



UNIVERSIDAD DE BUENOS AIRES
Facultad de Ciencias Exactas y Naturales
Departamento de Matemática

Desempeño asintótico de algoritmos secuenciales en grafos aleatorios

Tesis presentada para optar al título de Doctor de la
Universidad de Buenos Aires en el área Ciencias Matemáticas

Lic. Manuel Sáenz

Director de tesis: Dr. Matthieu T. Jonckheere
Consejero de estudios: Dr. Pablo Groisman

Fecha de defensa: 19 de Noviembre de 2019
Buenos Aires, 2019

Desempeño asintótico de algoritmos secuenciales en grafos aleatorios

Resumen

En esta tesis estudiamos la evolución, convergencia y optimalidad de algoritmos de búsqueda en grafos aleatorios a través de ejemplos y aplicamos estos resultados al estudio de cotas de desempeño para protocolos de comunicación en redes inalámbricas. En particular, estudiamos los siguientes problemas:

- La determinación de la optimalidad asintótica del algoritmo grado-egoísta en una amplia clase de Modelos de Configuraciones y el computo de sus respectivos cocientes de independencia. Aunque el problema de hallar conjuntos independientes de tamaño máximo es en el caso general una tarea NP-difícil, mostramos que en cierta clase de grafos aleatorios el algoritmo grado egoísta construye conjuntos independientes asintóticamente máximos en tiempo polinomial. Además, utilizamos estos resultados para caracterizar el cociente de independencia de grafos de Erdős-Rényi (de grado medio menor a e) y para encontrar procedimientos numéricos que permitan calcular cotas superiores en Modelos de Configuraciones generales.
- La evaluación del desempeño de variantes locales del algoritmo egoísta aplicados al problema de aproximar la capacidad de redes WiFi y su posible impacto en la justicia de la distribución de conexiones. Por otro lado, aplicamos estos resultados al estudio de una red de comunicación empírica: el sistema de antenas de celulares de Montevideo.
- El cálculo de los tiempos de convergencia para algoritmos de respuesta óptima en problemas de optimización distribuida donde las acciones se estructuran como grafos aleatorios raros. Usando estos resultados probamos, por medio de un lema de optimalidad, cotas inferiores generales para los tiempos de corrida de algoritmos locales de búsqueda.

La mayoría de nuestros resultados son aplicados tanto a grafos de Erdős-Rényi como a Modelos de Configuraciones y se centran en el régimen asintótico de grafos grandes. Para establecerlos utilizamos y extendemos teoremas existentes sobre los tamaños asintóticos de ciertos sub-grafos en modelos de grafos aleatorios, junto con límites de escala y concentraciones de variables aleatorias para estudiar las dinámicas en ellos.

Los problemas estudiados en esta tesis resultaron en los siguientes tres trabajos: [BJLS19][JS19b][JS19a]. De los cuales el último se encuentra aún en preparación.

Palabras clave: *Grafos Aleatorios, Optimización, Procesos Estocásticos, Límites de Escala.*

Asymptotic performance of sequential algorithms on random graphs

Abstract

In this thesis we study the evolution, convergence and optimality of search algorithms on random graphs through several examples and we use this analysis to give performance bounds of communication protocols for wireless networks. In particular, we study the following problems:

- The determination of the asymptotic optimality of the degree-greedy algorithm in a broad class of Configuration Model graphs, and the computation of their independence ratios. Although the problem of finding maximum size independent sets is an NP-hard task, we show that on this class of random graphs the degree-greedy algorithm can construct asymptotically maximum independent sets taking just a polynomial time. We also use these results to characterise the independence ratio of certain Erdős-Rényi graphs and to give numerical procedures to compute upper bounds for general Configuration Model graphs.
- The evaluation of the performance of variations of local greedy algorithms for approximating the capacity of WiFi networks and their possible impact on the fairness of the connection assignments. We use these results to study a real-world communication network: Montevideo's cellphone antennas.
- The computation of convergence times of best response algorithms for random distributed optimisation problems with actions structured as sparse random graphs. Using these results we prove, by means of an optimality lemma, general lower bounds for the running times of local search algorithms.

Most of our results are set either on sparse Erdős-Rényi random graphs, Configuration Models or both, while the questions we deal with correspond to the asymptotic regime of large graph size. To analyse them, we use and extend existing results on the asymptotic sizes of various sub-graphs on random graph models; together with scaling limits and concentrations of random variables, to study the dynamics on them.

The problems studied on this thesis resulted on three papers: [BJLS19][JS19b][JS19a]. The last of these still under preparation.

Keywords: *Random Graphs, Configuration Model, Stochastic Processes, Scaling Limits.*

A mis sobrines, Aurora y Pan.

Índice general

1. Introducción	1
1.1. Modelos de grafos aleatorios	1
1.2. Problemas de optimización difíciles	2
1.3. Algoritmos secuenciales	2
1.4. Contribuciones de la tesis	3
1.5. Estructura de la tesis	5
2. Modelos de grafos aleatorios	7
2.1. Grafos generales	7
2.2. Grafos aleatorios	10
2.2.1. Modelo de Erdős-Rényi	11
2.2.2. Modelo de Configuraciones	12
2.2.3. Otros modelos de grafos aleatorios	16
3. Límites de escala de procesos markovianos	17
3.1. Resultados previos	18
3.2. Convergencia de procesos de dimensión finita	20
3.2.1. Deriva globalmente Lipschitz	20
3.2.2. Ejemplo de límite de escala	22
3.3. Convergencia de procesos de dimensión numerable	22
3.3.1. Convergencia de tiempos de parada y valores terminales	24
4. CIM en grafos aleatorios	27
4.1. Conjuntos independientes y algoritmos secuenciales	28
4.1.1. Algoritmos secuenciales	28
4.1.2. Caracterización de conjuntos independientes máximos	32
4.2. Resultados y aplicaciones	32
4.2.1. Primera caracterización de la optimalidad del algoritmo grado-egoísta	32
4.2.2. Criterio de una aplicación	34
4.2.3. Criterio general	34
4.2.4. Aplicación a grafos de Erdős-Rényi	35
4.2.5. Más aplicaciones	36
4.3. Pruebas de los resultados de la sección anterior	37
4.3.1. Prueba del Lema 4.2.1 para grafos generales	37
4.3.2. Prueba de la Proposición 4.2.1	38
4.3.3. Prueba del Teorema 4.2.1	42

4.3.4.	Prueba del Teorema 4.2.2	47
4.3.5.	Prueba de la Proposición 4.2.2	52
4.4.	Posibles extensiones	54
5.	Capacidad de redes de comunicación WiFi	55
5.1.	Exploración dinámica	56
5.2.	Generalización de algoritmos egoístas	57
5.2.1.	Exploración grado-consciente dinámica	58
5.2.2.	Exploración grado-consciente estática	59
5.3.	Resultados de simulaciones	62
5.3.1.	Análisis de la aproximación al algoritmo grado-egoísta	62
5.3.2.	Conjuntos independientes máximos y redes inalámbricas	64
5.3.3.	Equidad y capacidad	68
5.4.	Discusión	70
6.	Mejor respuesta en optimización distribuida	73
6.1.	Introducción	74
6.1.1.	Algoritmo de mejor respuesta	74
6.1.2.	Descripción del proceso estudiado	76
6.2.	Resultados y discusión	78
6.2.1.	Estrategia de las pruebas	80
6.3.	Caracterización del proceso de exploración de costos	81
6.3.1.	Representación por medio de procesos de Poisson	81
6.3.2.	Estimación de los tiempos de saltos	82
6.3.3.	Estimación del tiempo total de corrida	84
6.4.	Pruebas de los resultados	86
6.4.1.	Prueba del Lema 6.2.1	87
6.4.2.	Prueba del Teorema 6.2.1	88
6.4.3.	Prueba de la Proposición 6.2.1	91
6.4.4.	Prueba del Teorema 6.2.2	93
A.	Procesos de ramificación	95
B.	Medidas puntuales de Poisson	99
C.	Topología de Skorohod y convergencia débil	103

Capítulo 1

Introducción

Una *red* es un sistema compuesto por objetos discretos y una relación entre algunos pares de ellos. Son herramientas de modelado ubicuas en muchos campos de investigación. Algunos ejemplos comunes incluyen redes sociales, neuronales, de proteínas y de comunicación. Recientes desarrollos en el poder de computo disponible han hecho que el estudio de redes y herramientas para analizarlas sean temas en alza dentro de la comunidad académica.

Esta tesis se centrará en el estudio de dinámicas asociadas a algoritmos en grafos aleatorios. En los problemas estudiados habrá dos fuentes de estocasticidad: por un lado, la aleatoriedad debida a la topología/conectividad del grafo; y por otro, la originada en la evolución del proceso estocástico que tiene lugar en el mismo. Es por ello que nuestros resultados requerirán algunas herramientas destinadas al estudio de estructuras dentro de los grafos y otras que aborden los procesos estocásticos en ellos.

1.1. Modelos de grafos aleatorios

Los grafos aleatorios son construcciones probabilísticas que permiten definir realizaciones aleatorias de grafos (ver [Bol01, VDH16]). Aunque primero fueron propuestos para el estudio de problemas combinatorios [ER59, Erd59], fueron rápidamente adoptados como herramientas de modelado. Esto se debe a que, por su naturaleza probabilística, generan topologías complicadas similares a las observadas en redes complejas.

De hecho, se han estudiado muchas propiedades observadas en redes empíricas por medio de grafos aleatorios. Por ejemplo, incluso modelos simples como los grafos de Erdős-Rényi rales presentan lo que se conoce como *propiedad de mundo pequeño* [VDH16]. Esto significa que las distancias típicas entre vértices en estos grafos son relativamente pequeñas en comparación con el tamaño del grafo completo, propiedad observada en muchas redes sociales y de comunicación.

Además, los grafos aleatorios pueden ser utilizados para estudiar *escenarios promedio* de algoritmos en grafos [FM97]. A diferencia del estudio de *peores casos posibles*, cuando se analizan escenarios promedio se trata de determinar cómo funcionará un algoritmo en una instancia promedio del problema que resuelve. En otras palabras, si uno estudiara el rendimiento de cierto algoritmo para calcular alguna cantidad en un grafo, uno podría estimar su comportamiento promedio describiendo cómo funciona en ciertos modelos de grafos aleatorios.

Esto puede ser mucho más informativo para algunas aplicaciones (por ejemplo, en redes de comunicaciones) que un análisis de peores casos posibles.

En esta tesis, nos centraremos en la mayoría de nuestras discusiones en grafos aleatorios con sucesiones de grado fijas (es decir, grafos del Modelo de Configuraciones). Una de las razones para hacer especial hincapié en este modelo es que es matemáticamente manejable, es útil en muchas aplicaciones y sirve como primer paso para comprender varias propiedades observadas en redes complejas. Además, muchos de los teoremas probados en este modelo pueden usarse para probar resultados análogos para grafos de Erdős-Rényi raros.

1.2. Problemas de optimización difíciles

La *optimización combinatoria* (ver [ACG⁺12, MM09]) es una familia de importantes problemas en informática y matemáticas que tienen muchas implicaciones prácticas. La idea detrás de ellos es encontrar la solución óptima entre todas las soluciones posibles de un problema combinatorio dado. Por ejemplo, el conocido problema del viajante, en el que se trata de encontrar una manera de visitar cada sitio en una colección de ubicaciones recorriendo la menor distancia posible. Estos son problemas de especial interés porque están presentes en muchos modelos de ingeniería (incluida la programación de tareas [Bur91], el encaminamiento [TV02], el diseño de redes [KN05], etc.). La dificultad de ellos proviene del hecho de que el número de posibles soluciones suele crecer muy rápidamente con el tamaño del sistema analizado.

Dentro de los problemas de optimización combinatoria, están los conocidos como *problemas de optimización difíciles*. Estos son aquellos para los cuales todos los algoritmos existentes que los resuelven tienen tiempos de ejecución que crecen más rápido que polinomialmente con el tamaño del problema. De hecho, se ha demostrado que muchos de ellos pertenecen a las clases de complejidad NP-difícil o NP-completo. Algunos ejemplos de problemas de optimización difíciles en grafos incluyen la búsqueda de árboles de expansión mínimos y conjuntos independientes máximos [FM97]. Estos problemas tienen muchas consecuencias prácticas; por ejemplo, las aproximaciones a árboles de expansión mínimos [GH85] son relevantes para el diseño de redes telefónicas, diseño de circuitos y modelos de taxonomía. Los conjuntos independientes máximos son importantes en aplicaciones que van desde la biología y la física hasta la ingeniería (por ejemplo, en redes de comunicaciones [BJLM14]). En muchos casos, esto se debe a la cercana relación que guardan con la adsorción secuencial [CAP07, Eva93] (modelos en los que partículas llegan una a la vez en un espacio limitado y ocupan un cierto volumen donde no se pueden colocar más partículas).

1.3. Algoritmos secuenciales

Una estrategia para resolver (o aproximar) problemas de optimización combinatoria en grafos es por medio de *algoritmos secuenciales*. Estos son algoritmos que construyen la estructura que resuelve el problema de a pasos, al mismo tiempo que van reduciendo el tamaño del grafo en el que trabajan. Su importancia radica en el hecho de que son fáciles de programar, analizar matemáticamente y, por lo general, tienen complejidad polinómica. Aunque, por otro lado, no suelen ser óptimos para casos generales de los problemas que resuelven.

Uno de los objetivos centrales de esta tesis será responder la siguiente pregunta: ¿se puede demostrar que ciertos algoritmos secuenciales se comportan de manera óptima para alguna clase de modelos de grafos aleatorios? Esto es importante porque, cuando la respuesta sea ‘sí’, estaremos demostrando que el problema se puede resolver en tiempo polinómico usando este algoritmo secuencial. Esto es especialmente relevante para problemas de optimización difíciles, ya que sus algoritmos de propósito general son prohibitivamente más lentos.

1.3.1. Rendimiento de algoritmos

Con el fin de estudiar el rendimiento de algoritmos secuenciales en grafos aleatorios, probaremos resultados que caractericen los límites de grafo grande de sus dinámicas asociadas. Estos resultados describirán los diferentes órdenes de crecimiento de los tamaños de las estructuras encontradas por los algoritmos y sus tiempos de ejecución, expresados como funciones del tamaño del grafo. Una parte importante de este análisis consistirá en determinar los diferentes regímenes de la dinámica para los cuales estas cantidades tienen comportamientos asintóticos cualitativamente distintos, junto con las transiciones de fase que tienen lugar entre ellos.

Una herramienta importante que utilizaremos para estudiar el comportamiento asintótico de las dinámicas de algoritmos secuenciales serán los *límites de escala* [Dar02, DN08, Wor95]. Estos son resultados análogos a la ley de grandes números pero para procesos estocásticos. Es decir, prueban la convergencia de procesos a funciones determinísticas cuando algún parámetro se hace grande (en nuestro caso, el tamaño del grafo).

1.4. Contribuciones de la tesis

1.4.1. Proporción de independencia en grafos aleatorios con grados fijos

Como ya señalamos, encontrar conjuntos independientes de tamaño máximo es un problema de optimización difícil el cual se ha demostrado que es NP-difícil en el caso general [FM97]. Por otro lado, los conjuntos independientes máximos y sus tamaños son cantidades importantes en muchos modelos físicos y de comunicación. Sin embargo, hay muy pocos resultados exactos que den sus tamaños (asintóticos) o formas eficientes de calcularlos. En particular, no hay caracterizaciones de las proporciones de independencia para grafos del Modelo de Configuraciones e incluso en el caso de los grafos de Erdős-Rényi raros, sólo existe una caracterización parcial de su proporción de independencia asintótica [KS81]. Una pregunta natural consiste en si, para grafos aleatorios, el problema de encontrar conjuntos independientes máximos sigue siendo NP-difícil; y si se pueden encontrar caracterizaciones (asintóticas) de sus proporciones de independencia.

Por medio de límites de escala, caracterizamos el comportamiento asintótico de la exploración grado-egoísta, un algoritmo secuencial. También proporcionamos condiciones suficientes para que este algoritmo encuentre conjuntos independientes de tamaño asintóticamente óptimo en grafos grandes del Modelo de Configuraciones. En estas condiciones, el algoritmo grado-egoísta construye en tiempo polinómico un conjunto independiente cuasi máximo. Utilizamos este hecho para derivar una caracterización de los tamaños de los conjuntos independientes máximos de esta clase de grafos. Además, al aplicar estos resultados a grafos de Erdős-Rényi raros, damos una prueba simple del “fenómeno e ” identificado para apareamientos por Karp

y Sipser [KS81], y damos una nueva caracterización de la proporción asintótica de vértices en sus conjuntos independientes máximos.

1.4.2. Capacidad de redes de comunicación WiFi

La capacidad de las redes WiFi se ha modelado teóricamente como el tamaño de los conjuntos independientes máximos de sus grafos de interferencia [BJLM14, BJLM16]. Estos son grafos donde cada vértice representa un dispositivo y las aristas entre ellos representan interferencia. Es decir, dos dispositivos no pueden transmitir información simultáneamente si sus vértices correspondientes están unidos por una arista. La capacidad [LCZW10] de las redes es una cantidad importante, ya que permite la evaluación de diferentes protocolos al describir el rendimiento máximo que pueden lograr.

Inspirados en los resultados de optimalidad para el algoritmo grado-egoísta, presentamos y estudiamos dos variantes de algoritmos de exploración secuencial egoístas: exploraciones grado-conscientes estáticas y dinámicas. Derivamos límites de escala para ambas familias, lo que a su vez nos permite calcular el tamaño de los conjuntos independientes construidos por ellos. Mientras que el primero tiene una dinámica límite más simple, el segundo puede usarse para aproximar arbitrariamente el algoritmo grado-egoísta. Ambas son dinámicas locales; es decir, algoritmos en los que las decisiones de los vértices de conectarse o no se toman utilizando sólo la información disponible de sus vecindarios (y no la información global de todo el grafo). Mostramos que los límites de escala correspondientes constituyen un método eficiente para calcular o aproximar la capacidad de una gran clase de grafos aleatorios raros. Como aplicación, mostramos cómo puede usarse nuestro método para estimar la capacidad de una red inalámbrica grande en la que se implementa el protocolo 802.11. Finalmente, también consideramos otros indicadores del rendimiento de estos algoritmos, como la equidad de la configuración resultante, y mostramos cómo se puede lograr un equilibrio entre la equidad y la capacidad de la red.

1.4.3. Optimización distribuida en grafos aleatorios

El objetivo de la optimización distribuida es estudiar dinámicas para las cuales el comportamiento de agentes individuales, cada uno de los cuales puede modificar algunas de las variables, lleva al estado del sistema a maximizar (o minimizar) cierta función de costo. Es un campo que encuentra muchas aplicaciones en ingeniería [AHK07, CMOP10, YZH14, MS15], entre las cuales se incluyen problemas de encaminamiento y diseño de redes eléctricas. Estos modelos están estrechamente relacionados con los juegos potenciales [DG16], para los cuales el *algoritmo de mejor respuesta* se ha mostrado que es óptimo, bajo ciertas condiciones, para la búsqueda de equilibrios de Nash.

En este contexto, estudiamos el algoritmo de mejor respuesta para problemas de optimización distribuida aleatoria donde las acciones que cada agente puede tomar se estructuran como grafos del Modelo de Configuraciones. De esta forma, damos una caracterización probabilística del algoritmo en términos de procesos de Poisson inhomogéneos, la cual utilizamos para estudiar la transición de fase de complejidad que tiene lugar a medida que crece el número de agentes. También extendemos pruebas existentes sobre la optimalidad de este algoritmo dentro del conjunto de algoritmos de búsqueda locales.

1.5. Estructura de la tesis

La estructura del resto de la tesis será la siguiente. En el Capítulo 2, proporcionaremos una breve introducción a los grafos en general y a los grafos aleatorios en particular. Haremos especial hincapié en los dos modelos de grafos aleatorios que serán los más relevantes a lo largo de la tesis: los grafos de Erdős-Rényi y el Modelo de Configuraciones. Mientras que en el Capítulo 3, presentamos algunos resultados referentes a límites de escala que luego serán utilizados para estudiar y caracterizar el comportamiento asintótico de dinámicas en grafos aleatorios. En el Capítulo 4 analizaremos nuestro primer problema dinámico: la optimalidad asintótica del algoritmo grado-egoísta para construir conjuntos independientes en grafos del Modelo de Configuraciones. Inspirados en estos resultados, en el Capítulo 5 estudiamos variaciones del algoritmo egoísta para determinar la capacidad de las redes de comunicación WiFi (es decir, el número máximo de personas capaces de transmitir información simultáneamente). Finalmente, en el Capítulo 6 analizaremos un algoritmo para encontrar soluciones de problemas de optimización distribuida aleatoria donde las acciones que los agentes pueden tomar se estructuran como grafos del Modelo de Configuraciones.

Los Apéndices A, B y C pueden usarse como referencia sobre los resultados utilizados en la tesis sobre procesos de ramificación, medidas aleatorias de Poisson y convergencia de procesos estocásticos.

Se agradecen los útiles comentarios de los jurades (Dra. Inés Armendáriz, Dr. Bernardo Nunes Borges de Lima y Dr. Remco van der Hofstad) que llevaron a esta versión revisada.

Capítulo 2

Modelos de grafos aleatorios

En este capítulo damos una breve introducción a los grafos en general y a los grafos aleatorios en particular. Se hace especial hincapié en los grafos de Erdős-Rényi y del Modelo de Configuraciones, ya que son los modelos más discutidos a lo largo de la tesis.

En la Sección 2.1 damos las definiciones básicas de grafos y describimos algunos de los tipos comunes de sub-grafos que trataremos. Luego, en la Sección 2.2 discutimos las construcciones que dan lugar a los grafos de Erdős-Rényi y los Modelos de Configuraciones. También incluimos algunos resultados relevantes con respecto a la transición de fase de los tamaños de sus componentes conexas más grandes y sobre percolación de sitios. Al final de la sección incluimos una descripción concisa de otros modelos de grafos aleatorios comunes.

2.1. Grafos generales

Los grafos son un modelo natural para las redes.

Definición 2.1.1. Un *grafo finito no-dirigido*¹ $G = (V, E)$ se compone de dos conjuntos: un conjunto finito V llamado *conjunto de vértices* y una colección E de pares de elementos no ordenados de V llamado *conjunto de aristas*.

El conjunto de vértices representará los objetos que componen la red mientras que las aristas a las relaciones entre ellos. Aunque el conjunto de vértices puede en principio ser cualquier conjunto finito, a partir de ahora asumiremos por simplicidad que está dado por $V = [n] := \{1, \dots, n\}$ (donde $n \in \mathbb{N}$). El *tamaño* de un grafo será el número de vértices en el mismo y se denotará con $|V| := n$.

Definición 2.1.2. Dado un grafo $G = (V, E)$, un *sub-grafo* es un nuevo grafo G' formado por subconjuntos $V' \subseteq V$ y $E' \subseteq E$ tales que si $(u, v) \in E'$ entonces $u, v \in V'$.

Esto es, un sub-grafo es un grafo formado a partir de restringirnos a un subconjunto de vértices y un subconjunto de aristas entre ellos.

Dos ejemplos triviales de grafos son:

- El *grafo vacío*, donde $E = \emptyset$. O sea, el grafo donde todos los vértices se encuentran aislados.

¹Generalmente omitiremos el *finito* y *no-dirigido*, ya que todos los gráficos en este trabajo serán de este tipo.

- El *grafo completo*, donde para todo par $u, v \in V$, $(u, v) \in E$. O sea, el grafo donde toda arista posible está presente. Notar que todo grafo finito es sub-grafo del grafo completo del mismo tamaño.

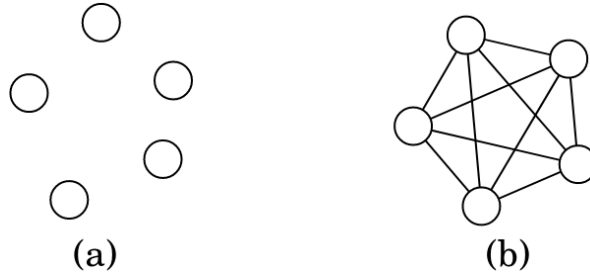


Figura 2.1: (a) Ejemplo de grafo vacío con 5 vértices. (b) Ejemplo de grafo completo con 5 vértices.

Dados dos vértices $u, v \in V$, diremos que u es un *vecino* de v si $(u, v) \in E$. Como los grafos que consideraremos no son dirigidos, la relación de vecindad será simétrica. Del mismo modo, llamaremos *vecindario de v* al conjunto de vértices que son sus vecinos, y lo denotaremos como $\mathcal{N}(v)$. Además, nos referiremos al número de vecinos que tiene un vértice v como *grado de v* , que se simbolizará con d_v .

Dado un grafo G , definiremos una variable aleatoria asociada D que da el grado de un vértice elegido uniformemente. Esta variable puede ser aleatoria incluso si el grafo es fijo. La distribución de la variable de grado se llamará *distribución de grados* y se denotará con $(p_k)_{k \geq 0}$ (o $(p_k^{(n)})_{k \geq 0}$, si deseamos enfatizar la dependencia con el tamaño del grafo). Es fácil ver que, para $k \geq 0$, el número N_k de vértices de grado k estará dado por p_k veces el tamaño del grafo n .

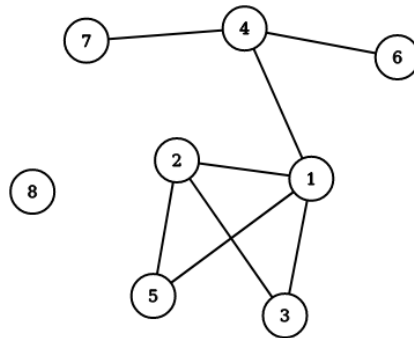


Figura 2.2: Ejemplo de una grafo de tamaño 8. Aquí la vecindad del vértice 1 es el conjunto $\mathcal{N}(1) = \{2, 3, 4, 5\}$, y por lo tanto su grado es 4. La distribución de grados para este grafo está dada por $p_0 = 1/8$, $p_1 = 1/4$, $p_2 = 1/4$, $p_3 = 1/4$ y $p_4 = 1/8$.

Tres tipos de estructuras dentro de los grafos serán relevantes en este trabajo: caminos, ciclos y componentes conexas.

Definición 2.1.3. Dado un par de vértices $u, v \in V$, diremos que una sucesión finita de vértices distintos $(v_i)_{i \in [l]}$ es un *camino entre u y v* si $v_1 = u$, $v_l = v$ y para todo $i \in [l - 1]$ tenemos que $(v_i, v_{i+1}) \in E$.

En palabras, un camino es una secuencia de vértices vecinos que conecta un par de vértices fijos. La *longitud* de un camino estará dada por el número de aristas en la sucesión de vértices que lo compone.

Definición 2.1.4. Un *ciclo* se define como un camino libre de intersecciones que conecta un vértice con sí mismo.

Definición 2.1.5. La *componente conexa* de un vértice $v \in V$ es el conjunto de vértices formado por todos aquellos $w \in V$ tales que existe un camino que conecta v con w . Se la simbolizará como $\mathcal{C}(v)$.

El *tamaño* $|\mathcal{C}(v)|$ de un componente conexa será el número de vértices dentro de ella. Las componentes conexas serán relevantes para nuestro trabajo, ya que en la mayoría de las dinámicas consideradas la evolución en el tiempo se podrá factorizar en las diferentes componentes conexas del grafo, lo que nos permitirá analizar los diferentes componentes por separado.

Además, se dice que un grafo es *conectado* si sólo hay un componente conexa en él. Un *árbol* puede entonces definirse como un grafo conectado sin ciclos.

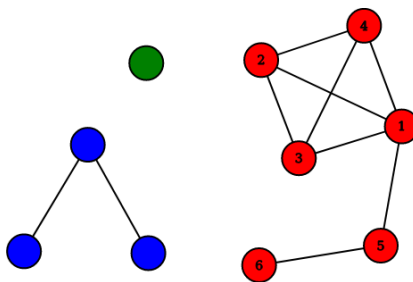


Figura 2.3: Ejemplo de un grafo de tamaño 10. Hay tres componentes conexas diferentes en este grafo, cada una representada por un color diferente: rojo, azul y verde. Los vértices 2 y 6 pertenecen a la misma componente conexa y, por lo tanto, existe al menos un camino entre ellos; por ejemplo, el camino $(2, 4, 1, 5, 6)$ cuya longitud es 5. Además, $(1, 4, 2, 3, 1)$ es un ciclo dentro del grafo.

Finalmente, definiremos otra estructura que será de importancia en los Capítulos 4 y 5.

Definición 2.1.6. Dado un grafo G , se dice que un subconjunto de vértices $I \subseteq V$ es un *conjunto independiente* si, para todo $u, v \in I$, $(u, v) \notin E$.

Es decir, un conjunto independiente es un subconjunto de vértices donde ningún par de ellos son vecinos entre sí. Esto significa que si un subconjunto I es un conjunto independiente, el único sub-grafo que contiene sus vértices es el grafo vacío.

Definición 2.1.7. Un conjunto independiente se dice que es *maximal* si no hay uno más grande que lo contenga como un subconjunto y *máximo* si no hay otro conjunto independiente de mayor tamaño.

El tamaño de los (posiblemente muchos) conjuntos independientes máximos de un grafo G se denominará *número de independencia* y se simbolizará con $\alpha(G)$. De manera similar, a la proporción de vértices en un conjunto independiente máximo se la llama *proporción de*

independencia del grafo. Como se discutió en el Capítulo 1, tanto la construcción de conjuntos independientes máximos como el cálculo de números de independencia son problemas notoriamente difíciles, los cuales se sabe son NP-difíciles [FM97].

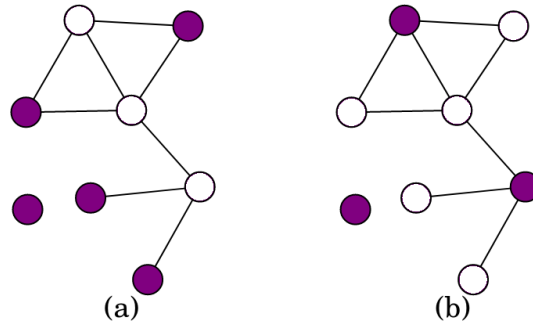


Figura 2.4: (a) En violeta, un ejemplo de un conjunto independiente máximo. (b) Ejemplo de un conjunto independiente que es maximal (no se le puede agregar más vértices) pero no máximo (el conjunto independiente que se muestra en (a) es más grande).

Definición 2.1.8. Una determinada propiedad A de grafos es *monótona creciente* si por cada par de grafos del mismo tamaño G y H tales que cada arista en G está presente también en H , si G tiene la propiedad A luego H también la tiene.

Se puede ver fácilmente que, por cada $k \geq 0$, la propiedad $\{\alpha(G) \leq k\}$ es monótona en este sentido. Llamando \mathcal{C}_{max} a la mayor componente conexa de un grafo, también se puede demostrar que la propiedad $\{|\mathcal{C}_{max}| \geq k\}$ también es monótona (donde $|\mathcal{C}_{max}|$ denota el tamaño de la mayor componente conexa).

2.2. Grafos aleatorios

El primer modelo de grafos aleatorios fue introducidos por P. Erdős en 1959 como una herramienta para aplicar el método probabilístico para probar una relación asintótica entre la cintura y el número cromático de grafos generales [Erd59]. Más tarde, junto con A. Rényi, continuó el estudio de este modelo [ER59, ER60]. Este es uno de los dos modelos conocidos hoy en día como *modelos de Erdős-Rényi*. Consiste en la construcción de grafos finitos con un número fijo de aristas distribuidas uniformemente a lo largo del conjunto de pares de vértices; mientras que el segundo modelo que comparte el mismo nombre, se debe a E. Gilbert [Gil59], y consiste en grafos finitos obtenidos al establecer, independientemente y con la misma probabilidad, una arista entre cada par de vértices. A partir de ahí, se desarrollaron muchos modelos de grafos aleatorios diferentes. Estos probaron ser no sólo construcciones interesantes, sino también muy útiles en aplicaciones: como herramientas para estudiar escenarios promedio en redes complejas y como una forma de explorar posibles mecanismos que dan lugar a muchas de las características distintivas de las redes observadas en el mundo real. Por ejemplo, incluso un modelo tan simple como el de Erdős-Rényi exhibe la propiedad del mundo pequeño (prevalencia de distancias relativamente cortas entre la mayoría de los pares de vértices) observada en muchas redes sociales y de comunicación.

En esta sección, nos centraremos principalmente en dos modelos de grafos aleatorios: grafos de Erdős-Rényi y grafos aleatorios con sucesiones de grados fijas (también conocidos como

grafos del Modelo de Configuraciones). Esta presentación será breve y sólo pretende dar una idea de las construcciones involucradas y los resultados necesarios para nuestro análisis. Para una discusión más exhaustiva sobre grafos aleatorios, el lector puede consultar [VDH16, Bol01, Mul68].

A lo largo de la tesis lidiaremos principalmente con propiedades asintóticas de grafos aleatorios en el límite de grafo grande. Para ello utilizaremos las notaciones habituales $\mathcal{O}(\cdot)$ y $o(\cdot)$ para describir el comportamiento asintótico de las funciones del tamaño de grafo n . También utilizaremos la notación probabilística $\mathcal{O}_{\mathbb{P}}(\cdot)$ y $o_{\mathbb{P}}(\cdot)$ en el siguiente sentido:

- Una sucesión de variables aleatorias X_n será $\mathcal{O}_{\mathbb{P}}(f_n)$ (para alguna función $f_n : \mathbb{N} \rightarrow \mathbb{R}_{>0}$) si para todo $\epsilon > 0$ existen $M > 0$ y $N \in \mathbb{N}$ tales que $\mathbb{P}(|X_n/f_n| > M) < \epsilon$ para todo $n \geq N$.
- De la misma forma, una sucesión de variables aleatorias X_n será $o_{\mathbb{P}}(f_n)$ si $X_n/f_n \xrightarrow{\mathbb{P}} 0$ cuando $n \rightarrow \infty$.

Además, diremos que un evento sucede *con alta probabilidad* (c.a.p.) siempre que su probabilidad sea alguna función $1 + o(1)$ del tamaño del grafo.

2.2.1. Modelo de Erdős-Rényi

Como se mencionó anteriormente, hay dos modelos de grafos aleatorios distintos (pero estrechamente relacionados) que se conocen como *modelo de Erdős-Rényi*. Ambos dan como resultado un grafo aleatorio de tamaño n . En uno, se selecciona un grafo uniformemente del conjunto de grafos con n vértices y m aristas entre ellos. En el otro, cada arista de un grafo de tamaño n está presente con una probabilidad $p \in (0, 1)$ e independientemente de las demás. Aquí, cuando nos referimos a un grafo de Erdős-Rényi, estaremos hablando de este último. Usaremos la notación $\text{ER}_n(p)$ para representar la distribución de un grafo aleatorio generado de acuerdo con esta construcción.

La construcción que da lugar al modelo Erdős-Rényi puede verse como una percolación de aristas con probabilidad p en el grafo completo de tamaño n . Suele ser útil cuando se comparan dos grafos $G \sim \text{ER}_n(p)$ y $H \sim \text{ER}_n(p')$ (con $p < p'$) tener en cuenta que ambos pueden ser *acoplados*² para que cualquier arista e esté presente en G si y sólo si también está presente en H .

Casi todos los resultados discutidos en esta tesis serán resultados asintóticos en el tamaño del grafo n . En este límite, discutiremos principalmente el caso de *grafo ralo* donde el grado de un vértice elegido uniformemente es una función acotada del tamaño del grafo. En el caso de los grafos de Erdős-Rényi, prestaremos especial atención al caso en el que $p(n) = \lambda/n$ (con $\lambda > 0$) que se corresponde con los grafos ralos no-triviales.

Un hecho importante con respecto a los grafos de Erdős-Rényi ralos es que el tamaño de la mayor componente conexa sufre una transición de fase al aumentar su grado medio. Existen muchos resultados que caracterizan esta transición. La prueba de la versión presentada aquí puede hallarse, por ejemplo, en [VDH16].

²Esto es así pues a cada arista en G se le puede asociar una variable independiente de Bernoulli de parámetro p que indica si la arista está presente en el grafo o no. Lo mismo se puede hacer con H y variables Bernoulli de parámetro p' . Luego, el acoplamiento entre ambos grafos se realiza por medio del acoplamiento usual de variables Bernoulli independientes.

Teorema 2.2.1 (Transición de fase del modelo de Erdős-Rényi ralo). *Sea $(G^{(n)})_{n \geq 1}$ una sucesión de grafos aleatorios distribuidos de acuerdo a $ER_n(\lambda/n)$. Si $\lambda \leq 1$ entonces $|\mathcal{C}_{max}^{(n)}| = o_{\mathbb{P}}(n)$. Mientras que, si $\lambda > 1$ entonces $|\mathcal{C}_{max}^{(n)}| = \chi n + o_{\mathbb{P}}(n)$, para algún $\chi > 0$ que depende de λ .*

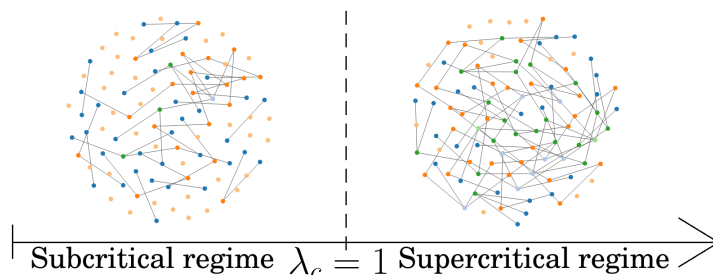


Figura 2.5: Visualización del resultado que se enuncia en el Teorema 2.2.1. Cuando el grado medio del grafo es menor o igual a 1, todas las componentes conexas son de un tamaño pequeño; pero cuando es mayor que 1, surge un componente gigante. (Gráficos creados en: www.networkpages.nl)

Lo que muestra este teorema es que, asintóticamente en probabilidad, hay una componente gigante (un componente conexo que contiene asintóticamente una proporción positiva de vértices) si y sólo si $\lambda > 1$. En cualquier otro caso, cada componente conexo será de tamaño $o_{\mathbb{P}}(n)$.

2.2.2. Modelo de Configuraciones

Un vértice elegido uniformemente en un grafo distribuido de acuerdo con $ER_n(\lambda/n)$ tendrá un grado que se distribuirá asintóticamente como una variable de Poisson de media λ . Muchas aplicaciones requieren grafos con otras distribuciones de grados. Hay varias formas de generalizar el modelo de Erdős-Rényi para permitir distintas distribuciones de grados [Bol01, VDH16], una de las más utilizadas es hacerlo por medio del Modelo de Configuraciones. Esta construcción dará como resultado un grafo donde cada vértice $v \in V$ tendrá un grado fijo $d_v \in \mathbb{N}_0$. La sucesión $\vec{d}^{(n)} \in \mathbb{N}_0^n$ que describe los grados en un grafo se llamará *sucesión de grados*³.

Aquí describiremos la construcción que da origen al Modelo de Configuraciones. Dada una sucesión de grados $\vec{d}^{(n)} \in \mathbb{N}_0^n$, formaremos un *multigrafo*⁴ de tamaño n con dicha sucesión de grados de la siguiente manera:

- asigne a cada vértice $v \in \{1, \dots, n\}$ un número d_v de *medias aristas*.
- aparezca uniformemente cada media arista con otra que no tenga pareja aún.
- repita hasta que no haya más medias aristas libres⁵. Luego, establezca una arista entre cada par de medias aristas apareadas.

³Una construcción alternativa para el Modelo de Configuraciones consiste en tomar los grados como variables aleatorias i.i.d. en lugar de asumirlos fijos. Para simplificar, siempre trabajaremos con la variante de grado fijo, pero la mayoría de los resultados de la tesis se pueden extender con poco trabajo al caso i.i.d.

⁴Es decir, un grafo en donde hay posiblemente aristas entre un vértice y sí mismo (*autoaristas*) y aristas múltiples entre un par de vértices (*multiaristas*).

⁵Si el número inicial de aristas era impar, deseche la restante. Esto difícilmente hará alguna diferencia cuando el tamaño del grafo sea grande.

La distribución de multigrafos aleatorios generados de acuerdo con este procedimiento se simbolizará con $CM_n(\bar{d}^{(n)})$. Debido a que todos los apareamientos se hacen de manera uniforme, el multigrafo resultante se distribuye de igual forma sin importar en qué orden se elijan las medias aristas. Este hecho permite, al analizar una dinámica en algún grafo, que el apareamiento se incorpore en la evolución temporal en cuestión.

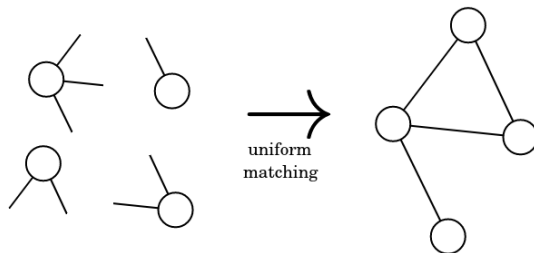


Figura 2.6: Ejemplo de la construcción de un grafo aleatorio uniforme con sucesión de grados $\bar{d} = (1, 2, 2, 3)$ a través del apareamiento uniforme de las medias aristas de cada uno de sus vértices.

