorogovtsev, J. F. F. Mendes, Evolution of networks, *Advances in Physics*, **51**, 4, pág. 1079, 2002.
- [59] S. N. Dorogovtsev, J. F. F. Mendes, *Evolution of networks: From biological nets to the Internet and WWW*, Oxford University Press, 2003.
- [60] S. N. Dorogovtsev, J. F. F. Mendes, A. N. Samukhin, Structure of Growing Networks with Preferential Linking, *Physical Review Letters*, **85**, 21, pág. 4633, 2000.
- [61] B. Drossel, Biological evolution and statistical physics, *Advances in Physics*, **50**, 2, págs. 209–295, 2001.
- [62] B. Drossel, Models for Food Webs, *Advances in Solid State Physics* (ed. B. Kramer), tomo 43, págs. 81–110, Springer, 2003.
- [63] J. A. Dunne, The network structure of food webs, *Ecological networks: Linking structure to dynamics in food webs* (eds. M. Pascual, J. A. Dunne), págs. 27–86, Oxford University Press, 2006.
- [64] J. A. Dunne, Food Webs, *Encyclopedia of Complexity and Systems Science* (ed. R. A. Meyers), págs. 3661–3682, Springer, New York, 2009.
- [65] J. A. Dunne, R. J. Williams, N. D. Martinez, 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*, **99**, 20, págs. 12917–12922, 2002.

- [66] H. Ebel, J. Davidsen, S. Bornholdt, Dynamics of social networks, *Complexity*, **8**, 2, págs. 24–27, 2002.
- [67] H. Ebel, L. Mielsch, S. Bornholdt, Scale-free topology of e-mail networks, *Physical Review E*, **66**, 3, pág. 035103, 2002.
- [68] V. M. Eguiluz, D. R. Chialvo, G. A. Cecchi, M. Baliki, A. V. Apkarian, Scale-Free Brain Functional Networks, *Physical Review Letters*, **94**, 1, pág. 018102, 2005.
- [69] E. Eisenberg, E. Y. Levanon, Preferential Attachment in the Protein Network Evolution, *Physical Review Letters*, **91**, 13, pág. 138701, 2003.
- [70] P. Erdős, A. Rényi, On random graphs I, *Publicationes Mathematicae Debrecen*, **6**, 26, págs. 290–297, 1959.
- [71] M. Faloutsos, P. Faloutsos, C. Faloutsos, On power-law relationships of the Internet topology, *SIGCOMM Computer Communication Review*, **29**, págs. 251–262, 1999.
- [72] I. J. Farkas, I. Derényi, A.-L. Barabási, T. Vicsek, Spectra of real-world” graphs: Beyond the semicircle law, *Physical Review E*, **64**, 2, pág. 026704, 2001.
- [73] L. Ferrarini, I. M. Veer, E. Baerends, M. van Tol, R. J. Renken, N. J. van der Wee, D. J. Veltman, A. Aleman, F. G. Zitman, B. W. Penninx, M. A. van Buchem, J. H. Reiber, S. A. Rombouts, J. Miles, Hierarchical functional modularity in the resting-state human brain, *Human Brain Mapping*, **30**, 7, págs. 2220–2231, 2009.
- [74] S. Fortunato, Community detection in graphs, *Physics Reports*, **486**, 3-5, págs. 75–174, 2010.
- [75] D. V. Foster, J. G. Foster, P. Grassberger, M. Paczuski, Clustering Drives Assortativity and Community Structure in Ensembles of Networks, *arXiv:1012.2384*, 2010.
- [76] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. V. Essen, M. E. Raichle, The human brain is intrinsically organized into dynamic, anticorrelated functional networks, *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 27, págs. 9673–9678, 2005.
- [77] M. Fukunaga, S. Horovitz, P. Vangelder, J. Dezwart, J. Jansma, V. Ikonomidou, R. Chu, R. Deckers, D. Leopold, J. Duyn, Large-amplitude, spatially correlated fluctuations in BOLD fMRI signals during extended rest and early sleep stages, *Magnetic Resonance Imaging*, **24**, 8, págs. 979–992, 2006.

-
- [78] M. R. Gardner, W. R. Ashby, Connectance of Large Dynamic (Cybernetic) Systems: Critical Values for Stability, *Nature*, **228**, 5273, pág. 784, 1970.
- [79] K. Goh, B. Kahng, D. Kim, Spectra and eigenvectors of scale-free networks, *Physical Review E*, **64**, 5, pág. 051903, 2001.
- [80] L. Goldwasser, J. Roughgarden, Construction and analysis of a large Caribbean food web, *Ecology*, **74**, págs. 1216–1233, 1993.
- [81] I. J. Gomez Portillo, P. M. Gleiser, An Adaptive Complex Network Model for Brain Functional Networks, *PLoS ONE*, **4**, 9, pág. e6863, 2009.
- [82] I. J. Gomez Portillo, P. M. Gleiser, A local adaptive algorithm for emerging scale-free hierarchical networks, *Journal of Physics: Conference Series*, **246**, pág. 012012, 2010.
- [83] R. Govindan, H. Tangmunarunkit, Heuristics for Internet Map Discovery, *Proceedings of the Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies*, págs. 1371–1380, Proc. IEEE, Tel Aviv, Israel, 2000.
- [84] C. Guill, A Model of Large-Scale Evolution of Complex Food Webs, *Mathematical Modelling of Natural Phenomena*, **5**, 6, pág. 20, 2010.
- [85] C. Guill, B. Drossel, Emergence of complexity in evolving niche-model food webs, *Journal of Theoretical Biology*, **251**, 1, págs. 108–120, 2008.
- [86] S. J. Hall, D. Raffaelli, Food-web patterns: lessons from a species-rich web, *The Journal of Animal Ecology*, **60**, 3, págs. 823–841, 1991.
- [87] H. M. Hastings, The May-Wigner stability theorem for connected matrices, *Bulletin (New Series) of the American Mathematical Society*, **7**, 2, págs. 387–388, 1982.
- [88] K. Havens, Scale and Structure in Natural Food Webs, *Science*, **257**, 5073, págs. 1107–1109, 1992.
- [89] B. A. Huberman, L. A. Adamic, Internet: Growth dynamics of the World-Wide Web, *Nature*, **401**, 6749, pág. 131, 1999.
- [90] M. Huxham, S. Beaney, D. Raffaelli, Do Parasites Reduce the Chances of Triangulation in a Real Food Web?, *Oikos*, **76**, 2, págs. 284–300, 1996.
- [91] R. F. i Cancho, C. Janssen, R. V. Solé, Topology of technology graphs: Small world patterns in electronic circuits, *Physical Review E*, **64**, 4, pág. 046119, 2001.

- [92] R. F. i Cancho, R. V. Solé, The small world of human language, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **268**, 1482, págs. 2261–2265, 2001.
- [93] T. C. Ings, J. M. Montoya, J. Bascompte, N. Blüthgen, L. Brown, C. F. Dormann, F. Edwards, D. Figuerola, U. Jacob, J. I. Jones, R. B. Lauridsen, M. E. Ledger, H. M. Lewis, J. M. Olesen, F. F. van Veen, P. H. Warren, G. Woodward, Review: Ecological networks - beyond food webs, *Journal of Animal Ecology*, **78**, 1, págs. 253–269, 2009.
- [94] A. R. Ives, S. R. Carpenter, Stability and diversity of ecosystems, *Science*, **317**, 5834, pág. 58, 2007.
- [95] H. Jeong, S. P. Mason, A.-L. Barabási, Z. N. Oltvai, Lethality and centrality in protein networks, *Nature*, **411**, 6833, págs. 41–42, 2001.
- [96] H. Jeong, Z. Néda, A.-L. Barabási, Measuring preferential attachment in evolving networks, *Europhysics Letters*, **61**, 4, págs. 567–572, 2003.
- [97] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, A.-L. Barabási, The large-scale organization of metabolic networks, *Nature*, **407**, 6804, págs. 651–654, 2000.
- [98] V. K. Jirsa, M. Ding, Will a Large Complex System with Time Delays Be Stable?, *Physical Review Letters*, **93**, 7, pág. 070602, 2004.
- [99] S. Johnson, J. J. Torres, J. Marro, M. A. Muñoz, Entropic Origin of Disassortativity in Complex Networks, *Physical Review Letters*, **104**, 10, pág. 108702, 2010.