Una característica importante del Modelo de Configuraciones es que se puede utilizar como un marco unificado para probar resultados en muchos tipos de grafos: grafos aleatorios regulares, redes libres de escala, grafos de Erdős-Rényi, etc. Por ejemplo, los resultados en [Kim07] implican que si la distribución de grado asintótica de una sucesión de grafos del Modelo de Configuraciones es Poisson de parámetro $\lambda > 0$ y se muestra que una propiedad vale c.a.p. en esta sucesión, esta propiedad vale también c.a.p. para grafos de Erdős-Rényi rales con grado medio asintótico λ .

Comportamiento asintótico del Modelo de Configuraciones

A lo largo de este trabajo nos centraremos en resultados asintóticos en el tamaño del grafo n . Para ello asumiremos, cuando se trate de Modelos de Configuraciones, la siguiente condición de convergencia⁶:

Condición de convergencia (CC): cuando tratemos con sucesiones de grafos $(G^{(n)})_{n \geq 1}$ con $G^{(n)} \sim CM_n(\bar{d}^{(n)})$, vamos a asumir que $D^{(n)} \xrightarrow{\mathbb{P}} D$ y $\mathbb{E}(D^{(n)2}) \xrightarrow{n \rightarrow \infty} \mathbb{E}(D^2)$.

Donde $D^{(n)}$ es la variable aleatoria que da el grado de un vértice elegido uniformemente en $G^{(n)}$ y D es una variable aleatoria con segundo momento finito. También usaremos $(p_k^{(n)})_{k \geq 0}$ para referirnos a la distribución de $D^{(n)}$ y $(p_k)_{k \geq 0}$ para la de la variable aleatoria de grado asintótico D . Cuando no haya ambigüedad, omitiremos el superíndice en $G^{(n)}$.

Aunque esta construcción da como resultado un multigrafo en lugar de un *grafo simple*⁷, como muestra el siguiente teorema, esto no es un problema ya que bajo la (CC) existe una probabilidad asintótica positiva de obtener un grafo simple.

⁶Para algunas aplicaciones, se pueden suponer condiciones más débiles. No discutiremos este asunto en detalle para evitar que la presentación se vuelva técnica. Puede encontrarse más información sobre esto en [VDH16].

⁷Uno sin auto ni multiaristas.

Teorema 2.2.2 (Probabilidad de simplicidad [Jan09b]). *Dada una sucesión de grafos $G \sim CM_n(\bar{d}^{(n)})$ que satisface la (CC) hacia una variable límite D , definimos $\nu := \mathbb{E}(D(D - 1))/\mathbb{E}(D)$. Entonces, la probabilidad de que G sea simple vendrá dada por $e^{-\nu/2 - \nu^2/4} + o(1)$.*

Esto significa que cualquier evento que se haya demostrado que vale c.a.p. para $G \sim CM_n(\bar{d}^{(n)})$, también vale c.a.p. para la construcción condicionada a dar como resultado un grafo simple. Además, condicionada a dar como resultado un grafo simple, el grafo obtenido se sigue distribuyendo uniformemente entre los grafos con la sucesión de grados $\bar{d}^{(n)}$.

Transición de fases del tamaño de la mayor componente conexa

Otra característica importante del modelo es que su mayor componente conexa contiene asintóticamente una proporción positiva de los vértices del grafo si $\nu > 1$ (por lo tanto, esta cantidad se denominará *parámetro de criticidad* del grafo). De manera similar al caso de Galton-Watson, cuando se cumple esta condición, diremos que el grafo es *supercrítico*; y cuando no lo hace, es *subcrítico*. Esto se resume en el siguiente teorema.

Teorema 2.2.3 (Transición de fase en el MC [MR98, JL09]). *Sea $G \sim CM_n(\bar{d}^{(n)})$ tal que satisface la (CC). Si $\nu \leq 1$, entonces $|\mathcal{C}_{max}^{(n)}| = o_{\mathbb{P}}(n)$. Por otro lado, si $\nu > 1$, existe $\chi > 0$ que depende de \bar{d} tal que $|\mathcal{C}_{max}^{(n)}| = \chi n + o_{\mathbb{P}}(n)$.*

La aparición de ν en este teorema no es una mera coincidencia. Los grafos del Modelo de Configuraciones se ven *localmente* como árboles de Galton-Watson⁸ con una distribución de grados conocida como *distribución sesgada por tamaño* [VDH16] cuya media es precisamente ν . En este sentido, lo que muestra este teorema es que hay un componente gigante si y sólo si el árbol de Galton-Watson que aproxima al grafo localmente es supercrítico.

También utilizando la idea de que los Modelos de Configuraciones se parecen localmente a árboles de Galton-Watson, si la cola de la distribución de grados de un grafo subcrítico está acotada por una ley de potencias, entonces el tamaño de la componente más grande se puede acotar:

Teorema 2.2.4 (Tamaño de la mayor componente subcrítica [Jan08]). *Sea $G \sim CM_n(\bar{d}^{(n)})$ tal que satisface la (CC) con parámetro de criticidad $\nu < 1$ y distribución de grado límite $p_k = \mathcal{O}(e^{-\gamma k})$, para algún $\gamma > 0$. Entonces existe $A > 0$ tal que $|\mathcal{C}_{max}^{(n)}| \leq An^{1/(\gamma-1)} + o_{\mathbb{P}}(n^{1/(\gamma-1)})$.*

En [Jan08], se presentan también versiones más fuertes de este resultado.

Percolación de sitios en el Modelo de Configuraciones

Dado un grafo G , existen dos tipos de percolación que se pueden estudiar en él: percolación de sitios y percolación de enlaces. El primero consiste en eliminar una muestra aleatoria de vértices; mientras que en el segundo se eliminan, en su lugar, aristas. Aquí discutiremos sólo la percolación de sitios, pero la mayoría de los resultados y construcciones se pueden extender fácilmente al caso de la percolación de enlaces.

La percolación se propuso en principio como un modelo para medios aleatorios [Gri99]: la presencia de caminos dentro del grafo percolado se interpretaba como posibles trayectorias

⁸Lo que significa que el límite local débil del grafo es un árbol de Galton-Watson. Para más detalles sobre este tipo de convergencia, consulte el Capítulo 3 de [Bor16].

de un elemento de fluido a través de un material poroso. El primer modelo estudiado fue \mathbb{Z}^2 , por Broadbent y Hammersley en 1957. En el contexto de grafos aleatorios, la percolación se puede usar para describir la capacidad de una red para seguir funcionando correctamente en una situación adversa. Por ejemplo, en comunicaciones, la percolación de sitios se puede usar para modelar el apagado de nodos dentro de una red como resultado de un ataque o falla.

En [Jan09a], Janson estudia la percolación en grafos del Modelo de Configuraciones. Su enfoque consiste en partir la remoción de vértices en dos pasos:

- (I) Primero, antes del apareo de las medias aristas de la construcción del Modelo de Configuraciones, cada vértice es seleccionado independientemente para ser eliminado con probabilidad $p > 0$. En lugar de simplemente eliminar los vértices y las aristas que están unidos a ellos⁹, los vértices se eliminan pero en el lugar de sus medias aristas se agrega un número igual de vértices de grado 1 de un tipo especial (a los que llamaremos, *vértices verdes*).
- (II) Después de esto, el apareo de las medias aristas se realiza de la forma habitual (incluidas las que pertenecen a los vértices verdes). De esta manera, se forma un grafo donde todas las aristas que provienen de vértices eliminados tienen al menos un vértice verde en uno de sus extremos. Luego, el grafo percolado se obtiene al eliminar los vértices verdes (o un número equivalente de vértices de grado 1) junto con sus aristas.

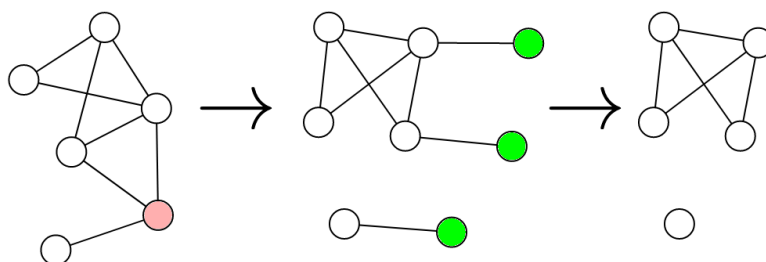


Figura 2.7: Construcción utilizada por Janson en [Jan09a] para describir la percolación de sitios en un Modelo de Configuraciones. Primero, los vértices a eliminar se seleccionan al azar; en este ejemplo, los rosas. Luego, se eliminan los vértices seleccionados y se agregan vértices de grado 1 en el lugar de cada una de sus aristas; en el ejemplo, estos nuevos vértices son los verdes. Finalmente, se eliminan todos los vértices de grado 1 agregados y se obtiene el grafo percolado.

En principio, esto puede parecer que no simplifica mucho el problema, sino que agrega un paso adicional innecesario. La utilidad de dividir la percolación en dos pasos proviene de la siguiente observación. Dado un Modelo de Configuraciones G con $p_1 > 0$, definimos el grafo \tilde{G} como el grafo obtenido al eliminar una cierta proporción de los vértices de grado 1 de G . Es fácil ver que el grafo \tilde{G} tiene una componente gigante si y sólo si G también tiene una. Esto es así porque la eliminación de cada uno de los vértices de grado 1 no divide la componente conexa más grande en otras más pequeñas¹⁰, como podría suceder con vértices de mayor grado.

⁹Tenga en cuenta que no sólo tenemos que eliminar las medias aristas que pertenecen al vértice, sino también las que están unidas a ellos en el grafo.

¹⁰Para hacer este argumento un poco más riguroso, uno necesitaría saber cuál es la distribución de grados de

Esto significa que, para responder a la pregunta de si hay un componente gigante o no en el grafo percolado, es suficiente analizar el grafo antes de eliminar los vértices verdes. De esta manera, la pregunta es equivalente a determinar si el grafo del Modelo de Configuraciones con los vértices verdes es supercrítico o no. Es decir, se reduce a estudiar un nuevo Modelo de Configuraciones con una distribución de grados modificada.

2.2.3. Otros modelos de grafos aleatorios

Aquí incluimos una breve descripción de otros modelos de grafos aleatorios que aparecen en la literatura. La lista no es exhaustiva y sólo trata de reflejar la diversidad de modelos existentes.

- *Grafos aleatorios generalizados*: es un modelo que generaliza los grafos de Erdős-Rényi. En él, las aristas también están presentes independientemente una de las otras, pero se permite que sus probabilidades dependan, de una manera específica, de los pares de vértices que conectan. Esto da como resultado grafos aleatorios con muchas de las propiedades *buenas* de Erdős-Rényi pero que tienen distribuciones de grados más diversas.
- *Adhesión preferencial*: es un modelo dinámico que se utiliza para estudiar el crecimiento de redes complejas. A cada tiempo se define un grafo de la siguiente manera: en un principio, el grafo consiste en un solo vértice; y más adelante en cada paso se agrega un vértice con un número fijo de aristas $m \in \mathbb{N}$ al grafo, con sus aristas unidas a vértices ya presentes con una probabilidad que depende de sus grados. Una propiedad importante del modelo es que, en condiciones robustas, el grafo resultante después de un número suficientemente grande de pasos es cercano a uno con una distribución de grados que sigue una ley de potencias. Estas distribuciones están presentes en muchas redes del mundo real.
- *Modelo de Configuraciones jerárquico*: el Modelo de Configuraciones tiene un coeficiente de agrupamiento bajo (no presenta grupos de vértices con una mayor tendencia a conectarse entre ellos). Esto puede ser un inconveniente para modelar muchas redes del mundo real donde la agrupación suele ser alta. Para superar esto, se puede utilizar el siguiente modelo. Construya un grafo del Modelo de Configuraciones y luego reemplace cada vértice por un grafo pequeño y altamente conectado. Después de eso, conecte cada uno de las aristas originales a uno de los miembros del nuevo grafo que reemplaza el vértice.
- *Grafos aleatorios geométricos*: estos grafos se construyen a partir de procesos puntuales en una caja en el plano. Por ejemplo, un proceso puntual de Poisson (ver Apéndice B). El grafo se obtiene a partir del proceso puntual estableciendo una arista entre cada par de puntos que estén más cerca el uno del otro que una cierta distancia fija. Los grafos resultantes comparten muchas de las propiedades que se ven en una amplia gama de redes complejas, pero son generalmente difíciles de estudiar analíticamente.

los vértices en la componente conexa más grande. Esta distribución es descrita por los resultados del Teorema 2.3 de [JL09]. Con esto, se puede demostrar que la componente gigante (si está presente) tiene una proporción positiva de vértices de grado mayor que 1 que permanecen conectados después de la eliminación de los de grado 1.

Capítulo 3

Límites de escala de procesos markovianos

En casi todos los cursos introductorios de probabilidad, uno de los primeros resultados que se discuten es la *ley de los grandes números*: cuando analizamos el promedio ($\bar{X}_n := \sum_{i=1}^n X_i/n$) de un número creciente de realizaciones de copias independientes de una variable aleatoria de media $\mu < \infty$, este converge a la media de las variables aleatorias ($\bar{X}_n \xrightarrow{n \rightarrow \infty} \mu$). Intuitivamente, lo que esto nos dice es que el factor $1/n$ en el promedio gana sobre las fluctuaciones aleatorias y hace que converja a un límite determinístico. El problema de los límites de escala de procesos estocásticos es completamente análogo, pero con respecto a trayectorias en lugar de números. Al reescalar el espacio de estados y el parámetro de tiempo de un proceso estocástico, mostraremos (si es este el caso) que este converge a un límite determinístico no trivial. En este sentido, pueden interpretarse como versiones funcionales de la ley de los grandes números.

En el Capítulo 2, dimos una introducción a los grafos aleatorios. Allí, presentamos sus definiciones y resultados básicos mientras que intentamos construir cierta intuición alrededor de ellos. Aunque este capítulo también es introductorio, su tono y contenido son de naturaleza diferente. Aquí no intentaremos dar una presentación amplia sobre la convergencia de procesos estocásticos a soluciones de ecuaciones diferenciales, sino que nos centraremos en definir y probar resultados pertinentes para los siguientes capítulos. Por ejemplo, los lemas presentes en la Sección 3.3 serán utilizados para el análisis de la dinámica estudiada en el Capítulo 4. Para una presentación más completa del tema, el lector puede consultar [Dar02, DN08].

El resto del capítulo está organizado de la siguiente manera. En la próxima sección damos algunas definiciones y resultados matemáticos que serán luego utilizados para derivar límites de escala. En la Sección 3.2, probamos un teorema simple que permite establecer límites de escala en procesos de saltos markovianos de dimensión finita, junto con un ejemplo. Finalmente, en la Sección 3.3, demostramos un resultado de límite de escala para procesos de saltos markovianos de dimensión numerable siguiendo la estrategia de Prohorov y presentamos también algunos lemas de tiempos de parada.

3.1. Resultados previos

Aquí damos algunas de las definiciones y resultados que se utilizarán en las secciones posteriores para establecer y analizar límites de escala. Aquí y en el resto del capítulo asumiremos que los procesos estocásticos en cuestión son procesos de saltos markovianos reales en algún espacio de estados contable E .

Desigualdad de Gronwall

El siguiente lema será fundamental para probar límites de escala. Su prueba se puede encontrar en muchos libros de ecuaciones diferenciales ordinarias (por ejemplo, [Chi06]).

Lema 3.1.1. *Sea $I = [a, b] \subseteq \mathbb{R}$ un intervalo y $\alpha, \beta, u : I \rightarrow \mathbb{R}$ funciones continuas, con α no-decreciente y β no-negativa. Si para todo $t \in I$ $u(t) \leq \alpha(t) + \int_a^t \beta(s)u(s)ds$, entonces*

$$u(t) \leq \alpha(t)e^{\int_a^t \beta(s)ds}, \text{ para todo } t \in I.$$

Este lema puede usarse para transformar desigualdades integrales en desigualdades explícitas. Nos será útil ya que el punto de partida para los límites de escala que presentaremos aquí será una representación integral del proceso estocástico en cuestión.

Martingalas aditivas

Dado un proceso de saltos markoviano $(X_t)_{t \geq 0}$ en la base $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$, tomando valores en E y con generador $L(\cdot)$ (de dominio $\mathcal{D}(L)$), vamos a definir una familia de martingalas que comparten una propiedad aditiva entre ellas.

Definición 3.1.1. *Sea $(X_t)_{t \geq 0}$ como arriba, y $f : E \rightarrow \mathbb{R}$ una función en $\mathcal{D}(L)$ tal que $f(X_t)$ es integrable. Entonces, para todo $t \geq u$ (donde $u \geq 0$ está fijo), al proceso*

$$M_t^u(f) := f(X_t) - f(X_u) - \int_u^t Lf(X_s)ds$$

se lo llama *martingala aditiva* de X_t .

Su nombre deriva del hecho de que, para todo $0 \geq u \geq v \geq t$, $M_v^u(f) + M_t^v(f)$ es igual a $M_t^u(f)$. De ahora en más, generalmente omitiremos el supraíndice y asumiremos que la martingala tiene como tiempo de inicio a 0. Además, si omitimos la función f , consideraremos que esta es la identidad ($f(x) = x$). En este caso, llamaremos a la martingala resultante como *martingala de Dynkin*. El siguiente lema [SW95] demuestra que son efectivamente martingalas.

Lema 3.1.2. *Sean X_t y f como antes, entonces $M_t(f)$ es una martingala.*

Demostración. Por hipótesis, $f(X_t)$ es integrable y \mathcal{F}_t -adaptada. Tenemos que demostrar que, para todo $0 \leq s \leq t$, $\mathbb{E}(M_{s+t}(f)|\mathcal{F}_s) = M_s$. Por el comentario anterior, $M_{s+t}(f) = M_s(f) + M_t^s(f)$. Entonces es suficiente mostrar que $\mathbb{E}(M_t^s(f)|\mathcal{F}_s) = 0$. Para ver esto, denotamos con L al generador de X_t y con $P(t)$ al elemento de su semi-grupo de transición correspondiente al tiempo t , entonces la derivada de $\mathbb{E}_i(f(X_t))$ se puede escribir como

$$\begin{aligned} \frac{d}{dt} \mathbb{E}_i(f(X_t)) &= \frac{d}{dt} P(t)f(i) = \lim_{u \searrow 0} \frac{P(t+u)f(i) - P(t)f(i)}{u} \\ &= P(t) \lim_{u \searrow 0} \frac{P(u)f(i) - f(i)}{u} = P(t)Lf(i). \end{aligned}$$

Por lo tanto, tenemos que para todo $t \geq 0$ (si $X_0 = i$)

$$\mathbb{E}_i(f(X_t)) - f(i) = \int_0^t P(s)Lf(i)ds. \quad (3.1)$$

Usando esto, la propiedad de Markov y el hecho de que $P(t)$ y L conmutan, obtenemos que $\mathbb{E}(M_t^s(f)|\mathcal{F}_s) = 0$. De lo cual se sigue el lema. \square

Definición 3.1.2. Dado un proceso estocástico X_t con generador infinitesimal $L(\cdot)$, definiremos la *deriva* del proceso como la función $\delta : E \rightarrow \mathbb{R}$ resultante de aplicar el generador a la función identidad $f(x) = x$.

Luego, dado un proceso X_t , su martingala de Dynkin estará dada por $X_t - X_0 - \int_0^t \delta(X_s)ds$. La deriva representará entonces la parte “predecible” del proceso estocástico, ya que determina su cambio de estado infinitesimal esperado; mientras que su martingala de Dynkin contendrá la información del ruido de su evolución temporal.

Es fácil ver que para un proceso de saltos markoviano de tasas de transición q_{ij} , la deriva vendrá dada por (para $i \in E$)

$$\delta(i) = \sum_{j \neq i \in E} q_{ij}(f(j) - f(i)) \quad (3.2)$$

Martingalas multiplicativas

De manera similar al caso de las martingalas aditivas, se puede definir una familia de martingalas asociadas a un proceso de Markov que compartan una propiedad de multiplicatividad.

Definición 3.1.3. Sea $f : E \rightarrow \mathbb{R}_{>0}$ en $\mathcal{D}(L)$ tal que $f(X_t)$ es integrable. Entonces, para todo $t \geq u$ (donde $u \geq 0$ está fijo), el proceso

$$R_t^u(f) := \frac{f(X_t)}{f(X_u)} e^{-\int_u^t Lf(X_s)/f(X_s)ds}$$

será denominado *martingala multiplicativa* de X_t .

Su nombre deriva del hecho de que, para todo $0 \leq u \leq v \leq t$, $R_v^u(f)R_t^v(f)$ es igual a $R_t^u(f)$. Al igual que con las martingalas aditivas, generalmente omitiremos el superíndice y asumiremos que la martingala tiene como tiempo de inicio a 0. El siguiente lema [SW95] demuestra que efectivamente son martingalas.

Lema 3.1.3. *Sea X_t y f como antes, entonces $R_t(f)$ es una martingala.*

Demostración. Que $R_t(f)$ sea integrable y \mathcal{F}_t -adaptada están nuevamente garantizados por hipótesis. Nos queda mostrar que para todo $0 \leq s < t$, $\mathbb{E}(R_t(f)|\mathcal{F}_s) = R_s(f)$. Ahora, por la propiedad de multiplicatividad tenemos que $\mathbb{E}(R_t(f)|\mathcal{F}_s) = R_s(f)\mathbb{E}(R_t^s(f)|\mathcal{F}_s)$. Debido a la propiedad de Markov, sólo necesitamos mostrar que para $t \geq 0$ $\mathbb{E}(R_t(f)) = 1$. Para esto, definimos un operador lineal (para $t \geq 0$) $\tilde{P}(t)$ de acuerdo a

$$\tilde{P}(t)f(i) := \mathbb{E}_i \left(f(X_t) e^{-\int_0^t Lf(X_s)/f(X_s)ds} \right).$$

Es fácil verificar que los operadores $\tilde{P}(t)$ son continuos y forman un semi-grupo de transición. Se puede ver que el generador infinitesimal asociado a este semi-grupo $\tilde{L}(\cdot)$ es nulo (es decir, $\tilde{L} = 0$). Luego, por (3.1) tendremos que (para todo $i \in E$ y $t \geq 0$)

$$f(i) = \mathbb{E}_i \left(f(X_t) e^{-\int_0^t Lf(X_s)/f(X_s) ds}, \right), \quad (3.3)$$

como queríamos probar. \square

Estas martingalas son útiles para derivar desigualdades de grandes desvíos, ya que pueden usarse para definir cambios de medida que preserven la propiedad de Markov. Además, como veremos más adelante, también se pueden usar para estimar momentos exponenciales de los procesos y, de ese modo, derivar cotas exponenciales para las fluctuaciones del proceso alrededor de su límite de escala.

3.2. Convergencia de procesos de dimensión finita

En esta sección demostraremos un teorema que servirá para probar límites de escala para procesos de saltos markovianos de dimensión finita. Aunque las pruebas son casi idénticas en dimensiones más grandes, nos centraremos por simplicidad en el caso unidimensional. Como antes, asumiremos que los procesos en cuestión son reales y tienen un espacio de estados contable. Sólo probaremos el teorema para sucesiones de procesos que cumplan una condición global de Lipschitz. Si bien en la mayoría de las situaciones prácticas esta es una condición demasiado restrictiva, la prueba del teorema ayuda a la comprensión de los principios subyacentes y será suficiente para el presente trabajo.

3.2.1. Deriva globalmente Lipschitz

En esta sección vamos a determinar condiciones suficientes para que una sucesión de procesos de saltos markovianos $(X_t^{(n)})_{t \geq 0}$ converja a un límite determinístico. Denotaremos por $q_{ij}^{(n)}$ a la matriz de transición del n -ésimo proceso estocástico y por $\delta^{(n)}(\cdot)$ a su deriva correspondiente. Además, asumiremos (por simplicidad) que todos los procesos de la sucesión tienen como espacio de estados a \mathbb{N} .

Definimos procesos de Markov reescalados como $\bar{X}_t^{(n)} := \frac{X_{\lfloor tn \rfloor}^{(n)}}{n}$ y sus derivas correspondientes como $\bar{\delta}^{(n)}(\cdot)$ (donde las extendemos continuamente a \mathbb{R}). Además, supongamos que las derivas son tales que $\bar{\delta}^{(n)} \rightarrow \delta$ (donde δ es alguna función localmente Lipschitz). Por supuesto, en muchas situaciones prácticas, el escalamiento correcto de los procesos para el cual se obtiene un límite de escala no trivial puede no ser lineal en el tiempo y el espacio; pero aquí nos limitaremos a las situaciones en las que esta es la escala correcta.

En este caso, queremos probar que $\bar{X}_t^{(n)}$ converge en algún sentido a $y(t)$. Donde $y(t)$ es la solución a la ecuación diferencial

$$\dot{y} = \delta(y),$$

con condición inicial $y(0) = x_0$; donde x_0 es tal que $\bar{X}_0^{(n)} \rightarrow x_0$. Para esto definiremos la distancia uniforme esperada como

$$\bar{E}_n(t) := \mathbb{E} \left(\sup_{0 \leq s \leq t} |X_s^{(n)} - y(s)| \right).$$

En la siguiente proposición daremos condiciones suficientes para que haya convergencia en este sentido.

Proposición 3.2.1. *Sea $X_t^{(n)}$ una sucesión de procesos de Markov como la descrita arriba tal que también cumple que:*

- (I) *para algún $x_0 \in \mathbb{R}$, $\bar{X}_0^{(n)} \rightarrow x_0$.*
- (II) *para todo $n \geq 1$, $u, v \in \mathbb{R}$, la siguiente condición global de Lipschitz $|\bar{\delta}^{(n)}(u) - \delta(v)| \leq K|u - v| + \eta_n$ se cumple para alguna constante $K > 0$ y una sucesión $(\eta_n)_{n \geq 1}$ que tiende a 0,*
- (III) *para todo $i \in E$, $\frac{1}{n^2} \mathbb{E}_i \left(\int_0^{nt} \sum_{j \in E} q_{X_s^{(n)}, j}^{(n)} |j - X_s^{(n)}|^2 ds \right) \leq \epsilon_n$ para alguna otra sucesión $(\epsilon_n)_{n \geq 1}$ que también tiende a 0.*

Entonces tendremos que $\bar{E}_n(t) \leq (|\bar{X}_0^{(n)} - y(0)| + \eta_n t + 2\sqrt{\epsilon_n})e^{Kt}$; y por lo tanto, $\bar{X}_t^{(n)} \xrightarrow{L^1} y(t)$, uniformemente en conjuntos compactos.

Demostración. Si definimos $M_t^{(n)}$ como la martingala de Dynkin asociada a $X_t^{(n)}$, podemos escribir

$$\bar{X}_t^{(n)} = \bar{X}_0^{(n)} + \frac{M_{[nt]}^{(n)}}{n} + \int_0^t \bar{\delta}^n(\bar{X}_s^{(n)}) ds. \quad (3.4)$$

Usando Cauchy-Schwarz y la desigualdad de Doob obtenemos que

$$\mathbb{E} \left[\left(\sup_{0 \leq s \leq t} |M_{[ns]}^{(n)}|/n \right)^2 \right] \leq \frac{1}{n^2} \mathbb{E} \left(\sup_{0 \leq s \leq t} (M_{[ns]}^{(n)})^2 \right) \leq \frac{4}{n^2} \mathbb{E} \left((M_{[nt]}^{(n)})^2 \right) \leq 4\epsilon_n.$$

Donde usamos que

$$\frac{1}{n^2} \mathbb{E} \left((M_{[nt]}^{(n)})^2 \right) = \frac{1}{n^2} \mathbb{E} \langle M_{[nt]}^{(n)} \rangle = \frac{1}{n^2} \mathbb{E} \left(\int_0^{nt} \sum_{j \in E} q_{X_s^{(n)}, j}^{(n)} (j - X_s^{(n)})^2 ds \right) \leq \epsilon_n,$$

con $\langle M_t^{(n)} \rangle$ la variación cuadrática de $M_t^{(n)}$. Por otro lado, la función $y(t)$ es tal que

$$y(t) = y(0) + \int_0^t \delta(y(s)) ds. \quad (3.5)$$

Si hacemos (3.4) menos (3.5), tomamos valor absoluto y luego supremo sobre $[0, t]$ y valor medio, llegamos a que

$$\begin{aligned} \bar{E}_n(t) &\leq |\bar{X}_0^{(n)} - y(0)| + 2\sqrt{\epsilon_n} + \int_0^t |\bar{\delta}^n(\bar{X}_s^{(n)}) - \delta(y(s))| ds, \\ &\leq |\bar{X}_0^{(n)} - y(0)| + 2\sqrt{\epsilon_n} + \eta_n t + K \int_0^t |\bar{X}_s^{(n)} - y(s)| ds. \end{aligned}$$

Y entonces

$$\bar{E}_n(t) \leq |\bar{X}_0^{(n)} - y(0)| + 2\sqrt{\epsilon_n} + \eta_n t + K \int_0^t \bar{E}_n(s) ds.$$

Finalmente, terminamos la prueba por medio de la desigualdad de Gronwall. □

Notar que en la prueba la condición global de Lipschitz puede ser reemplazada fácilmente por el requisito de que se cumpla sólo en un intervalo cerrado y que la sucesión de procesos c.s. no se escape de este antes del tiempo t . Además, la escala correcta del espacio y el tiempo puede no ser lineal en n , pero en ese caso la prueba sería completamente análoga.

3.2.2. Ejemplo de límite de escala

Ahora consideraremos una aplicación fácil de la Proposición 3.2.1. Sean $(N_t^{(n)})_{t \geq 0}$ y $(\tilde{N}_t^{(n)})_{t \geq 0}$ dos sucesiones de procesos de Poisson homogéneos independientes y de intensidades λ_n y μ_n , de modo tal que $\lambda_n \xrightarrow{n \rightarrow \infty} \lambda$ y $\mu_n \xrightarrow{n \rightarrow \infty} \mu$. Ahora, definamos una sucesión de procesos dada por la diferencia de ellos como $X_t^{(n)} := N_t^{(n)} - \tilde{N}_t^{(n)}$.

Como antes, consideraremos los procesos reescalados según $\bar{X}_t^{(n)} := \frac{X_{\lfloor tn \rfloor}^{(n)}}{n}$. Necesitamos calcular las derivas $\bar{\delta}^{(n)}(\cdot)$ de estos procesos. Para ello, vamos a tener en cuenta que las transiciones de las dos sucesiones de procesos de Poisson son $q_{ij}^{(n)} = \lambda_n \delta_{j(i+1)}$ y $\tilde{q}_{ij}^{(n)} = \mu_n \delta_{j(i+1)}$, y por lo tanto sus derivas son constantes y están dadas por $\delta^{(n)}(i) = \lambda_n$ y $\tilde{\delta}^{(n)}(i) = -\mu_n$. Esto implica que $\bar{\delta}^{(n)}(i) = \lambda_n - \mu_n \xrightarrow{n \rightarrow \infty} \lambda - \mu$, que son (obviamente) globalmente Lipschitz.

Debido a que convergen a $\lambda - \mu$, satisfacen trivialmente la condición (ii) de la Proposición 3.2.1. Además, las variaciones cuadráticas de sus martingalas de Dynkin escaladas estarán dadas por

$$\mathbb{E} \left(\int_0^{nt} (\lambda_n - \mu_n) ds \right) = (\lambda_n - \mu_n)nt,$$

lo que implica que la condición (iii) también se cumple. Luego, si las condiciones iniciales de los procesos se fijan de manera que $\bar{X}_0^{(n)} \xrightarrow{n \rightarrow \infty} x_0$, se cumplen todas las hipótesis de la Proposición 3.2.1 y entonces (para todo $T > 0$)

$$\bar{X}_t^{(n)} \xrightarrow{L^1} y(t)$$

en $[0, T]$. Donde $y(t) = (\lambda - \mu)t + x_0$; o sea, es la solución de la ecuación $\dot{y} = \lambda - \mu$ con condición inicial $y(0) = x_0$.

3.3. Convergencia de procesos de dimensión numerable

En esta sección presentamos los resultados que usaremos en el Capítulo 4 para mostrar la convergencia de procesos y tiempos de parada hacia límites determinísticos. Como en el resto del capítulo, estos resultados no se presentan necesariamente de manera general, sino en la forma más conveniente para el resto del trabajo.

Aunque, en principio, un resultado similar al de la Proposición 3.2.1 puede derivarse para procesos de dimensión numerable siguiendo un razonamiento análogo, en muchas situaciones prácticas no es posible darle una norma al espacio de estados para la cual las derivas involucradas sean globalmente Lipschitz. En lugar de hacer esto, un enfoque alternativo es seguir la metodología de Prohorov. La idea es demostrar que cada subsucesión de la sucesión de procesos tiene una subsucesión convergente, y que todas ellas tienden al mismo límite. Para lograr esto, primero se prueba que la sucesión es apretada, estableciendo así para cada subsucesión una subsucesión convergente.

Este enfoque clásico se sigue, por ejemplo, en la prueba de los límites en el teorema principal de [BJL16]. El siguiente lema es una adaptación de estos resultados.

Lema 3.3.1. *Sea $(X_t^{(n)}(1), X_t^{(n)}(2), \dots) \in D[0, \infty)^{\mathbb{N}}$ una sucesión numerable de procesos de saltos markovianos donde (para todo $k \in \mathbb{N}$) $X_t^{(n)}(k)$ tiene deriva $\delta_k[X_t^{(n)}(1), X_t^{(n)}(2), \dots]$. Si (para todo $k \in \mathbb{N}$):*

- (I) $\delta_k[X_t^{(n)}(1), X_t^{(n)}(2), \dots] = \sum_{i \geq 1}^{i_k} \alpha_i(k) X_t^{(n)}(i)$, donde $i_k \in \mathbb{N}$ y $\alpha_i(k) \in \mathbb{R}$,
- (II) $\delta_k[X_t^{(n)}(1), X_t^{(n)}(2), \dots]/n$ es uniformemente acotada,
- (III) $X_0^{(n)}(k)/n \xrightarrow{n \rightarrow \infty} y_0(k)$ (para alguna constante $y_0(k)$),
- (IV) y las martingalas de Dynkin asociadas $M_t^{(n)}(k)$ tienen variación cuadrática de orden $o_{\mathbb{P}}(n^2)$

entonces, si además el sistema de ecuaciones integrales definido por

$$\begin{aligned} y_t(1) &= y_0(1) + \int_0^t \delta_1(y_s(1), y_s(2), \dots) ds \\ y_t(2) &= y_0(2) + \int_0^t \delta_2(y_s(1), y_s(2), \dots) ds \\ &\dots \end{aligned}$$

tiene solución única, los procesos $X_t^{(n)}(1), X_t^{(n)}(2), \dots$ convergen en probabilidad hacia las funciones continuas $y_t(1), y_t(2), \dots$ que son solución de este sistema.

Demostración. Dividiendo la fórmula de Dynkin por n obtenemos que

$$\begin{aligned} \frac{X_t^{(n)}(1)}{n} &= \frac{X_0^{(n)}(1)}{n} + \int_0^t \frac{\delta_1^{(n)}[X_s^{(n)}(1), X_s^{(n)}(2), \dots]}{n} ds + \frac{M_t^{(n)}(1)}{n} \\ \frac{X_t^{(n)}(2)}{n} &= \frac{X_0^{(n)}(2)}{n} + \int_0^t \frac{\delta_2^{(n)}[X_s^{(n)}(1), X_s^{(n)}(2), \dots]}{n} ds + \frac{M_t^{(n)}(2)}{n} \\ &\dots \end{aligned}$$

Como estas martingalas de Dynkin tienen variación cuadrática de orden $o_{\mathbb{P}}(n^2)$, por la desigualdad de Doob obtenemos que, para todo $k \in \mathbb{N}$, $\sup_{s \leq t} |M_t^{(n)}|/n$ converge débilmente en $D[0, \infty)$ a 0. Y como $\delta_k^{(n)}(\cdot)/n$ son uniformemente acotadas, para todo $k \in \mathbb{N}$ y $T > 0$, la sucesión de procesos $(X_t^{(n)}(k) - M_t^{(n)})/n$ es uniformemente de Lipschitz y uniformemente acotada en $[0, T]$. Por lo tanto, por el Teorema de Arzela-Ascoli, para todo $T > 0$ estas familias de procesos son apretadas en $C[0, T]$. Esto implica que para toda subsucesión $(n_k)_{k \geq 1}$ existe una subsucesión $(n_{k_l})_{l \geq 1}$ tal que

$$\begin{aligned} \frac{X_t^{(n_{k_l})}(1) - M_t^{(n_{k_l})}(1)}{n_{k_l}} &\Rightarrow y_t(1) \\ \frac{X_t^{(n_{k_l})}(2) - M_t^{(n_{k_l})}(2)}{n_{k_l}} &\Rightarrow y_t(2) \\ &\dots \end{aligned}$$

en conjuntos compactos, para algunas funciones continuas $y_t(1)$, $y_t(2)$, etc. Como los procesos son contables, podemos tomar una subsubsucesión común donde se den todas estas convergencias simultáneamente. Además, según el Teorema de representación de Skorohod (Teorema C.2.1), existe un espacio de probabilidad tal que todos estos límites y la convergencia (para todo $k \in \mathbb{N}$) de $\sup_t |M_t^{(n)}(k)|/n$ hacia 0 se dan casi seguramente en conjuntos compactos. Entonces, para esta subsubsucesión

$$\begin{aligned} \frac{X_t^{(n_{k_l})}(1)}{n_{k_l}} &\xrightarrow{a.s.} y_t(1) \\ \frac{X_t^{(n_{k_l})}(2)}{n_{k_l}} &\xrightarrow{a.s.} y_t(2) \\ &\dots \end{aligned}$$

Y por hipótesis $\delta_k^{(n)}[X_s^{(n)}(1), X_s^{(n)}(2), \dots]/n = \sum_{i \geq k} \alpha_i(k) X_t^{(n)}(i)/n$, entonces

$$\lim_l \delta_k \left[X_t^{(n_{k_l})}(1), X_t^{(n_{k_l})}(2), \dots \right] / n_{k_l} = \delta_k(y_t(1), y_t(2), \dots).$$

Y debido a que todas las derivas sobre n están acotadas de manera uniforme, acotando esta subsubsucesión y usando que $X_t^{(n)}(k)/n \rightarrow y_0(k)$ y convergencia dominada

$$\begin{aligned} y_t(1) &= y_0(1) + \int_0^t \delta_1(y_s(1), y_s(2), \dots) ds \\ y_t(2) &= y_0(2) + \int_0^t \delta_2(y_s(1), y_s(2), \dots) ds \\ &\dots \end{aligned}$$

La convergencia hacia la solución de este sistema de ecuaciones integrales está bien definida ya que, por hipótesis, tiene solución única¹. Ahora tenemos que mostrar que esta convergencia no es sólo en esta subsubsucesión sino más bien en toda la sucesión original. Para esto, vamos a tener en cuenta que debido a que cada subsubsucesión tiene una subsubsucesión que converge al mismo límite (porque por hipótesis el sistema tiene solución única), la sucesión original converge débilmente a ella. Además, como el límite es determinístico, la convergencia se puede considerar equivalentemente en probabilidad. \square

3.3.1. Convergencia de tiempos de parada y valores terminales

En nuestras aplicaciones será importante poder, no sólo determinar la convergencia de los procesos, sino que en muchas situaciones tendremos que establecer la convergencia de ciertos tiempos de parada y de los procesos detenidos de acuerdo a ellos.

En el siguiente lema, estableceremos criterios de convergencia para los tiempos de parada de sucesiones de procesos decrecientes que convergen a un límite de escala.

Lema 3.3.2. *Sea $(X_t^{(n)}(1), X_t^{(n)}(2), \dots) \in D[0, \infty)^\mathbb{N}$ una sucesión de procesos de saltos markovianos donde $X_t^{(n)}(1)$ es decreciente y tiene tasas de transición $q_{ij}^{(n)}$ y definamos el tiempo de parada $T^{(n)} := \inf\{t \geq 0 : X_t^{(n)}(1) = 0\}$ y el tiempo determinístico $T := \inf\{s \geq 0 : y_s(1) = 0\}$. Bajo las mismas hipótesis del lema anterior y asumiendo también que:*

¹Observemos que, en general, esto no es una declaración trivial para un sistema de dimensión infinita y tiene que ser probado por separado.

(I) para todo $t \leq T^{(n)}$ y (si $X_t^{(n)}(1) = i \geq 0$), $\sum_{j \leq i} q_{ij}^{(n)}(X_t^{(n)}(2), \dots) \geq C^{(n)}n$ con $C^{(n)} \xrightarrow{n \rightarrow \infty} C > 0$,

(II) la función $y_t(1)$ es continuamente diferenciable con $\dot{y}_t(1) \leq -C'$ (para algún $C' > 0$ y $t \leq T$),

entonces $T^{(n)} \xrightarrow{\mathbb{P}} T$.