- [100] F. Jordan, I. Scheuring, Searching for keystones in ecological networks, *Oikos*, **99**, 3, págs. 607–612, 2002.
- [101] B. Kartascheff, C. Guill, B. Drossel, Positive complexity-stability relations in food web models without foraging adaptation, *Journal of Theoretical Biology*, **259**, 1, págs. 12–23, 2009.
- [102] T. H. Keitt, P. A. Marquet, The Introduced Hawaiian Avifauna Reconsidered: Evidence for Self-Organized Criticality?, *Journal of Theoretical Biology*, **182**, 2, págs. 161–167, 1996.
- [103] J. Kim, P. L. Krapivsky, B. Kahng, S. Redner, Infinite-order percolation and giant fluctuations in a protein interaction network, *Physical Review E*, **66**, 5, pág. 055101, 2002.

- [104] M. G. Kitzbichler, M. L. Smith, S. R. Christensen, E. Bullmore, Broadband Criticality of Human Brain Network Synchronization, *PLoS Computational Biology*, **5**, 3, 2009.
- [105] K. Klemm, V. M. Eguíluz, Growing scale-free networks with small-world behavior, *Physical Review E*, **65**, 5, pág. 057102, 2002.
- [106] D. E. Knuth, *The Stanford GraphBase: a platform for combinatorial computing*, ACM, New York, NY, USA, 1993.
- [107] M. Kondoh, Foraging Adaptation and the Relationship Between Food-Web Complexity and Stability, *Science*, **299**, 5611, págs. 1388–1391, 2003.
- [108] M. Kondoh, Does foraging adaptation create the positive complexity-stability relationship in realistic food-web structure?, *Journal of Theoretical Biology*, **238**, 3, págs. 646–651, 2006.
- [109] P. L. Krapivsky, S. Redner, Organization of growing random networks, *Physical Review E*, **63**, 6, pág. 066123, 2001.
- [110] P. L. Krapivsky, S. Redner, F. Leyvraz, Connectivity of Growing Random Networks, *Physical Review Letters*, **85**, 21, pág. 4629, 2000.
- [111] R. Kumar, P. Raghavan, S. Rajagopalan, A. Tomkins, Extracting Large-Scale Knowledge Bases from the Web, *Proceedings of the 25th International Conference on Very Large Data Bases*, págs. 639–650, Morgan Kaufmann, Edinburg, Scotland, United Kingdom, 1999.
- [112] V. Kulin, J. B. Pereira-Leal, C. A. Ouzounis, Functional Evolution of the Yeast Protein Interaction Network, *Molecular Biology and Evolution*, **21**, 7, págs. 1171–1176, 2004.
- [113] K. D. Lafferty, S. Allesina, M. Arim, C. J. Briggs, G. D. Leo, A. P. Dobson, J. A. Dunne, P. T. J. Johnson, A. M. Kuris, D. J. Marcogliese, N. D. Martinez, J. Memmott, P. A. Marquet, J. P. McLaughlin, E. A. Mordecai, M. Pascual, R. Poulin, D. W. Thieltges, Parasites in food webs: the ultimate missing links, *Ecology Letters*, **11**, 6, págs. 533–546, 2008.
- [114] K. D. Lafferty, A. P. Dobson, A. M. Kuris, Parasites dominate food web links, *Proceedings of the National Academy of Sciences*, **103**, 30, págs. 11211–11216, 2006.
- [115] S. Li, C. M. Armstrong, N. Bertin, H. Ge, S. Milstein, M. Boxem, P. O. Vidalain, J. D. Han, A. Chesneau, T. Hao, *et al.*, A map of the interactome network of the metazoan *C. elegans*, *Science*, **303**, 5657, pág. 540, 2004.

- [116] S. Light, P. Kraulis, A. Elofsson, Preferential attachment in the evolution of metabolic networks, *BMC Genomics*, **6**, 1, pág. 159, 2005.
- [117] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, Y. Aberg, The web of human sexual contacts, *Nature*, **411**, 6840, págs. 907–908, 2001.