Demostración. Dado $\delta > 0$, queremos probar que la probabilidad del evento $\{|T - T^{(n)}| \geq \delta\}$ tiende a 0. Para ello, supongamos que $T > T^{(n)}$. Ahora, supongamos también que el evento $\{\sup_{t \leq T} |X_t^{(n)}(1)/n - y_t(1)| \leq \delta C'\}$ vale. Entonces, $y_{T^{(n)}}(1)$ vale a lo sumo $\delta C'$. Por hipótesis tenemos que para todo $t \leq T$ la derivada de $y_t(1)$ es menor a $-C'$, entonces $y_{T^{(n)}+\delta}(1) \leq y_{T^{(n)}}(1) - C'\delta \leq 0$. Por lo tanto, $\{T - T^{(n)} \geq \delta\} \subseteq \{\sup_{t \leq T} |X_t^{(n)}(1)/n - y_t(1)| \geq \delta C'\}$. Y porque $X_t^{(n)}(1)/n \xrightarrow{\mathbb{P}} y_t(1)$, la probabilidad de este último evento tiende a 0.

Por otro lado, ahora supongamos que $T^{(n)} > T$. Si el evento $\{\sup_{t \leq T} |X_t^{(n)}(1)/n - y_t(1)| \leq C\delta/4\}$ vale, el valor de $X_T^{(n)}/n$ será a lo sumo $C\delta/4$. Por hipótesis $X_t^{(n)}(1)$ es decreciente y tiene transiciones a estados menores con tasa mayor a $nC^{(n)} = nC + o(n)$. Si definimos $(Z_t)_{t \geq T}$ como un proceso de muerte puro con valor inicial $Z_T = nC\delta/4$ y tasa de muerte $nC/2$, podemos acoplar (para n mayor a cierto $n_0 \geq 0$) el proceso $X_t^{(n)}$ para que sea menor a Z_t para todo $t \geq T$. Entonces, para $n \geq n_0$, $\{T^{(n)} - T \geq \delta\} \subseteq \{Z_{T+\delta} \geq 0\}$. Pero definiendo $Y \sim \text{Pois}(nC\delta/2)$, la probabilidad de este último evento es igual a $\mathbb{P}(Y \leq nC\delta/4)$ y por la desigualdad de Chebychev obtenemos que

$$\begin{aligned} \mathbb{P}(Y \leq nC\delta/4) &= \mathbb{P}(nC\delta/2 - Y \geq nC\delta/4) \leq \mathbb{P}(|Y - nC\delta/2| \geq nC\delta/4) \\ &\leq \frac{nC\delta/2}{n^2C^2\delta^2/16} = \frac{8}{nC\delta} \end{aligned}$$

En resumen,

$$\begin{aligned} \mathbb{P}(|T^{(n)} - T| \geq \delta) &= \mathbb{P}(T - T^{(n)} > \delta) \\ &\quad + \mathbb{P}(\sup_{t \leq T} |X_t^{(n)}(1)/n - y_t(1)| \leq C\delta/4, T^{(n)} - T > \delta) \\ &\quad + \mathbb{P}(\sup_{t \leq T} |X_t^{(n)}(1)/n - y_t(1)| \geq C\delta/4, T^{(n)} - T > \delta) \\ &\leq \mathbb{P}(\sup_{t \leq T} |X_t^{(n)}(1)/n - y_t(1)| \geq \delta C') \\ &\quad + \mathbb{P}(\sup_{t \leq T} |X_t^{(n)}(1)/n - y_t(1)| \leq C\delta/4, T^{(n)} - T > \delta) \\ &\quad + \mathbb{P}(\sup_{t \leq T} |X_t^{(n)}(1)/n - y_t(1)| \geq C\delta/4) \\ &\leq \mathbb{P}(\sup_{t \leq T} |X_t^{(n)}(1)/n - y_t(1)| \geq \delta C') + \mathbb{P}(Y \leq nC\delta/4) \\ &\quad + \mathbb{P}(\sup_{t \leq T} |X_t^{(n)}(1)/n - y_t(1)| \geq C\delta/4) \xrightarrow{n \rightarrow \infty} 0 \end{aligned}$$

donde el último término $\mathbb{P}(\sup_{t \leq T} |X_t^{(n)}(1)/n - y_t(1)| \geq C\delta/4)$ tiende a 0 porque estamos bajo las hipótesis del lema anterior. \square

También necesitaremos establecer la convergencia de las coordenadas de los procesos de saltos en algunos tiempos de parada específicos hacia los valores correspondientes de sus límites de escala. Para esto, utilizaremos el siguiente corolario.

Corolario 3.3.1. Sea $(X_t^{(n)}(1), X_t^{(n)}(2), \dots) \in D[0, \infty)^{\mathbb{N}}$ como en el lema anterior. Entonces, para todo $k \in \mathbb{N}$ vamos a tener que $X_{T^{(n)}}^{(n)}(k)/n \xrightarrow{\mathbb{P}} y_T(k)$.

Demostración. Tenemos que

$$|X_{T^{(n)}}^{(n)}(k)/n - y_T(k)| \leq |X_{T^{(n)}}^{(n)}(k)/n - y_{T^{(n)}}| + |y_{T^{(n)}} - y_T(k)|$$

donde el primer término del lado derecho tiende c.a.p. a 0 ya que por el Lema 3.3.1 $X_t^{(n)}(k)/n \xrightarrow{\mathbb{P}} y_t(k)$ uniformemente en conjuntos compactos (y como $T^{(n)} \xrightarrow{\mathbb{P}} T$, los $T^{(n)}$ son c.a.p. acotados). Para el segundo término alcanza con notar que tenemos que $T^{(n)} \xrightarrow{\mathbb{P}} T$ por el Lema 3.3.2 y que $y_t(k)$ es continua. \square

Capítulo 4

Conjuntos independientes máximos en grafos aleatorios

Los conjuntos independientes son relevantes en el estudio de diversos modelos físicos y de comunicación. Por ejemplo, en física, donde dinámicas que generan conjuntos independientes se utilizan para estudiar la deposición de partículas en superficies [CAP07, Eva93], así como para modelar el número de excitaciones en gases ultrafríos [SJK15]. Como discutiremos con más detalle en el siguiente capítulo, también se usan procesos estocásticos similares [BJLM14, DvLM16] para modelar teóricamente el posible número de transmisiones simultáneas de información dentro de una red WiFi.

Encontrar conjuntos independientes de tamaño máximo en grafos fijos es un problema de optimización difícil que se ha demostrado es NP-difícil [FM97]. Tanto la caracterización de números de independencia como de algoritmos que construyen conjuntos independientes en tiempos polinómicos en grafos aleatorios raros han recibido mucha atención recientemente, pero son aún problemas abiertos para la mayoría de los casos. Unas de las pocas excepciones son las caracterizaciones de los números de independencia de los grafos regulares de grado grande [DSS16] y la optimalidad del algoritmo grado-egoísta para grafos de Erdős-Rényi raros con grado medio $\lambda < e$ [KS81]¹.

Aquí abordamos la pregunta de si existe alguna clase de grafos aleatorios raros para los que la tarea, a pesar de la complejidad del problema general, se puede realizar en tiempo polinómico. Usando límites de escala en exploraciones secuenciales que seleccionan sólo vértices de grado 1, descomponemos la exploración en diferentes pasos y mostramos que estos pasos pueden describirse como una combinación de dos mapas que actúan sobre la distribución de grados del grafo. Utilizando estos resultados, mostramos que para cierta clase de grafos aleatorios raros, la exploración grado-egoísta es asintóticamente óptima. Primero damos una condición suficiente, que se puede verificar fácilmente en la práctica, para que esta exploración sea óptima *en un solo paso*. Luego mostramos cómo generalizar esta condición suficiente caracterizando el grafo restante después de *varios pasos*. Finalmente, estudiamos el caso de las distribuciones de Poisson (que son asintóticamente equivalentes a los grafos de Erdős-Rényi) y mostramos, de manera alternativa, que cuando el grado medio es menor que e la exploración

¹De hecho, en este trabajo, los autores no abordan directamente la optimalidad de este algoritmo para la construcción de conjuntos independientes máximos. En cambio, lo prueban para apareos, pero el resultado para conjuntos independientes máximos se puede derivar de él.

es asintóticamente óptima. Utilizamos este hecho para dar una nueva caracterización de su número de independencia en estas circunstancias.

El capítulo está estructurado de la siguiente manera. En la Sección 4.1 presentamos la bibliografía existente sobre el problema de conjuntos independientes máximos en especial la concerniente a grafos aleatorios. Luego, en la Sección 4.2, presentamos los principales resultados del capítulo junto con algunas aplicaciones de los mismos. Las pruebas de estos resultados se pueden encontrar en la Sección 4.3, mientras que en la Sección 4.4 proponemos algunas extensiones posibles a los resultados del capítulo.

4.1. Conjuntos independientes y algoritmos secuenciales

Principalmente se han seguido dos tipos de enfoques para obtener conjuntos independientes máximos en grafos aleatorios: por un lado, utilizando dinámicas estocásticas reversibles (generalmente la dinámica de Glauber, discutida con cierto detalle en el Capítulo 5); y por otro lado, usando algoritmos secuenciales. Aquí limitaremos nuestra discusión a este segundo tipo.

4.1.1. Algoritmos secuenciales

El nombre de estos algoritmos proviene del hecho de que exploran los grafos secuencialmente (y por lo tanto terminan la exploración en menos de n pasos). Para esto, en cada paso $k \geq 0$ el conjunto de vértices se dividirá en tres conjuntos: el de los *vértices inexplorados* \mathcal{U}_k , el de los *vértices activos* \mathcal{A}_k y el de los *vértices bloqueados* \mathcal{B}_k . Un algoritmo secuencial típico funciona de la siguiente manera. Inicialmente, establece $\mathcal{U}_0 = V$, $\mathcal{A}_0 = \emptyset$ y $\mathcal{B}_0 = \emptyset$. Para explorar el grafo, en el paso $k + 1$ selecciona un vértice $v_{k+1} \in \mathcal{U}_k$ (teniendo posiblemente en cuenta su grado actual o pasado hacia otros vértices inexplorados) y cambia su estado a activo. Después de esto, toma todos sus vecinos inexplorados, es decir, el conjunto $\mathcal{N}(v_{k+1}) := \{w \in \mathcal{U}_k \mid (v_{k+1}, w) \in E\}$, y cambia sus estados a bloqueado. Esto significa que, si en el paso $k + 1$ se selecciona el vértice v_{k+1} , el conjunto resultante de vértices estará dado por $\mathcal{U}_{k+1} = \mathcal{U}_k \setminus \{v_{k+1} \cup \mathcal{N}(v_{k+1})\}$, $\mathcal{A}_{k+1} = \mathcal{A}_k \cup \{v_{k+1}\}$ y $\mathcal{B}_{k+1} = \mathcal{B}_k \cup \mathcal{N}(v_{k+1})$. Notar que en cada paso, el conjunto de vértices activos define un conjunto independiente. El algoritmo sigue este procedimiento hasta el paso k_n^* en el que todos los vértices están activos o bloqueados (o de manera equivalente $\mathcal{U}_{k_n^*} = \emptyset$). El conjunto de vértices activos en el paso k_n^* define luego un conjunto independiente maximal (aunque, en general, no máximo).

En el *algoritmo egoísta*, durante el paso k se selecciona el vértice a ser activado v_k uniformemente del sub-grafo de los vértices restantes G_k (es decir, el sub-grafo formado por los vértices inexplorados). Si G es un grafo, llamaremos $\sigma_{G_r}(G)$ al tamaño del conjunto independiente obtenido por el algoritmo egoísta ejecutado en G . Este algoritmo ha sido ampliamente estudiado, especialmente en el contexto de los grafos de Erdős-Rényi. Hay muchas maneras de abordar el problema de determinar el valor asintótico del conjunto independiente obtenido por el algoritmo egoísta en un grafo de Erdős-Rényi ralo. Por ejemplo, por medio de un análisis combinatorio [GM75, Kar76] del problema. Más parecido al enfoque del resto del capítulo, en [BJS17] se demuestra un límite de escala para un proceso de Markov unidimensional asociado al algoritmo. Sin embargo, la descripción de este proceso de exploración en un Modelo de Configuraciones no puede describirse por medio de un proceso de Markov unidimensional,

ya que los vértices inexplorados tienen grados (hacia otros vértices inexplorados) que no son intercambiables y que dependen de manera complicada de la evolución del proceso. Esto hace que el análisis sea mucho más complicado que en el caso de un grafo Erdős-Rényi. Ha habido dos trabajos en la literatura que describen un límite de escala para este proceso. En [BJM17], se obtiene un límite de escala para la distribución en grados del grafo restante. Mientras que en [BJL16], se muestra un límite de escala similar pero con una dinámica modificada que permite una simplificación de las ecuaciones diferenciales límites. A través de estos resultados, se determina el tamaño del conjunto independiente obtenido.

El algoritmo *grado-egoísta* es una variación del algoritmo egoísta que tiene en cuenta el grado de los vértices en el grafo restante. Durante el paso k se selecciona un vértice inexplorado v_k uniformemente de entre los vértices de *grado mínimo* (hacia otros vértices inexplorados) dentro del sub-grafo restante G_k . Si G es un grafo, denotaremos por $\sigma_{DG}(G)$ al tamaño (posiblemente aleatorio) de vértices en el conjunto independiente obtenido por el algoritmo grado-egoísta corrido en G . Aunque este algoritmo se ha estudiado en profundidad en la comunidad informática (por ejemplo, en [HR97]), hay pocos resultados matemáticos exactos que caractericen o acoten los conjuntos independientes encontrados por él. Una notable excepción es [KS81, AFP98] donde, aunque se considera un problema completamente diferente (a saber, el de apareos máximos), los resultados implican que el algoritmo grado-egoísta es asintóticamente óptimo para grafos de Erdős-Rényi cuando el grado medio es $\lambda < e$. A este resultado se lo conoce como *fenómeno e*. Los resultados de Wormald [Wor95] también describen un límite de escala para el proceso generado por el algoritmo grado-egoísta cuando se lo corre en un grafo d -regular pero no discuten la optimalidad asintótica.

Algoritmo egoísta en grafos de Erdős-Rényi ralos

Sea $G \sim \text{ER}_n(\lambda/n)$. El análisis de estos grafos es bastante simple debido a la simetría e independencia de las conexiones entre sus vértices. De hecho, el proceso que describe la dinámica del algoritmo egoísta se puede modelar como un proceso unidimensional de Markov, esto es así porque después del paso k del algoritmo, si el número de vértices inexplorados es $Z_k^{(n)}$, entonces el sub-grafo inexplorado sigue siendo un grafo de Erdős-Rényi ralo $G_k \sim \text{ER}_{Z_k^{(n)}}(\lambda/n)$. Además, en un grafo de Erdős-Rényi, el grado de los vértices es intercambiable (cada vértice tiene, en distribución, el mismo número de vecinos) y sólo depende del parámetro λ y el tamaño del grafo.

Entonces se puede probar por medio de límites de escala (ver [BJS17]) que:

$$\mathbb{E} \left(\sup_{t \in [0, T]} \left| \frac{Z_{\lfloor tn \rfloor}^{(n)}}{n} - z(t) \right| \right) \xrightarrow{n \rightarrow \infty} 0,$$

donde $z(t)$ es la solución de la ecuación diferencial $\dot{z} = -\lambda(1 - z) - 1$ con condición inicial $z(0) = 1$.

Luego, se puede demostrar que en este caso $\sigma_{Gr}(G^{(n)})$ converge en L^1 a τ^* (definido como la solución más pequeña de $z(\tau^*) = 0$), y por lo tanto el número de estancamiento de grafos de Erdős-Rényi ralos está arbitrariamente cerca de $\tau^* = \frac{\log(1+\lambda)}{\lambda}$. En [BJS17] también se prueba un teorema central del límite para esta convergencia.

Exploración egoísta en Modelos de Configuraciones

Los procesos de exploración de grafos del Modelo de Configuraciones no pueden describirse simplemente por el número de vértices en el sub-grafo inexplorado. Sin embargo, como se muestra en [BJM17], las propiedades del Modelo de Configuraciones permiten la descripción de la evolución del algoritmo egoísta por medio de un proceso markoviano sin seguir exactamente la evolución de toda la estructura del grafo, sino sólo de la distribución de grados de los vértices inexplorados hacia otros inexplorados. Esto se hace construyendo simultáneamente el grafo aleatorio y el proceso de exploración asociado. Como veremos, esto hace posible el cálculo de la constante de estancamiento resultante.

En contraste con la discusión anterior, ahora será conveniente describir el algoritmo como un proceso a tiempo continuo en lugar de uno discreto. Para esto, le asociaremos a cada vértice inexplorado un reloj exponencial aleatorio de tasa 1. Cada vértice inexplorado se activará cuando suene su reloj, en cuyo momento sus medias aristas son apareadas y sus vecinos inexplorados son bloqueados.

Denotando por $\mu_t^{(n)}(i)$ ($i = 1, \dots, n$) el número de vértices inexplorados de grado i hacia otros vértices inexplorados (cuyo valor inicial $\mu_0^{(n)}(i)$ es simplemente cuántos elementos en $\bar{d}^{(n)}$ son iguales a i), la evolución temporal de esta medida empírica es suficiente para caracterizar el proceso de exploración en grafos del Modelo de Configuraciones [BJM17]. Para ello, el proceso de saltos de Markov asociado debe escalarse de acuerdo a

$$\bar{\mu}_t^{(n)}(i) = \frac{1}{n} \mu_t^{(n)}(i), \quad t \geq 0,$$

para todo $t \geq 0$ y $i = 1, \dots, n$.

Luego, bajo el supuesto de que la distribución de grados inicial satisface la (CC), y llevando n al infinito, se puede probar un límite de escala para la evolución de esta distribución de grados [BJM17]. En este trabajo también se muestra que el límite es único y que viene dado por la solución de un sistema no-lineal de ecuaciones diferenciales de dimensión numerable.

Para expresar estas ecuaciones en forma concisa, llamaremos

$$\alpha_t(i) := \frac{\bar{\mu}_t(i)}{\sum_{j=0}^{\infty} \bar{\mu}_t(j)}, \quad (4.1)$$

$$\beta_t(i) := \frac{i \bar{\mu}_t(i)}{\sum_{j=0}^{\infty} j \bar{\mu}_t(j)}, \quad (4.2)$$

para todo $t \geq 0$ y $i \geq 1$. Las medidas de probabilidad $\alpha_t(\cdot)$ y $\beta_t(\cdot)$ tienen interpretaciones intuitivas: la primera describe la distribución de grados de un vértice inexplorado elegido al azar (y uniformemente) en el tiempo t ; mientras que la segunda da la distribución sesgada por tamaño asociada a $\alpha_t(\cdot)$ y representa la distribución de grados de un vecino de un vértice elegido al azar. Esta interpretación será útil al generalizar estos resultados en el siguiente capítulo.

Luego, el sistema de EDO que describe la evolución del número de vértices inexplorados con grado igual a i hacia otros inexplorados puede escribirse como

$$\frac{d}{dt} \bar{\mu}_t(i) = - \sum_{j=0}^{\infty} \bar{\mu}_t(j) [\alpha_t(i) + F_t(i)], \quad (4.3)$$

donde $F_t(i) := \sum_{k=0}^{\infty} k\alpha_t(k) (\beta_t(i+1) + (\beta_t(i) - \beta_t(i+1)) \sum_{l=0}^{\infty} l\beta_t(l))$. Estas ecuaciones también tienen una interpretación simple. La primera suma en el lado derecho es la tasa de transición total del proceso de Markov. El primer término entre corchetes representa la activación de un vértice de grado i en el momento t , en cuyo caso el número de vértices inexplorados con grado igual a i disminuye en uno. Mientras que el término $F_t(i)$ se corresponde con el bloqueo de vecinos del vértice seleccionado y la eliminación de sus aristas del sub-grafo inexplorado. Observar que el vértice activado tiene una distribución de grados $\alpha_t(\cdot)$, mientras que los vecinos bloqueados y sus vecinos tienen una distribución de grados $\beta_t(\cdot)$. Esta representación será útil para analizar los algoritmos grado-conscientes presentados en el Capítulo 5.

La principal consecuencia de este límite es la caracterización de la proporción asintótica de vértices en el conjunto independiente obtenido por el algoritmo egoísta corrido en grafos del Modelo de Configuraciones [BJM17]:

Teorema 4.1.1. *Para todo $n \geq 1$, sea $G^{(n)} \sim CM_n(d^{(n)})$. Bajo ciertas condiciones sobre los momentos (ver Teorema 3.1 en [BJM17]), se cumple la siguiente convergencia:*

$$\mathbb{E}(|\sigma_{Gr}(G^{(n)}) - c_{\bar{\mu}_0}|) \xrightarrow{n \rightarrow \infty} 0,$$

donde

$$c_{\bar{\mu}_0} = \int_0^{\infty} \sum_{i=0}^{\infty} \bar{\mu}_t(i) dt. \quad (4.4)$$

En un trabajo posterior [BJL16] se muestra un resultado más explícito para la constante de atascamiento por medio de una dinámica modificada que permite la simplificación del sistema de ecuaciones diferenciales límites. En este trabajo, los autores estudiaron un límite de escala diferente que da como resultado ecuaciones más simples que pueden integrarse directamente:

Teorema 4.1.2. *Bajo la (CC), y llamando τ_{∞} al único valor en $(0, \infty]$ tal que*

$$\int_0^{\tau_{\infty}} \frac{e^{-2h}}{\sum_{i=0}^{\infty} i\bar{\mu}_0(i)e^{-ih}} dh = 1,$$

entonces

$$\sigma_{Gr}(G^{(n)}) \xrightarrow[n \rightarrow \infty]{\mathbb{P}} \int_0^{\tau_{\infty}} \frac{e^{-2h} \sum_{i=0}^{\infty} \bar{\mu}_0(i)e^{-ih}}{\sum_{i=0}^{\infty} i\bar{\mu}_0(i)e^{-ih}} dh. \quad (4.5)$$

Algoritmo grado-egoísta en grafos regulares

El resultado principal en esta dirección es mostrado por Wormald en [Wor95]. En este trabajo, se muestra un límite de escala para el proceso generado por el algoritmo grado-egoísta corrido en un grafo d -regular (un grafo construido por el Modelo de Configuraciones donde cada vértice tiene el mismo grado $d \geq 1$). También se proporcionan estimaciones numéricas basadas en este límite de escala para diferentes valores de d en [Wor95].

4.1.2. Caracterización de conjuntos independientes máximos

En el caso de los grafos de Erdős-Rényi, sus números de independencia se entienden mejor en el caso de probabilidades de conexión grandes que en el caso ralo. En este caso, un argumento de segundo momento [Bol01] muestra su convergencia en probabilidad hacia $-\frac{2 \log np}{\log q}$. En este mismo contexto, también se sabe que la relación entre el número de independencia y el tamaño asintótico del conjunto independiente obtenido por el algoritmo egoísta es $1/2$. En el caso de grafos de Erdős-Rényi ralos, como consecuencia de los resultados probados en [BE76, Fri90], se pueden mostrar resultados similares (pero más débiles). Estos muestran que el algoritmo egoísta tiene, para G un grafo de Erdős-Rényi ralo de grado medio λ , una relación de rendimiento $\sigma_{Gr}(G)/\alpha(G)$ que es asintóticamente $1/2$ para n y λ grandes. Como se mencionó anteriormente, de [KS81, AFP98] se desprende que el algoritmo grado-egoísta encuentra asintóticamente c.s. conjuntos independientes máximos en grafos de Erdős-Rényi de grado medio menor que e . En los mismos trabajos, también caracterizan el número de independencia asintótica de estos grafos como la combinación de las raíces de ciertas funciones.

La prueba de la existencia de una proporción de independencia límite para grafos aleatorios d -regulares se da en [BGT10]. En [LW07], Wormald muestra un límite inferior para el número de independencia de un grafo d -regular. En el mismo trabajo, Wormald (y Gamanik y Goldberg independientemente en [GG10]) muestran que la proporción de vértices en el conjunto independiente encontrado por un algoritmo egoísta en un grafo d -regular es (para $d \geq 3$), asintóticamente (en probabilidad) y para valores de cintura² $g \rightarrow \infty$, dada por una cierta función de d . Además, este límite también es demostrado en [Bol81, McK87] y otras cotas inferiores son probadas en [FL92]. Finalmente, en un trabajo reciente [DSS16], se caracteriza el límite del número de independencia de grafos d -regulares por medio de la solución de una ecuación polinómica.

4.2. Resultados y aplicaciones

En esta sección, primero establecemos un lema de optimalidad para grafos deterministas. Luego, enunciaremos resultados equivalentes para grafos aleatorios, para los casos en que una exploración secuencial sólo selecciona vértices de grado 1. Por medio de este resultado, mostramos que esta propiedad es satisfecha por una amplia clase de grafos aleatorios (incluidos todos los grafos estrictamente subcríticos). Finalmente, obtenemos una caracterización de esta clase de grafos.

4.2.1. Primera caracterización de la optimalidad del algoritmo grado-egoísta

Criterio para grafos generales

En el Lema 4.2.1 presentado a continuación, mostramos que cualquier algoritmo secuencial que solamente selecciona vértices de grado 1 (o 0) es óptimo en el sentido de que encuentra un conjunto independiente máximo. Una versión análoga de este lema para el problema de apareos se presenta en [KS81].

Para enunciar esto de forma más precisa, necesitamos introducir una serie de definiciones.

²Longitud del ciclo más corto.

Definición 4.2.1. Dado un grafo $G = (V, E)$, llamaremos *secuencia de selección* a toda secuencia finita de vértices distintos $W = \{w_1, w_2, \dots, w_m\}$ (con $m \leq n$) tal que ningún vértice en W es vecino de otro y todo vértice en V está o bien en W o es vecino de un vértice en W .

Notar que las condiciones en esta definición aseguran que los vértices en W definen un conjunto independiente maximal. Entonces, por definición, los algoritmos secuenciales definen secuencias de selección aleatorias.

Definición 4.2.2. Sea $W = \{w_1, \dots, w_m\}$ una secuencia de selección. Entonces, para todo $1 \leq i \leq m$, vamos a llamar *i -ésimo grafo restante* al sub-grafo formado por los vértices que no están en $\{w_1, \dots, w_i\}$ ni son vecinos de algún vértice que lo está. A este lo simbolizaremos como G_i , y definimos $G_0 := G$.

Cuando no hay ambigüedad sobre a qué valor de i corresponde el grafo restante, simplemente lo llamamos *grafo restante*. Al analizar el algoritmo grado-egoísta, los grafos restantes se referirán a los grafos restantes con respecto a la secuencia de selección definida por el algoritmo. Por supuesto, como una secuencia de selección $W = \{w_1, \dots, w_m\}$ siempre determina un conjunto independiente maximal, $G_m = \emptyset$.

Un algoritmo grado-egoísta corrido en un grafo finito G puede considerarse como una secuencia de selección aleatoria W_{DG} , construida inductivamente de la siguiente manera: dado $\{w_1, \dots, w_k\}$ (los primeros $k \geq 1$ vértices de W_{DG}), w_{k+1} es un vértice elegido uniformemente entre los vértices de menor grado de G_k .

Definición 4.2.3. Sea W una secuencia de selección. Decimos que W tiene la propiedad T_1 si por cada $1 \leq i \leq m$ el grado de w_i en G_{i-1} es igual o menor que 1.

Luego, una secuencia de selección tiene la propiedad T_1 si en cada paso *selecciona* un vértice que tiene grado 0 o 1 en el grafo restante correspondiente.

Ahora podemos enunciar el lema de optimalidad.

Lema 4.2.1. *Sea G un grafo finito y W una secuencia de selección de G . Entonces, si W tiene la propiedad T_1 , $|W| = \alpha(G)$.*

Por medio de este lema podemos establecer una condición suficiente para que el algoritmo grado-egoísta encuentre c.a.p. un conjunto independiente que contenga asintóticamente la misma proporción de vértices que uno máximo:

Proposición 4.2.1 (Condición de optimalidad asintótica en el MC). *Sea $G \sim CM_n(\bar{d}^{(n)})$ una sucesión de grafos que satisfagan la (CC) hacia una distribución límite $(p_k)_{k \in \mathbb{N}}$. Si el algoritmo grado-egoísta define c.a.p. una secuencia de selección que elige sólo vértices de grado 1 o 0 hasta que el grafo restante sea subcrítico y tenga una distribución de grados que sea $\mathcal{O}(e^{-\gamma k})$ (para algún $\gamma > 0$), luego (para todo $\alpha > 0$) $\sigma_{DG}(G) = \alpha(G) + \mathcal{O}_{\mathbb{P}}(n^\alpha)$.*

Esto es así porque un grafo subcrítico se parece (hasta una diferencia lo suficientemente pequeña) a una colección de árboles. Luego podemos acoplar el algoritmo que se ejecuta en el grafo subcrítico con uno que se ejecuta en la colección de árboles de expansión de sus componentes. Este acoplamiento sólo diferirá en las componentes que no sean árboles las cuales, como se muestra en la Proposición 4.3.1, contienen (para todo $\alpha > 0$) $\mathcal{O}_{\mathbb{P}}(n^\alpha)$ vértices

si la distribución del grado límite tiene una cola exponencialmente delgada. Por lo tanto, ambos algoritmos encuentran un conjunto independiente de aproximadamente el mismo tamaño.

En lo que sigue presentaremos criterios explícitos para mostrar que una distribución límite dada cumple con las hipótesis de la Proposición 4.2.1. Primero estableceremos un criterio que puede ser fácilmente aplicado en la práctica. Luego refinaremos este criterio y daremos una forma general de caracterizar las distribuciones de grados para las cuales aplica la proposición.

4.2.2. Criterio de una aplicación

Para caracterizar cuándo el algoritmo grado-egoísta sólo selecciona vértices de grado 1 o 0 hasta que el sub-grafo restante sea subcrítico, será útil dividir la evolución del proceso en intervalos de tiempo discretos para los cuales sabemos con certeza que los únicos vértices seleccionados tienen estos grados.

La observación clave es que si el grafo inicialmente tiene $np_1^{(n)} + o(n)$ vértices de grado 1³, entonces el grado-egoísta seleccionará vértices de grado 1 al menos hasta que haya explorado un número equivalente de vértices de grado 1. Luego definimos el mapa $M_1^{(n)} : \mathbb{R}_{\geq 0}^{\mathbb{N}} \rightarrow \mathbb{R}_{\geq 0}^{\mathbb{N}}$ como el mapa que, cuando se evalúa en una distribución de grados $(p_k^{(n)})_{k \geq 0}$ de un grafo de tamaño n , da la medida de grados normalizada resultante después de que $np_1^{(n)}$ vértices de grado 1 se hayan activado o bloqueado (y sus vecinos bloqueado). Este es en principio un mapa estocástico pero, como demostraremos, la exploración grado-egoísta converge a un límite determinístico lo cual implica que $M_1^{(n)}(\cdot)$ también converge a un mapa límite determinístico $M_1(\cdot)$.

En esta sección, determinamos cuándo la distribución de grados obtenida después de una aplicación del mapa $M_1^{(n)}$ es c.a.p. subcrítica, y por lo tanto se cumple la hipótesis de la Proposición 4.2.1. Nuestro principal resultado aquí es el siguiente teorema:

Teorema 4.2.1. *Dado $G \sim CM_n(\bar{d}^{(n)})$ que cumple la (CC) hacia una distribución límite $(p_k)_{k \geq 0}$ de media $\lambda > 0$. Si*

$$\tilde{\nu} := G_D''(Q)/\lambda < 1 \tag{4.6}$$

donde $Q := (1-p_1/\lambda)$ y $G_D(z)$ es la función generadora de probabilidad de la variable aleatoria límite D ; entonces, (para todo $\alpha > 0$) $\sigma_{DG}(G) = \alpha(G) + \mathcal{O}_{\mathbb{P}}(n^\alpha)$.

4.2.3. Criterio general

El Teorema 4.2.1 establece una condición asintótica para que el grafo restante obtenido después de una aplicación del mapa $M_1^{(n)}(\cdot)$ sea subcrítico, y por lo tanto para que el grado-egoísta sea asintóticamente óptimo. Aquí calculamos la distribución de grados asintótica del grafo restante después de una aplicación de $M_1^{(n)}(\cdot)$ y al hacerlo permitimos el estudio de sucesivas aplicaciones de $M_1^{(n)}(\cdot)$. Esto se puede utilizar para establecer condiciones más generales que determinen la optimalidad asintótica del algoritmo grado-egoísta. Para lograr esto, determinamos un límite de escala para la segunda fase de $M_1^{(n)}(\cdot)$ y resolvemos las ecuaciones obtenidas.

³Por simplicidad, los vértices de grado 0 se omitirán del análisis porque, cuando se seleccionan, no bloquean vértices y entonces no modifican el número de vértices inexplorados de otros grados. Podemos pensar que el algoritmo los selecciona inmediatamente después de que se generan.

Teorema 4.2.2. *Definimos (para todo $i, j \geq 1$) $\eta_j(i) := (-1)^{j_i} \mathbb{I}_{i \leq j}$. Luego, bajo los mismos supuestos del Teorema 4.2.1 y si llamamos $(a_j)_{j \in \mathbb{N}}$ a las componentes de la distribución $(Q^k p_k \mathbb{I}_{\{k \geq 2\}})_{k \in \mathbb{N}}$ en la base $\{\eta_j(\cdot)\}_{j \in \mathbb{N}}$, tenemos que el grafo restante después de una aplicación del mapa es un grafo del Modelo de Configuraciones con una medida de grado normalizada dada por*

$$M_1^{(n)}(p_k^{(n)})(i) \xrightarrow{\mathbb{P}} M_1(p_k)(i) := \sum_{j \geq i} a_j (-1)^{j-i} \tilde{Q}^j \binom{j}{i}, \text{ for } i \geq 1$$

donde $\tilde{Q} := \sum_{i \geq 2} i Q^i p_i / Q^2 \lambda$.

Como se mencionó anteriormente, este resultado permite generalizar la condición de optimalidad dada en el Teorema 4.2.1:

Criterio general de optimalidad asintótica (CG): *dada una distribución de grados límite, si después de un número finito de aplicaciones del mapa $M_1(\cdot)$ la distribución de grados obtenida es subcrítica, entonces el algoritmo grado-egoísta es asintóticamente óptimo para un grafo del Modelo de Configuraciones con esa distribución límite.*

La prueba de este criterio es una consecuencia directa de la Proposición 4.2.1. Entonces, el Teorema 4.2.2 puede usarse para verificarlo. En la siguiente sección, presentamos los cálculos de algunas distribuciones que cumplen con este criterio.

4.2.4. Aplicación a grafos de Erdős-Rényi

Aquí analizamos el caso especial de grafos con distribuciones asintóticas de grados de Poisson. Son de particular importancia porque, por [Kim07], los resultados asintóticos para ellos pueden extenderse directamente a los grafos aleatorios de Erdős-Rényi. Aquí damos una prueba alternativa y más simple del llamado “fenómeno e ”, identificado para apareos en [KS81, AFP98]. También damos una caracterización más explícita del número de independencia asintótica que la presente en estos trabajos: demostramos que ambas constantes en la expresión obtenida en [AFP98] son de hecho iguales y las caracterizamos en términos de la función de Lambert.

Proposición 4.2.2. *Sea $G \sim ER_n(\lambda)$. Si $\lambda < e$, entonces $\sigma_{DG}(G) = \sigma(G) + o_{\mathbb{P}}(n)$; de lo contrario, la secuencia de selección resultante no tiene la propiedad T_1 . Además, en el caso en que $\lambda < e$,*

$$\alpha(G) = n \left(z(\lambda) + \frac{\lambda}{2} z(\lambda)^2 \right) + o_{\mathbb{P}}(n),$$

donde $z(\lambda) := e^{-W(\lambda)}$ con $W(x)$ la función de Lambert.

La prueba de esta proposición se presenta en la Sección 4.3.5 y se deduce de la aplicación del Teorema 4.2.2 y la Proposición 4.2.3, junto con la manipulación de algunos mapas iterativos obtenidos a partir de ellos.

4.2.5. Más aplicaciones

Ahora aplicamos nuestros resultados a las distribuciones de leyes de potencias, que dan lugar a redes libres de escala. Además, explicamos cómo calcular números de independencia utilizando nuestros resultados y probamos cotas superiores generales para ellos.

Distribuciones libres de escala

Aquí estudiaremos el caso en que la distribución de grados corresponde a una ley de potencias de parámetro $\alpha > 3$. Como la función generadora de una distribución de ley de potencias $p_k = C_\alpha k^{-\alpha}$ viene dada por $C_\alpha \text{Li}_\alpha(z)$ (donde $\text{Li}_\alpha(z)$ es el polilogaritmo de orden α):

$$\tilde{\nu}(\alpha) = \frac{(C_\alpha \text{Li}_\alpha(z))''|_{Q(\alpha)}}{\sum_{i \geq 1} i C_\alpha i^{-\alpha}} = \frac{\text{Li}_{\alpha-2}(1 - \zeta(\alpha-1)^{-1}) - \text{Li}_{\alpha-1}(1 - \zeta(\alpha-1)^{-1})}{\zeta(\alpha-1)}.$$

Donde $\zeta(z)$ es la función zeta de Riemann y en la última línea usamos ese $Q(\alpha) = 1 - \zeta(\alpha-1)^{-1}$ y que $\text{Li}_\alpha(z)' = \text{Li}_{\alpha-1}(z)/z$. Se puede ver que esta última expresión es menor que 1 para todo $\alpha > 3$; lo que significa que para toda distribución de ley de potencias con segundo momento finito, el algoritmo grado-egoísta es asintóticamente óptimo. En particular, siempre que tenga segundo momento finito y $\zeta(\alpha-2) > 2\zeta(\alpha-1)$ (o $3 < \alpha \lesssim 3,478$), esta distribución será supercrítica y, sin embargo, el grado-egoísta será asintóticamente óptimo.

Cálculo de números de independencia

Como consecuencia del Teorema 4.2.2, cada vez que una sucesión de grafos cumple el (CG), el número de independencia asintótico puede ser obtenido calculando el número de vértices en el conjunto independiente construido por el algoritmo grado-egoísta.

Proposición 4.2.3. *Dados $G \sim CM_n(\bar{d}^{(n)})$ que satisfacen la (CC) hacia una distribución de grados límite $(p_k)_{k \geq 0}$. Entonces, si vale el (CG), tendremos que*

$$\alpha(G) = n \left(1 - \sum_{i=1}^{\infty} \frac{\mu^{(i)}(1)(1 - Q_i)}{2} - \sum_{j=2}^{\infty} (1 - Q_i^j) \mu^{(i)}(j) \right) + o_{\mathbb{P}}(n). \quad (4.7)$$

Donde $\mu^{(i)}(j)$ es el número restante de vértices de grado j dividido por n y Q_i es el parámetro correspondiente definido en el Teorema 4.2.1, después de i aplicaciones del mapa $M_1(\cdot)$ sobre la distribución de grados límite de la sucesión de grafos.

Demostración. Durante la i -ésima aplicación del mapa $M_1(\cdot)$, el número de vértices de grado j (donde $j \geq 2$) que se bloquea es⁴ $(1 - Q_i^j) \mu^{(i)}(j)$. Además, por cada par de vértices de grado 1 que se conecta a otro vértice de grado 1, se bloquea uno. El número de vértices de grado 1 que se bloquean es entonces $(1 - Q_i) \mu^{(i)}(1)/2$. Entonces, el tamaño del conjunto independiente obtenido será n menos el número total de vértices bloqueados. Además, dado que suponemos que vale la optimalidad asintótica del algoritmo, el tamaño normalizado del conjunto independiente obtenido será el número de independencia del grafo. \square

⁴Porque, como veremos en la Sección 4.3.4, Q_i es la probabilidad de que una media arista se conecte a un vértice activado.

Cotas al número de independencia

Para sucesiones de grafos donde no se cumple la condición de optimalidad asintótica, se puede sin embargo utilizar el Teorema 4.2.1 para construir cotas superiores para el número de independencia límite.

Proposición 4.2.4. *Dados $G \sim CM_n(\bar{d}^{(n)})$ donde la (CC) se satisface hacia una distribución límite $(p_k)_{k \geq 0}$; si se define $c_1 := \inf\{a > 0 : G''_D(1 - (a + p_1)/(a + \lambda)) < 1\}$ y (para $k \geq 0$) $p_k^* := (c_1 \delta_{1k} + p_k)/(c_1 + 1)$, tendremos que*

$$\alpha(G) \leq n \left(1 - \sum_{i=1}^{\infty} \frac{\mu^{*(i)}(1)(1 - Q_i^*)}{2} - \sum_{j=2}^{\infty} (1 - Q_i^{*j}) \mu^{*(i)}(j) \right) + o_{\mathbb{P}}(n).$$

Donde $\mu^{*(i)}(j)$ es el número restante de vértices de grado j sobre n y Q_i^* es el parámetro correspondiente definido en el Teorema 4.2.1, después de i aplicaciones del mapa $M_1(\cdot)$ sobre la distribución de grados $(p_k^*)_{k \geq 0}$.

Demostración. Definimos una nueva sucesión de grafos $G^* \sim CM_n(\bar{d}^{*(n)})$ donde la (CC) se cumple hacia la distribución de grados límite $(p_k^*)_{k \geq 0}$. Esta sucesión puede considerarse como los grafos formados al cambiar los grados de una cierta proporción de los vértices originales de G a 1, de modo tal que la desigualdad en el Teorema 4.2.1 se cumpla. Los grafos resultantes se distribuirán como los originales con algunas de sus aristas eliminadas. Entonces, debido a que el número de independencia es monótono en la eliminación de aristas (ver Capítulo 2), tendremos que $\alpha(G) \leq \alpha(G^*)$. Finalmente, usando la ecuación (4.7), se puede calcular $\alpha(G^*)$ y así obtener la cota superior deseada. \square

4.3. Pruebas de los resultados de la sección anterior

En esta sección, vamos a presentar las pruebas de los resultados enunciados anteriormente.

4.3.1. Prueba del Lema 4.2.1 para grafos generales

Probaremos el Lema 4.2.1 por inducción en el tamaño del grafo $|G|$. Para $|G| = 1$ y $|G| = 2$ el resultado es trivialmente verdadero.

Ahora supongamos que es válido para $|G| = n$, mostraremos que, por lo tanto, también es cierto para $|G| = n + 1$. Supongamos que no es el caso que $|W| = \alpha(G)$, entonces hay al menos un conjunto independiente A tal que $|W| < |A| = \alpha(G)$. Llamando a $W' = (w_i)_{i=2}^{|W|}$, por la hipótesis de inducción (porque W' tiene la propiedad T_1 en G_1) sabemos que $|W'| = \alpha(G_1)$.

Llamando n_1 al vértice adyacente a w_1 ⁵, debido a que el conjunto independiente A es de tamaño máximo, tiene que contener a w_1 o a n_1 (si no, uno podría construir un conjunto independiente aún mayor al agregar w_1 a los vértices de A , lo que sería una contradicción). Esto implica que

$$|A| = |A \setminus \{w_1, n_1\}| + 1 > |W| = |W'| + 1,$$

lo que significa que $|A \setminus \{w_1, n_1\}| > |W'|$ lo cual es una contradicción porque $A \setminus \{w_1, n_1\}$ define un conjunto independiente de G_1 y por hipótesis tenemos que el conjunto independiente

⁵Aquí asumiremos que $d_{w_1} = 1$. La prueba es muy similar en el caso donde w_1 es un vértice aislado.

definido por W' es un conjunto independiente de G_1 de tamaño máximo. Luego tenemos que W define un conjunto independiente de tamaño máximo de G , avanzando así la inducción.

Corolario 4.3.1. *Si H es una colección de árboles, entonces el algoritmo grado-egoísta corriendo en H encuentra c.s. un conjunto independiente máximo.*

Demostración. Debido a que H es una colección de árboles, por cada hoja eliminada por el algoritmo, se crearán más hojas (o vértices aislados). Por lo tanto, el algoritmo tendrá la propiedad T_1 ya que siempre habrá hojas (o vértices aislados) para seleccionar. Por el Lema 4.2.1 llegamos entonces a la conclusión. \square

4.3.2. Prueba de la Proposición 4.2.1

Primero probaremos el caso en el que G es un grafo subcrítico.

Para mostrar la condición de optimalidad de la Proposición 4.2.1, estudiaremos el número de veces que un proceso de exploración en anchura de las componentes une dos vértices ya explorados (para más detalles sobre la exploración en anchura de las componentes, vea el Apéndice A) formando un ciclo. Aquí llamaremos a N_u la cantidad de veces que esto sucede durante la exploración de la componente conexa asociada al vértice $u \in V$ y N a la cantidad total de veces que sucede en la exploración de todos las componentes en el grafo. Para cada componente $C(u)$ tendremos que si $N_u = 0$, entonces la componente es exactamente un árbol (ya que no se forman ciclos durante el proceso de exploración). La idea será demostrar que en un grafo subcrítico con una distribución de grados con una cola exponencialmente delgada, casi todos los vértices están en una componente que es un árbol. Notar que esto no es una consecuencia de los resultados en [JL07], ya que no es suficiente mostrar que el 2-núcleo⁶ es $o_{\mathbb{P}}(n)$ para concluir esto.

Llamando $v(i)$ al vértice visitado durante el paso i del proceso de exploración de la componente del vértice 1 y T al tiempo de parada en el que finaliza el proceso, podemos escribir

$$N_1 = \sum_{i=1}^T \sum_{j>i}^T \mathbb{I}_{\{v(i)=v(j)\}}. \quad (4.8)$$

Con esto, podemos establecer una cota a la esperanza condicional de N_1 que usaremos para probar la proposición.

Lema 4.3.1. *Sea $G \sim CM_n(\bar{d})$ tal que la (CC) se satisface hacia una distribución límite $(p_k)_{k \geq 0}$. Si $C(1)$ es la componente conexa del vértice 1, entonces condicionando al número T de aristas en la componente tenemos que*

$$\mathbb{E}(N_1|T) \leq \frac{(\nu^{(n)} + 1)n\lambda^{(n)}T(T-1)}{2(n\lambda^{(n)} - T)^2}$$

donde $\lambda^{(n)} := \sum_{u \in V^{(n)}} d_u^{(n)} / n \xrightarrow{n \rightarrow \infty} \lambda$ y $\nu^{(n)} := \sum_{u \in V^{(n)}} d_u^{(n)}(d_u^{(n)} - 1) / n\lambda^{(n)} \xrightarrow{n \rightarrow \infty} \nu$.

Demostración. El límite se deriva haciendo uso del límite correspondiente para la probabilidad de que (condicional a T) un vértice particular sea visitado por el proceso de exploración en el

⁶El 2-núcleo de un grafo es el mayor sub-grafo con grado mínimo 2.

paso i . Por lo tanto, en primer lugar, queremos mostrar que para $u \in V^{(n)}$ y $i \leq T$

$$\mathbb{P}(v(i) = u|T) \leq \frac{d_u^{(n)}}{n\lambda^{(n)} - T}. \quad (4.9)$$

Esto se puede ver usando que la variable aleatoria T es c.s. positiva y acotada por n^2 . Luego, para cualquier conjunto A $\sigma(T)$ -medible tenemos eso

$$\begin{aligned} \mathbb{E}(\mathbb{I}_A \mathbb{I}_{\{v(i)=u\}}) &= \mathbb{P}(A, v(i) = u) \\ &= \sum_{t \leq n^2} \mathbb{P}(A|v(i) = u, T = t) \mathbb{P}(v(i) = u|T = t) \mathbb{P}(T = t) \\ &\leq \sum_{t \leq n^2} \mathbb{I}_A(t) \frac{d_u^{(n)}}{n\lambda^{(n)} - t} \mathbb{P}(T = t) = \mathbb{E} \left(\mathbb{I}_A \frac{d_u^{(n)}}{n\lambda^{(n)} - T} \right). \end{aligned}$$

En donde en la tercera línea hemos usado que $\mathbb{P}(A|v(i) = u, T = t) = \mathbb{I}_A(t)$ porque A es $\sigma(T)$ -medible y que como el apareo de medias aristas se realiza de manera uniforme entre todos las medias aristas no apareadas en el paso i tenemos que $\mathbb{P}(v(i) = u|T = t) = \frac{d_u(i)^{(n)}}{n\lambda^{(n)} - (i-1)} \leq \frac{d_u^{(n)}}{n\lambda^{(n)} - t}$ (con $d_u(i)^{(n)}$ el número de medias aristas no apareadas de u durante el paso i de la exploración).

Usando que

$$N_1 = \sum_{i=1}^T \sum_{j>i}^T \mathbb{I}_{\{v(i)=v(j)\}} = \sum_{i=1}^T \sum_{j>i}^T \sum_{u \in V} \mathbb{I}_{\{v(i)=u\}} \mathbb{I}_{\{v(j)=u\}},$$

y tomando esperanza condicional a T a ambos lados de la desigualdad, se obtiene que

$$\begin{aligned} \mathbb{E}(N_1|T) &= \sum_{i=1}^T \sum_{j>i}^T \sum_{u \in V^{(n)}} \mathbb{P}(v(i) = u, v(j) = u|T) \leq \sum_{i=1}^T \sum_{j>i}^T \sum_{u \in V^{(n)}} \left(\frac{d_u^{(n)}}{n\lambda^{(n)} - T} \right)^2 \\ &\leq \frac{n\lambda^{(n)}}{(n\lambda^{(n)} - T)^2} \left(\sum_{i=1}^T \sum_{j>i}^T 1 \right) \left(\sum_{u \in V^{(n)}} d_u^{(n)} \frac{d_u^{(n)}}{n\lambda^{(n)}} \right) = \frac{(\nu^{(n)} + 1)n\lambda^{(n)}T(T-1)}{2(n\lambda^{(n)} - T)^2} \end{aligned}$$

Donde para la primera desigualdad se usó que para todo $i < j$

$$\begin{aligned} \mathbb{P}(v(i) = u, v(j) = u|T) &= \mathbb{P}(v(j) = u|v(i) = u, T) \mathbb{P}(v(i) = u|T) \\ &\leq \mathbb{P}(v(j) = u|T) \mathbb{P}(v(i) = u|T) \end{aligned}$$

□

Por medio de este lema, si denotamos por B el número de *vértices malos* que se encuentran en componentes que no son árboles, podemos demostrar que (bajo ciertos supuestos) este crece más lento que cualquier potencia positiva del tamaño del grafo n .

Proposición 4.3.1. *Sea $G \sim CM_n(\bar{d})$ tal que la (CC) se cumple hacia una distribución de grados límite $(p_k)_{k \geq 0}$ de media λ , de parámetro de criticalidad $\nu < 1$ y tal que existe $\gamma > 0$ para el cual $p_k = \mathcal{O}(e^{-\gamma k})$. Entonces, para todo $0 < \alpha < 1$ tenemos que $B = \mathcal{O}_{\mathbb{P}}(n^\alpha)$.*

Demostración. La prueba aplicará el Lema 4.3.1 y un acoplamiento presentado por Janson et al. en [Jan08], que describimos ahora.

Supongamos que estamos en el paso $i \leq \sqrt{n}$ (la potencia exacta de n es irrelevante, sólo se necesita que sea $o(n)$ y de un orden menor que los tamaños de las componentes) del proceso de exploración. Entonces la probabilidad de encontrar un vértice de grado k (excluyendo los vértices ya explorados) será

$$\frac{nkp_k^{(n)}}{n\lambda^{(n)} - \mathcal{O}(\sqrt{n})} = \frac{kp_k^{(n)}}{\lambda^{(n)}}(1 + \mathcal{O}(n^{-1/2})).$$

Definimos una variable aleatoria X distribuida de acuerdo a

$$\mathbb{P}(X \geq x) = \min \left(1, \frac{\nu'}{\nu} \sum_{k \geq x} \frac{(k+1)p_{k+1}}{\lambda} \right). \quad (4.10)$$

Para $\nu' = \nu + \epsilon'$ fijo y con $0 < \epsilon' < 1 - \nu$, entonces para todo n lo suficientemente grande X domina estocásticamente el tamaño del paso de la caminata aleatoria asociada al proceso de exploración. Como consecuencia, tendremos que si \tilde{T} es el tiempo de parada en el cual una caminata aleatoria con tamaño de paso X que parte de 1 alcanza el 0, $T \leq \tilde{T}$ c.s. siempre que $\tilde{T} \leq \sqrt{n}$. La variable \tilde{T} también puede considerarse como la progenie total de un proceso de ramificación con descendencia dada por copias i.i.d. de X . Al sumar la expresión en (4.10) obtenemos que $\mathbb{E}(X) \leq \frac{\nu'}{\nu} \nu < 1$, lo que significa que el proceso de ramificación asociado es subcrítico.

Haciendo uso de este acoplamiento, podemos probar la proposición. Tomando esperanza a N_1 y separando de acuerdo a diferentes valores de T obtenemos que

$$\mathbb{E}(N_1) = \mathbb{E} \left(N_1 \mathbb{I}_{\{T \leq n^\beta\}} \right) + \mathbb{E} \left(N_1 \mathbb{I}_{\{T > n^\beta\}} \right).$$

Donde $0 < \beta < 1/2$ (su valor exacto será fijado luego). Por la cota superior en la esperanza condicional en el Lema 4.3.1 obtenemos para el primer término que

$$\mathbb{E} \left(N_1 \mathbb{I}_{\{T \leq n^\beta\}} \right) \leq \frac{(\nu^{(n)} + 1)n\lambda^{(n)}n^\beta(n^\beta - 1)}{2(n\lambda^{(n)} - n^\beta)^2} \mathbb{P}(T \leq n^\beta) \leq \frac{(\nu^{(n)} + 1)n\lambda^{(n)}n^\beta(n^\beta - 1)}{2(n\lambda^{(n)} - n^\beta)^2}.$$

Ya que por hipótesis $\lambda^{(n)} \rightarrow \lambda$ y $\nu^{(n)} \rightarrow \nu$, el lado derecho es entonces $\mathcal{O}(n^{2\beta-1})$.

Para el segundo término, usaremos que $N_1 \leq T$ c.s. para obtener que

$$\mathbb{E} \left(N_1 \mathbb{I}_{\{T > n^\beta\}} \right) \leq \mathbb{E} \left(T \mathbb{I}_{\{T > n^\beta\}} \right) \leq n\mathbb{P}(T > n^\beta).$$

Debido al acoplamiento descrito anteriormente, también tendremos que $\mathbb{P}(T > n^\beta) \leq \mathbb{P}(\tilde{T} > n^\beta)$ (porque $n^\beta < \sqrt{n}$), donde \tilde{T} es la progenie total de un proceso de ramificación subcrítico con descendencia X , donde X se distribuye de acuerdo con (4.10). Por hipótesis, para algún $\gamma' > 0$, $\mathbb{P}(X = k) = \mathcal{O}(e^{-\gamma'k})$. Entonces, X tendrá momento exponencial finito y por el Teorema A.2.1, también lo tendrá \tilde{T} . Por lo tanto, por la desigualdad de Markov, tendremos que para todo $\theta > 0$

$$\mathbb{P}(\tilde{T} > n^\beta) \leq \frac{\mathbb{E}(\tilde{T}^\theta)}{n^{\beta\theta}} = \mathcal{O}(n^{-\beta\theta}).$$

Poniendo todas estas cotas juntamos (y tomando θ suficientemente grande)

$$\mathbb{E}(N_1) = \mathcal{O}(n^{2\beta-1}) + \mathcal{O}(n^{1-\beta\theta}) = \mathcal{O}(n^{2\beta-1}).$$

Con $0 < \alpha < 1$. Usando esto y la desigualdad de Markov, obtenemos que

$$\mathbb{P}(N > n^{\alpha/2}) \leq \frac{\mathbb{E}(N)}{n^{\alpha/2}} \leq \frac{\mathbb{E}(\sum_{u \in V} N_u)}{n^{\alpha/2}} = \frac{n\mathbb{E}(N_1)}{n^{\alpha/2}} = \mathcal{O}(n^{-\delta_1}).$$

Donde $\delta_1 := \alpha/2 - 2\beta$. Tomando $\beta < \alpha/4$ llegamos a que $\delta_1 > 0$.

Por otro lado, si llamamos a C_{max} a la componente más grande del grafo, por el Teorema 1.1 de [Jan08]⁷ también tenemos que

$$\mathbb{P}(|C_{max}| > n^{\alpha/2}) = \mathcal{O}(n^{-\delta_2})$$

para algún $\delta_2 > 0$. Entonces, tomando $\delta := \min(\delta_1, \delta_2) > 0$, tenemos que

$$\mathbb{P}(B > n^\alpha) \leq \mathbb{P}(N|C_{max}| > n^\alpha) \leq \mathbb{P}(N > n^{\alpha/2}) + \mathbb{P}(|C_{max}| > n^{\alpha/2}) = \mathcal{O}(n^{-\delta}).$$

□

Esta última proposición muestra que el número de vértices en componentes que no son árboles es (para todo $\alpha > 0$) $\mathcal{O}_{\mathbb{P}}(n^\alpha)$. Ahora, definamos T_G como un grafo formado por los árboles de expansión de cada una de las componentes de G (no es importante cuáles en particular se eligen). Llamaremos $W_{DG}(G)$ y $W_{DG}(T_G)$ a las respectivas secuencias de selección definidas por el algoritmo grado-egoísta ejecutado en G y T_G . Observemos que las componentes que son árboles se ven exactamente iguales en G y T_G .

Luego podemos acoplar de la siguiente manera las realizaciones del algoritmo grado-egoísta en G y T_G para que coincidan en las componentes que son árboles:

- Llamemos a las componentes de G como C_1, C_2, \dots, C_l y a las de T_G como C'_1, C'_2, \dots, C'_l .
- Ejecutamos una realización del algoritmo grado-egoísta en cada una de estas componentes. Esto genera las secuencias de selección W_1, W_2, \dots, W_l para las componentes de G y W'_1, W'_2, \dots, W'_l para las de T_G . Si para algún $j \leq l$ la componente C_j es un árbol, entonces $C_j = C'_j$ y las ejecuciones respectivas del algoritmo pueden acoplarse trivialmente para dar $W_j = W'_j$. Acoplemos de esta manera todas las secuencias de selección de todas las componentes que sean árboles.
- Ahora, construyamos $W_{DG}(G)$ inductivamente de la siguiente manera: en cada paso $i \geq 1$, contemos el número de vértices de grado mínimo en cada componente $j \leq l$ de G_{i-1} y llamemos a este número $d_j^{(i)}$. Seleccionemos la componente $j \leq l$ con probabilidad $d_j^{(i)} / \sum_{k=1}^m d_k^{(i)}$. Fijemos w_i (el vértice i -ésimo de $W_{DG}(G)$) como el primer vértice de W_j que aún no está en $\{w_1, \dots, w_{i-1}\}$.
- Finalmente, construyamos $W_{DG}(T_G)$ de manera análoga pero usando secuencias de selección W'_1, W'_2, \dots, W'_l .

⁷Aquí usamos que las colas de la distribución de grados son exponencialmente delgadas y, por lo tanto, $\mathcal{O}(k^{-\gamma''})$ para todo $\gamma'' > 0$.

Es claro que esta construcción se distribuye como una exploración secuencial grado-egoísta en cada una de los dos grafos.

La construcción fue hecha de tal manera que para cada componente que es un árbol, los mismos vértices terminan en $W_{DG}(G)$ que en $W_{DG}(T_G)$. Esto significa por la Proposición 4.3.1 que, como máximo, $|W_{DG}(G)|$ y $|W_{DG}(T_G)|$ diferirán en tamaño en $\mathcal{O}_{\mathbb{P}}(n^\alpha)$ (el número de vértices en componentes que no son árboles). Además, T_G es una colección de árboles y, por lo tanto, según el Corolario 4.3.1, el grado-egoísta definirá una secuencia de selección con la propiedad T_1 , y luego por el Lema 4.2.1 $|W_{DG}(T_G)| = \alpha(T_G)$. Pero todas las aristas en T_G también están presentes en G y, por lo tanto, T_G tendrá un conjunto independiente máximo más grande que G^8 . Esto implica que

$$|W_{DG}(T_G)| = |W_{DG}(G)| + \mathcal{O}_{\mathbb{P}}(n^\alpha) = \alpha(T_G) \geq \alpha(G),$$

lo cual significa que (porque $|W_{DG}(G)| = \sigma_{DG}(G)$)

$$|\alpha(G) - |W_{DG}(G)|| \leq \mathcal{O}_{\mathbb{P}}(n^\alpha).$$

Lo que prueba la proposición en el caso subcrítico.

Ahora damos una prueba para el caso en el que G es un grafo supercrítico. Llamemos W_{DG} a la secuencia de selección definida por el algoritmo grado-egoísta ejecutado en G . Queremos mostrar que $|W_{DG}| = \alpha(G) + \mathcal{O}_{\mathbb{P}}(n^\alpha)$. C.a.p. tenemos que esta secuencia selecciona vértices de grado 1 o 0 al menos hasta que el grafo restante sea subcrítico. Supongamos que esto es así, entonces existe algún valor $k_0 \geq 1$ tal que para todo $k \geq k_0$ el grafo restante G_k (ver 4.2.2) es subcrítico y para todo $l \leq k_0$ el vértice $W_{DG}(l)$ tiene grado 1 o 0 en G_l .

La idea es definir una secuencia de selección \tilde{W} que sea similar a W_{DG} , que tenga aproximadamente el mismo tamaño y para la cual $|\tilde{W}| \geq \alpha(G)$. Llamando a $T_{G_{k_0}}$ el grafo formado por los árboles de expansión de G_{k_0} , definimos el grafo \tilde{G} como una copia de G en el que el sub-grafo G_{k_0} ha sido reemplazado por $T_{G_{k_0}}$. Luego podemos definir una secuencia de selección \tilde{W} para \tilde{G} que coincida con W_{DG} hasta el paso k_0 . Para $k \geq k_0$, porque G_{k_0} es subcrítico y el grafo restante de \tilde{G} es una colección de árboles de expansión de G_{k_0} (es decir, $T_{G_{k_0}}$), podemos hacer que \tilde{W} tenga la propiedad T_1 y difiera como máximo en $\mathcal{O}_{\mathbb{P}}(n^\alpha)$ vértices de W_{DG} exactamente de la misma manera que en el caso subcrítico. Debido a que \tilde{W} tiene la propiedad T_1 , usando el Lema 4.2.1, $|\tilde{W}| = \alpha(\tilde{G})$. Además, debido a que todas las aristas presentes en \tilde{G} están también presentes en G tenemos ese $\alpha(\tilde{G}) \geq \alpha(G)$. Entonces,

$$|\tilde{W}| = |W_{DG}| + \mathcal{O}_{\mathbb{P}}(n^\alpha) = \alpha(\tilde{G}) \geq \alpha(G)$$

lo que significa que

$$|\alpha(G) - \sigma_{DG}(G)| \leq \mathcal{O}_{\mathbb{P}}(n^\alpha),$$

como queríamos probar.

4.3.3. Prueba del Teorema 4.2.1

Para el análisis del efecto de $M_1^{(n)}(\cdot)$ partiremos la dinámica grado-egoísta en dos. En la *primera fase*, sólo conectaremos los vértices de grado 1 durante lo cual registramos también el número de medias aristas libres que pertenecen a vértices bloqueados en la variable

⁸Esto se debe a que el número de independencia decrece monótonamente al agregar aristas a un grafo.

$B_t^{(n)}$. Haremos esto hasta que hayamos explorado (y/o bloqueado) $np_1^{(n)}$ vértices de grado 1. Mientras que en la *segunda fase*, tomaremos el valor final de $B_t^{(n)}$ de la fase 1 y uniremos secuencialmente cada una de estas medias aristas y eliminaremos las aristas así formadas del grafo.

Debido a que el Modelo de Configuraciones no es sensible al orden en que se realiza el apareo de sus medias aristas, esto no afectará la distribución final de grados obtenida (que será la misma que la obtenida aplicando el mapa $M_1^{(n)}(\cdot)$) pero, sin embargo, facilitará el análisis del límite resultante. Como veremos, para la prueba del Teorema 4.2.1 será suficiente analizar sólo la primera de estas dos fases.

La prueba consta de tres partes. Primero, damos una descripción estocástica de un proceso de Markov que da la evolución de la primera fase. Luego, establecemos la concentración de los valores asintóticos del número de medias aristas no apareadas, el número de medias aristas libres que pertenecen a vértices bloqueados y la medida de grados normalizada de vértices no conectados a ningún vértice de grado 1 activado. Usando estas convergencias, demostramos que el grafo resultante después de la fase 1 puede tratarse como un Modelo de Configuraciones con distribución límite conocida. Finalmente, estos límites combinados con una observación de percolación de sitios para Modelos de Configuraciones (ver Capítulo 2 para una introducción al tema), nos permiten establecer los criterios dados por el Teorema 4.2.1.

(i) *Descripción estocástica de la primera fase:* el proceso estocástico utilizado para modelar la primera fase del apareo de los vértices de grado 1 será similar en espíritu al utilizado en [BJL16] para estudiar el algoritmo egoísta. Aquí, también llevaremos la cuenta, en la variable aleatoria $B_t^{(n)}$, del número de medias aristas libres que pertenecen a vértices bloqueados. Las otras variables que se usarán para describir el proceso serán: el número $U_t^{(n)}$ de medias aristas no apareadas, el número $A_t^{(n)}$ de vértices restantes de grado 1 a aparear y (para $k \in \mathbb{N}$) el número $\mu_t^{(n)}(k)$ de vértices con grado k hacia otros vértices inexplorados.

Luego, en cada momento $t \geq 0$ el estado del proceso será descrito por el vector de dimensión infinita $(U_t^{(n)}, A_t^{(n)}, B_t^{(n)}, \mu_t^{(n)}(2), \mu_t^{(n)}(3), \dots)$. El proceso en cuestión evoluciona de la siguiente manera: a todo $t \geq 0$ cada vértice de grado 1 en $A_t^{(n)}$ tendrá un reloj exponencial de tasa 1. Cuando suena uno de los relojes de algunos de estos vértices, el vértice se elimina de $A_t^{(n)}$ y su arista se aparea uniformemente con otra arista libre. El estado del vértice con la media arista seleccionada para el apareo se declara *bloqueado* y sus medias aristas libres se agregan a $B_t^{(n)}$. El proceso continúa hasta que no haya más de los vértices originales de grado 1 para aparear. Debido a que está definido por tasas de transición que se comportan bien, es sencillo ver que el proceso resultante es markoviano. El tiempo de parada en el que finaliza el proceso estará dado por $T_1^{(n)} := \inf\{t \geq 0 : A_t^{(n)} = 0\}$ y será c.s. finito.

(ii) *Convergencia de los valores finales de las coordenadas:* para analizar la convergencia del valor final del número de medias aristas libres, representaremos con $\{e_i \not\rightarrow (1)\}$ al evento en el que la media arista e_i no se conecta con ninguna otra que provenga de alguno de los vértices de grado 1 originales. Luego definiremos la variable aleatoria $Y^{(n)} := \sum_{i=1}^{n\lambda^{(n)}} \mathbb{I}_{\{e_i \not\rightarrow (1)\}} = U_{T_1^{(n)}}^{(n)}$ que da el número de medias aristas no apareadas a ningún vértice de grado 1 al final de la primera fase. Ahora mostraremos que esta variable se concentra alrededor de su media. Para

esto, primero calcularemos su valor medio correspondiente

$$\mathbb{E}(Y^{(n)}) = \mathbb{E}\left(\sum_{i=1}^{n\lambda^{(n)}} \mathbb{I}_{\{e_i \not\leftrightarrow (1)\}}\right).$$

Dado que las medias aristas son intercambiables, esto es igual a $n\lambda^{(n)}\mathbb{P}(e_1 \not\leftrightarrow (1))$. La probabilidad de que el evento $\{e_1 \not\leftrightarrow (1)\}$ suceda puede mostrarse que vale $\frac{n\lambda^{(n)} - np_1^{(n)} - 1}{n\lambda^{(n)} - 1} \frac{n\lambda^{(n)} - np_1^{(n)} - 2}{n\lambda^{(n)} - 2} = (1 - p_1/\lambda)^2 + o(1)$.

Con lo cual tenemos que

$$\mathbb{E}(Y^{(n)}) = n\lambda^{(n)}(1 - p_1/\lambda)^2 + o(n). \quad (4.11)$$

Por lo tanto, para mostrar la concentración acotaremos la varianza de esta variable

$$\text{Var}(Y^{(n)}) = \mathbb{E}\left[(Y^{(n)})^2\right] - \left[\mathbb{E}(Y^{(n)})\right]^2. \quad (4.12)$$

Donde el primer término en (4.12) vendrá dado por

$$\mathbb{E}\left(\sum_{i \neq j}^{n\lambda^{(n)}} \mathbb{I}_{\{e_i \not\leftrightarrow (1)\}} \mathbb{I}_{\{e_j \not\leftrightarrow (1)\}}\right) + \mathbb{E}(Y^{(n)}).$$

Entonces, definiendo $A := \mathbb{E}\left(\sum_{i \neq j}^{n\lambda^{(n)}} \mathbb{I}_{\{e_i \not\leftrightarrow (1)\}} \mathbb{I}_{\{e_j \not\leftrightarrow (1)\}}\right)$, $B := \mathbb{E}(Y^{(n)})$ y $C := \left[\mathbb{E}(Y^{(n)})\right]^2$, tendremos que $\text{Var}(Y^{(n)}) = A + B - C$.

El término A puede ser acotado por

$$n\lambda^{(n)}(n\lambda^{(n)} - 1) (\mathbb{P}(e_1 \not\leftrightarrow e_2, e_1 \not\leftrightarrow (1), e_2 \not\leftrightarrow (1)) + \mathbb{P}(e_1 \leftrightarrow e_2)).$$

Con $\{e_1 \leftrightarrow e_2\}$ el evento en donde la media arista e_1 no se aparea con e_2 y $\{e_1 \not\leftrightarrow e_2\}$ su negación. La primera de estas probabilidades vendrá dada por

$$\mathbb{P}(e_1 \not\leftrightarrow (1), e_2 \not\leftrightarrow (1) | e_1 \not\leftrightarrow e_2) \mathbb{P}(e_1 \not\leftrightarrow e_2).$$

Se puede ver fácilmente que $\mathbb{P}(e_1 \not\leftrightarrow (1), e_2 \not\leftrightarrow (1) | e_1 \not\leftrightarrow e_2) = \prod_{k=1}^4 \left(1 - \frac{np_1^{(n)}}{n\lambda^{(n)} - k}\right)$, que $\mathbb{P}(e_1 \not\leftrightarrow e_2) = 1 + o(1)$ y que $\mathbb{P}(e_1 \leftrightarrow e_2) = o(1)$.

Con lo cual obtenemos que $A \leq n\lambda(n\lambda - 1)(1 - p_1/\lambda)^4 + o(n^2)$. También, por (4.11), $B = n\lambda(1 - p_1/\lambda)^2 + o(n)$ y $C = n^2\lambda^2(1 - p_1/\lambda)^4 + o(n^2)$. Lo cual muestra que la varianza de $Y^{(n)}$ es $o(n^2)$.

Por medio de la desigualdad de Chebychev obtenemos que (para todo $\epsilon > 0$)

$$\mathbb{P}\left(|Y^{(n)} - \mathbb{E}(Y^{(n)})| > \epsilon n\right) \leq \frac{\text{Var}(Y^{(n)})}{\epsilon^2 n^2} \rightarrow 0. \quad (4.13)$$

Lo que prueba que $U_{T_1^{(n)}}^{(n)}/n$ converge en probabilidad a $u_1 := (1 - p_1/\lambda)^2 \lambda = Q^2 \lambda$.

El cálculo para la convergencia del valor final de la distribución de grados de los vértices no explorados será completamente análogo. Representaremos con $\{v_j \not\leftrightarrow (1)\}$ al evento en el que el vértice v_j no se conecta a ninguno de los vértices de grado 1 originales. Entonces, (para

$i \geq 2$) la variable aleatoria $Z_i^{(n)} := \sum_{j=1}^n \mathbb{I}_{\{v_j \not\leftrightarrow (1)\}} \mathbb{I}_{\{d_{v_j}=i\}} = \mu_{T_1^{(n)}}^{(n)}(i)$ da el número de vértices de grado i no conectados a vértices de grado 1 al final de la primera fase. En otras palabras, $Z_i^{(n)}$ da el número de vértices de grado i que quedan después de la primera fase. Probaremos entonces que estas variables convergen en probabilidad.

Para ello, calculamos sus valores medios

$$\mathbb{E} \left(Z_i^{(n)} \right) = \mathbb{E} \left(\sum_{j=1}^n \mathbb{I}_{\{v_j \not\leftrightarrow (1)\}} \mathbb{I}_{\{d_{v_j}=i\}} \right).$$

Que, como los vértices son intercambiables, es igual a $n\mathbb{P}(v_1 \not\leftrightarrow (1), d_{v_1} = i)$. Esta última probabilidad es fácil de calcular y da $p_i^{(n)} \prod_{l=1}^i \left(1 - \frac{np_1^{(n)}}{n \sum_{i \geq 1} ip_i^{(n)} - (2l-1)} \right)$, que converge a $(1 - p_1/\lambda)^i p_i$ cuando $n \rightarrow \infty$.

Por lo tanto, tenemos que

$$\mathbb{E} \left(Z_i^{(n)} \right) = n(1 - p_1/\lambda)^i p_i + o(n) \quad (4.14)$$

Ahora acotaremos sus varianzas:

$$\text{Var} \left(Z_i^{(n)} \right) = \mathbb{E} \left[\left(Z_i^{(n)} \right)^2 \right] - \left[\mathbb{E} \left(Z_i^{(n)} \right) \right]^2. \quad (4.15)$$

Donde el primer término de (4.15) vendrá dado por

$$\mathbb{E} \left(\sum_{j \neq k}^n \mathbb{I}_{\{v_j \not\leftrightarrow (1)\}} \mathbb{I}_{\{v_k \not\leftrightarrow (1)\}} \mathbb{I}_{\{d_{v_j}=d_{v_k}=i\}} \right) + \mathbb{E} \left(Z_i^{(n)} \right).$$

Entonces, definiendo $A' := \mathbb{E} \left(\sum_{j \neq k}^n \mathbb{I}_{\{v_j \not\leftrightarrow (1)\}} \mathbb{I}_{\{v_k \not\leftrightarrow (1)\}} \mathbb{I}_{\{d_{v_j}=d_{v_k}=i\}} \right)$, $B' := \mathbb{E} \left(Z_i^{(n)} \right)$ y $C' := \left[\mathbb{E} \left(Z_i^{(n)} \right) \right]^2$, obtenemos que $\text{Var} \left(Z_i^{(n)} \right) = A' + B' - C'$.

Ahora, el término A' puede ser acotado según

$$n(n-1) \left(\mathbb{P}(v_1 \not\leftrightarrow v_2, v_1 \not\leftrightarrow (1), v_2 \not\leftrightarrow (1), d_{v_1} = d_{v_2} = i) + \mathbb{P}(v_1 \leftrightarrow v_2, d_{v_1} = d_{v_2} = i) \right).$$

Donde $\{v_1 \leftrightarrow v_2\}$ representa el evento en el cual v_1 no se conecta con v_2 (con v_1 y v_2 dos vértices distintos elegidos uniformemente) y $\{v_1 \not\leftrightarrow v_2\}$ es su negación. La primera de estas probabilidades viene dada por

$$\mathbb{P}(v_1 \not\leftrightarrow (1), v_2 \not\leftrightarrow (1) | v_1 \not\leftrightarrow v_2, d_{v_1} = d_{v_2} = i) \mathbb{P}(v_1 \not\leftrightarrow v_2 | d_{v_1} = d_{v_2} = i) \mathbb{P}(d_{v_1} = d_{v_2} = i).$$

En donde cálculos simples muestran que

$$\mathbb{P}(v_1 \not\leftrightarrow (1), v_2 \not\leftrightarrow (1) | v_1 \not\leftrightarrow v_2, d_{v_1} = d_{v_2} = i) = \prod_{l=1}^i \left(1 - \frac{np_1^{(n)}}{n \sum_{i \geq 1} ip_i^{(n)} - (i+2l-1)} \right) \times \left(1 - \frac{np_1^{(n)}}{n \sum_{i \geq 1} ip_i^{(n)} - (2i+2l-1)} \right),$$

que $\mathbb{P}(v_1 \not\leftrightarrow v_2 | d_{v_1} = d_{v_2} = i) = 1 + o(1)$, que $\mathbb{P}(v_1 \leftrightarrow v_2 | d_{v_1} = d_{v_2} = i) = o(1)$ y que $\mathbb{P}(d_{v_1} = d_{v_2} = i) = p_i^{(n)2} + o(1)$. Poniendo todas estas cotas juntas obtenemos que $A' \leq n(n-1)(1 - p_1/\lambda)^{2i} p_i^2 + o(n^2)$. Finalmente, por (4.14), $B' = n(1 - p_1/\lambda)^i p_i + o(n)$ y $C' = n^2(1 - p_1/\lambda)^{2i} p_i^2 + o(n^2)$. Lo que prueba que la varianza de $Z_i^{(n)}$ es $o(n^2)$.

De vuelta, la desigualdad de Chebychev implica que

$$\mathbb{P}\left(|Z_i^{(n)} - \mathbb{E}(Z_i^{(n)})| > \epsilon n\right) \leq \frac{\text{Var}\left(Z_i^{(n)}\right)}{\epsilon^2 n^2} \rightarrow 0. \quad (4.16)$$

Lo que a su vez significa que para cualquier distribución inicial de grados asintótica que tenga soporte compacto, todos los $Z_i^{(n)}/n$ convergerán conjuntamente en probabilidad a $(1 - p_1/\lambda)^i p_i$. Pero debido a que $(Z_i^{(n)})_{n \text{ geq } 0}$ está acotada por $(p_i^{(n)})_{n \geq 0}$ y esta última sucesión es uniformemente sumable, entonces

$$\left(Z_2^{(n)}/n, \dots, Z_i^{(n)}/n, \dots\right) \xrightarrow{\mathbb{P}} (\mu_1(2), \dots, \mu_1(i), \dots) := (Q^2 p_2, \dots, Q^i p_i, \dots). \quad (4.17)$$

Además, como tenemos que (para todo $t \geq 0$) $U_t^{(n)} = A_t^{(n)} + B_t^{(n)} + \sum_{i=2}^{\infty} i \mu_t^{(n)}(i)$ c.s., el número normalizado $B_{T_1^{(n)}}/n$ de medias aristas libres que pertenecen a vértices bloqueados al final de la primera fase convergerá en probabilidad a $b_1 := u_1 - \sum_{i=2}^{\infty} i \mu_1(i)$.

Finalmente, como $\sum_{k \geq 1} k^2 p_k^{(n)}$ converge hacia $\sum_{k \geq 1} k^2 p_k$, entonces estas sumas son uniformemente sumables. Y como, para todo $k \geq 2$, tenemos que $\mu_{T_1^{(n)}}^{(n)}(k)/n \leq p_k^{(n)}$, las sumas $\sum_{k \geq 2} k^2 \mu_{T_1^{(n)}}^{(n)}/n$ también serán uniformemente sumables. Esto implica que

$$\sum_{k \geq 2} k^2 p_k^{(n)} \mu_{T_1^{(n)}}^{(n)}/n \xrightarrow{\mathbb{P}} \sum_{k \geq 2} k^2 \mu_1(k).$$

Luego recuperamos la (CC) para la distribución de grados del grafo restante al final de la primera fase. Esto significa que, si consideramos las aristas libres de vértices bloqueados como vértices de grado 1, el grafo obtenido cuando finaliza la primera fase puede tratarse como un Modelo de Configuraciones⁹ con distribución de grados límite $\tilde{p}_1 = b_1/K = (u_1 - \sum_{i \geq 2} i \mu_1(i))/K$, y (para $k \geq 2$) $\tilde{p}_k = \mu_1(k)/K$ (donde K es una constante de normalización).

(iii) *Criterio de subcriticalidad:* aquí estableceremos en qué circunstancias el grafo obtenido después de aplicar el mapa $M_1^{(n)}(\cdot)$ una vez es c.a.p. subcrítico. Observemos que, para hacerlo, uno en principio tiene que analizar qué sucede con la distribución de grados durante la segunda fase de la dinámica y luego determinar si la distribución obtenida es subcrítica o no. Pero la segunda fase de la dinámica es equivalente a aparear los $B_{T_1^{(n)}}^{(n)}$ vértices de grado 1 y luego eliminarlos a ellos junto con sus aristas. De la misma manera que se discutió en el Capítulo 2, aparear vértices de grado 1 y luego eliminarlos junto con sus aristas no modifica la criticidad de un grafo del Modelo de Configuraciones con segundo momento finito. Luego podemos establecer la subcriticalidad del grafo después de aplicar $M_1^{(n)}(\cdot)$ simplemente calculando el parámetro de criticidad $\tilde{\nu}$ de un grafo de distribución de grados límite $(\tilde{p}_k)_{k \geq 1}$.

Calculando explícitamente este parámetro, obtenemos

$$\tilde{\nu} = \frac{\sum_{i \geq 2} i(i-1) \tilde{p}_i}{\sum_{i \geq 1} i \tilde{p}_i} = \frac{1}{\lambda} \sum_{i \geq 2} i(i-1) Q^{i-2} p_i = G_D''(Q)/\lambda. \quad (4.18)$$

Según el Teorema 2.3 [JL09], el grafo obtenido es subcrítico cuando este parámetro es estrictamente inferior a 1. La conclusión se sigue debido a la Proposición 4.2.1¹⁰.

⁹Para esto también necesitamos que el número de vértices restantes tienda a infinito cuando $n \rightarrow \infty$, pero esto puede verificarse fácilmente en el caso en que $p_1 < 1$.

¹⁰A primera vista, uno sólo podría aplicar esta proposición para distribuciones con colas asintóticamente

4.3.4. Prueba del Teorema 4.2.2

La estructura de la prueba es la siguiente. Primero damos la descripción del proceso estocástico asociado a la *segunda fase* de la dinámica. Después de esto, usamos los resultados de la Sección 3.3 para establecer límites de escala para este proceso. Finalmente, usamos estos límites para probar la el teorema.

(i) *Descripción estocástica de la segunda fase:* esta fase de la dinámica consiste en aparear secuencialmente las medias aristas de los vértices bloqueados y eliminar las aristas que se forman. Esto se hace hasta que no queden más medias aristas libres de bloqueados. En esta fase, los estados de los vértices no cambian, sólo sus grados; es por ello que no se agrega ningún vértice al conjunto independiente.