- [118] J. Link, Does food web theory work for marine ecosystems?, *Marine Ecology Progress Series*, **230**, págs. 1–9, 2002.
- [119] N. D. Martinez, Artifacts or Attributes? Effects of Resolution on the Little Rock Lake Food Web, *Ecological Monographs*, **61**, 4, pág. 367, 1991.
- [120] N. D. Martinez, Constant Connectance in Community Food Webs, *The American Naturalist*, **139**, 6, págs. 1208–1218, 1992.
- [121] N. D. Martinez, B. A. Hawkins, H. A. Dawah, B. P. Feifarek, Effects of sampling effort on characterization of food-web structure, *Ecology*, **80**, 3, págs. 1044–1055, 1999.
- [122] S. Maslov, K. Sneppen, Specificity and Stability in Topology of Protein Networks, *Science*, **296**, 5569, págs. 910–913, 2002.
- [123] S. Maslov, K. Sneppen, A. Zaliznyak, Detection of topological patterns in complex networks: correlation profile of the internet, *Physica A*, **333**, págs. 529–540, 2004.
- [124] R. Matei, A. Iamnitchi, P. Foster, Mapping the Gnutella network, *Internet Computing, IEEE*, **6**, 1, págs. 50–57, 2002.
- [125] R. M. May, Will a Large Complex System be Stable?, *Nature*, **238**, 5364, págs. 413–414, 1972.
- [126] K. McCann, A. Hastings, G. R. Huxel, Weak trophic interactions and the balance of nature, *Nature*, **395**, 6704, págs. 794–798, 1998.
- [127] K. S. McCann, The diversity-stability debate, *Nature*, **405**, 6783, págs. 228–233, 2000.
- [128] A. J. McKane, B. Drossel, Models of food web evolution, *Ecological networks: Linking structure to dynamics in food webs* (eds. M. Pascual, J. Dunne), págs. 223–243, Oxford University Press, 2005.
- [129] C. J. Melián, J. Bascompte, Food Web Cohesion, *Ecology*, **85**, 2, págs. 352–358, 2004.

- [130] J. Memmott, N. Martinez, J. Cohen, Predators, parasitoids and pathogens: species richness, trophic generality and body sizes in a natural food web, *Journal of Animal Ecology*, **69**, 1, págs. 1–15, 2000.
- [131] A. Mislove, H. S. Koppula, K. P. Gummadi, P. Druschel, B. Bhattacharjee, Growth of the flickr social network, *Proceedings of the first workshop on Online social networks*, págs. 25–30, ACM, Seattle, Washinton, USA, 2008.
- [132] J. M. Montoya, S. L. Pimm, R. V. Solé, Ecological networks and their fragility, *Nature*, **442**, 7100, págs. 259–264, 2006.
- [133] J. M. Montoya, R. V. Solé, Small World Patterns in Food Webs, *Journal of Theoretical Biology*, **214**, 3, págs. 405–412, 2002.
- [134] J. M. Montoya, R. V. Solé, Topological properties of food webs: from real data to community assembly models, *Oikos*, **102**, 3, págs. 614–622, 2003.
- [135] J. C. Nacher, N. Ueda, M. Kanehisa, T. Akutsu, Flexible construction of hierarchical scale-free networks with general exponent, *Physical Review E*, **71**, 3, págs. 036132, 2005.
- [136] A. Neutel, J. A. P. Heesterbeek, P. C. de Ruiter, Stability in Real Food Webs: Weak Links in Long Loops, *Science*, **296**, 5570, págs. 1120–1123, 2002.
- [137] M. Newman, The structure and function of networks, *Computer Physics Communications*, **147**, 1-2, págs. 40–45, 2002.
- [138] M. E. J. Newman, Self-Organized Criticality, Evolution and the Fossil Extinction Record, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **263**, 1376, págs. 1605–1610, 1996.
- [139] M. E. J. Newman, A Model of Mass Extinction, *Journal of Theoretical Biology*, **189**, 3, págs. 235–252, 1997.
- [140] M. E. J. Newman, Clustering and preferential attachment in growing networks, *Physical Review E*, **64**, 2, págs. 025102, 2001.
- [141] M. E. J. Newman, Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality, *Physical Review E*, **64**, 1, págs. 016132, 2001.