Inicialmente tenemos $B_{T_1}^{(n)}$ medias aristas libres de bloqueados y (para $k \geq 2$) $\mu_{T_1}^{(n)}(k)$ vértices no explorados de grado k . Por los resultados de la parte (ii) de la prueba del Teorema 4.2.1 tenemos que $B_{T_1}^{(n)}/n \xrightarrow{\mathbb{P}} Q^2\lambda - \sum_{i \geq 2} iQ^i p_i$ y (para todo $k \geq 2$) que $\mu_{T_1}^{(n)}(k)/n \xrightarrow{\mathbb{P}} Q^k p_k$.

Luego, a cada tiempo $t \geq 0$ el estado del proceso será descrito por el vector de dimensión infinita $(U_t^{(n)}, B_t^{(n)}, \mu_t^{(n)}(0), \mu_t^{(n)}(1), \dots)$. El proceso en cuestión evolucionará de la siguiente manera: en cada tiempo $t \geq 0$ cada media arista bloqueada libre tendrá un reloj exponencial con tasa $U_t^{(n)}/B_t^{(n)}$ (donde, cuando $B_t^{(n)} = 0$, definiremos la tasa de transición como 0). Cuando uno de los relojes de alguna de estas medias aristas suene, se la apareará uniformemente con otra media arista libre y se eliminará la arista del grafo. El proceso continuará hasta que no haya más medias aristas bloqueadas libres para aparear. Debido a que está definido por medio de tasas de transición que se comportan bien, es sencillo ver que el proceso resultante sea markoviano.

(ii) *Límite de escala de la segunda fase:* aquí establecemos la convergencia del proceso asociado a la segunda fase de la dinámica hacia soluciones de un conjunto numerable de ecuaciones diferenciales. Para esto, primero encontramos las derivas asociadas a cada una de las coordenadas del vector de estado:

- Con tasa $U_t^{(n)}$, sonará el reloj de una de las medias aristas bloqueadas libres, en cuyo caso se apareará con otra media arista libre. Entonces, $U_t^{(n)}$ tendrá como deriva a $\delta(U_t^{(n)}) := -2U_t^{(n)}$.
- Con tasa $U_t^{(n)}$, sonará el reloj de una de las medias aristas bloqueadas libres, en cuyo caso pueden suceder dos cosas: con probabilidad $\frac{B_t^{(n)}}{U_t^{(n)}}$ la media arista se apareará con otra media arista bloqueada libre y, por lo tanto, $B_t^{(n)}$ se reduce en 2; o con probabilidad $\frac{U_t^{(n)} - B_t^{(n)}}{U_t^{(n)}}$ se apareará con una media arista que pertenece a un vértice inexplorado y $B_t^{(n)}$ solo se reduce en 1. Entonces, $B_t^{(n)}$ tendrá como deriva a $\delta'(U_t^{(n)}, B_t^{(n)}) := -(B_t^{(n)} + U_t^{(n)})$.
- Finalmente, con tasa $U_t^{(n)}$ sonará el reloj de una de las medias aristas bloqueadas libres, en cuyo caso si se apareará con una media arista inexplorada, se seleccionará un vértice inexplorado de acuerdo con la distribución sesgada por tamaño y se eliminará una de

exponencialmente delgadas. Pero, como se muestra en el Lema 4.3.2, esta condición será cierta para toda distribución de grados luego de aplicar el mapa $M_1(\cdot)$ una vez.

sus medias aristas. Esto significa que (para toda $k \geq 0$) con probabilidad $\frac{k\mu_t^{(n)}(k)}{U_t^{(n)}}$ se seleccionará para el apareo un vértice de grado k y, por lo tanto, $\mu_t^{(n)}(k)$ se reducirá en 1 y $\mu_t^{(n)}(k-1)$ se incrementará en 1. Entonces, $\mu_t^{(n)}(k)$ tendrá como deriva a $\delta_k(\mu_t^{(n)}(k)) := -k\mu_t^{(n)}(k) + (k+1)\mu_t^{(n)}(k+1)$.

Fijemos $\delta > 0$, luego las sucesiones de procesos $\frac{U_t^{(n)}}{n}$, $\frac{B_t^{(n)}}{n}$ y (para todo $k \geq 0$) $\frac{\mu_t^{(n)}(k)}{n}$, están (para n suficientemente grandes) acotadas uniformemente por $\lambda + \delta$. Luego están uniformemente acotadas. Esto, a su vez, significa que todas las derivas asociadas a las coordenadas de $(U_t^{(n)}, B_t^{(n)}, \mu_t^{(n)}(0), \mu_t^{(n)}(1), \dots)$ están uniformemente acotadas.

Por las convergencias presentes en la prueba del Teorema 4.2.1, $U_0^{(n)}/n \rightarrow \lambda Q^2$, $B_0^{(n)}/n \rightarrow \lambda Q^2 - \sum_{i \geq 2} i Q^i p_i$ y (para todo $k \geq 0$) $\mu_0^{(n)}(k)/n \rightarrow Q^k p_k \mathbb{I}_{k \geq 2}$.

Aquí llamaremos a las martingalas de Dynkin asociadas a $U_t^{(n)}$, $B_t^{(n)}$ y (para todo $k \geq 0$) $\mu_t^{(n)}(k)$ como $M_t^{(n)}$, $M_t^{\prime(n)}$ y $N_t^{(n)}(k)$, respectivamente. Estas son todas martingalas de variación localmente acotada. Por lo tanto, sus variaciones cuadráticas vendrán dadas por

$$[M_t^{(n)}]_t = \sum_{0 \leq s \leq t} (\Delta M_s^{(n)})^2 = \sum_{0 \leq s \leq t} (\Delta U_s^{(n)})^2 \leq \sum_{s \geq 0} (\Delta U_s^{(n)})^2 \leq 4B_0^{(n)}n = \mathcal{O}(n) \quad (4.19)$$

$$[M_t^{\prime(n)}]_t = \sum_{0 \leq s \leq t} (\Delta M_s^{\prime(n)})^2 = \sum_{0 \leq s \leq t} (\Delta B_s^{(n)})^2 \leq \sum_{s \geq 0} (\Delta B_s^{(n)})^2 \leq 2B_0^{(n)}n = \mathcal{O}(n) \quad (4.20)$$

$$[N_t^{(n)}(k)]_t = \sum_{0 \leq s \leq t} (\Delta N_s^{(n)}(k))^2 = \sum_{0 \leq s \leq t} (\Delta \mu_s^{(n)}(k))^2 \leq \sum_{s \geq 0} (\Delta \mu_s^{(n)}(k))^2 \leq Q^k p_k^{(n)}n = \mathcal{O}(n). \quad (4.21)$$

El sistema de ecuaciones diferenciales correspondiente (tal como se lo presenta en el Lema 3.3.1) vendrá dado por

$$u_t = \lambda - \int_0^t 2u_s ds \quad (4.22)$$

$$b_t = b_0 - \int_0^t (b_s + u_s) ds \quad (4.23)$$

$$\mu_t(k) = Q^k p_k \mathbb{I}_{\{k \geq 2\}} + \int_0^t (k+1)\mu_s(k+1) - k\mu_s(k) ds \quad (4.24)$$

Las primeras dos de estas ecuaciones pueden ser integradas directamente y resultan en

$$u_t = \lambda Q^2 e^{-2t} \quad (4.25)$$

$$b_t = Q^2 \lambda e^{-2t} - e^{-t} \sum_{i \geq 2} i Q^i p_i. \quad (4.26)$$

Mientras que para las ecuaciones (4.24) con $k \geq 1$, puede verse que tienen los siguientes modos normales

$$\eta_k(i) = (-1)^{k-i} \binom{k}{i} \mathbb{I}_{\{i \leq k\}} \quad (4.27)$$

con autovalor asociado $\omega_k = -k$.

Además, la unicidad de las soluciones de este sistema se puede probar desacoplándolo y escribiéndolo en la base de los modos normales. Esto da como resultado un número numerable de ecuaciones independientes con derivadas Lipschitz, de lo cual se desprende la unicidad por medio de la teoría estándar de EDO.

Estamos entonces bajo las hipótesis del Lema 3.3.1 y por lo tanto tendremos que, uniformemente en conjuntos compactos, $U_t^{(n)}/n \xrightarrow{\mathbb{P}} u_t$, $B_t^{(n)}/n \xrightarrow{\mathbb{P}} b_t$ y (para todo $k \geq 0$) $\mu_t^{(n)}(k)/n \xrightarrow{\mathbb{P}} \mu_t(k)$.

Y si además se definen el tiempo de parada $T_2^{(n)} := \inf\{t \geq 0 : B_t^{(n)} = 0\}$ y el tiempo determinístico $T_2 := \inf\{t \geq 0 : b_t = 0\}$, se puede probar que las condiciones del Lema 3.3.2 valen. Para ello, primero notemos que $B_t^{(n)}$ es decreciente (sólo salta a estados de menor valor) y que para todo $i \in \mathbb{N}$ si $B_t^{(n)} = i$ entonces las tasas de salto vienen dadas por

$$\sum_{j \leq i} Q_{ij} = U_t^{(n)}.$$

Porque por cada $t \leq T_2^{(n)}$ tenemos ese $U_t^{(n)} \geq U_0^{(n)} - 2B_0^{(n)}$, si la proporción inicial de vértices bloqueados es menor que $1/2$, entonces $U_t^{(n)}$ estará acotada de forma uniforme en el tiempo por un valor positivo. Pero debido al Lema 4.3.3 que se puede consultar al final de la sección, la proporción de vértices bloqueados caerá c.a.p. (para cualquier valor inicial) por debajo de $1/2$ en un tiempo en donde $U_t^{(n)}$ seguirá siendo una proporción positiva de n . De esta forma, la estricta positividad de $U_t^{(n)}/n$ seguirá siendo cierta.

Por otro lado, si definimos $\tilde{\lambda} := \sum_{i \geq 2} iQ^i p_i$, T_2 puede mostrarse que viene dada por $\log(Q^2 \lambda / \tilde{\lambda}) = -\log \tilde{Q}$. Ya que (para todo $t \leq T_2$)

$$\dot{b}_t = -2Q^2 \lambda e^{-2t} + e^{-t} \tilde{\lambda} = e^{-t} (\tilde{\lambda} - 2Q^2 \lambda e^{-t}).$$

Para $t \leq T_2$, $e^{-t} \geq Q^2 \lambda / \tilde{\lambda}$ y $\tilde{\lambda} - 2Q^2 \lambda e^{-t} \leq -\tilde{\lambda}$. Entonces, la mano derecha de esta expresión es menor o igual a $-Q^2 \lambda$ para este rango de tiempos y entonces podemos aplicar el Lema 3.3.2 para probar que $T_2^{(n)} \xrightarrow{\mathbb{P}} T_2$. Y por el Corolario 3.3.1, tenemos que

$$U_{T_2^{(n)}}/n \xrightarrow{\mathbb{P}} u_{T_2}$$

$$B_{T_2^{(n)}}/n \xrightarrow{\mathbb{P}} b_{T_2}$$

$$\mu_{T_2^{(n)}}(1)/n \xrightarrow{\mathbb{P}} \mu_{T_2}(1)$$

$$\mu_{T_2^{(n)}}(2)/n \xrightarrow{\mathbb{P}} \mu_{T_2}(2)$$

...

Finalmente, como $\sum_{k \geq 1} k^2 p_k^{(n)}$ converge a $\sum_{k \geq 1} k^2 p_k$, entonces estas sumas son uniformemente sumables. Y por el Lema 4.3.2, tenemos que $\mu_{T_2^{(n)}}^{(n)}(k)/n$ es $\mathcal{O}_{\mathbb{P}}(e^{-\gamma k})$ (para algún

$\gamma > 0$), las sumas $\sum_{k \geq 2} k^2 \mu_{T_2}^{(n)}/n$ serán c.a.p. eventualmente uniformemente sumables. Esto implica que

$$\sum_{k \geq 1} k^2 p_k^{(n)} \mu_{T_2}^{(n)}/n \xrightarrow{\mathbb{P}} \sum_{k \geq 2} k^2 \mu_{T_2}(k).$$

Luego recuperamos la (CC) para la distribución de grados del grafo restante después de la segunda fase. Esto significa que el grafo obtenido cuando finaliza la segunda fase se puede tratar como un Modelo de Configuraciones con distribución de grados límite (para $k \geq 1$) $\tilde{p}_k = \mu_{T_2}(k)/Z$, donde Z es una constante de normalización.

Finalmente, mostraremos que la condición de colas exponencialmente delgadas en la Proposición 4.2.1 no es realmente restrictiva, ya que toda distribución inicial resulta en una distribución con esta propiedad después de una aplicación del mapa $M_1^{(n)}(\cdot)$.

Lema 4.3.2. *Bajo las hipótesis del Teorema 4.2.1 y asumiendo que $p_1 > 0$, entonces $M_1(p_i)(k)$ es $\mathcal{O}(e^{-\gamma k})$ con $\gamma := -\log(Q) > 0$.*

Demostración. Recordemos que $M_1(p_i)(k)$ viene dado por la coordenada k de la solución del sistema (4.24) a tiempo T_2 . Entonces se puede ver que este valor será más pequeño o igual que la coordenada k de la solución del sistema modificado:

$$\begin{cases} \tilde{\mu}_t(i) = Q^i p_i + \int_0^t (i+1) \tilde{\mu}_s(i+1) - i \tilde{\mu}_s(i) ds, & \text{if } i > k \\ \tilde{\mu}_t(i) = Q^i p_i + \int_0^t (i+1) \tilde{\mu}_s(i+1) ds, & \text{if } i = k \\ \tilde{\mu}_t(i) = Q^i p_i, & \text{if } 1 \leq i < k \end{cases}$$

a tiempo T_2 .

Además, la k -ésima coordenada de este sistema puede verse que converge monótonamente a $\tilde{\mu}_t(k) \nearrow \sum_{i=k}^{\infty} \tilde{\mu}_0(i)$ cuando $t \rightarrow \infty$. Entonces,

$$\mu_{T_2}(k) \leq \tilde{\mu}_{T_2}(k) \leq \sum_{i=k}^{\infty} \tilde{\mu}_0(i) = \sum_{i=k}^{\infty} \mu_0(i) \leq C \sum_{i=k}^{\infty} Q^i = \mathcal{O}(Q^k),$$

donde $C > 0$ es alguna constante. □

Urna de apareo

Supongamos que tenemos el siguiente modelo de urna, al que nos referiremos como *modelo de apareo*. Inicialmente tenemos una urna con k bolas rojas y $n - k$ bolas blancas. En cada paso del proceso, se retira una bola roja y luego se elige una segunda bola uniformemente de la urna y también se retira. El proceso continúa hasta que no queden más bolas rojas. Lo que probaremos aquí es que, si $k(n) = X_0 n + o(n)$ (para $0 < X_0 < 1$), entonces la proporción de bolas rojas decae c.a.p. por debajo de $1/2$ en un momento donde todavía hay una proporción positiva de las bolas en la urna. Denotaremos por $(R_i)_{i \in \mathbb{N}}$ al proceso que da en cada paso $i \geq 1$ la cantidad de bolas rojas eliminadas hasta el momento. También definiremos el tiempo de parada $T := \inf\{i \geq 0 : (k - R_i)/(n - 2i) < 1/2\}$ cuando la proporción de bolas rojas cae por debajo de $1/2$.

Lema 4.3.3. *Sea $0 < X_0 < 1$ la proporción inicial asintótica de bolas rojas en una urna de tamaño $n \in \mathbb{N}$, entonces c.a.p. $T < n/2$.*

Demostración. Suponga que $k > n/2$ (si no, la conclusión es trivial). Compararemos $(R_i)_{i \in \mathbb{N}}$ con un segundo proceso $(\tilde{R}_i(l))_{i \in \mathbb{N}}$, donde $l(n)$ es alguna función de n que se definirá más adelante. En cada paso $i \geq 1$, $\tilde{R}_i(l)$ dará el número total de bolas rojas eliminadas hasta el momento en una urna sin reemplazo con inicialmente $k - l$ bolas rojas y $n - k$ blancas. Denotaremos por X_i y $\tilde{X}_i(l)$ la proporción correspondiente de bolas rojas en las urnas para ambos procesos.

En cada momento $j \geq 1$, la probabilidad de sacar una bola roja para la urna de apareo viene dada por una variable Bernoulli de parámetro $X_j = (k - R_j)/(n - 2j)$ y la probabilidad correspondiente para la urna sin reemplazo también será una Bernoulli pero de parámetro $\tilde{X}_j(l) = (k - l - \tilde{R}_j(l))/(n - l - j)$. Luego podemos acoplar ambas probabilidades de selección mediante el acoplamiento habitual para dos variables de Bernoulli. Definiendo $T_l := \inf\{i \geq 1 : X_i \leq \tilde{X}_i(l)\}$, tendremos que para todo $i \leq T_l$, $R_i - i \geq \tilde{R}_i(l)$. Ahora, supongamos que $T_l > l$, luego en el paso l tendremos que $R_l - l \geq \tilde{R}_l(l)$ que implica que la proporción de vértices obedece que

$$X_l = \frac{k - R_l}{n - 2l} \leq \frac{k - l - \tilde{R}_l(l)}{n - 2l} = \tilde{X}_l(l). \quad (4.28)$$

Lo que contradice la hipótesis de que $T_l > l$. Entonces tendremos que $T_l \leq l$ c.s. En cada paso $i \geq 1$, el valor correspondiente de $\tilde{X}_i(l)$ se distribuirá de acuerdo con

$$\mathbb{P}(\tilde{R}_i(l) = n - l - m) = \frac{\binom{k-i}{(k-i)-m} \binom{n-k}{m-(k-2i)}}{\binom{n-i}{i}}. \quad (4.29)$$

Si $1 - l(n)/n$ está acotado por un número positivo, por el Teorema 2.4 en [BM15] tendremos que

$$\mathbb{P}\left(\max_{i \leq l} \tilde{X}_i - \tilde{X}_0 \geq \delta\right) \xrightarrow{n \rightarrow \infty} 0. \quad (4.30)$$

Porque $T_l \leq l$, al fijar $l = 2k + \delta' - n$ (con $\delta' > 0$ suficientemente pequeño), da $\tilde{X}_0(l) < 1/2$; lo que implica que la proporción de bolas rojas para el primer proceso caerá en un tiempo finito por debajo de $1/2$. Entonces, si existe un δ' tal que $2l < n$, el número de bolas restantes en la urna al tiempo l será una proporción positiva de n (es decir, $T < n/2$) ya que en cada paso se eliminan exactamente dos bolas. Es fácil comprobar que existirán tales δ' siempre que $X_0 < 3/4$.

Veamos que cuando la proporción inicial de vértices rojos es mayor a $3/4$, el enunciado sigue siendo cierto. Para probar esto, trabajaremos inductivamente. Llamemos a $a_1 := 1/2$ y (para $i \geq 2$) $a_i := 1 - (2/3)^i$. Supondremos que inicialmente $X_0 \in [a_j, a_{j+1})$. Vamos a mostrar que el proceso llega al intervalo $[a_{j-1}, a_j)$ en un tiempo finito y con una proporción positiva de bolas restantes en la urna. Esto se seguirá del mismo argumento que en el caso anterior fijando $l(n) = ((X_0 - a_j)/(1 - a_j) + \delta'')n$, para algunos $\delta'' > 0$ suficientemente pequeño. Este valor de $l(n)$ dará entonces $\tilde{X}_0(l) < a_j$ y $l(n) < (1/3 + \delta'')n$, asegurando que el mismo razonamiento que antes puede ser usado. Debido a que estos intervalos son una partición de $(1/2, 1)$, si la urna de apareo comienza con cualquier proporción inicial de bolas rojas superior a $1/2$, eventualmente caerá (después de pasar por un número finito de intervalos) debajo de $1/2$ en un momento en el que todavía hay una proporción positiva de bolas en la urna. Lo que equivale a decir que $T < n/2$. \square

4.3.5. Prueba de la Proposición 4.2.2

Por la siguiente igualdad para polinomios de Bernstein [Lor12],

$$\binom{n}{i} x^i (1-x)^{n-i} = \sum_{j=i}^n \binom{n}{j} x^j \eta_j(i),$$

la expansión en la base $\{\eta_k(\cdot)\}_{k \in \mathbb{N}}$ se puede calcular explícitamente para distribuciones binomiales, lo que a su vez da la transformación para las distribuciones de Poisson tomando el límite habitual. Usando esto, se puede ver fácilmente que después de i aplicaciones del mapa $M_1(\cdot)$ a una distribución de Poisson, la distribución resultante es una combinación lineal de una distribución de Poisson de media μ_i y un término $\delta_1(\cdot)$ (aquí ignoramos a los vértices de grado 0 ya que no juegan un papel en la dinámica), con los coeficientes respectivos A_i y B_i .

Esta transformación puede ser usada para derivar expresiones iterativas para λ_i , A_i y B_i :

$$\begin{cases} \mu_{i+1} = \tilde{Q}_i Q_i \mu_i, \\ A_{i+1} = e^{-(1-Q_i)\mu_i} A_i, \\ B_{i+1} = -A_i \tilde{Q}_i Q_i \mu_i e^{-\mu_i}, \end{cases}$$

donde Q_i y \tilde{Q}_i son los coeficientes correspondientes definidos en los Teoremas 4.2.1 y 4.2.2 para la distribución después de la i -ésima aplicación del mapa. Al escribir explícitamente los coeficientes Q_i y \tilde{Q}_i , se llega al mapa iterativo de tres dimensiones

$$\begin{cases} \mu_{i+1} = \left(e^{-\frac{a_i \mu_i}{\lambda_i}} - e^{-\mu_i} \right) \frac{A_i \mu_i^2}{\lambda_i}, \\ A_{i+1} = e^{-\frac{a_i \mu_i}{\lambda_i}} A_i, \\ B_{i+1} = -e^{-\mu_i} A_i \mu_{i+1}, \end{cases} \quad (4.31)$$

donde $a_i := A_i \mu_i e^{-\mu_i} + B_i$ es el número de vértices restantes de grado 1 y $\lambda_i := A_i \mu_i + B_i$ es el número de aristas restantes. Además, la coordenada B_i puede eliminarse llegando a un mapa iterativo cerrado de dos dimensiones

$$\begin{cases} \mu_{i+1} = \left[\exp \left(-\frac{A_i e^{-\mu_i} - A_{i-1} e^{-\mu_{i-1}}}{A_i - A_{i-1} e^{-\mu_{i-1}}} \mu_i \right) - e^{-\mu_i} \right] \frac{A_i \mu_i}{A_i - A_{i-1} e^{-\mu_{i-1}}}, \\ A_{i+1} = \exp \left(-\frac{A_i e^{-\mu_i} - A_{i-1} e^{-\mu_{i-1}}}{A_i - A_{i-1} e^{-\mu_{i-1}}} \mu_i \right) A_i. \end{cases}$$

Realizando el cambio de coordenadas $v_i := e^{-\mu_i}$ y $w_i := v_i^{A_i/(A_i - v_{i-1} A_{i-1})}$ el mapa puede ser reescrito como

$$\begin{cases} w_{i+1} = w_i^{(v_i w_i^{v_i - 1})}, \\ v_{i+1} = w_{i+1} w_i^{-v_i}, \end{cases} \quad (4.32)$$

donde las condiciones iniciales para este sistema discreto son $v_0 = w_0 = e^{-\lambda}$. Ahora mostraremos que para las condiciones iniciales en la recta identidad (v_0, v_0) esta recursión se puede resolver explícitamente. Primero mostraremos por inducción que en este caso

$$w_i = h_{2i+1}(v_0),$$

$$v_i = \frac{h_{2i+1}(v_0)}{h_{2i}(v_0)},$$

donde hemos definido $h_m(a) := \underbrace{a^{a^{\cdot^{\cdot^{\cdot^a}}}}_m$; esto es, $h_m(a)$ es la m -ésima tetración de a . Primero notemos que aplicando el mapa (4.32) se obtiene que $w_1 = v_0^{(v_0^{v_0})} = h_3(v_0)$ y que $v_1 = w_1 v_0^{-v_0} = h_3(v_0) h_2(v_0)^{-1}$, lo que inicia la inducción. Para avanzarla, supongamos $w_i = h_{2i+1}(v_0)$ y que $v_i = h_{2i+1}(v_0) h_{2i}(v_0)^{-1}$. Entonces, aplicando el mapa se obtiene que

$$\begin{aligned} w_{i+1} &= h_{2i+1}(v_0) \left(h_{2i+1}(v_0) h_{2i}(v_0)^{-1} h_{2i+1}(v_0)^{(h_{2i+1}(v_0) h_{2i}(v_0)^{-1} - 1)} \right), \\ &= h_{2i+1}(v_0) \left(h_{2i}(v_0)^{-1} h_{2i+1}(v_0)^{h_{2i+1}(v_0) h_{2i}(v_0)^{-1}} \right), \\ &= \left(h_{2i+1}(v_0) h_{2i}(v_0)^{-1} \right) \left(h_{2i+1}(v_0) h_{2i}(v_0)^{-1} \right)^{h_{2i+1}(v_0)}, \\ &= v_0^{\left(v_0^{h_{2i+1}(v_0)} \right)} = h_{2(i+1)+1}(v_0), \end{aligned}$$

donde usamos que $h_{2i+1}(v_0) h_{2i}(v_0)^{-1} = v_0$. De forma análoga, podemos probar que $v_{i+1} = \frac{h_{2(i+1)+1}(v_0)}{h_{2(i+1)}(v_0)}$, lo que avanza la inducción.

Por el Teorema 5 de [Bar36], $h_i(v_0)$ converge (cuando $i \rightarrow \infty$) si y sólo si $v_0 \in (e^{-e}, e^{1/e})$. Y en estos casos, debido a (4.32), v_i convergerá a 1 cuando $i \rightarrow \infty$. Al deshacer el cambio de coordenadas, esto implica que si $\lambda < e$, tendremos que $\mu_i \xrightarrow{i \rightarrow \infty} 0$. Esto a su vez prueba que el número de vértices restantes en el grafo $(1 - e^{-\mu_i})A_i + B_i \xrightarrow{i \rightarrow \infty} 0$, lo que significa que el grafo restante se vuelve nulo asintóticamente, lo que demuestra que el algoritmo grado-egoísta es asintóticamente óptimo. Además, también por el Teorema 5 de [Bar36], si $\lambda > e$, $h_{2i}(v_0)$ y $h_{2i+1}(v_0)$ convergerán a diferentes límites dando como resultado que $v_i \xrightarrow{i \rightarrow \infty} t < 1$. Es decir, en estos casos, la secuencia de selección definida por el algoritmo grado-egoísta no tendrá (c.a.p.) la propiedad T_1 .

Además, haciendo uso de la Proposición 4.2.3, podemos calcular el número de independencia. Porque en este caso el número de vértices de grado 2 o mayor vendrá dado por una densidad de Poisson de media μ_i multiplicada por A_i ,

$$\sum_{j \geq 2} (1 - Q_i^j) A_i \frac{\mu_i^j}{j!} e^{-\mu_i} = A_i (1 - e^{-\mu_i(1-Q_i)}) - A_i \mu_i e^{-\mu_i} (1 - Q_i).$$

Y usando la solución explícita para el mapa (4.31), podemos probar que

- $A_i = h_{2i}(e^{-\lambda})$
- $Q_i = \frac{(h_{2i}(e^{-\lambda}) - h_{2i+1}(e^{-\lambda}))}{(h_{2i}(e^{-\lambda}) - h_{2i-1}(e^{-\lambda}))}$
- $\mu_i = \lambda(h_{2i}(e^{-\lambda}) - h_{2i-1}(e^{-\lambda}))$
- $a_i = \lambda(h_{2i+1}(e^{-\lambda}) - h_{2i-1}(e^{-\lambda}))(h_{2i}(e^{-\lambda}) - h_{2i-1}(e^{-\lambda}))$

donde a_i es la cantidad normalizada de vértices de grado 1 restantes. Usando esto, la Proposición 4.2.3 nos da que

$$\begin{aligned} \alpha(G) &= n \left(1 - \sum_{i \geq 0} h_{2i}(e^{-\lambda}) - h_{2i+2}(e^{-\lambda}) + \sum_{i \geq 0} \frac{\lambda}{2} h_{2i-1}^2(e^{-\lambda}) - \frac{\lambda}{2} h_{2i+1}^2(e^{-\lambda}) \right) + o_{\mathbb{P}}(n) \\ &= n(z(\lambda) + \lambda/2z(\lambda)^2) + o_{\mathbb{P}}(n), \end{aligned}$$

donde usamos que (para $m \in \mathbb{N}$ y $a > 0$) $h_{-m}(a) = 0$, $h_0(a) = 1$ y que $\lim_{i \rightarrow \infty} h_i(a) = e^{-W(-\log(a))}$.

4.4. Posibles extensiones

En este capítulo mostramos que, para un grafo aleatorio con grados dados, si el algoritmo grado-egoísta selecciona sólo vértices de grado 1 o 0 hasta que el grafo restante sea subcrítico, entonces el conjunto independiente obtenido por él es del mismo tamaño que uno máximo hasta un término de error menor que cualquier potencia positiva del tamaño del grafo. Luego caracterizamos para qué distribuciones asintóticas de grados esto sucede y proporcionamos una forma de calcular su número de independencia asintótico.

Todavía es un problema abierto mostrar si el conjunto independiente encontrado es siempre c.s. asintóticamente máximo como en el caso Erdős-Rényi; y si no, bajo qué condiciones lo es.

En la Sección 4.2.5 explicamos cómo, cambiando algunos vértices de grado superior por vértices de grado 1, se pueden obtener cotas superiores para el número de independencia de grafos generales. Sería posible, en principio, obtener cotas más finas si se logra encontrar una manera óptima de dominar los grafos estudiados por grafos en los que el algoritmo grado-egoísta es asintóticamente óptimo.

Además, el Lema 4.3.2 parece sugerir que el apareo de vértices de grado 1 genera rápidamente una cola exponencialmente delgada en la distribución de grados resultante. Luego conjeturamos que la condición de segundo momento finito en los Teoremas 4.2.1 y 4.2.2 podría de hecho evitarse, extendiendo el resultado a distribuciones de cola pesada.

Finalmente, se sabe que la medida invariante de la dinámica de Glauber (bajo ciertos límites) se concentra alrededor de conjuntos independientes máximos. Sin embargo, sus tiempos de mezcla son en muchos casos exponenciales en el tamaño del grafo. Los resultados de este capítulo podrían ayudar a mostrar que el tiempo de mezcla podría reducirse iniciando la dinámica desde un conjunto independiente encontrado por un algoritmo grado-egoísta. Exploraremos numéricamente esta conjetura en el próximo capítulo.

Capítulo 5

Capacidad de redes de comunicación WiFi

En el capítulo anterior, mostramos que un algoritmo egoísta de baja complejidad es asintóticamente óptimo en una amplia clase de grafos aleatorios raros. Animados por este resultado, presentamos y estudiamos dos variantes de algoritmos de exploración secuenciales: algoritmos grado-conscientes estáticos y dinámicos. Derivamos límites de escala para ambos, que a su vez nos permiten calcular el tamaño de los conjuntos independientes obtenidos por ellos. Mientras que el primero es más simple de calcular, el segundo se puede usar para aproximar arbitrariamente al algoritmo grado-egoísta. Ambos se pueden implementar de manera distribuida. Como veremos, los límites de escala correspondientes constituyen un método eficiente para calcular o estimar el número de independencia para una amplia clase de grafos aleatorios raros. Como aplicación, mostramos cómo se pueden usar estos métodos para estimar la capacidad de una red inalámbrica grande en donde se implementa el protocolo 802.11. Finalmente consideramos otros indicadores de estas dinámicas, como la equidad de la configuración resultante, y mostramos cómo se puede lograr un equilibrio entre su equidad y capacidad.

Hay cuatro resultados principales en este capítulo. Primero, mostramos cómo se puede implementar y evaluar una versión descentralizada del algoritmo grado-egoísta, generalizando límites de escala conocidos para otros algoritmos secuenciales [BJM17, BJL16] (discutidos en el capítulo anterior) a algoritmos grado-conscientes. Más específicamente, definimos dos tipos de mecanismos: uno que tiene en cuenta sólo el grado inicial de los vértices (que llamaremos *algoritmo de exploración grado-consciente estático*) y otro que tiene en cuenta el grado de forma adaptativa (*algoritmo de exploración grado-consciente dinámico*).

Con respecto al algoritmo de exploración grado-consciente estático, aunque es subóptimo, su límite de escala resultante se puede resolver exactamente, es decir, en forma cerrada. Esto nos permite, por ejemplo, calcular el tamaño del conjunto independiente correspondiente, así como también calcular el límite de la proporción de vértices con cierto grado dado en el conjunto independiente. Por otro lado, el algoritmo de exploración grado-consciente dinámico puede usarse para aproximar hasta una precisión arbitraria el rendimiento del algoritmo grado-egoísta. Aunque su límite de escala resultante no puede ser resuelto analíticamente y se debe recurrir a estimaciones numéricas, constituye un método de estimación eficiente para el número de independencia para una gran familia de grafos aleatorios. Como ejemplo de aplicación, comparamos la capacidad de redes 802.11 con la estimación propuesta. Mediante

varias simulaciones, mostramos que el rendimiento de este último algoritmo es muy cercano al óptimo, incluso cuando el grafo aleatorio subyacente no pertenece a la clase de optimalidad del grado-egoísta. Es importante tener en cuenta que, si bien el algoritmo grado-egoísta no puede implementarse de manera distribuida (utilizando sólo información local del grafo), los algoritmos grado-conscientes dinámicos y estáticos sí pueden serlo, lo que significa que estas dinámicas podrían implementarse en principio como protocolos de comunicación.

Una segunda contribución consiste en estudiar numéricamente las ventajas de un algoritmo que combina las dinámicas grado-egoísta y Glauber para lograr resultados cuasi óptimos en menor tiempo. Esto se hace para todos los casos de conectividad (sea esta tanto alta como baja).

Como tercer aporte, mostramos que uno de los algoritmos secuenciales estudiados permite alcanzar un equilibrio entre la equidad (igualdad de oportunidades para vértices de diferentes grados de acceder al canal de comunicación) y eficiencia: la equidad puede mejorarse significativamente perdiendo poca capacidad.

Nuestra última contribución consiste en comparar numéricamente nuestros hallazgos para grafos aleatorios grandes con resultados preexistentes que no habían sido comparados entre ellos anteriormente, ni siquiera en el caso de Erdős-Rényi. Además, mostramos cómo estas ideas permiten predecir indicadores de rendimiento de redes inalámbricas reales.

El resto del capítulo está estructurado de la siguiente manera. En la Sección 5.1 damos una breve presentación de la dinámica de Glauber y algunos resultados relevantes. Luego, en la Sección 5.2, presentamos y analizamos dos variantes de algoritmos secuenciales. Allí, probamos sus límites de escala y caracterizamos los tamaños de sus conjuntos independientes. Después de esto, en la Sección 5.3, llevamos a cabo un análisis de rendimiento por medio de simulaciones en diferentes grafos aleatorios. En esta sección, también estudiamos el equilibrio entre la equidad y eficiencia. Finalmente, en la Sección 5.4 discutimos nuestros resultados en perspectiva.

5.1. Exploración dinámica

A diferencia del carácter secuencial de los algoritmos discutidos en el capítulo anterior, se pueden considerar también dinámicas más complejas donde los vértices se incorporan y abandonan el conjunto independiente a distintos tiempos. La más simple de este tipo es la conocida como *dinámica de Glauber* que, como discutiremos más adelante en la sección de simulaciones, es muy similar a la que implementan los vértices en una red 802.11.

La dinámica de Glauber consiste en una dinámica de Markov reversible en grafos donde los vértices se activan (con una tasa fija llamada tasa de activación) cuando ninguno de sus vecinos lo está; y, si está activado, pasa a desactivarse con tasa 1. Cuando la tasa de activación tiende a infinito, se puede mostrar fácilmente que esta dinámica se concentra en configuraciones que son conjuntos independientes de tamaño máximo. Aunque esta es una propiedad muy útil, al igual que en muchos problemas de optimización discreta, estas dinámicas pueden no servir en la práctica ya que la convergencia hacia una configuración de tamaño máximo puede ser extremadamente lenta cuando la tasa de activación es grande.

Dado un grafo G , la dinámica de Glauber se define como un proceso de Markov a tiempo discreto en el espacio de estados dado por los subconjuntos de vértices $S = \mathcal{P}(V)$. En cada paso $k \in \mathbb{N}_0$, el conjunto $A_k \subseteq V$ se interpretará como el conjunto de vértices *transmitiendo*

información en ese momento (es decir, los vértices activos) y será denominado como una *configuración*. Dada una configuración A_k , la configuración en el siguiente paso A_{k+1} se construirá de acuerdo con las siguientes reglas:

- Elija un vértice v uniformemente de V .
- Con probabilidad $\frac{\beta}{1+\beta} > 0$, si ningún vecino de v pertenece a A_k , fije $A_{k+1} = A_k \cup \{v\}$.
- O con probabilidad $\frac{1}{1+\beta} > 0$, fije $A_{k+1} = A_k \setminus \{v\}$.

Es fácil ver que si la configuración inicial A_0 es un conjunto independiente, A_k será un conjunto independiente para todo $k \in \mathbb{N}_0$. Es sabido que la medida invariante viene dada por (para cada configuración $A \subseteq V$) [Vig01]:

$$\mu(A) = \frac{\beta^{|A|}}{Z}, \quad (5.1)$$

donde Z es una constante de normalización. Esto significa que en el límite $\beta \rightarrow \infty$, la medida invariante se concentra en los conjuntos independientes máximos. En algunas situaciones especiales, el tiempo de mezcla ha sido caracterizado. Por ejemplo, en [Vig01], se muestra que cuando la distribución de grados está acotada por $\Delta \geq 0$ y $\beta < \frac{2}{\Delta-2}$ el tiempo de mezcla es $\mathcal{O}(n \log(n))$. O en el caso de un grafo regular bipartito, ha sido mostrado [GT06] que para β lo suficientemente grande el tiempo de mezcla es exponencial en n .

En el límite $\beta \rightarrow \infty$, la dinámica de Glauber se comporta de la siguiente manera:

- En una primera fase (tomando A_0 como el conjunto vacío), los vértices se activarán hasta que se alcance un conjunto independiente maximal. La distribución del tamaño del conjunto independiente resultante coincidirá con la del algoritmo egoísta, ya que los vértices son elegidos de manera uniforme.
- Debido a que los intentos de activación ocurren mucho más rápido que las desactivaciones, después de esta fase inicial habrá muchos intentos fallidos de activación (que pueden omitirse en el análisis) seguidos de una desactivación única de algún vértice v .
- Luego de ocurrida esta desactivación, los vértices en $\mathcal{N}(v) \cup \{v\}$ que no tienen vecinos activos intentarán activarse de manera uniforme hasta que no sean posibles más activaciones y se alcance un nuevo conjunto independiente maximal.
- Estos últimos dos pasos son repetidos indefinidamente.

Debido a que la medida invariante en este límite se concentra en conjuntos independientes de tamaño máximo, esta dinámica se aproximará asintóticamente a un conjunto independiente máximo, pero esto puede tomar un tiempo prohibitivamente largo. En la Sección 5.3.2, discutiremos la aplicación de este tipo de dinámicas límite a redes de comunicación modeladas por el Modelo de Configuraciones.

5.2. Generalización de algoritmos egoístas

En esta sección adaptamos algunos de los resultados presentados en la Sección 4.1 para el algoritmo egoísta en el Modelo de Configuraciones con el fin de obtener resultados más

generales que permitan incorporar el grado de los vértices a la dinámica. A través de estos resultados, podremos proponer esquemas descentralizados con un rendimiento muy cercano al del algoritmo grado-egoísta.

En la Sección 5.2.1, analizamos el caso dinámico. Como veremos, el algoritmo grado-egoísta puede considerarse como un proceso límite para esta familia de dinámicas, permitiendo que estos procesos se utilicen para aproximarlos. Por otro lado, en la Sección 5.2.2 analizaremos el caso estático que resulta en un comportamiento asintótico más simple que puede describirse resolviendo una única EDO, permitiendo simulaciones e implementación más sencillas.

5.2.1. Exploración grado-consciente dinámica

En la Sección 4.1.1, se presentó la dinámica egoísta junto con un conjunto de ecuaciones diferenciales que describen su comportamiento límite. Aquí consideramos una generalización en la que cada vértice con grado $i \geq 0$ hacia otros vértices inexplorados tiene un reloj exponencial de parámetro $0 < \lambda(i) < \infty$. El bloqueo de los vecinos de los vértices seleccionados y el apareo de sus medias aristas dentro del conjunto de vértices inexplorados se realizan exactamente como antes (es decir, de manera uniforme). Llamaremos a este proceso como *exploración grado-consciente dinámica*.

Notar que el cambio de las tasas de estos relojes exponenciales es equivalente a cambiar la probabilidad de que durante una transición a tiempo $t \geq 0$ se active un vértice de grado $i \geq 0$, de $\alpha_t(i)$ (probabilidad correspondiente al algoritmo egoísta, definida en la Sección 4.1.1) a

$$\gamma_t(i) = \frac{\lambda(i)\bar{\mu}_t^{(n)}(i)}{\sum_{j=0}^{\infty} \lambda(j)\bar{\mu}_t^{(n)}(j)}, \quad (5.2)$$

donde $\{\bar{\mu}_t^{(n)}(i)\}_{i \geq 0}$ es (como en la Sección 4.1.1) el número escalado de vértices inexplorados con grado i hacia otros inexplorados.