- [142] M. E. J. Newman, Scientific collaboration networks. I. Network construction and fundamental results, *Physical Review E*, **64**, 1, 2001.

- [143] M. E. J. Newman, The structure of scientific collaboration networks, *Proceedings of the National Academy of Sciences of the United States of America*, **98**, 2, págs. 404–409, 2001.
- [144] M. E. J. Newman, Assortative Mixing in Networks, *Physical Review Letters*, **89**, 20, pág. 208701, 2002.
- [145] M. E. J. Newman, Mixing patterns in networks, *Physical Review E*, **67**, 2, pág. 026126, 2003.
- [146] M. E. J. Newman, The Structure and Function of Complex Networks, *SIAM Review*, **45**, 2, págs. 167–256, 2003.
- [147] M. E. J. Newman, S. Forrest, J. Balthrop, Email networks and the spread of computer viruses, *Physical Review E*, **66**, 3, pág. 035101, 2002.
- [148] M. E. J. Newman, M. Girvan, Mixing Patterns and Community Structure in Networks, *Statistical Mechanics of Complex Networks*, págs. 66–87, 2003.
- [149] M. E. J. Newman, C. Moore, D. J. Watts, Mean-Field Solution of the Small-World Network Model, *Physical Review Letters*, **84**, 14, pág. 3201, 2000.
- [150] M. E. J. Newman, R. G. Palmer, Models of Extinction: A Review, *arXiv:adap-org/9908002*, 1999.
- [151] M. E. J. Newman, J. Park, Why social networks are different from other types of networks, *Physical Review E*, **68**, 3, pág. 036122, 2003.
- [152] M. E. J. Newman, B. W. Roberts, Mass Extinction: Evolution and the Effects of External Influences on Unfit Species, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **260**, 1357, págs. 31–37, 1995.
- [153] M. E. J. Newman, S. H. Strogatz, D. J. Watts, Random graphs with arbitrary degree distributions and their applications, *Physical Review E*, **64**, 2, pág. 026118, 2001.
- [154] M. E. J. Newman, D. J. Watts, Renormalization group analysis of the small-world network model, *Physics Letters A*, **263**, 4-6, págs. 341–346, 1999.
- [155] T. Ohgushi, Herbivore-induced indirect interaction webs on terrestrial plants: the importance of non-trophic, indirect, and facilitative interactions, *Entomologia Experimentalis et Applicata*, **128**, 1, págs. 217–229, 2008.

- [156] J. M. Olesen, J. Bascompte, H. Elberling, P. Jordano, Temporal Dynamics in a Pollination Network, *Ecology*, **89**, 6, págs. 1573–1582, 2008.
- [157] S. Opitz, *Trophic interactions in Caribbean coral reefs*, Informe Técnico 43, ICLARM, 1996.
- [158] R. Pan, S. Sinha, Modular networks with hierarchical organization: The dynamical implications of complex structure, *Pramana*, **71**, 2, págs. 331–340, 2008.
- [159] R. K. Pan, S. Sinha, Modular networks emerge from multiconstraint optimization, *Physical Review E*, **76**, 4, pág. 045103, 2007.
- [160] J. Park, M. E. J. Newman, Origin of degree correlations in the Internet and other networks, *Physical Review E*, **68**, 2, pág. 026112, 2003.
- [161] R. Pastor-Satorras, E. Smith, R. V. Solé, Evolving protein interaction networks through gene duplication, *Journal of Theoretical biology*, **222**, 2, págs. 199–210, 2003.
- [162] R. Pastor-Satorras, A. Vespignani, Epidemic Spreading in Scale-Free Networks, *Physical Review Letters*, **86**, 14, pág. 3200, 2001.
- [163] R. Pastor-Satorras, A. Vázquez, A. Vespignani, Dynamical and Correlation Properties of the Internet, *Physical Review Letters*, **87**, 25, pág. 258701, 2001.
- [164] J. I. Perotti, O. V. Billoni, F. A. Tamarit, S. Cannas, Stability as a natural selection mechanism on interacting networks, *Papers in Physics*, **2**, 0, 2010.
- [165] J. I. Perotti, O. V. Billoni, F. A. Tamarit, D. R. Chialvo, S. A. Cannas, Emergent Self-Organized Complex Network Topology out of Stability Constraints, *Physical Review Letters*, **103**, 10, pág. 108701, 2009.
- [166] S. L. Pimm, The complexity and stability of ecosystems, *Nature*, **307**, 5949, págs. 321–326, 1984.
- [167] S. L. Pimm, *The balance of nature?: Ecological issues in the conservation of species and communities*, University of Chicago Press, 1991.
- [168] M. Piraveenan, M. Prokopenko, A. Y. Zomaya, Local assortativeness in scale-free networks, *Europhysics Letters*, **84**, 2, pág. 28002, 2008.

- [169] M. Piraveenan, M. Prokopenko, A. Y. Zomaya, Local assortativity and growth of Internet, *The European Physical Journal B*, **70**, págs. 275–285, 2009.
- [170] M. Piraveenan, M. Prokopenko, A. Y. Zomaya, Assortativity in cyber-physical networks, *Preprint at The First Australasian Workshop on Computation in Cyber-Physical Systems (CompCPS-2010)*, Sydney, Australia, 2010.
- [171] M. Piraveenan, M. Prokopenko, A. Y. Zomaya, Local assortativeness in scale-free networks – addendum, *Europhysics letters*, **89**, 4, pág. 49901, 2010.
- [172] D. Plenz, D. R. Chialvo, Scaling properties of neuronal avalanches are consistent with critical dynamics, *arXiv:0912.5369*, 2009.
- [173] G. A. Polis, Complex Trophic Interactions in Deserts: An Empirical Critique of Food-Web Theory, *The American Naturalist*, **138**, 1, pág. 123, 1991.
- [174] P. Pons, Post-Processing Hierarchical Community Structures: Quality Improvements and Multi-scale View, *arXiv:cs/0608050*, 2006.
- [175] E. Ravasz, A.-L. Barabási, Hierarchical organization in complex networks, *Physical Review E*, **67**, 2, pág. 026112, 2003.
- [176] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, A.-L. Barabási, Hierarchical Organization of Modularity in Metabolic Networks, *Science*, **297**, 5586, págs. 1551–1555, 2002.
- [177] D. P. Reagan, R. B. Waide, *The food web of a tropical rain forest*, University of Chicago Press, 1996.
- [178] S. Redner, How popular is your paper? An empirical study of the citation distribution, *The European Physical Journal B*, **4**, 2, págs. 131–134, 1998.
- [179] A. Rossberg, H. Matsuda, T. Amemiya, K. Itoh, Food webs: Experts consuming families of experts, *Journal of Theoretical Biology*, **241**, 3, págs. 552–563, 2006.
- [180] M. Sales-Pardo, R. Guimerà, A. A. Moreira, L. A. N. Amaral, Extracting the hierarchical organization of complex systems, *Proceedings of the National Academy of Sciences*, **104**, 39, págs. 15224–15229, 2007.
- [181] P. Sen, S. Dasgupta, A. Chatterjee, P. A. Sreeram, G. Mukherjee, S. S. Manna, Small-world properties of the Indian railway network, *Physical Review E*, **67**, 3, pág. 036106, 2003.

- [182] S. Sinha, Complexity vs. stability in small-world networks, *Physica A*, **346**, 1-2, págs. 147–153, 2005.
- [183] S. Sinha, S. Sinha, Evidence of universality for the May-Wigner stability theorem for random networks with local dynamics, *Physical Review E*, **71**, 2, pág. 020902, 2005.
- [184] S. Milgram, The small world problem, *Psychology Today*, **2**, 1, págs. 60–67, 1967.
- [185] S. N. Soffer, A. Vázquez, Network clustering coefficient without degree-correlation biases, *Physical Review E*, **71**, 5, pág. 057101, 2005.
- [186] R. V. Solé, D. Alonso, A. McKane, Scaling in a network model of a multispecies ecosystem, *Physica A*, **286**, 1-2, págs. 337–344, 2000.
- [187] R. V. Solé, J. Bascompte, Are Critical Phenomena Relevant to Large-Scale Evolution?, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **263**, 1367, págs. 161–168, 1996.