Entonces, $\bar{\mu}_t^{(n)}(\cdot)$ tiene el siguiente límite:

Proposición 5.2.1. *Bajo las hipótesis del Teorema 4.1.1 y fijando $(\lambda(i))_{i \geq 0}$ acotada y positiva, entonces la sucesión de procesos $\bar{\mu}_t^{(n)}(\cdot)$ converge en probabilidad y sobre conjuntos compactos de tiempo a la única medida $\bar{\mu}_t$ que es solución del siguiente sistema de ecuaciones diferenciales:*

$$\frac{d}{dt}u_t(i) = - \sum_{j=0}^{\infty} \lambda(j)u_t(j) [\gamma_t(i) + G_t(i)], \quad (5.3)$$

donde $G_t(i) := \sum_{k=0}^{\infty} k\gamma_t(k) (\beta_t(i+1) + (\beta_t(i) - \beta_t(i+1)) \sum_{l=0}^{\infty} l\beta_t(l))$ con $\gamma_t(\cdot)$ dado por (5.2) y $\beta_t(\cdot)$ definido como en (4.2) (ambas funciones, asociadas a $(u_t(i))_{i \geq 0}$).

Demostración. La principal diferencia con estos procesos es que, en el caso del algoritmo egoísta, un vértice inexplorado se activa cuando suena su reloj exponencial de tasa 1; mientras que en estos, las tasas de los relojes no son necesariamente las mismas y pueden depender de sus grados.

Notar que estas ecuaciones son adaptaciones de (4.3) donde sustituimos $\alpha_t(i)$ por $\gamma_t(i)$ y la tasa de transición total ahora es $\sum_{j=0}^{\infty} \lambda(j)u_t(j)$. El resto de las ecuaciones permanecen

sin cambios ya que, como se explicó anteriormente, los vecinos de los vértices activados se seleccionan en ambos casos de la misma manera. La prueba es entonces, mutatis mutandi, la misma que la prueba del Teorema 4.1.1 de [BJM17], que es complicada y por tanto no reproducimos aquí. Se remite al lector interesado a [BJM17]. \square

Una vez resueltas las ecuaciones (5.3), se puede calcular la constante de atascamiento correspondiente como

$$c_{\bar{\mu}_0} = \int_0^\infty \sum_{i=0}^\infty \lambda(i) \bar{\mu}_t(i) dt, \quad (5.4)$$

donde $\bar{\mu}_0(\cdot)$ es el límite escalado de la distribución inicial de grados. Como antes, llamaremos $\sigma_{DDA}(G)$ al tamaño del conjunto independiente obtenido, el cual dividido por n convergerá en L^1 a $c_{\bar{\mu}_0}$ cuando n tiende a infinito.

En este contexto, se puede considerar la dinámica grado-egoísta fijando $\gamma_t(i)$ como

$$\gamma_t^*(i) = \begin{cases} 1, & \text{if } \bar{\mu}_t(i) > 0 \text{ y } \bar{\mu}_t(j) = 0 \forall j < i; \\ 0, & \text{else.} \end{cases} \quad (5.5)$$

Sin embargo, entre otras dificultades, estas probabilidades no tienen la forma (5.2) y, por lo tanto, el límite descrito por la proposición anterior (ecuaciones (5.3)) no aplica. No obstante, podemos considerar familias de frecuencias de reloj $\lambda(\cdot)$ que seleccionarán en cada transición un vértice de grado mínimo con alta probabilidad. Por ejemplo, $\lambda(i) = (i+1)^{-L}$ (con $L > 0$). En este caso, las probabilidades asociadas vendrán dadas por

$$\gamma_t(i) = \frac{(i+1)^{-L} \bar{\mu}_t(i)}{\sum_{i=0}^\infty (i+1)^{-L} \bar{\mu}_t(i)}.$$

Observemos que estos $\gamma_t(i)$ tienden a $\gamma_t^*(i)$ cuando L va a infinito, por lo que esta familia de procesos ofrece una forma de aproximar el tamaño del conjunto independiente descubierto por el algoritmo grado-egoísta. En la Sección 5.3 compararemos esta aproximación con los resultados obtenidos usando las probabilidades $\gamma_t^*(\cdot)$. Como veremos, la aproximación propuesta presenta un excelente rendimiento para diferentes distribuciones de grados iniciales.

Como consecuencia de los resultados del Capítulo 4, esta aproximación del algoritmo grado-egoísta puede usarse para estimar el número de independencia de una amplia familia de grafos aleatorios. Además, tengamos en cuenta que este algoritmo puede implementarse de manera distribuida, ya que sólo utiliza los grados hacia vértices inexplorados (que es información local del grafo).

Observación: ni las técnicas ni los resultados de [BJL16], que serán de utilidad en la siguiente sección, pueden simplificar el límite presentado aquí. Ya que, en los procesos estudiados, la información de la distribución de grados de los vértices inexplorados hacia otros vértices inexplorados es necesaria en cada momento para derivar las ecuaciones límite, lo que hace que la estrategia de [BJL16] sea inaplicable.

5.2.2. Exploración grado-consciente estática

Aquí discutiremos otra variante de la dinámica egoísta que conduce a un límite más simple. De manera similar a la situación discutida en la sección anterior, cada vértice inexplorado se

activará cuando suene un reloj exponencial, pero en lugar de que cada vértice tenga un reloj que dependa dinámicamente de su grado en el sub-grafo inexplorado, ahora su tasa sólo dependerá del grado inicial del vértice en cuestión y, por lo tanto, será constante en el tiempo.

Al igual que antes, el bloqueo de vértices y apareo de medias aristas se realiza como en el algoritmo egoísta. Nos vamos a referir a este proceso como *exploración grado-consciente estática* y denotaremos la cantidad de vértices en el conjunto independiente que este encuentra en un grafo G como $\sigma_{SDA}(G)$. En este contexto, se cumple el siguiente resultado:

Proposición 5.2.2. *Para $n \geq 1$, sea $G \sim CM_n(d^{(n)})$ una sucesión de grafos que satisface la (CC) hacia una distribución de grados límite $(p_k)_{k \geq 1}$ de media $m > 0$. Entonces, si $(\lambda(k))_{k \in \mathbb{N}_0}$ es tal que $\sum_{k=0}^{\infty} k\lambda(k)p_k^{(n)}$ es uniformemente sumable, tendremos que*

$$\frac{\sigma_{SDA}(G)}{n} \xrightarrow[\mathbb{P}]{n \rightarrow \infty} \int_0^{\infty} \sum_{k=0}^{\infty} \lambda(k)p_k e^{-\lambda(k)t} e^{-k\tau(t)} dt,$$

donde $\tau(t)$ es la solución de la siguiente EDO

$$\frac{d\tau}{dt} = \frac{1}{m} \sum_{k=1}^{\infty} k\lambda(k)p_k e^{-\lambda(k)t} e^{-(k-2)\tau}$$

con condición inicial $\tau(t) = 0$.

Demostración. La prueba es una modificación de la del teorema principal de [BJL16].

Primero se describe el proceso como un proceso de Markov que involucra las coordenadas de la distribución de grados de vértices inexplorados $(\bar{\mu}_t(k))$, para $k \geq 0$; donde $\bar{\mu}_0(k) = p_k$, el número total de medias aristas no apareadas (U_t) y el número de vértices activos (A_t) .

Luego se encuentran explícitamente las derivas de todas las coordenadas del proceso. En nuestro caso esto nos da que

$$\delta(A_t)/n = \sum_{k=0}^{\infty} \lambda(k)\bar{\mu}_t(k),$$

$$\delta(U_t)/n = - \sum_{k=0}^{\infty} k\lambda(k)\bar{\mu}_t(k) \left(2 - \frac{k-1}{U_t-1} \right),$$

$$\text{y } \delta(\bar{\mu}_t(k)) = -\lambda(k)\bar{\mu}_t(k) - \sum_{j=0}^{\infty} np_{jk}\lambda(j)\bar{\mu}_t(j)(\bar{\mu}_t(k) - \delta_{jk}/n).$$

Aquí p_{jk} representa la probabilidad de conexión entre un vértice de grado j y k .

De manera similar al Lema 3.3.1, encontramos las formas límite de las derivas y acotamos la variación cuadrática correspondiente de las martingalas de Dynkin asociadas. Con esto mostramos la convergencia de cada uno de los límites de escala estableciéndolo primero para algunas subsucesiones de cada subsucesión.

Las cotas para las expresiones límite de las derivas son análogas a las encontradas en [BJL16]. Sin embargo, para probar que $\delta(U_t)/n \rightarrow -2 \sum_{k=0}^{\infty} k\lambda(k)\bar{\mu}_t(k)$ se necesita un argumento distinto si $\lambda(k)$ no es acotada (el caso acotado se maneja como en [BJL16]).

Para ver esto, notemos que como $\sum_{k=1}^{\infty} k\lambda(k)p_k^{(n)}$ es uniformemente sumable, existe una sucesión $(\epsilon_N)_{N \geq 1}$ tal que (para todo $n \geq 1$) $\sum_N^{\infty} k\lambda(k)\bar{\mu}_t(k) \leq \epsilon_N$ y $\epsilon_N \xrightarrow{N \rightarrow \infty} 0$. Entonces

usando la martingala de Dynkin,

$$U_t/n \geq U_0/n - 2 \int_0^t \sum_{k=1}^{\infty} k \lambda(k) \bar{\mu}_s(k) ds \geq U_0/n - 2\bar{\lambda}_N \int_0^t \sum_{k=1}^N k \bar{\mu}_s(k) ds - 2\epsilon_N t \quad (5.6)$$

donde $\bar{\lambda}_N := \max_{k \in \{1, \dots, N\}} \lambda(k)$. Llamando u_t al límite de U_t/n en la subsubsucesión correspondiente, por la ecuación (5.6) y el hecho de que $U_t/n \geq \sum_{k=1}^{\infty} k \bar{\mu}_t(k)$, tendremos que para todo $t \geq 0$ y $\bar{N} \geq 1$ suficientemente grande $u_t \geq \bar{u}_t$, donde \bar{u}_t es la solución de la ecuación diferencial $\bar{u}'_t = -2\bar{\lambda}_{\bar{N}} \bar{u}_t - 2\epsilon_{\bar{N}}$ con condición inicial $\bar{u}_0 = m$. Si se integra esto se obtiene que $\bar{u}_t = \left(m + \epsilon_{\bar{N}}/\bar{\lambda}_{\bar{N}}\right) e^{-2\bar{\lambda}_{\bar{N}}t} - \epsilon_{\bar{N}}/\bar{\lambda}_{\bar{N}}$. Esto entonces implica que $u_t \geq \bar{u}_t > 0$ para todo $t \in [0, \bar{t}_{\bar{N}})$ con

$$\bar{t}_{\bar{N}} = \log \left[\left(1 + \frac{m\bar{\lambda}_{\bar{N}}}{\epsilon_{\bar{N}}}\right)^{1/2\bar{\lambda}_{\bar{N}}} \right] = \frac{m}{2\epsilon_{\bar{N}}} + o(1).$$

En donde usamos que $\bar{\lambda}_{\bar{N}} \xrightarrow{\bar{N} \rightarrow \infty} \infty$ porque $\lambda(k)$ no es acotada. Debido a que $\epsilon_{\bar{N}} \xrightarrow{\bar{N} \rightarrow \infty} 0$, tendremos que $\bar{t}_{\bar{N}} \xrightarrow{\bar{N} \rightarrow \infty} \infty$. Esto finalmente muestra que para todo $t \geq 0$ fijo, $u_t > 0$ que a su vez implica que $U_t \xrightarrow{n \rightarrow \infty} \infty$.

Finalmente, mediante argumentos similares a los del Lema 3.3.1, la convergencia que es válida en estas subsubsucesiones se extiende a toda la sucesión de procesos. Luego, el límite de escala de la distribución de grados de vértices inexplorados en una dinámica grado-consciente estática estará dada por (para $k \geq 1$)

$$\bar{\mu}_t(k) = \bar{\mu}_0(k) e^{-\lambda(k)t} e^{-k\tau(t)}, \quad (5.7)$$

donde $\tau(t)$ está definido como en el enunciado de la proposición. Gracias a la convergencia de la distribución de grados, tendremos que

$$\sigma_{SDA}(G) = \lim_{t \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{A_t}{n} = \int_0^{\infty} \sum_{k=0}^{\infty} \lambda(k) \bar{\mu}_t(k) dt \quad (5.8)$$

$$= \int_0^{\infty} \sum_{k=0}^{\infty} \lambda(k) \bar{\mu}_0(k) e^{-\lambda(k)t} e^{-k\tau(t)} dt, \quad (5.9)$$

más un término que se anula en probabilidad. \square

Observar que esta proposición reduce el problema de analizar la nueva dinámica a integrar una única EDO, que es mucho más simple que resolver el sistema presentado en la sección anterior. Sin embargo, el conjunto independiente resultante puede ser más pequeño que en el caso de exploración grado-consciente dinámica.

Por otro lado, si llamamos $c_{\bar{\mu}_0}$ a la proporción de atascamiento límite, la proporción de vértices de grado $i \geq 1$ en el conjunto independiente vendrá dada por

$$q_i = \frac{1}{c_{\bar{\mu}_0}} \int_0^{\infty} \lambda(i) \bar{\mu}_0(i) e^{-\lambda(i)t} e^{-i\tau(t)} dt. \quad (5.10)$$

La Proposición 5.2.2 y la ecuación (5.10) se utilizarán en la Sección 5.3.3 para analizar la distribución de grados de los vértices que pertenecen al conjunto independiente y derivar estrategias para mejorar la igualdad de probabilidad de acceso (es decir, la equidad).

5.3. Resultados de simulaciones

5.3.1. Análisis de la aproximación al algoritmo grado-egoísta

En esta sección estudiaremos numéricamente las aproximaciones del algoritmo grado-egoísta presentadas en las Secciones 5.2.1 y 5.2.2. A menos que se indique lo contrario, las simulaciones corresponderán a 10 grafos aleatorios G , con $n = 1000$, generados de acuerdo con el Modelo de Configuraciones correspondiente al escenario considerado. Para cada uno de estos grafos calcularemos su distribución de grados empírica $\mu(\cdot)$, resolveremos numéricamente (5.3) o (5.7) (límites correspondientes a las variantes de algoritmos dinámicos y estáticos, respectivamente) con parámetros de reloj $\lambda(i) = (i + 1)^{-L}$ (con $L > 0$ suficientemente grande) usando como condición inicial $\bar{\mu}_0(\cdot) = \mu(\cdot)/n$, y calcularemos el tamaño del conjunto independiente obtenido por medio de (5.4) o (5.9).

También ejecutaremos un algoritmo grado-egoísta en cada uno de estos 10 grafos y señalaremos el tamaño de los conjuntos independientes resultantes con una caja. Además, también ejecutamos una dinámica de Glauber lo suficientemente larga como para que alcance un valor cercano al de un conjunto independiente máximo. Sin embargo, como se señaló anteriormente, aunque esta dinámica proporciona asintóticamente el número de independencia, este puede llegar a alcanzarse muy lentamente en una simulación, ya que el tiempo de mezcla es exponencial en n en muchas circunstancias. Esto se discutirá más a fondo en la Sección 5.3.2.

Finalmente, también incluiremos el tamaño del conjunto independiente correspondiente al algoritmo egoísta (el cual se puede obtenerse mediante los resultados presentados en la Sección 4.1).

Grafos aleatorios regulares

Primero, consideraremos grafos aleatorios d -regulares. Como discutimos en la Sección 4.1.2, este es uno de los modelos más estudiados en la literatura. También incluiremos en nuestra comparación el límite inferior en el número de independencia demostrado en [Wor95] (se recomienda consultar los resultados numéricos presentados en la Tabla 1 de ese artículo), y el valor exacto presentado en [DSS16].

Los resultados se muestran en la Figura 5.1. Observar que, como se esperaba en este caso, nuestro análisis de la exploración grado-consciente dinámica proporciona una aproximación muy precisa del tamaño del conjunto independiente obtenido por el algoritmo grado-egoísta. Además, los límites inferiores proporcionados por Wormald en [Wor95] dan esencialmente los mismos valores que nuestra aproximación. Esto se debe probablemente al hecho de que, como lo sugiere Wormald [Wor99], es muy probable que este límite inferior sea en realidad el valor límite exacto.

Además, al observar los resultados correspondientes a [DSS16], podemos concluir que posiblemente todavía haya espacio para aumentar el tamaño del conjunto independiente del algoritmo grado-egoísta. Sin embargo, debe tenerse en cuenta que este resultado es exacto sólo para d mayor que cierto d_0 no especificado por los autores. Nuestras simulaciones de la dinámica de Glauber respaldan aún más que, al menos para estos valores de d , esos resultados parecen ser límites superiores. En cualquier caso, esta mejora sería relativamente modesta, y ciertamente más pequeña que la distancia entre los algoritmos grado-egoísta y egoísta. Finalmente, notar que en este caso (donde todos los vértices tienen el mismo grado inicial),

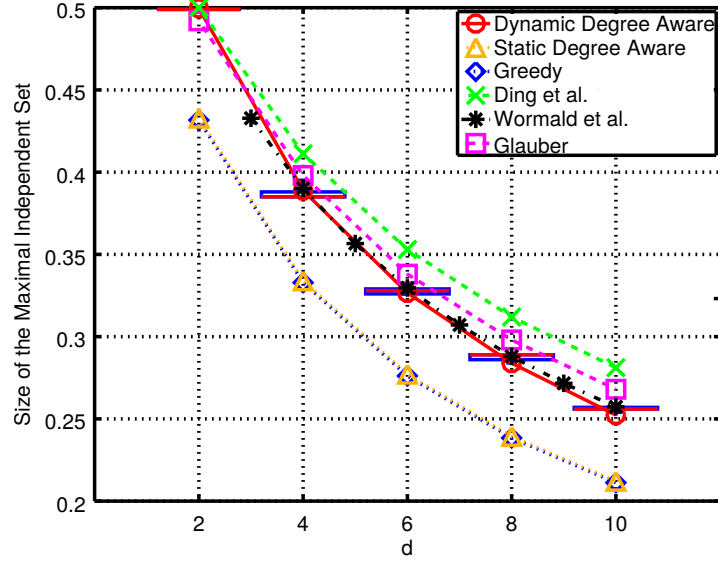


Figura 5.1: *Tamaño del conjunto independiente obtenido para distintos valores de d para grafos aleatorios d -regulares. Los resultados del algoritmo grado-egoísta se marcan con cajas (que se ven como líneas horizontales), mientras que los círculos rojos se corresponden con los resultados calculados a partir de (5.4), los triángulos amarillos con los de (5.9), los diamantes azules con los de (4.4) o (4.5), las cruces verdes con los valores de [DSS16], las estrellas negras con las cotas de [Wor95] y los cuadrados magentas con la dinámica de Glauber.*

las variantes estáticas y dinámicas de nuestros algoritmos son equivalentes y, por lo tanto, obtienen el mismo resultado.

Grafos de Erdős-Rényi ralos

Ahora consideraremos grafos de Erdős-Rényi ralos. Aquí sólo estudiaremos los resultados de nuestras aproximaciones, junto con simulaciones de la dinámica de Glauber con tasa de activación grande. Los resultados se muestran en la Figura 5.2. Como era de esperar, nuestra aproximación es extremadamente cercana al algoritmo grado-egoísta, que a su vez es óptimo¹ para valores de λ más pequeños que e (notar que la dinámica de Glauber no produce un número de independencia mayor en estos casos). Finalmente, el algoritmo grado-consciente estático (es decir, la ecuación relativamente simple de resolver (5.9)) en este caso produce una aproximación razonable al número de independencia.

Grafos geométricos

Ahora consideraremos un caso íntimamente relacionado con el modelo de Erdős-Rényi: grafos generados a partir de un proceso puntual de Poisson en el plano. En estas simulaciones, generamos un proceso puntual de Poisson en un círculo de un tamaño fijado de tal forma que el número medio de vértices sea $n = 1000$. Con el objetivo de modelar redes inalámbricas, dados dos vértices i y j , consideraremos que están conectados si:

$$P(i, j) = X_{i,j} d(i, j)^{-a} > T,$$

¹Como se discutió en el Capítulo 4.

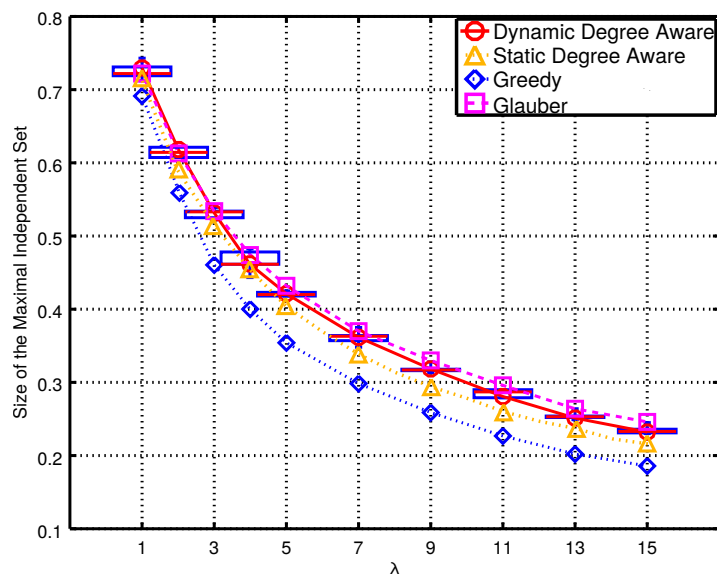


Figura 5.2: *Tamaño de los conjuntos independientes máximos obtenidos para grafos de Erdős-Rényi raros de grado medio λ .*

donde $d(i, j)$ es la distancia euclídea entre los vértices i y j , $X_{i,j}$ es una variable aleatoria que modela el desvanecimiento (y que asumiremos que se distribuye como una log-normal con una media de 1, una varianza θ y simétrica en i, j), $a = 2$ y T , de modo que el número medio de vecinos cuando no hay desvanecimiento sea igual a 2. Los resultados se muestran en la Figura 5.3.

Notar que el caso donde $\theta = 0$ estos grafos se corresponden con una variante del denominado modelo Matèrn, y que la distribución inicial de grados es la misma que en el caso de Erdős-Rényi (con $\lambda = 2$ en este caso en particular). Sin embargo, el origen geométrico del grafo genera correlaciones entre los grados de los vértices, que el Modelo de Configuraciones no posee. Esto hace que nuestras aproximaciones (es decir, las ecuaciones (5.4) y (5.9)) sobreestimen el tamaño del conjunto independiente encontrado por el algoritmo grado-egoísta. Sin embargo, a medida que aumenta el desvanecimiento, estas correlaciones se vuelven cada vez más pequeñas, lo que resulta en una mejor aproximación.

Es interesante notar que en el caso en que $\theta = 1$ y $\theta = 0$, la dinámica de Glauber no es capaz de mejorar los resultados del algoritmo grado-egoísta. La distribución de grados resultante para ambos casos es (comprobada empíricamente para $\theta = 1$) una binomial de media menor que 2,4. Esto parece indicar que los resultados probados en el capítulo anterior podrían extenderse a una clase aún más amplia de grafos aleatorios.

5.3.2. Conjuntos independientes máximos y redes inalámbricas

Ahora centraremos nuestra atención en el problema del cálculo (o aproximación) de la capacidad de una red inalámbrica grande que implementa el protocolo 802.11 (o WiFi como se conoce comercialmente). Este tipo de redes implementan un protocolo de acceso medio llamado algoritmo CSMA/CA (por sus nombre en inglés: *carrier sense multiple access with collision avoidance*).

En pocas palabras, el CSMA/CA funciona de la siguiente manera. Cuando un vértice

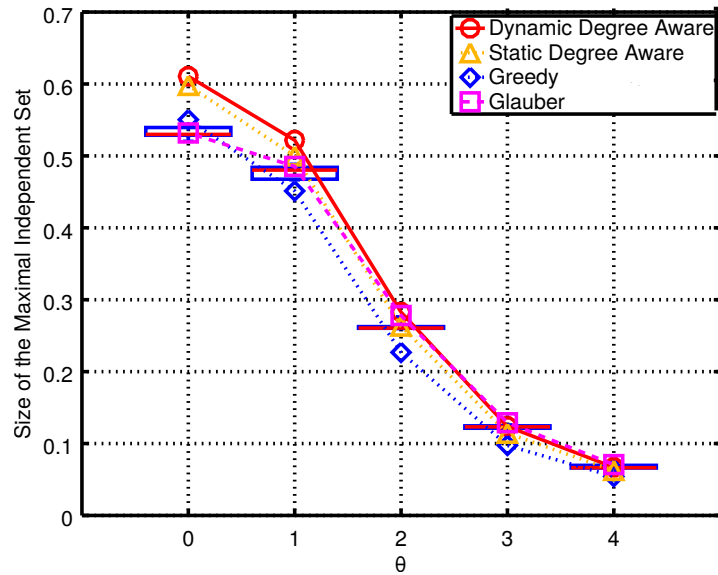


Figura 5.3: *Tamaño del conjunto independiente obtenido en función de θ para grafos generados a partir de un proceso puntual y con desvanecimiento.*

quiere transmitir un paquete, primero establece un tiempo aleatorio de espera, el cual se reduce siempre que ninguno de sus vecinos esté transmitiendo. Cuando este tiempo llega a cero, se transmite el paquete. Aquí asumiremos que los vértices siempre intentan transmitir paquetes, por lo que este proceso se repite indefinidamente.

Como se muestra en [LK16], cuando el tiempo de espera es mucho menor que el tiempo medio de transmisión (que, si el sistema está diseñado para maximizar la eficiencia, debería ser el caso), entonces el número de vértices transmitiendo luego de un tiempo lo suficientemente largo será cercano al tamaño de un conjunto independiente máximo ¿Por qué es esto así? Intuitivamente, cuando un vértice deja de transmitir, sus vecinos y ese mismo vértice comienzan a competir por el medio. Por lo tanto, tenemos que cualquier disminución de uno en los vértices activos o transmisores es seguida rápidamente por un aumento en *al menos* uno. De hecho, si tanto el tiempo de espera como el tiempo de transmisión se distribuyen exponencialmente, este proceso se comporta como una versión a tiempo continuo de la dinámica de Glauber.

La discusión anterior implica que la capacidad de una red inalámbrica puede ser aproximada por nuestros métodos. En esta subsección, sumaremos al análisis que realizamos en la subsección anterior simulaciones de la dinámica CSMA/CA sobre dos familias de grafos y mostraremos cómo evoluciona el número de vértices activos con el tiempo y qué tan rápido converge al número de independencia.

Grafos de Erdős-Rényi ralos

Primero nos centraremos en el caso de Erdős-Rényi. Como mostramos en el Capítulo 4, el algoritmo grado-egoísta produce un conjunto independiente asintóticamente máximo en el caso en que $\lambda < e$. Para ilustrar este resultado (que se presenta en la Figura 5.2), presentamos la evolución del CSMA/CA (con tiempos de espera y transmisión distribuidos exponencialmente

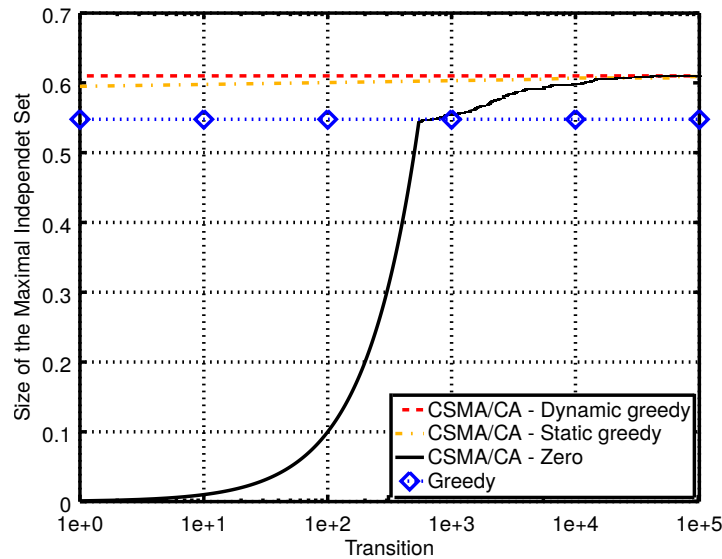


Figura 5.4: Número de vértices transmitiendo mientras corre el algoritmo CSMA/CA (comenzando desde tres posibles condiciones iniciales) en un grafo de Erdős-Rényi de grado medio $\lambda = 2$. La abscisa se grafica en escala logarítmica.

con medias iguales a 1×10^{-6} y 1×10^2 respectivamente) sobre grafos de Erdős-Rényi con grado medio λ .

En particular, ejecutamos el algoritmo CSMA/CA con tres condiciones iniciales diferentes: sin ningún vértice transmitiendo, con el conjunto de vértices que transmiten proveniente de un algoritmo de exploración grado-consciente estático y con el conjunto de vértices que transmiten proveniente de un algoritmo de exploración grado-consciente dinámico. Los resultados correspondientes a $\lambda = 2$ se muestran en la Figura 5.4 (donde la abscisa está en escala logarítmica para mejor visualización), junto con el tamaño del conjunto independiente obtenido por el algoritmo egoísta.

Consideremos el caso donde todos los vértices comienzan esperando transmitir. Como se discutió en la Sección 5.1, la evolución del CSMA/CA en este caso se puede separar en dos etapas. Durante la primera etapa, el CSMA/CA alcanza rápidamente una configuración que se distribuye como el conjunto independiente obtenido por algoritmo egoísta (ya que el CSMA/CA está diseñado para elegir aleatoriamente entre los vértices a activarse). Una vez que se alcanza este conjunto independiente maximal, el CSMA/CA todavía puede aumentar el número de transmisiones simultáneas, pero lo hace muy lentamente: se necesitan aproximadamente las primeras 20,000 transiciones para converger a un valor cercano al del algoritmo grado-consciente dinámico.

Vale la pena verificar cómo, incluso después de 80,000 transiciones adicionales (e incluso cuando la configuración inicial se modifica para partir de un conjunto independiente proveniente de algoritmos grado-conscientes), el CSMA/CA no logra aumentar el número de vértices activos. Esto verifica que (5.4) se puede usar para calcular la capacidad de una red inalámbrica cuando se cumple la condición de optimización asintótica del Capítulo 4. Si este no es el caso, aún puede resultar muy útil, como lo discutiremos a continuación.

Consideremos entonces un grafo de Erdős-Rényi, pero ahora con un grado medio de 10. En este caso, dado que no se cumplen las condiciones de optimalidad asintótica, para estimar el

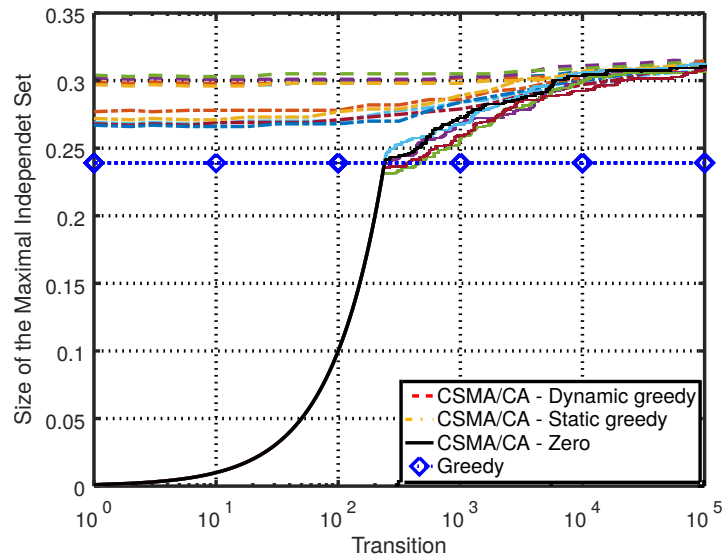


Figura 5.5: *Número de vértices transmitiendo durante una corrida del algoritmo CSMA/CA (comenzando desde tres configuraciones iniciales distintas) en grafos de Erdős-Rényi de grado medio $\lambda = 10$. La abscisa se encuentra en escala logarítmica.*

número de independencia (o la capacidad de la red inalámbrica) uno tiene que recurrir a otros algoritmos como el de Glauber (o equivalente al CSMA/CA), que eventualmente convergen a un conjunto independiente máximo. Sin embargo, según las simulaciones anteriores y nuestros resultados, para “acelerar” esta estimación, proponemos comenzar una dinámica de Glauber a partir de una configuración dada por uno de los algoritmos secuenciales que discutimos (grado-conscientes).

Los resultados de la simulación se muestran en la Figura 5.5, donde hemos incluido cinco ejecuciones de un CSMA/CA para cada una de las configuraciones iniciales. Es importante tener en cuenta que estas configuraciones iniciales se pueden encontrar en una cantidad de pasos igual al número de vértices en el conjunto independiente obtenido y, por lo tanto, en un tiempo que es una fracción del número total de vértices en el grafo ($n = 1000$ en estas simulaciones). Notar que la tendencia observada en nuestras simulaciones anteriores todavía es válida: si inicialmente no hay vértices transmitiendo, el número de transiciones que necesita CSMA/CA para alcanzar un conjunto independiente de tamaño similar al del algoritmo grado-consciente dinámico es (al menos) un orden de magnitud mayor que el número de vértices en el grafo. Además, incluso después de 100,000 iteraciones, la diferencia entre las configuraciones iniciales es no despreciable, y los mejores resultados se obtienen iniciando el algoritmo CSMA/CA a partir de la configuración encontrada por el algoritmo grado-consciente dinámico.

Grafos geométricos

Ahora discutiremos un escenario un tanto más realista. En lugar de posicionar arbitrariamente los vértices de acuerdo con un proceso puntual de Poisson, consideraremos sus ubicaciones según son provistas por un conjunto de datos público (en este caso por OpenCellID [Unw18], una base de datos abierta de ubicaciones de torres celulares de todo el mundo). La

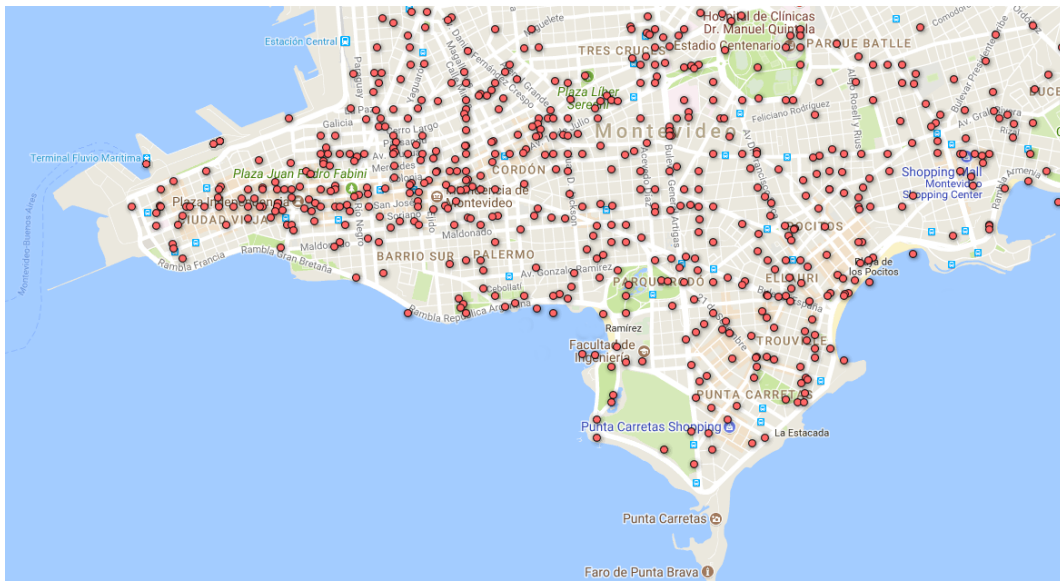


Figura 5.6: *Los vértices y sus posiciones. Los datos se corresponden con la ciudad de Montevideo.*

Figura 5.6 muestra el área considerada en este ejemplo, que consiste en una región urbana de alrededor de 20 km^2 e incluye 579 vértices.

Consideramos el mismo modelo que antes, con desvanecimiento log-normal de varianza θ y un número promedio de vecinos igual a 2.4 cuando $\theta = 0$ (equivalente a un rango de 150 m). Los resultados correspondientes a $\theta = 3$ se muestran en la Figura 5.7, junto con los tres resultados teóricos.

Discutamos estos resultados. Primero, si el desvanecimiento no es despreciable como en este caso (y como en la Figura 5.3), nuestra aproximación basada en los resultados para el Modelo de Configuraciones es muy precisa. Simulaciones adicionales, no informadas aquí para no extendernos de más, indican que esta tendencia es cierta para otros escenarios, como suburbano o urbano denso. Esto ilustra la flexibilidad de los análisis basados en el Modelo de Configuraciones en general, y de nuestra aproximación en particular. En segundo lugar, las tres condiciones iniciales dan como resultado una diferencia no despreciable en el número de vértices activos, incluso después de 100,000 iteraciones. Esta observación ilustra además que la dinámica de Glauber puede atascarse en los mínimos locales, pero comenzarla desde configuraciones ventajosas puede conducir a conjuntos independientes más grandes en tiempos más cortos.

5.3.3. Equidad y capacidad

Hasta ahora, no hemos discutido cuán *equitativa* es la dinámica que establece las conexiones en la red; pero en una dinámica egoísta (y aún más en una grado-egoísta), los vértices con grado más alto tienden a bloquearse más rápido que los vértices con uno más pequeño. Esto es así porque, si el grado de vértice es grande, tiene una mayor probabilidad de que alguno de sus vecinos se active antes que él. Aquí discutimos cómo compensar esto implementando una familia de algoritmos grado-conscientes estáticos.

Como una medida de la *inequidad* usaremos la distancia de *variación total* $\|\cdot\|_{TV}$ entre

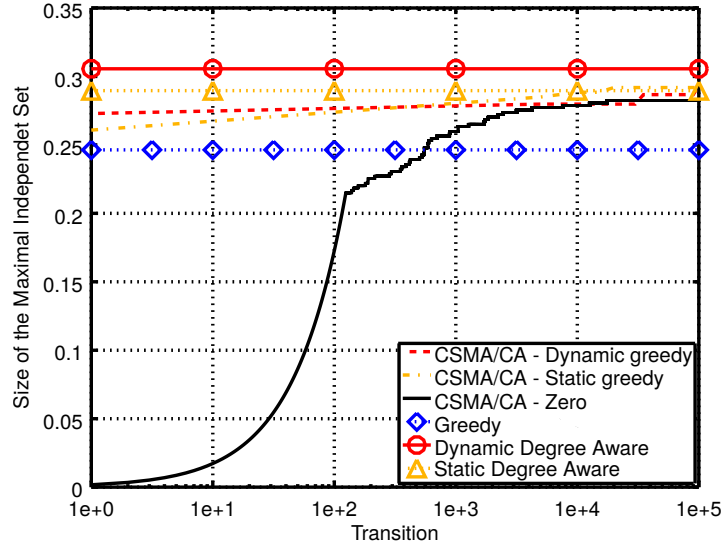


Figura 5.7: Número de vértices transmitiendo durante una corrida del algoritmo CSMA/CA (comenzando desde tres configuraciones iniciales distintas) para el grafo formado por los vértices ilustrados en la Figura 5.6. La abscisa se encuentra en escala logarítmica.

la distribución de grados del grafo original $(p_i)_{i \geq 0}$ y la del conjunto independiente construido $(q_i)_{i \geq 0}$, dada por

$$\|p - q\|_{TV} = \frac{1}{2} \sum_{i=1}^{\infty} |p_i - q_i|. \quad (5.11)$$

Es fácil ver que cuando $\|p - q\|_{TV} = 0$, la probabilidad de que cada vértice se active en algún punto de la evolución del proceso es la misma. Por otro lado, valores más altos de la variación total se corresponderán a situaciones en las que los vértices de algunos grados se conecten con mayor probabilidad que otros; es decir, con situaciones menos equitativas.

Recordemos que para describir el límite de una exploración grado-consciente estática es suficiente integrar una sola EDO. Después de esto, la proporción asintótica de vértices del conjunto independiente encontrado se puede determinar por (5.9) mientras que la distribución de grados del mismo $(q_i)_{i \geq 0}$ por (5.10).

Usando esta relación, se puede calcular la inequidad de los procesos grado-conscientes estáticos con diferentes conjuntos de relojes. Como ejemplo, en la Figura 5.8, presentamos la inequidad y la proporción de vértices en el conjunto independiente para relojes dados por $\lambda(i) = (i + 1)^L$ con $L \in \mathbb{R}$ en función de la potencia L para una distribución de grados de Poisson de media 16.

Como se puede ver en la figura, existe un punto en donde se llega a un equilibrio entre la equidad y el tamaño de conjunto independiente obtenido. Las potencias negativas obtienen un conjunto independiente más grande pero resultan en una mayor inequidad que la dinámica egoísta (que se corresponde con $L = 0$). Por otro lado, se puede lograr una mayor equidad asignando relojes que dependan del grado como potencias positivas a expensas de un conjunto independiente más pequeño. Intuitivamente, esto es así porque con potencias negativas, los vértices con menor grado tienden a conectarse más rápido, lo que resulta en menos vértices bloqueados y, en última instancia, en un conjunto independiente más grande. Por otro lado, las potencias positivas hacen que los vértices con mayor grado se conecten más rápido, lo que

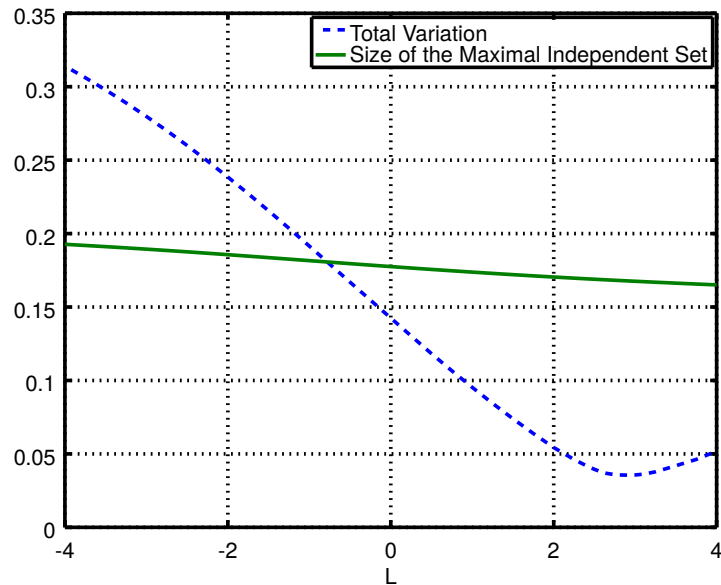


Figura 5.8: *Inequidad y proporción de vértices en el conjunto independiente obtenido en función de la potencia L de la distribución de relojes para una distribución de grados de Poisson de media 16.*

tiende a compensar la inequidad natural de las dinámicas del tipo egoísta discutidas en este capítulo.

De la figura también podemos observar que hay un mínimo de inequidad en el intervalo $(2,5, 3)$. En él, se da una reducción en la inequidad que se logra a costas de una reducción de menos del 10% en el tamaño del conjunto independiente. Esto significa que la inequidad en las transmisiones puede mejorarse considerablemente sin reducir mucho el número de vértices que logran transmitir. Se observa una situación similar para un rango de distribuciones de Poisson con medias entre 4 y 30.

5.4. Discusión

Hemos extendido límites de escala existentes para grafos del Modelo de Configuraciones a dos variantes de algoritmos grado-conscientes: los dinámicos y estáticos. Como vimos, el algoritmo dinámico es capaz de aproximar con precisión arbitraria el algoritmo grado-egoísta. Por lo tanto, a través de la evaluación numérica de un sistema de ecuaciones diferenciales, logramos estimar la constante de atascamiento de algoritmos de exploración grado-egoístas para grafos aleatorios mucho más generales que en el capítulo anterior, además de caracterizar el número de independencia de esta familia de grafos.

La dinámica estática no sólo proporciona resultados de forma cerrada para el tamaño del conjunto independiente descubierto, sino que también lo hace para la distribución de grados de sus vértices. Aunque la dinámica es subóptima, esta propiedad nos permitió analizar el equilibrio entre la equidad y el tamaño del conjunto independiente correspondiente y proponer una selección de parámetros que garantice la mínima inequidad junto con una reducción leve de la capacidad.

También ilustramos la utilidad de nuestros resultados al estimar la capacidad de una red

basada en el protocolo 802.11, tanto para redes teóricas como reales. Es interesante observar que muchos de los resultados anteriores existentes sobre la capacidad de estas redes se han centrado en situaciones de campo medio donde todos los vértices se “ven” entre sí (por ejemplo, en el trabajo seminal de Bianchi [Bia00]), o en pequeños grafos donde es factible calcular el número de independencia (como en [LKLW10, LK16]). Por lo tanto, los métodos de grafos aleatorios ofrecen una alternativa viable a la geometría estocástica, cuya complejidad puede ser prohibitiva. Para una discusión más exhaustiva y una comparación con la geometría estocástica, ver [RLBB17].

Capítulo 6

Mejor respuesta en optimización distribuida

La *optimización distribuida* es un campo que encuentra muchas aplicaciones importantes en sistemas de ingeniería (véase, por ejemplo, [AHK07, CMOP10, YZH14, MS15]). En un problema de optimización distribuida, muchas agentes tienen control sobre las variables del sistema y cada una puede cambiar el valor de algunas de ellas. Además, se define una función de costo sobre los valores de los estados. El objetivo de la optimización distribuida es describir conjuntos de comportamientos de las agentes que resultan en estados que tienen valores altos (o bajos) de esta función.

Estos problemas están íntimamente relacionados con una familia de modelos de teoría de juegos conocidos como *juegos potenciales*. Estos fueron estudiados primero en [Ros73] como un ejemplo de juegos no cooperativos con equilibrios de Nash de estrategia pura. Son juegos para los cuales los cambios de las ganancias inducidos por el ajuste de las estrategias de las jugadoras pueden describirse mediante una función global conocida como *potencial*. Una de las preguntas centrales de la teoría algorítmica de juegos es el estudio de los algoritmos de búsqueda de equilibrios de Nash. Con respecto a esto, los algoritmos de búsqueda de equilibrios en juegos potenciales se han estudiado ampliamente. Por ejemplo, en [ARV08] se demostró que este problema pertenece a la clase de complejidad PLS (siglas en inglés de *búsqueda local polinómica*). Y en [DG16], se estudiaron la optimalidad y la complejidad de la dinámica de mejor respuesta para juegos potenciales aleatorios. Aquí analizamos una generalización de la dinámica inducida por el algoritmo de mejor respuesta definido para problemas de optimización distribuida.

Este capítulo sigue en espíritu el análisis de [DG16]. En este trabajo, los autores analizan la optimalidad y el rendimiento asintóticos (en el número de jugadoras) del algoritmo de mejor respuesta para juegos potenciales aleatorios en los que los potenciales de cada estado del juego están dados por variables aleatorias uniformes $[0, 1]$. Nuestro principal objetivo también será caracterizar la complejidad de la dinámica de mejor respuesta mediante el estudio de la cantidad de pasos y cambios de estados que se requieren antes de llegar a una solución. Pero mientras en [DG16] estudian la dinámica asumiendo que cada mejor respuesta se calcula entre todas las estrategias posibles, nosotros consideraremos el caso en el que se calcula a partir de conjuntos aleatorios de tamaño $\mathcal{O}(1)$. Más específicamente, asumiremos que los cambios de las variables que cada agente puede realizar se modelan como un grafo aleatorio del Modelo

de Configuraciones¹. Además, nuestras técnicas son diferentes de las utilizadas en [DG16] y contextualizamos el problema en términos de problemas de optimización distribuida, lo que permite un rango diferente de relaciones entre el número de estados y agentes.

En este capítulo estudiamos la transición de fase del tiempo necesario para alcanzar un máximo local/global de la función de costo cuando el número de agentes y de variables va a infinito. Para esto, damos estimaciones del tiempo que tarda el algoritmo en ejecutarse en cada uno de tres regímenes (subcrítico, crítico y supercrítico). En el caso especial del régimen subcrítico (al que pertenece el algoritmo de mejor respuesta para juegos potenciales), deducimos el número medio de pasos necesarios para lograr un máximo, damos una descripción probabilística precisa de la dinámica en términos de una familia de procesos de Poisson inhomogéneos y demostramos que el tiempo que tarda el algoritmo en encontrar un máximo es sublineal en el tamaño del grafo. Un resultado sorprendente es que el número medio de pasos que toma el algoritmo de mejor respuesta en este modelo ralo es el mismo que el probado para el caso no ralo en [DG16]. Para el régimen crítico, establecemos una relación entre las cantidades relevantes que determinan la complejidad del algoritmo. Y para el régimen supercrítico demostramos que el tiempo empleado hasta encontrar un máximo es lineal con el tamaño del grafo. Además, con argumentos de acoplamiento, demostramos que la dinámica estudiada es el algoritmo de búsqueda local más rápido.

6.1. Introducción

6.1.1. Algoritmo de mejor respuesta

Definición 6.1.1. Un *problema de optimización distribuida* es una tupla $(A, \mathcal{V}, \mathcal{D}, \chi(\cdot), f(\cdot))$ donde A es el número de agentes, $\mathcal{V} := \{v_1, \dots, v_V\}$ es un conjunto de variables cuyos dominios son $\mathcal{D} := \{D_1, \dots, D_V\}$, $\chi : [A] \rightarrow \mathcal{P}(\mathcal{V})$ (con $[A] := \{1, \dots, A\}$ y $\mathcal{P}(\cdot)$ partes de un conjunto) es la función que determina qué variables puede cambiar qué agente y $f : D_1 \times \dots \times D_V \rightarrow \mathbb{R}$ es una función de costo.

Notemos que, según nuestra definición, muchas agentes pueden cambiar una misma variable. Esto puede parecer redundante a primera vista, pero como nuestro modelo será ralo, dos agentes a cargo de la misma variable pueden cambiar sus valores de diferentes maneras.

Como se señaló anteriormente, el objetivo de la optimización distribuida es encontrar dinámicas que generen estados para los cuales la función de costo sea tan alta (o baja) como sea posible como resultado de las acciones individuales de las agentes. El *algoritmo de mejor respuesta* es una de esas dinámicas que encuentra un máximo local de la función de costo.

Aquí presentamos una versión local del algoritmo. Es decir, en cada estado $\mathbf{v} = (v_1, \dots, v_i, \dots)$, si la agente $j \in [A]$ está llevando a cabo una acción que cambia el valor de v_i , sólo podrá cambiar su valor a otro en un subconjunto determinado $V_{ij}(\mathbf{v}) \subseteq D_i$.

Dado un estado $\mathbf{v} \in D_1 \times \dots \times D_V$ y una agente $j \in [A]$, su correspondiente *mejor respuesta* $b_r^{(j)}(\mathbf{v})$ vendrá dada por el estado \mathbf{v}' para el cual sólo se cambian las variables $v_{j_1}, \dots, v_{j_k} \in \chi(j)$ y da el mayor incremento de $f(\cdot)$ dentro de los posibles $V_{j_i}(\mathbf{v}) \subseteq D_{j_i}$ valores (con $i \in [k]$). Es

¹Como veremos más adelante, cada uno de estos grafos se visitará eventualmente un $\mathcal{O}(n^\epsilon)$ (con $\epsilon > 0$ arbitrariamente pequeño) número de veces por lo que los vecindarios visitados son c.a.p. independientes. Esto implica que la estructura de grafo subyacente no desempeñará un papel central, excepto a través de su distribución de grados. En cualquier caso, esta representación es una buena forma de visualizar la dinámica.

decir,

$$b_r^{(j)}(\mathbf{v}) = \arg \max_{v'_{j_1} \in V_{j_1}(\mathbf{v}) \dots v'_{j_k} \in V_{j_k}(\mathbf{v})} \{f(v_1, \dots, v'_{j_1}, \dots, v'_{j_k}, \dots)\}.$$

Luego, el *algoritmo de mejor respuesta* definirá una sucesión de estados $(\mathbf{v}_i)_{i \geq 1}$ de acuerdo con la siguiente regla: si durante el paso $i \geq 1$ el estado pasa a ser \mathbf{v}_i como resultado de una acción de la agente $j \in [A]$, luego en el siguiente paso la agente $j + 1$ realizará una acción y el estado se cambiará a $\mathbf{v}_{i+1} = b_r^{(j+1)}(\mathbf{v}_i)$. Después de seleccionar la última agente, el algoritmo comienza de nuevo con la primera. Para representar este comportamiento cíclico, denotamos (para $N \geq 1$) al resto de N/A como \bar{N} . Se puede ver fácilmente que este algoritmo converge a un máximo de la función de costo.

En los casos que vamos a analizar, diremos que el algoritmo de mejor respuesta será ralo, ya que los conjuntos $V_j(\mathbf{v})$ de los cuales las agentes pueden elegir los nuevos valores de sus variables serán funciones acotadas del tamaño del sistema.

Juegos potenciales

Los juegos potenciales son modelos muy similares a los problemas de optimización distribuida. Nos centraremos en el caso de los juegos potenciales puros, aunque existen muchas otras generalizaciones.

Definición 6.1.2. Un *juego* es una tupla $\mathcal{G} := (P, S, \mathbf{u})$ donde P es el número de jugadoras, $S = \{s_1, \dots, s_k\}$ es el conjunto de posibles estrategias y $\mathbf{u} : S^P \rightarrow \mathbb{R}^P$ es la función de ganancia.

Dado un juego, sus posibles estados serán los elementos de S^P . Cuando el juego está en un estado dado $\mathbf{s} \in S^P$, diremos que la coordenada k -ésima de $\mathbf{u}(\mathbf{s})$ es la ganancia de la jugadora k en ese estado.

Definición 6.1.3. Un juego es un *juego potencial* si existe una función $\phi : S^P \rightarrow \mathbb{R}$ tal que, para todo $k \in [P]$, si el juego se encuentra en un estado $\mathbf{s} = (s_1, \dots, s_k, \dots, s_P) \in S^P$ y su k -ésima coordenada se cambia a un nuevo valor $s'_k \in S$ tal que $\mathbf{s}' = (s_1, \dots, s'_k, \dots, s_P)$, entonces $u_k(\mathbf{s}') - u_k(\mathbf{s}) = \phi(\mathbf{s}') - \phi(\mathbf{s})$.

Es decir, en un juego potencial cada vez que una jugadora cambia de estrategia, el cambio de su ganancia viene dado por el cambio del potencial inducido por su movimiento.

Definición 6.1.4. Una estrategia $\mathbf{s} := (s_1, \dots, s_P) \in S^P$ es un *equilibrio de Nash* si para todo $s'_k \in S$ (con $k \in [P]$) y $\mathbf{s}' := (s_1, \dots, s'_k, \dots, s_P)$ se tiene que, para todo $l \in [P]$, $u_l(\mathbf{s}) \geq u_l(\mathbf{s}')$.

Una característica importante de los juegos potenciales es que los equilibrios de Nash se pueden definir fácilmente en términos de su potencial. Más precisamente, dado un juego potencial \mathcal{G} con potencial $\phi(\cdot)$, un estado $\mathbf{s} \in S^P$ es un equilibrio de Nash si cualquier cambio en una coordenada de \mathbf{s} disminuye el valor de $\phi(\cdot)$.

El algoritmo de mejor respuesta se define de manera análoga a como se hizo para los problemas de optimización distribuida. De hecho, originalmente proviene de la teoría de juegos. Observemos que existe una estrecha relación entre los juegos potenciales y los problemas de optimización distribuida. En efecto, el potencial en ellos se puede tomar como el análogo de la función de costo. Una de las principales diferencias será que en los juegos potenciales hay una

única jugadora que puede cambiar cada variable. En los modelos de optimización distribuida, por otro lado, puede haber más de una agente controlando una variable o muchas variables controladas por una única agente. Por conveniencia, elegimos usar los nombres y la notación de la optimización distribuida.

6.1.2. Descripción del proceso estudiado

Como se señaló anteriormente, estudiaremos una variante local del algoritmo de mejor respuesta en la que en cada estado \mathbf{v} la mejor respuesta de cada agente no se calcula entre todos los valores posibles de sus variables, sino que es dada por una mejor respuesta local que se calcula a partir de conjuntos aleatorios $V_{j_i}(\mathbf{v}) \subseteq D_{j_i}$ que son de tamaño $\mathcal{O}(1)$. En nuestro modelo, cada agente $i \in [A]$ tendrá un grafo aleatorio G_i asociado que le dará las posibles acciones que puede realizar. Todos los grafos se construirán en el mismo conjunto de vértices $[n]$, donde cada uno de los vértices representa un posible estado del sistema. Luego, cuando una agente realiza una acción y cambia el valor de sus variables, el vértice que representa el estado actual del sistema cambia como consecuencia. Supongamos que el sistema está en el estado \mathbf{v} que está asociado al vértice $v \in [n]$ y la agente $i \in [A]$ va a tomar una acción. El conjunto de acciones posibles (nuevos estados a los que la agente puede mover el sistema) vendrá dado por la vecindad de v en el grafo G_i ; es decir, $\mathcal{N}_i(v)$. A cada vértice v' asociamos un valor $W_{v'}$ de la función de costo. Entonces, la mejor respuesta de la agente i en el vértice v será mover el sistema al vértice $v' \in \mathcal{N}_i(v)$ con el valor más alto de $W_{v'}$. Más precisamente, la dinámica considerada será descrita en términos de un proceso estocástico sobre una construcción de grafos aleatorios. Primero definimos un conjunto de vértices $[n]$ y fijamos el valor del costo de cada vértice $v \in [n]$ por medio de una variable aleatoria $W_v \sim \text{Unif}[0, 1]$ independiente de los otros valores. Para cada agente $i \in [A]$ construimos un grafo del Modelo de Configuraciones² $G_i^{(n)} = ([n], E_i^{(n)})$, todos ellos independientemente de los demás pero en el mismo conjunto de vértices y con la misma distribución de grados $(p_k^{(n)})_{k \geq 1}$ y grado medio $\lambda^{(n)}$. Como en los capítulos anteriores, suponemos que estos grafos del Modelo de Configuraciones satisfacen la (CC) presentada en el Capítulo 2. En cada paso de la dinámica, se dirá que un vértice de la colección de grafos es *inexplorado* si no pertenece a un vecindario ya visitado por el proceso o *explorado* si sí lo hace. La dinámica en cuestión estará dada por una cadena de Markov

$$\left(X_i^{(n)}, U_i^{(n)}, (E_i^{(n)}(k))_{k \in [A]}, (\mu_i^{(n)}(k))_{k \in [A]} \right)_{i \geq 1}$$

definida sobre la colección de grafos $(G_i^{(n)})_{i=1}^A$, donde $X_i^{(n)}$ da el costo del vértice en el que se encuentra el algoritmo durante el paso i , $U_i^{(n)}$ el número de vértices inexplorados, $E_i^{(n)}(k)$ el número de medias aristas libres de vértices explorados en el k -ésimo grafo y $\mu_i^{(n)}(k)$ es la distribución de grados de vértices inexplorados en el k -ésimo grafo. Inicialmente, el proceso se encuentra en un vértice elegido uniformemente entre $[n]$, $X_1^{(n)}$ se establece como el costo de ese vértice, $U_1^{(n)} = n - 1$, $E_1^{(n)}(k) = 0$ y las distribuciones $\mu_1^{(n)}(k) = p_k^{(n)}$. Entonces, si en el paso $i > 1$ el proceso está en algún vértice v_i con un costo W_{v_i} , en el paso $i + 1$ tendremos que $X_{i+1} = \max_{v \in (\mathcal{N}_{i+1}^{(n)}(v_i) \cup \{v_i\})} \{W_v\}$, $U_{i+1}^{(n)}$ será igual a $U_i^{(n)}$ menos una cantidad igual al número

²Aquí suponemos que existe alguna secuencia de grados $\bar{d}^{(n)}$ tal que las secuencias de grados de estos grafos son permutaciones independientes de esta. Entonces, todos tienen la misma distribución de grados, pero cada vértice tiene diferentes grados en diferentes grafos.

de vértices inexplorados en $\mathcal{N}_{\bar{i}+1}(v_i)$ y $E_{\bar{i}+1}^{(n)}(\cdot)$ y $\mu_{\bar{i}+1}^{(n)}(\cdot)$ se actualizarán de acuerdo con los grados de los vértices recién explorados en cada grafo.

Supongamos que después del paso i , el proceso se encuentra en el vértice v_i . En este caso, el número de nuevos valores de costo explorados durante el siguiente paso será $\bar{D}_{i+1}^{(n)} := \sum_{j=1}^{D_i} \mathbb{I}_{\{v_{ij} \in \mathcal{U}_{ij}\}}$, donde $D_i^{(n)}$ es el grado de v_i en el \bar{i} -ésimo grafo, v_{ij} es el vértice apareado con la j -ésima media arista de v_i y \mathcal{U}_{ij} es el conjunto de vértices inexplorados después de que se haya apareado la $(j-1)$ -ésima media arista de v_i . Luego, condicional a la cantidad de vértices inexplorados durante el paso i , $\bar{D}_i^{(n)}$ estará asintóticamente cerca de una variable $\text{Binom}(U_i^{(n)}/n, D_i)$.

Entonces, el número de vértices ya explorados vendrá dado por

$$F_i^{(n)} := n - U_i^{(n)} = \sum_{j=1}^{i-1} \bar{D}_j^{(n)}. \quad (6.1)$$

Aunque para que el proceso represente la dinámica estudiada tenemos que seguir la evolución de las variables $(E^{(n)}(k))_{k \in [A]}$ y $(\mu_i^{(n)}(k))_{k \in [A]}$, como demostraremos más adelante, cada uno de los grafos se visita a lo sumo n^ϵ (con $\epsilon > 0$ arbitrariamente pequeño) veces y como los grafos son independientes entre sí, tendremos que (por cada $k \in [A]$) $E^{(n)}(k) = \lambda^{(n)}(1 - U_i^{(n)})(1 + o(1))$ y $\mu_i^{(n)}(k) = p_k^{(n)}(1 + o(1))$. De esta forma, podremos describir el comportamiento asintótico del proceso asumiendo que estas variables tienen estos valores fijos.

También definiremos una serie de tiempos de saltos inductivamente según $\tau_0^{(n)} := 0$ y para $j \geq 1$, $\tau_{j+1}^{(n)} := \inf\{i \geq 0 : X_{\tau_j+i}^{(n)} > X_{\tau_j}^{(n)}\}$. Estos tiempos dan la cantidad de pasos que el proceso pasa en cada vértice visitado.

Si definimos el *grafo global* $G_g^{(n)} := ([n], \cup_{i=1}^A E_i^{(n)})$ asociado³, las soluciones del problema de optimización vendrán dadas por los máximos locales del costo en él. Notemos que el j -ésimo vértice visitado es un máximo local de la función de costo si $\tau_j^{(n)} = A$. Según la ley de los grande números, $G_g^{(n)}$ será un grafo donde sus vértices tendrán grados iguales a $\lambda A(n) + o_{\mathbb{P}}(\sqrt{A(n)})$. Notar que los máximos locales de $G_g^{(n)}$ son estados absorbentes del algoritmo. Aquí supondremos que la dinámica se ejecuta hasta que se alcanza un vértice que es un máximo local y el algoritmo ha explorado todo su vecindario en $G_g^{(n)}$, comprobando efectivamente que es un máximo. Luego, el algoritmo se ejecuta hasta el tiempo de parada

$$\mathcal{S}^{(n)} := \inf\{i \geq A(n) : X_i^{(n)} = X_{i-A(n)}^{(n)}\}. \quad (6.2)$$

Aunque la mayoría de nuestros resultados caracterizan $\mathcal{S}^{(n)}$, también estudiaremos

$$\mathcal{V}^{(n)} := \sum_{i=1}^{\mathcal{S}^{(n)}} \mathbb{I}_{\{X_i^{(n)} > X_{i-1}^{(n)}\}}, \quad (6.3)$$

es decir, el número de vértices visitados por el proceso; y el número total de vértices explorados por el proceso

$$\mathcal{E}^{(n)} := \sum_{i=1}^{\mathcal{S}^{(n)}} \bar{D}_i. \quad (6.4)$$

³Donde $E_i^{(n)}$ es el conjunto de aristas de $G_i^{(n)}$.

Comparación con la dinámica de mejor respuesta

Existe una diferencia entre esta dinámica y el algoritmo de mejor respuesta discutido anteriormente. Primero, en los problemas de optimización distribuida, los estados tienen una estructura subyacente heredada del hecho de que son elementos del producto $D_1 \times \cdots \times D_A$. Nuestro modelo no tiene en cuenta esta estructura ya que, en un grafo dado, cualquier vértice puede estar conectado a cualquier otro. No obstante, de nuestros resultados se deduce que cada grafo se visita como máximo $\mathcal{O}(n^\epsilon)$ (con $\epsilon > 0$ arbitrariamente pequeño) de veces durante toda la ejecución del algoritmo, lo que hace que las posibles inconsistencias entre las dos dinámicas no sucedan.

También tengamos en cuenta que en el caso de un juego potencial, cada jugador tiene un cierto número fijo K de posibles estrategias. Por lo tanto, el tamaño del grafo se escalará en este caso como función del número de grafos como $n(A) = K^A$. Pero nuestra dinámica se define para cualquier relación entre el número de grafos A y sus tamaños n . Por lo tanto, exploramos las posibles diferencias que surgen cuando uno escala estas dos cantidades de diferentes maneras. En el resto del capítulo, seguimos la convención de tomar el tamaño de los grafos n como parámetro de escala en lugar de A . En este caso, la escala que corresponde al modelo de juegos potenciales tendrá la forma $A(n) = \lceil \log^{-1}(K) \log(n) \rceil$. Como veremos más adelante, esta escala pertenece al régimen subcrítico en nuestro análisis.

Acoplamiento de la sucesión de procesos

Aquí y en el resto del capítulo asumiremos que $(X_i^{(n)})_{i \geq 1}$ es una sucesión de procesos aleatorios como el definido anteriormente, cada uno de los cuales tiene lugar sobre una colección $A(n)$ de grafos construidos independientemente sobre el mismo conjunto de vértices $[n]$ y de acuerdo con un Modelo de Configuraciones de distribución de grados $(p_k^{(n)})_{k \geq 1}$ y media $\lambda^{(n)}$. Supondremos además que estos grafos satisfacen la (CC) del Capítulo 2 hacia una distribución de grados límite $(p_k)_{k \geq 1}$ de media λ .

Sea $(W_j)_{j \geq 1}$ una sucesión de variables aleatorias uniformes independientes $[0, 1]$. Acoplaremos la sucesión de procesos $(X_i^{(n)})_{i \geq 1}$ de tal manera que cada vez que se explore un nuevo vértice, su valor de costo se extraiga de $(W_j)_{j \geq 1}$ secuencialmente. Es decir, en toda la secuencia $(X_i^{(n)})_{i \geq 1}$, el décimo valor de costo explorado vendrá dado por W_{10} independientemente del valor de n . Tengamos en cuenta que esto no afecta las distribuciones individuales de los procesos, sino que hace que sus distribuciones conjuntas se correlacionen de una manera específica. Esto se usará luego para acoplar toda la sucesión a una dinámica más simple.

6.2. Resultados y discusión

Nuestro primer resultado muestra que la dinámica en cuestión es más rápida, en el sentido de orden estocástico, que cualquier otro *algoritmo de búsqueda local*. La palabra *local* aquí significa que en cada paso los algoritmos en cuestión sólo pueden decidir *permanecer* en el vértice en el que se encuentran actualmente o *moverse* a alguno de los vértices recién descubiertos en el vecindario actual que están explorando. Denotamos por \mathcal{L} al conjunto de algoritmos de búsqueda local en este sentido.

Lema 6.2.1. *Sea $(X_i)_{i \geq 1}$ el proceso estocástico presentado en la sección anterior y sea $(\bar{X}_i)_{i \geq 1}$ una realización de un algoritmo en \mathcal{L} , ambos definidos sobre los grafos $(G_i)_{i=1}^{A(n)}$ construidos como antes. Entonces, estos dos procesos pueden ser acoplados de tal forma que (para todo $i \geq 1$) $X_i \geq \bar{X}_i$. Además, si definimos \mathcal{S} y $\bar{\mathcal{S}}$ como los tiempos que les toma a $(X_i)_{i \geq 1}$ y $(\bar{X}_i)_{i \geq 1}$ encontrar un máximo, entonces $\mathcal{S} \leq_{st} \bar{\mathcal{S}}$.*

En [DG16] se establece un resultado de optimalidad análogo, pero un poco más general, para juegos potenciales no ralos. Este lema implica que nuestros resultados caracterizan la complejidad de las búsquedas locales sobre problemas de optimización distribuidos aleatorios y ralos. Los otros resultados del capítulo caracterizan los diferentes tiempos de ejecución posibles para algoritmos con diferentes relaciones de escala para el número de grafos $A(n)$. Parametrizamos esta relación como $A(n) = \lceil n^\alpha \rceil$ con $\alpha > 0$. Con esta parametrización, distinguimos entre tres regímenes diferentes del proceso: el *subcrítico* cuando $\alpha < 1$, el *crítico* cuando $\alpha = 1$ y el *supercrítico* cuando $\alpha > 1$.

Ahora sugeriremos una relación entre esta transición de fase y el número de máximos locales del sistema. Para esto, definamos $\mathcal{M} \subseteq [n]$ como el conjunto de vértices que son máximos locales del costo en $G_g^{(n)}$. Luego, si ordenamos los vértices desde el valor de costo más alto v_1 hasta el más bajo v_n , tenemos ese $v_k \in \mathcal{M}$ si $A_k := \{v_k \not\leftrightarrow \{v_1, \dots, v_{k-1}\}\}$ vale (es decir, el vértice no está conectado a ningún vértice de mayor valor de costo). Es fácil ver que, para $\alpha < 1$, $\mathbb{P}(A_k) = (e^{-\lambda n^{\alpha-1}})^k (1 + o(1))$, lo que hace que $\mathbb{E}(|\mathcal{M}|) = n^{1-\alpha}/\lambda(1 + o(1))$. Y también se puede mostrar que si $\alpha > 1$, $\mathbb{E}(|\mathcal{M}|) = 1 + o(1)$.

Heurísticamente, los diferentes regímenes se corresponden a diferentes posibilidades para el número de máximos presentes en $G_g^{(n)}$. Es decir, en el régimen subcrítico, el número medio de soluciones diverge con el tamaño del grafo. Esto contrasta con el régimen supercrítico en el que el número medio de soluciones es $1 + o(1)$. Esto significa que en el régimen subcrítico hay muchas más soluciones que el algoritmo puede encontrar, lo que sugiere que su tiempo de ejecución debería ser menor que en los casos críticos o supercríticos.

Nuestra principal contribución es la caracterización de los tiempos de ejecución asintóticos para el algoritmo de búsqueda de mejor respuesta. Como se indicó anteriormente, debido a la optimalidad de la dinámica estudiada en el conjunto de algoritmos de búsqueda local, estos resultados determinan la complejidad de la tarea en los tres regímenes diferentes.

Como se muestra en el siguiente teorema, en el régimen subcrítico el algoritmo tarda un tiempo sublineal en encontrar un máximo. Es interesante observar que el comportamiento asintótico del tiempo medio de ejecución es exactamente el mismo que se muestra en [DG16] para juegos potenciales aleatorios no ralos. Esto sugiere que la rareza tiene dos efectos que se compensan exactamente entre sí: por un lado, aumenta el número de máximos con respecto al caso no ralo; mientras que, por otro lado, la exploración lleva más tiempo ya que en cada paso sólo se explora un número de estados $\mathcal{O}(1)$.

En el siguiente teorema conectaremos el número total de pasos en el régimen subcrítico con una construcción probabilística dada en términos de una serie de procesos de Poisson inhomogéneos. Para esto, definamos $K(n) := \lfloor \lambda^{(n)} A(n) \rfloor$. Sea (para $i < K(n)$) $\lambda_i(t) := (1 - e^{-t})^{i-1} e^{-t}$ y $\lambda_K := (1 - e^{-t})^{K(n)}$. Para todo $i \in [K(n)]$ sea $N_t^{(i)}$ un proceso de Poisson independiente e inhomogéneo de intensidad $\lambda_i(t)$. Para estos procesos de Poisson, definamos las variables aleatorias asociadas $\tilde{S}^{(n)} := K(n) + \sum_{i=1}^{K(n)-1} i N_{T_1}^{(i)}$, donde $T_1 := \inf\{t \geq 0 : N_t^{(K)} > 0\}$.

Teorema 6.2.1. *Si $A(n) = \mathcal{O}(n^\alpha)$ (con $\alpha < 1$) y $A(n) \xrightarrow{n \rightarrow \infty} \infty$, entonces*

$$\mathbb{E} \left(\left| \frac{\mathcal{S}^{(n)}}{A(n)} - \frac{\tilde{\mathcal{S}}^{(n)}}{\lambda^{(n)} A(n)} \right| \right) \xrightarrow{n \rightarrow \infty} 0.$$

En particular, $\mathcal{S}^{(n)}/A(n) = \tilde{\mathcal{S}}^{(n)}/(\lambda^{(n)} A(n)) + o_{\mathbb{P}}(1)$. Además, $\mathbb{E}(\mathcal{S}^{(n)}) = e^\gamma A(n)(1 + o(1))$, donde γ es la constante de Euler-Mascheroni.

Como discutimos en las secciones anteriores, este resultado para el régimen subcrítico caracteriza la complejidad asintótica del algoritmo de mejor respuesta para juegos potenciales aleatorios malos porque en este caso $A(n) = \lfloor C \log(n) \rfloor = o(n^\alpha)$ (por todo $\alpha < 1$ y algún $C > 0$). En el caso de un juego potencial aleatorio, la media del tiempo de ejecución total coincide con los resultados de [DG16] pero aquí proporcionamos también una construcción probabilística asintótica para este tiempo.

Proposición 6.2.1. *Si $A(n) = n$, entonces $\mathcal{S}^{(n)}/n = -\frac{1}{\lambda} \log(1 + o(1) - \mathcal{E}^{(n)}/n)$.*

La caracterización de este régimen es aún parcial porque (como es habitual en probabilidad) el régimen crítico es más difícil de estudiar que los demás. Sin embargo, esta proposición puede ser útil ya que la variable $\mathcal{E}^{(n)}$ es más fácil de caracterizar que $\mathcal{S}^{(n)}$. De hecho, de manera similar a la prueba del Teorema 6.2.1, los resultados de la Sección 6.3.1 podrían quizás usarse para aproximar $\mathcal{E}^{(n)}$ y de esta manera puedan derivarse límites para $\mathcal{S}^{(n)}$. Tenemos la intención de hacer esto en el futuro.

Finalmente, el último teorema caracteriza el régimen supercrítico.

Teorema 6.2.2. *Si $A(n) = \lceil n^\alpha \rceil$ con $\alpha > 1$, entonces el número medio de pasos vendrá dado por $\mathbb{E}(\mathcal{S}^{(n)}) = n^\alpha + n/\lambda + o(n)$. Por otro lado, c.a.p. el máximo encontrado es el máximo global y se lo encuentra en un número medio n/λ de pasos.*

Esto implica que en el régimen supercrítico, no importa cuál sea la potencia α del número de grafos, se necesita un tiempo lineal en n para alcanzar un máximo. Esto es una consecuencia del hecho presente en la prueba del teorema de que en este régimen $G_g^{(n)}$ es eventualmente completo. Por lo tanto, en este caso tenemos que el único máximo es el máximo global.

Todos estos resultados están en acuerdo con la heurística discutida anteriormente: en el régimen subcrítico hay muchos máximos y, por lo tanto, la tarea de encontrar uno es más rápida que en los otros regímenes.

6.2.1. Estrategia de las pruebas

La prueba del Teorema 6.2.1 se basa en una caracterización de una dinámica más simple, que llamaremos *proceso de exploración de costos*, estudiada en la sección 6.3. Como veremos, $(X_i)_{i \geq 1}$ se puede acoplar con este proceso de tal forma que venga dado por un proceso de exploración de costos evaluado en un conjunto aleatorio de tiempos. Luego, el número de pasos se deriva de los cálculos para el proceso de exploración de costos y los límites de las probabilidades de ciertos eventos que permiten relacionar ambos procesos.

Para la Proposición 6.2.1, damos un límite de escala para el proceso $U_i^{(n)}$ que da el número de vértices inexplorados. Luego, por medio de una estimación de grandes desvíos para este límite, demostramos la relación establecida en la proposición.

Finalmente, la prueba del Teorema 6.2.2 se basa en una proposición que muestra que el grafo $G_g^{(n)}$ es eventualmente completo en el régimen supercrítico. En ese caso, la única solución estará dada por el máximo global de la función de costo y los resultados del teorema provienen de estimaciones sobre el tiempo que lleva encontrarlo.

6.3. Caracterización del proceso de exploración de costos

En esta sección estudiamos una dinámica simplificada que luego se utilizará en las pruebas de nuestros resultados. Es análoga a la *dinámica libre de intersección* presentada en [DG16], pero donde ellos sólo estudian la media de algunos tiempos de parada, nosotros damos descripciones probabilísticas completas de ellos. Dada una sucesión $(W_i)_{i \geq 1}$ de variables aleatorias independientes y uniformes en $[0, 1]$, definimos su *proceso de exploración de costos* asociado como un proceso estocástico $(\tilde{X}_i)_{i \geq 1}$ donde, para todo $i \geq 1$, $\tilde{X}_i = \max\{W_1, \dots, W_i\}$ y $\tilde{X}_0 = 0$.

Asociado a este proceso, definimos la sucesión de *tiempos de salto* inductivamente como $\tilde{\tau}_0 = 0$ y, para $j \geq 1$, $\tilde{\tau}_{j+1} := \inf\{i \geq 0 : \tilde{X}_{\tilde{\tau}_j+i} > \tilde{X}_{\tilde{\tau}_j}\}$. Aquí estudiamos el tiempo de parada $\tilde{\mathcal{S}}^{(n)} := \inf\{i \geq A(n) : \tilde{X}_i = \tilde{X}_{i-A(n)}\}$ y $\tilde{\mathcal{V}}^{(n)} := \sup\{j \geq 1 : \sum_{k=1}^j \tilde{\tau}_k \leq \tilde{\mathcal{S}}^{(n)}\}$. Estas dos cantidades serán asociadas más tarde a las correspondientes de la dinámica de mejor respuesta.

Si fijamos $t_j := \sum_{k=1}^j \tilde{\tau}_k$, podemos definir el *proceso de saltos* $(\tilde{Y}_j)_{j \geq 1}$ de $(\tilde{X}_i)_{i \geq 1}$ según $\tilde{Y}_j = \tilde{X}_{t_j}$. El proceso \tilde{Y}_j es simplemente el proceso \tilde{X}_i restringido a los tiempos en los que cambia de valor (salta). Este proceso será útil porque muchos cálculos pueden ser realizados explícitamente para él.

6.3.1. Representación por medio de procesos de Poisson

Ahora daremos una caracterización del proceso de saltos $(\tilde{Y}_j)_{j \geq 1}$. Se basa en el hecho de que, en el proceso de saltos, \tilde{Y}_{j+1} se distribuye igual que

$$\tilde{Y}_j + W_{j+1}(1 - \tilde{Y}_j), \quad (6.5)$$

o sea, en cada paso, el salto realizado será uniforme en el intervalo $(\tilde{Y}_j, 1)$. Usando esto, podemos encontrar una caracterización explícita del proceso de saltos en términos de una serie de procesos de Poisson inhomogéneos.

Lema 6.3.1. *Dado el proceso $(\tilde{Y}_j)_{j \geq 1}$, tenemos que $\tilde{Y}_j = 1 - \tilde{W}_1 \cdots \tilde{W}_j$, donde $(\tilde{W}_j)_{j \geq 1}$ variables independientes y uniformes en $[0, 1]$. En particular, $-\ln(1 - \tilde{Y}_j)$ está distribuido de acuerdo a un proceso puntual de Poisson de intensidad 1.*

Demostración. Reagrupando los términos de la relación de recurrencia (6.5) obtenemos que

$$\tilde{Y}_{j+1} = (1 - W_{j+1})\tilde{Y}_j + W_{j+1}. \quad (6.6)$$

Ahora, si definimos $\tilde{W}_j := 1 - W_j$, obtenemos que (para $j \geq 1$) las variables \tilde{W}_j también son independientes y uniformes en $[0, 1]$. Entonces la recurrencia queda como $\tilde{Y}_{j+1} = \tilde{W}_{j+1}\tilde{Y}_j + (1 - \tilde{W}_{j+1})$ que puede verse que se resuelve con $\tilde{Y}_j = 1 - \tilde{W}_1 \cdots \tilde{W}_j$ como queríamos mostrar.

Finalmente, tomando menos logaritmo

$$-\ln(1 - \tilde{Y}_i) = \sum_{j=1}^i -\ln(\tilde{W}_j).$$

Y como las variables $-\ln(\tilde{W}_j)$ se distribuyen como exponenciales de parámetro 1, se sigue el enunciado. \square

Llamemos $(N_t)_{t \geq 0}$ al proceso de Poisson asociado al proceso de saltos. Notar que el j -ésimo punto en el proceso (T_j) está asociado al j -ésimo valor del proceso \tilde{Y}_j según $T_j = -\ln(1 - \tilde{Y}_j)$. Además, (para todo $j \geq 1$) condicionalmente a \tilde{Y}_j , los $\tilde{\tau}_j$ se distribuyen como variables aleatorias geométricas de parámetro $1 - \tilde{Y}_j$, es decir

$$\mathbb{P}(\tilde{\tau}_j \geq k | \tilde{Y}_j) = \tilde{Y}_j^{k-1}. \quad (6.7)$$

Luego tenemos que para cada punto T_j en el proceso su tiempo de salto viene dado por una variable $\tilde{\tau}_j \sim \text{Geom}(e^{-T_j})$. Estos tiempos son, condicionalmente a T_j , también independientes entre sí. Luego podemos describir la distribución de los puntos y los tiempos de salto juntos mediante adelgazamientos del proceso (para una introducción sobre adelgazamientos de procesos de Poisson, consulte el Apéndice B).