- [188] R. V. Solé, S. C. Manrubia, Extinction and self-organized criticality in a model of large-scale evolution, *Physical Review E*, **54**, 1, pág. R42, 1996.
- [189] R. V. Solé, M. Montoya, Complexity and fragility in ecological networks, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **268**, 1480, págs. 2039–2045, 2001.
- [190] R. V. Solé, R. Pastor-Satorras, E. Smith, T. B. Kepler, A model of large-scale proteome evolution, *Advances in Complex Systems*, **5**, 1, pág. 43–54, 2008.
- [191] H. J. Sommers, A. Crisanti, H. Sompolinsky, Y. Stein, Spectrum of Large Random Asymmetric Matrices, *Physical Review Letters*, **60**, 19, pág. 1895, 1988.
- [192] R. Tanaka, T. Yi, J. Doyle, Some protein interaction data do not exhibit power law statistics, *FEBS Letters*, **579**, 23, págs. 5140–5144, 2005.
- [193] S. A. Teichmann, M. M. Babu, Gene regulatory network growth by duplication, *Nature Genetics*, **36**, 5, págs. 492–496, 2004.
- [194] Townsend, Thompson, McIntosh, Kilroy, Edwards, Scarsbrook, Disturbance, resource supply, and food-web architecture in streams, *Ecology Letters*, **1**, 3, págs. 200–209, 1998.
- [195] S. Valverde, R. F. Cancho, R. V. Solé, Scale-free networks from optimal design, *Europhysics Letters*, **60**, 4, págs. 512–517, 2002.

- [196] E. A. Variano, J. H. McCoy, H. Lipson, Networks, Dynamics, and Modularity, *Physical Review Letters*, **92**, 18, pág. 188701, 2004.
- [197] A. Vázquez, R. Pastor-Satorras, A. Vespignani, Large-scale topological and dynamical properties of the Internet, *Physical Review E*, **65**, 6, pág. 066130, 2002.
- [198] J. L. Vincent, G. H. Patel, M. D. Fox, A. Z. Snyder, J. T. Baker, D. C. V. Essen, J. M. Zempel, L. H. Snyder, M. Corbetta, M. E. Raichle, Intrinsic functional architecture in the anaesthetized monkey brain, *Nature*, **447**, 7140, págs. 83–86, 2007.
- [199] A. Vázquez, Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations, *Physical Review E*, **67**, 5, pág. 056104, 2003.
- [200] A. Wagner, D. A. Fell, The small world inside large metabolic networks, *Proceedings of the Royal Society of London. Series B: Biological Sciences*, **268**, 1478, págs. 1803–1810, 2001.
- [201] P. H. Warren, Spatial and temporal variation in the structure of a freshwater food web, *Oikos*, **55**, págs. 299–311, 1989.
- [202] P. H. Warren, Variation in Food-Web Structure: The Determinants of Connectance, *The American Naturalist*, **136**, 5, págs. 689–700, 1990.
- [203] S. Wasserman, K. Faust, *Social network analysis: Methods and applications*, Cambridge University Press, 1994.
- [204] D. J. Watts, S. H. Strogatz, Collective dynamics of ‘small-world’ networks, *Nature*, **393**, 6684, págs. 440–442, 1998.
- [205] E. Weiher, P. Keddy, *Ecological assembly rules: Perspectives, advances, retreats*, Cambridge University Press, 2001.
- [206] E. P. Wigner, Random matrices in physics, *SIAM Review*, **9**, 1, págs. 1–23, 1967.
- [207] R. J. Williams, N. D. Martinez, Simple rules yield complex food webs, *Nature*, **404**, 6774, págs. 180–183, 2000.
- [208] P. Yodzis, Local trophodynamics and the interaction of marine mammals and fisheries in the Benguela ecosystem, *Journal of Animal Ecology*, **67**, 4, págs. 635–658, 1998.
- [209] S. Yook, Z. N. Oltvai, A.-L. Barabási, Functional and topological characterization of protein interaction networks, *PROTEOMICS*, **4**, 4, págs. 928–942, 2004.

-
- [210] M. Ángeles Serrano, M. Boguñá, Tuning clustering in random networks with arbitrary degree distributions, *Physical Review E*, **72**, 3, pág. 036133, 2005.