Si definimos $N_t^{(A)}$ como el sub-proceso puntual de N_t donde los puntos tienen un tiempo de salto mayor o igual a $A(n)$, $N_t^{(A)}$ es un adelgazamiento de N_t con parámetro de adelgazamiento $\lambda_A(t) = (1 - e^{-t})^{A(n)-1}$. Del mismo modo, al definir (para $i \in [A]$) $N_t^{(i)}$ como el sub-proceso puntual de N_t que tiene puntos con tiempo de salto igual a i , $N_t^{(i)}$ es un adelgazamiento de N_t con parámetro de adelgazamiento $\lambda_i(t) = (1 - e^{-t})^{i-1} e^{-t}$. Por la Proposición B.2.1 del Apéndice B, los procesos resultantes de estos adelgazamientos son procesos de Poisson independientes con intensidades dadas por sus respectivos parámetros de adelgazamiento.

Si llamamos T_1 al primer punto del proceso $N_t^{(A)}$, podemos describir los tiempos de parada que nos interesan como $\tilde{\nu}^{(n)} = 1 + \sum_{i=1}^{A(n)-1} N_{T_1}^{(i)}$ y $\tilde{\mathcal{S}}^{(n)} = A(n) + \sum_{i=1}^{A(n)-1} i N_{T_1}^{(i)}$.

6.3.2. Estimación de los tiempos de saltos

Usando la caracterización dada por el Lema 6.3.1 se puede calcular la distribución de los tiempos de salto $\tilde{\tau}_j$.

Lema 6.3.2. *Para $j \geq 1$, el tiempo de salto $\tilde{\tau}_j$ tiene una distribución dada por*

$$\mathbb{P}(\tilde{\tau}_j > k) = \frac{1}{(j-1)!} \int_0^\infty t^{j-1} e^{-t} (1 - e^{-t})^k dt$$

Demostración. Debido a (6.7), tenemos que

$$\mathbb{P}(\tilde{\tau}_j > k) = \mathbb{E}(\tilde{Y}_j^k).$$

Que de acuerdo con el Lema 6.3.1 está dado por

$$\begin{aligned} \mathbb{P}(\tilde{\tau}_j > k) &= \int_{[0,1]^j} (1 - x_1 \cdots x_j)^k dx_1 \cdots dx_j = \int_0^1 \cdots \int_0^1 (1 - x_1 \cdots x_j)^k dx_1 \cdots dx_j \\ &= \sum_{l=0}^k (-1)^l \binom{k}{l} \int_0^1 x_1^l dx_1 \cdots \int_0^1 x_j^l dx_j \\ &= \sum_{l=0}^k (-1)^l \binom{k}{l} \frac{1}{(l+1)^j}. \end{aligned}$$

Donde en la segunda igualdad utilizamos el binomio de Newton. La última línea no es otra cosa que la k -transformada binomial⁴ de la sucesión $1/(l+1)^j$. Entonces, por la igualdad (8.52) de [Boy18] obtenemos que

$$\mathbb{P}(\tilde{\tau}_j > k) = \frac{1}{(j-1)!} \int_0^\infty t^{j-1} e^{-t} (1 - e^{-t})^k dt,$$

que era lo que queríamos probar. □

Utilizando este lema, probaremos que los valores de los tiempos de salto $\tilde{\tau}_j$ pertenecen, para $\delta > 0$ y con muy alta probabilidad, al intervalo determinístico $(m_j(\delta), M_j(\delta))$.

Proposición 6.3.1. *Para todo $\delta > 0$, sea $m_j(\delta) := e^{j(1-\delta)}$ y $M_j(\delta) := e^{j(1+\delta)}$. Entonces, $\mathbb{P}(\tilde{\tau}_j \leq m_j(\delta)) = \mathcal{O}(e^{-j})$ y $\mathbb{P}(\tilde{\tau}_j \geq M_j(\delta)) = \mathcal{O}(e^{-j})$.*

Demostración. En la prueba utilizaremos que para dos sucesiones positivas $(a_j)_{j \geq 1}, (b_j)_{j \geq 1} \subseteq \mathbb{R}_{\geq 0}$ tales que $a_j, b_j \xrightarrow{j \rightarrow \infty} \infty$ y $a_j^2/b_j \xrightarrow{j \rightarrow \infty} 0$, entonces

$$\left(1 - \frac{a_j}{b_j}\right)^{b_j} = e^{-a_j} (1 + \mathcal{O}(a_j^2/b_j)). \quad (6.8)$$

Para la cota superior, fijemos δ' tal que $0 < \delta' < \delta$ y definamos $\tilde{M}_j(\delta') := j e^{j(1+\delta')}$ y $t_j^*(\delta') := j(1+\delta')$. Entonces, por la Proposición 6.3.2 tenemos que

$$\begin{aligned} \mathbb{P}(\tilde{\tau}_j \geq M_j) &\leq \mathbb{P}(\tilde{\tau}_j \geq \tilde{M}_j) = \frac{1}{(j-1)!} \int_0^\infty t^{j-1} e^{-t} (1 - e^{-t})^{\tilde{M}_j} dt \\ &\leq \left(1 - e^{-t_j^*}\right)^{\tilde{M}_j} \frac{1}{(j-1)!} \int_0^{t_j^*} t^{j-1} e^{-t} dt + \frac{1}{(j-1)!} \int_{t_j^*}^\infty t^{j-1} e^{-t} dt, \end{aligned}$$

donde usamos que $(1 - e^{-t})^{\tilde{M}_j}$ es una función creciente acotada por 1. Ahora, si definimos una variable $Z_j \sim \Gamma(1, j)$, podemos reescribir esto como

$$\mathbb{P}(\tilde{\tau}_j \geq M_j) \leq \left(1 - e^{-t_j^*}\right)^{\tilde{M}_j} \mathbb{P}(Z_j \leq j(1+\delta')) + \mathbb{P}(Z_j \geq j(1+\delta')).$$

Pero las variables $\Gamma(1, j)$ se distribuyen como la suma de j variables exponenciales de parámetro 1. Entonces, podemos aplicar el Teorema de Cramer [DZ10] a Z_j , del cual obtenemos que

$$\mathbb{P}(\tilde{\tau}_j \geq M_j) \leq \left(1 - \frac{j}{\tilde{M}_j}\right)^{\tilde{M}_j} + e^{-I(\delta')j} = e^{-j} (1 + \mathcal{O}(j^2 e^{-j(1+\delta')})) + e^{-I(\delta')j}. \quad (6.9)$$

Donde usamos (6.8) y definimos $I(\cdot)$ como la función de tasa de una variable exponencial de parámetro 1. La prueba para la cota superior de $\mathbb{P}(\tilde{\tau}_j \leq m_j(\delta))$ es completamente análoga. □

Como las variables de la Proposición 6.3.1 son sumables, se sigue el siguiente corolario.

Corolario 6.3.1. *Para todo $\delta > 0$, casi seguro y eventualmente $m_j(\delta) \leq \tilde{\tau}_j \leq M_j(\delta)$.*

⁴Consultar [Boy18] para una introducción a la transformada binomial.

6.3.3. Estimación del tiempo total de corrida

Debido a la independencia de T_1 respecto de los procesos $N_t^{(i)}$ tenemos el siguiente lema, que es una variación de la fórmula de Campbell (ver Apéndice B).

Lema 6.3.3. Sean T_1 y $N_t^{(i)}$ como antes, entonces tenemos que para todo $i \in [A(n) - 1]$

$$\mathbb{E} \left(N_{T_1}^{(i)} \right) = \int_0^\infty \lambda_i(t) e^{-\int_0^t \lambda_A(s) ds} dt.$$

Demostración. Definamos $l_{jk} := [(k-1)/j, k/j)$. Fijando $j \geq 1$, entonces para todo $i \in [A(n) - 1]$,

$$\sum_{k \geq 1} \mathbb{E} \left(N_{(k-1)/j}^{(i)} \right) \mathbb{P}(T_1 \in l_{jk}) \leq \mathbb{E} \left(N_{T_1}^{(i)} \right) \leq \sum_{k \geq 1} \mathbb{E} \left(N_{k/j}^{(i)} \right) \mathbb{P}(T_1 \in l_{jk}).$$

Donde usamos la independencia de $N_{k/j}^{(i)}$ respecto de T_1 . Y como $\mathbb{E} \left(N_{k/j}^{(i)} \right) = \int_0^{k/j} \lambda_i(t) dt$, dejando que $j \rightarrow \infty$ se obtiene que

$$\begin{aligned} \mathbb{E} \left(N_{T_1}^{(i)} \right) &= \int_0^\infty \left(\int_0^s \lambda_i(t) dt \right) \mathbb{P}_{T_1}(ds) \\ &= \int_0^\infty \lambda_i(t) \mathbb{P}(T_1 > t) dt. \end{aligned}$$

En donde en la última igualdad cambiamos el orden de integración. □

Usando este lema, podemos calcular el valor medio de $\tilde{\mathcal{S}}^{(n)}$.

Proposición 6.3.2. Dada la construcción detallada más arriba, tenemos que $\mathbb{E} \left(\tilde{\mathcal{S}}^{(n)} \right) = e^{H_{A(n)-1}}$, donde (para $n \in \mathbb{N}_0$) H_n es el n -ésimo número armónico.

Demostración. Primero, notemos que

$$\sum_{i=1}^{A(n)-1} i \lambda_i(t) = e^t [1 - (1 - e^{-t})^{A(n)-1}] - (A(n) - 1)(1 - e^{-t})^{A(n)-1}.$$

Entonces, por el Lema 6.3.3 tenemos que

$$\mathbb{E} \left(\tilde{\mathcal{S}}^{(n)} \right) = I_1 + (A(n) - 1)I_2 + A(n),$$

donde

$$I_1 = \int_0^\infty e^t [1 - (1 - e^{-t})^{A(n)-1}] e^{-\int_0^t (1 - e^{-s})^{A(n)-1} ds} dt = e^{\int_0^\infty 1 - (1 - e^{-s})^{A(n)-1} ds} - 1,$$

y

$$I_2 = - \int_0^\infty (1 - e^{-t})^{A(n)-1} e^{-\int_0^t (1 - e^{-s})^{A(n)-1} ds} dt = -1.$$

La primera integral fue calculada realizando el cambio de variables $u = e^{\int_0^t 1 - (1 - e^{-s})^{A(n)-1} ds}$, mientras que para la segunda se utilizó $v = e^{-\int_0^t (1 - e^{-s})^{A(n)-1} ds}$. Entonces,

$$\mathbb{E} \left(\tilde{\mathcal{S}}^{(n)} \right) = e^{\int_0^\infty 1 - (1 - e^{-s})^{A(n)-1} ds}.$$

Para probar el resultado, alcanza con ver que $\int_0^\infty 1 - (1 - e^{-s})^{A(n)-1} ds = H_{A(n)-1}$. Por el binomio de Newton,

$$\begin{aligned} \int_0^\infty 1 - (1 - e^{-s})^{A(n)-1} ds &= \int_0^\infty \sum_{l=1}^{A(n)-1} (-1)^l \binom{A(n)-1}{l} \int_0^\infty e^{-ls} ds \\ &= (A(n)-1) \sum_{l'=0}^{A(n)-2} (-1)^{l'} \binom{A(n)-2}{l'} \frac{1}{(l'+1)^2} \end{aligned}$$

en donde en la segunda igualdad se usó que (para $0 \leq k \leq n$) $\binom{n}{k} = n/k \binom{n-1}{k-1}$ y se definió $l' = l - 1$. Además, la suma en la última igualdad es la transformada binomial de la sucesión $1/(l'+1)^2$ que, por la ecuación (8.39) en [Boy18], es igual a $H_{A(n)-1}/(A(n)-1)$. \square

Utilizando esta construcción, también podemos derivar una cota superior para el segundo momento del número total de pasos.

Proposición 6.3.3. *Dada la construcción anterior, tenemos que*

$$\mathbb{E} \left[\left(\tilde{\mathcal{S}}^{(n)} \right)^2 \right] \leq 2A^2(n) + A(n)e^{H_{A(n)-1}} - 1 = \mathcal{O}(A^2(n)).$$

Demostración. Primero notemos que, para todo $i \geq 1$, $N_{T_1}^{(i)} \leq N_\infty^{(i)}$. Además, el valor medio del lado derecho de la desigualdad puede calcularse explícitamente según

$$\begin{aligned} \mathbb{E} \left(N_\infty^{(i)} \right) &= \int_0^\infty (1 - e^{-t})^{i-1} e^{-t} dt = \sum_{k=0}^{i-1} (-1)^k \binom{i-1}{k} \int_0^\infty e^{-(k+1)t} dt \\ &= \sum_{k=0}^{i-1} (-1)^k \binom{i-1}{k} \frac{1}{k+1} = \frac{1}{i} \end{aligned} \tag{6.10}$$

donde en la última igualdad utilizamos la relación (9.25) de [Boy18]. Utilizando (6.10), podemos acotar el segundo momento de $\tilde{\mathcal{S}}^{(n)}$ según

$$\begin{aligned} \mathbb{E} \left[\left(\tilde{\mathcal{S}}^{(n)} \right)^2 \right] &\leq \mathbb{E} \left[\left(A(n) + \sum_{i=1}^{A(n)-1} i N_\infty^{(i)} \right)^2 \right] \\ &= \sum_{i=1}^{A(n)-1} i^2 \text{Var} \left(N_\infty^{(i)} \right) + \sum_{i=1}^{A(n)-1} i^2 \mathbb{E} \left(N_\infty^{(i)} \right)^2 \\ &\quad + \sum_{j < i}^{A(n)-1} ij \mathbb{E} \left(N_\infty^{(i)} N_\infty^{(j)} \right) + A(n)e^{H_{A(n)-1}} + A^2(n) \\ &= \sum_{i=1}^{A(n)-1} i^2 \mathbb{E} \left(N_\infty^{(i)} \right) + \sum_{i=1}^{A(n)-1} i^2 \mathbb{E} \left(N_\infty^{(i)} \right)^2 \\ &\quad + \sum_{j < i}^{A(n)-1} ij \mathbb{E} \left(N_\infty^{(i)} \right) \mathbb{E} \left(N_\infty^{(j)} \right) + A(n)e^{H_{A(n)-1}} + A^2(n) \\ &= \sum_{i=1}^{A(n)-1} i + \sum_{i=1}^{A(n)-1} 1 + \sum_{j < i}^{A(n)-1} 1 + A(n)e^{H_{A(n)-1}} + A^2(n) \\ &= 2A^2(n) + A(n)e^{H_{A(n)-1}} - 1, \end{aligned}$$

donde en la primera igualdad usamos la Proposición anterior y en la segunda utilizamos la independencia de $N_\infty^{(i)}$ respecto de $N_\infty^{(j)}$ y que, como los procesos son de Poisson, $\text{Var}(N_\infty^{(i)}) = \mathbb{E}(N_\infty^{(i)})$. \square

6.4. Pruebas de los resultados

En esta sección presentamos las pruebas de los resultados de la Sección 6.2. A lo largo de muchos de ellos, hacemos uso del hecho de que podemos acoplar una realización del proceso de exploración de costos y el algoritmo de mejor respuesta para que este último sea igual al primero restringido a un conjunto aleatorio de tiempos.

Como se mencionó anteriormente, definiendo una sucesión $(W_i)_{i \geq 1}$ de variables aleatorias independientes uniformes $[0, 1]$, acoplamos la sucesión de exploraciones de mejor respuesta $(X_i^{(n)})_{i \geq 1}$ para que cada vez que se explore un nuevo vértice, su valor se tome secuencialmente de esa sucesión. Entonces también podemos acoplar toda la sucesión de exploraciones de mejor respuesta a una exploración de costo única: la asociada a $(W_i)_{i \geq 1}$, que denotaremos como $(\tilde{X}_i)_{i \geq 1}$. Bajo este acoplamiento tenemos que para todo $n, i \geq 1$, $X_i^{(n)} = \tilde{X}_{F_i^{(n)}}$ (donde $F_i^{(n)}$ es el número de vértices explorados hasta el paso i tal como fue definido en (6.1)).

Una observación importante que se utilizará en las pruebas es que cada vértice se visita como máximo durante $A(n)$ pasos (porque después de eso, se declara una solución y el algoritmo se detiene). Entonces, tenemos la siguiente cota:

$$\mathcal{S}^{(n)} \leq \mathcal{V}^{(n)} A(n). \quad (6.11)$$

Donde ambas cantidades han sido definidas en (6.2) y (6.3), respectivamente. Y como el número de vértices visitados es menor o igual a n , tenemos que

$$\mathcal{S}^{(n)} \leq A(n)n. \quad (6.12)$$

Además, combinando la desigualdad (6.11) con la siguiente proposición se llega a una cota más ajustada para $\mathcal{S}^{(n)}$.

Proposición 6.4.1. *Para todo $\epsilon > 0$, existe algún $\gamma > 0$ tal que $\mathbb{P}(\mathcal{V}^{(n)} \geq n^\epsilon) = \mathcal{O}(e^{-\gamma n^\epsilon})$.*

Demostración. Para cada vértice $v_j \in [n]$ visitado en el paso $i_j \geq 1$, definamos c_j como el número de vértices inexplorados en la vecindad de v_j en $G_g^{(n)}$ al momento en que se explora. Claramente, v_j es una solución si su costo es mayor que todos estos c_j valores nuevos. Entonces, v_j es una solución si su tiempo de salto asociado en el proceso de exploración de costos $\tilde{\tau}_{F_{i_j}^{(n)}} = \tilde{\tau}_{i_j+k_j}$ (para algún $k_j \geq 0$) es mayor que c_j . Por lo tanto, para todo $\delta > 0$, por el Lema 6.3.1, tenemos que

$$\mathbb{P}(\mathcal{V}^{(n)} > n^\epsilon) \leq \mathbb{P}(\tilde{\tau}_{n^\epsilon+k_n^\epsilon} < c_{n^\epsilon}) \leq \mathbb{P}(\tilde{\tau}_{n^\epsilon+k_n^\epsilon} < n) = \mathcal{O}(e^{-n^\epsilon(1-\delta)}), \quad (6.13)$$

donde en la primera desigualdad usamos que $\{\mathcal{V}^{(n)} > n^\epsilon\} \subseteq \{\tilde{\tau}_{n^\epsilon} < c_{n^\epsilon}\}$ y para la segunda que $c_{n^\epsilon} \leq n$. \square

Esta última proposición muestra que cada uno de los grafos $A(n)$ se visita a lo sumo un número n^ϵ de veces (con $\epsilon > 0$ arbitrariamente pequeño). Esto implica que cada vecindario de

los grafos individuales está c.a.p. desconectado de los demás y que la distribución de grados de los grafos no cambia significativamente durante el proceso (el cambio es de orden $o(1)$). En este sentido, las estructuras de grafos individuales subyacentes no son tan relevantes como sus distribuciones de grados.

6.4.1. Prueba del Lema 6.2.1

Para probar este lema, acoplamos $(X_i)_{i \geq 1}$ y $(\bar{X}_i)_{i \geq 1}$ de tal manera que durante cada paso, ambas exploraciones revelen el mismo número de vértices y con los mismos valores de costo. Esto es posible porque si combinamos ambas exploraciones para descubrir el mismo número de vértices y nuevos valores de costo, todas las demás variables del proceso seguirán siendo las mismas. Esto se logra construyendo los grafos y descubriendo los valores de los costos simultáneamente con la ejecución de los algoritmos. En la literatura de algoritmos aleatorios, esto se conoce como *principio de decisión diferida* [MU17].

Supongamos que el algoritmo de mejor respuesta es descrito por el siguiente proceso

$$\left(X_i, U_i, (E_i(k))_{k=1}^A, (\mu_i(k))_{k=1}^A \right)_{i \geq 1}$$

(tal como se lo presentó en la Sección 6.1) y definamos el proceso análogo para la otra exploración como

$$\left(\bar{X}_i, \bar{U}_i, (\bar{E}_i(k))_{k=1}^A, (\bar{\mu}_i(k))_{k=1}^A \right)_{i \geq 1}.$$

Cada dinámica se definirá sobre una colección de grafos $(G_i)_{k=1}^A$ y $(\bar{G}_i)_{k=1}^A$, respectivamente; donde ambas colecciones están compuestas por grafos independientes del Modelo de Configuraciones con la misma distribución de grados $(p_k)_{k \geq 1}$. Queremos mostrar que podemos acoplar la realización de ambos procesos (hasta que uno de ellos encuentre una solución) para que para todo $1 \leq i \leq \mathcal{S} \wedge \bar{\mathcal{S}}$ y $k \in [A]$ tengamos que $X_i \geq \bar{X}_i$, $U_i = \bar{U}_i$, $E_i(k) = \bar{E}_i(k)$ y $\mu_i(k) = \bar{\mu}_i(k)$. Para esto, supongamos que inicialmente ambos procesos tienen los mismos valores. Mostraremos que si los procesos en el paso $i \geq 1$ satisfacen estas relaciones, todavía las satisfarán en el paso $i + 1$. De hecho, debido a la hipótesis inductiva, los grafos explorados durante el paso $i + 1$ tendrán el mismo número de vértices explorados ($U_i = \bar{U}_i$), las mismas distribuciones de grados sin explorar ($\mu_i(k) = \bar{\mu}_i(k)$) y la misma cantidad de medias aristas libres de vértices explorados ($E_i(k) = \bar{E}_i(k)$), entonces las exploraciones se pueden acoplar en este paso para que los vértices de los vecindarios explorados por ambos tengan los mismos grados y la misma cantidad de vértices explorados/no explorados en ellos. Esto garantiza que $U_{i+1} = \bar{U}_{i+1}$, $E_{i+1}(k) = \bar{E}_{i+1}(k)$ y $\mu_{i+1}(k) = \bar{\mu}_{i+1}(k)$. Además, los valores de los nuevos costos explorados pueden combinarse para ser iguales para ambos procesos. Y debido a que el algoritmo de mejor respuesta siempre está en el vértice con el valor de costo más alto explorado hasta el momento, esto significa que $X_{i+1} \geq \bar{X}_{i+1}$. Lo que avanza la inducción, demostrando la primera parte del lema.

Para la segunda parte, definamos T_j como el tiempo, en el paso $j \geq 1$, desde el último cambio de valor de X_i y \bar{T}_j de forma análoga para \bar{X}_i . El proceso X_i ha encontrado una solución en el paso $i \geq 1$ sii $T_i = A(n)$. Análogamente, el proceso \bar{X}_i ha encontrado una solución en el paso $i \geq 1$ sii $\bar{T}_i = A(n)$ y no se encontró ningún vértice de mayor valor de costo desde el último salto. Notar que, bajo este acoplamiento, si $\bar{T}_i > T_i$, entonces el vértice

\bar{v}_i asociado al costo \bar{X}_i no es una solución. En efecto, si $\bar{T}_i > T_i$, significa que el proceso X_i ha saltado después del último salto de \bar{X}_i . Pero esto implica que se ha encontrado algún vértice con un costo superior a \bar{X}_i en alguno de los vecindarios de \bar{v}_i , y por lo tanto no es una solución. Entonces, $\mathcal{S} \leq \bar{\mathcal{S}}$ bajo este acoplamiento, lo que demuestra el lema.

6.4.2. Prueba del Teorema 6.2.1

Como se señaló anteriormente, la dinámica de mejor respuesta es la misma que la exploración de costos discutida en la sección anterior pero restringida a una colección aleatoria de tiempos. Este hecho será importante a lo largo de la prueba.

Ahora definiremos cuatro eventos que nos permitirán reducir el comportamiento del algoritmo en el régimen subcrítico al discutido en la sección anterior:

- $A_\epsilon := \{\mathcal{V}^{(n)} \geq n^\epsilon\}$, o sea, el evento bajo el cual *demasiados vértices son visitados*;
- $B_\epsilon := \left\{ \left| \frac{1}{\mathcal{S}^{(n)}} \sum_{i=1}^{\mathcal{S}^{(n)}} D_i^{(n)} - \lambda^{(n)} \right| > \epsilon \right\}$, o sea, el evento bajo el cual *las vecindades visitadas son o bien demasiado grandes o bien demasiado chicas*;
- $B'_\epsilon := \left\{ \exists j \leq \mathcal{V}^{(n)} : \left| \frac{1}{A^{(n)}} \sum_{i=1}^{A^{(n)}} D_{ij}^{(n)} - \lambda^{(n)} \right| > \epsilon \right\}$ (donde $D_{ij}^{(n)}$ es el grado del j -ésimo vértice visitado en el grafo $G_i^{(n)}$), o sea, el evento bajo el cual *las vecindades en $G_g^{(n)}$ de los vértices visitados son o bien demasiado grandes o bien demasiado chicas*;
- y $C_\epsilon := \{N_r \geq \epsilon n^\alpha\}$, o sea, el evento bajo el cual *demasiados vértices explorados son explorados múltiples veces*.

Aquí $N_r := \sum_{i=1}^{\mathcal{S}^{(n)}} D_i^{(n)} - \bar{D}_i^{(n)}$ es el número total de veces que vértices ya explorados se visitan durante la exploración y $\mathcal{V}^{(n)}$ fue definido en (6.3). Al estimar las probabilidades de estos cuatro eventos “malos”, podemos determinar el comportamiento asintótico del número total de pasos durante el proceso. Pero primero probaremos que, en el complemento de estos eventos, el número de pasos $\mathcal{S}^{(n)}$ puede estimarse por el tiempo de detención correspondiente del proceso de exploración de costos asociado.

Dado el proceso de exploración de costos \tilde{X}_i acoplado con la sucesión de procesos de mejor respuesta $X_i^{(n)}$, definimos los siguientes tiempos de parada (para $\epsilon > 0$):

$$\tilde{\mathcal{S}}^{(n)} := \inf\{j \geq 1 : \tilde{X}_j = \tilde{X}_{j - \lceil \lambda^{(n)} A^{(n)} \rceil}\}$$

$$\tilde{\mathcal{S}}_{\epsilon^+}^{(n)} := \inf\{j \geq 1 : \tilde{X}_j = \tilde{X}_{j - \lceil (\lambda^{(n)} + \epsilon) A^{(n)} \rceil}\}$$

$$\tilde{\mathcal{S}}_{\epsilon^-}^{(n)} := \inf\{j \geq 1 : \tilde{X}_j = \tilde{X}_{j - \lfloor (\lambda^{(n)} - 2\epsilon) A^{(n)} \rfloor}\}.$$

Lema 6.4.1. *Condiciona a los eventos B_ϵ^c , $B'_\epsilon{}^c$ y C_ϵ^c , $\frac{\tilde{\mathcal{S}}_{\epsilon^-}^{(n)}}{\lambda^{(n)} + \epsilon} \leq \mathcal{S}^{(n)} \leq \frac{\tilde{\mathcal{S}}_{\epsilon^+}^{(n)}}{\lambda^{(n)} - 2\epsilon}$.*

Demostración. Primero notemos que

$$\mathcal{S}^{(n)} = \frac{\mathcal{E}^{(n)}}{\frac{1}{\mathcal{S}^{(n)}} \mathcal{E}^{(n)}} = \frac{\mathcal{E}^{(n)}}{\frac{1}{\mathcal{S}^{(n)}} \sum_{i=1}^{\mathcal{S}^{(n)}} \bar{D}_i^{(n)}}.$$

Pero condicionalmente al evento C_ϵ^c , $\sum_{i=1}^{\mathcal{S}^{(n)}} \bar{D}_i^{(n)} \geq \sum_{i=1}^{\mathcal{S}^{(n)}} D_i^{(n)} - \epsilon \mathcal{S}^{(n)}$. Y condicionalmente a B_ϵ^c , $\sum_{i=1}^{\mathcal{S}^{(n)}} D_i^{(n)} \geq (\lambda^{(n)} - \epsilon) \mathcal{S}^{(n)}$. Entonces

$$\mathcal{S}^{(n)} \leq \frac{\mathcal{E}^{(n)}}{\lambda^{(n)} - 2\epsilon}.$$

Es más, en el evento $B_\epsilon'^c$ ninguna vecindad en $G_g^{(n)}$ de un vértice visitado es más grande que $(\lambda^{(n)} + \epsilon)A(n)$, entonces por el acoplamiento presentado al principio de la sección tendremos que $\mathcal{E}^{(n)} \leq \tilde{\mathcal{S}}_{\epsilon^+}^{(n)}$. De lo que se obtiene que

$$\mathcal{S}^{(n)} \leq \frac{\tilde{\mathcal{S}}_{\epsilon^+}^{(n)}}{\lambda^{(n)} - 2\epsilon}.$$

Mientras que la cota inferior puede ser obtenida de una manera análoga. □

Prueba del Teorema 6.2.1. Primero daremos estimaciones a las probabilidades de los eventos A_ϵ , B_ϵ , B_ϵ' y C_ϵ . La Proposición 6.4.1 y (6.11) implican que

$$\mathbb{P}(A_\epsilon) = \mathcal{O}(e^{-n^\epsilon}). \tag{6.14}$$

Para la estimación del segundo evento, asumimos que el evento A_ϵ^c vale. Por las desigualdades (6.11) y (6.12) en este evento $A(n) = n^\alpha \leq \mathcal{S}^{(n)} \leq n^{\alpha+\epsilon}$, de lo que se llega a que

$$\begin{aligned} \mathbb{P}(B_\epsilon | A_\epsilon^c) &\leq n^\epsilon \mathbb{P}\left(\left|\sum_{i=1}^{n^{\alpha+\epsilon}} (D_i^{(n)} - \lambda^{(n)})\right| > \epsilon n^\alpha\right) \\ &\leq n^\epsilon \frac{n^{\alpha+\epsilon} \text{Var}(|D - \lambda|)}{\epsilon^2 n^{2\alpha}} (1 + o(1)) = \mathcal{O}(n^{2\epsilon-\alpha}). \end{aligned} \tag{6.15}$$

Donde en la última línea usamos la desigualdad de Chebychev. El cálculo respectivo para B_ϵ' es muy similar y lleva a que

$$\mathbb{P}(B_\epsilon' | A_\epsilon^c \cap B_\epsilon^c) \leq \mathbb{P}(B_\epsilon' | A_\epsilon^c) (1 + o(1)) = \mathcal{O}(n^{\epsilon-\alpha}). \tag{6.16}$$

Finalmente, para el último evento observemos que

$$\begin{aligned} \mathbb{E}(N_r | A_\epsilon^c \cap B_\epsilon^c \cap B_\epsilon'^c) &= \sum_{i=1}^{\mathcal{S}^{(n)}} \sum_{j=1}^{D_i} \mathbb{P}(v_{ij} \in \mathcal{U}_{ij}^c | A_\epsilon^c \cap B_\epsilon^c \cap B_\epsilon'^c) \\ &\leq \sum_{i=1}^{\mathcal{S}^{(n)}} \sum_{j=1}^{D_i} \frac{(\lambda^{(n)} + \epsilon)A(n)n^\epsilon}{n} \\ &\leq \frac{[(\lambda^{(n)} + \epsilon)A(n)n^\epsilon]^2}{n} = \mathcal{O}(n^{2(\alpha+\epsilon)-1}). \end{aligned}$$

Donde usamos que bajo el evento $A_\epsilon^c \cap B_\epsilon^c$ el número total de vértices explorados es a lo sumo $(\lambda^{(n)} + \epsilon)A(n)$ veces $\mathcal{V}^{(n)}$ (que a su vez es a lo sumo n^ϵ). Entonces, por la desigualdad de Markov tenemos que

$$\mathbb{P}(N_r \geq \epsilon n^\alpha | A_\epsilon^c \cap B_\epsilon^c \cap B_\epsilon'^c) = \mathcal{O}(n^{\alpha+2\epsilon-1}). \tag{6.17}$$

Definiendo el evento $D_\epsilon = A_\epsilon^c \cap B_\epsilon^c \cap B_\epsilon'^c \cap C_\epsilon^c$, la diferencia en L^1 entre $\mathcal{S}^{(n)}$ y $\tilde{\mathcal{S}}^{(n)}$ es dada por

$$\begin{aligned} \mathbb{E} \left(\left| \frac{\mathcal{S}^{(n)}}{A(n)} - \frac{\tilde{\mathcal{S}}^{(n)}}{\lambda^{(n)} A(n)} \right| \right) &= \mathbb{E} \left(\left| \frac{\mathcal{S}^{(n)}}{A(n)} - \frac{\tilde{\mathcal{S}}^{(n)}}{\lambda^{(n)} A(n)} \right| \mathbb{I}_{D_\epsilon} \right) + \mathbb{E} \left(\frac{\mathcal{S}^{(n)}}{A(n)} \mathbb{I}_{A_\epsilon} \right) \\ &\quad + \mathbb{E} \left(\frac{\mathcal{S}^{(n)}}{A(n)} \mathbb{I}_{A_\epsilon^c \cap B_\epsilon} \right) + \mathbb{E} \left(\frac{\mathcal{S}^{(n)}}{A(n)} \mathbb{I}_{A_\epsilon^c \cap B_\epsilon^c \cap B_\epsilon'} \right) \\ &\quad + \mathbb{E} \left(\frac{\mathcal{S}^{(n)}}{A(n)} \mathbb{I}_{A_\epsilon^c \cap B_\epsilon^c \cap B_\epsilon'^c \cap C_\epsilon^c} \right) + \mathbb{E} \left(\frac{\tilde{\mathcal{S}}^{(n)}}{\lambda^{(n)} A(n)} \mathbb{I}_{D_\epsilon^c} \right) \end{aligned} \quad (6.18)$$

Donde usamos que $D_\epsilon^c = A_\epsilon \cup (A_\epsilon^c \cap B_\epsilon) \cup (A_\epsilon^c \cap B_\epsilon^c \cap B_\epsilon') \cup (A_\epsilon^c \cap B_\epsilon^c \cap B_\epsilon'^c \cap C_\epsilon^c)$. Ahora acotaremos los distintos términos en el lado derecho de esta expresión. Para el primero notemos que

$$\mathbb{E} \left(\left| \frac{\mathcal{S}^{(n)}}{A(n)} - \frac{\tilde{\mathcal{S}}^{(n)}}{\lambda^{(n)} A(n)} \right| \mathbb{I}_{D_\epsilon} \right) \leq \mathbb{E} \left(\left| \frac{\tilde{\mathcal{S}}_{\epsilon^+}^{(n)}}{(\lambda^{(n)} - 2\epsilon) A(n)} - \frac{\tilde{\mathcal{S}}_{\epsilon^-}^{(n)}}{(\lambda^{(n)} + \epsilon) A(n)} \right| \right)$$

ya que $\tilde{\mathcal{S}}_{\epsilon^-}^{(n)} \leq \tilde{\mathcal{S}}^{(n)} \leq \tilde{\mathcal{S}}_{\epsilon^+}^{(n)}$ y por el Lema 6.4.1 bajo D_ϵ tenemos que $\frac{\tilde{\mathcal{S}}_{\epsilon^-}^{(n)}}{\lambda^{(n)} + \epsilon} \leq \mathcal{S}^{(n)} \leq \frac{\tilde{\mathcal{S}}_{\epsilon^+}^{(n)}}{\lambda^{(n)} - 2\epsilon}$. Además, porque $\tilde{\mathcal{S}}_{\epsilon^-}^{(n)} \leq \tilde{\mathcal{S}}_{\epsilon^+}^{(n)}$ entonces el módulo puede ser omitido del lado derecho de la desigualdad. Entonces, por la Proposición 6.3.2 tenemos que

$$\mathbb{E} \left(\left| \frac{\mathcal{S}^{(n)}}{A(n)} - \frac{\tilde{\mathcal{S}}^{(n)}}{\lambda^{(n)} A(n)} \right| \mathbb{I}_{D_\epsilon} \right) \leq \lambda^{(n)} e^\gamma \left(\frac{1}{\lambda^{(n)} - 2\epsilon} - \frac{1}{\lambda^{(n)} + \epsilon} \right) (1 + o(1)) = \mathcal{O}(\epsilon). \quad (6.19)$$

Los siguientes cuatro términos en (6.18) pueden ser acotados usando las desigualdades (6.11)-(6.12) y las cotas (6.14)-(6.17):

- $\mathbb{E} \left(\frac{\mathcal{S}^{(n)}}{A(n)} \mathbb{I}_{A_\epsilon} \right) \leq n \mathbb{P}(A_\epsilon) = \mathcal{O}(e^{-n^\epsilon(1-\delta)}) = o(1)$.
- $\mathbb{E} \left(\frac{\mathcal{S}^{(n)}}{A(n)} \mathbb{I}_{A_\epsilon^c \cap B_\epsilon} \right) \leq n^\epsilon \mathbb{P}(B_\epsilon | A_\epsilon^c) = \mathcal{O}(n^{3\epsilon-\alpha}) = o(1)$.
- $\mathbb{E} \left(\frac{\mathcal{S}^{(n)}}{A(n)} \mathbb{I}_{A_\epsilon^c \cap B_\epsilon^c \cap B_\epsilon'} \right) = n^\epsilon \mathbb{P}(B_\epsilon' | A_\epsilon^c \cap B_\epsilon^c) = \mathcal{O}(n^{2\epsilon-\alpha}) = o(1)$.
- $\mathbb{E} \left(\frac{\mathcal{S}^{(n)}}{A(n)} \mathbb{I}_{A_\epsilon^c \cap B_\epsilon^c \cap B_\epsilon'^c \cap C_\epsilon^c} \right) \leq n^\epsilon \mathbb{P}(C_\epsilon | A_\epsilon^c \cap B_\epsilon^c \cap B_\epsilon'^c) = \mathcal{O}(n^{\alpha+3\epsilon-1}) = o(1)$.

Finalmente, para el último término observemos que por las cotas (6.14)-(6.17) $\mathbb{P}(D_\epsilon^c) = o(1)$. Y por la Proposición 6.3.3,

$$\mathbb{E} \left(\left(\frac{\tilde{\mathcal{S}}^{(n)}}{A(n)} \right)^2 \right) = \mathcal{O}(1).$$

Por lo tanto, por la desigualdad de Cauchy-Schwarz obtenemos que $\mathbb{E} \left(\frac{\tilde{\mathcal{S}}^{(n)}}{(\lambda^{(n)} A(n))} \mathbb{I}_{D_\epsilon^c} \right) = o(1)$. De lo cual se concluye que

$$\mathbb{E} \left(\left| \frac{\mathcal{S}^{(n)}}{A(n)} - \frac{\tilde{\mathcal{S}}^{(n)}}{\lambda^{(n)} A(n)} \right| \right) = \mathcal{O}(\epsilon) + o(1).$$

Y como ϵ es arbitrariamente chico, llegamos a que $\mathbb{E} \left(\left| \frac{\mathcal{S}^{(n)}}{A(n)} - \frac{\tilde{\mathcal{S}}^{(n)}}{\lambda^{(n)} A(n)} \right| \right) \xrightarrow{n \rightarrow \infty} 0$. \square

6.4.3. Prueba de la Proposición 6.2.1

En el régimen crítico, tenemos que el número de agentes viene dado por $A(n) = n$. En este caso, el número de pasos será de orden $\mathcal{O}(n)$. Como subproducto, el número de vértices explorados durante la exploración se acumulará hasta un punto donde ya no serán despreciables. Luego tenemos que tener en cuenta cuántos vértices ya se han explorado para poder saber cuántos vértices nuevos se explorarán en cada paso. Para esto, derivamos un límite de escala para la proporción de vértices explorados y una estimación de grandes desvíos para las fluctuaciones alrededor de este.

Por comodidad, modelaremos el número de vértices inexplorados a tiempo $t \geq 0$ ($U_t^{(n)}$) como un proceso de saltos markoviano a tiempo continuo en el que con tasa 1 se ejecuta cada paso del algoritmo. Notar que el esqueleto de este proceso será el proceso inexplorado $(U_i^{(n)})_{i \geq 1}$ descrito en la Sección 6.1. Cuando se produzca un paso del algoritmo, se explorará la vecindad del vértice actual en el siguiente grafo durante lo cual se explorará un número cercano a $\text{Binom}(U_t^{(n)}/n, D^{(n)})$ de nuevos vértices. Para ver esto, tengamos en cuenta que el número de vértices recién explorados está en este caso acotado por debajo por $\text{Binom}(U_t^{(n)}/n - D^{(n)}/n, D^{(n)})$ y por arriba por $\text{Binom}(U_t^{(n)}/n, D^{(n)})$. Como en el Capítulo 3, consideraremos los procesos escalados $\bar{U}_t^{(n)} := U_{\lfloor nt \rfloor}^{(n)}/n$.

Definimos, para $u \in [0, 1]$

$$\bar{p}_i^{(n)}(u) := \sum_{j=i}^{\infty} p_j^{(n)} \binom{j}{i} u^i (1-u)^{j-i},$$

o sea, la distribución de $\text{Binom}(u, D^{(n)})$; y

$$\bar{\bar{p}}_i^{(n)}(u) := \sum_{j=i}^{\infty} p_j^{(n)} \binom{j}{i} (u - j/n)^i (1 - u + j/n)^{j-i}$$

o sea, la distribución de $\text{Binom}(u - D^{(n)}/n, D^{(n)})$. Donde $D^{(n)}$ es un variable distribuida como el grado de un vértice uniforme en los grafos $G_i^{(n)}$. Entonces, si dejamos que $f : [0, 1] \rightarrow \mathbb{R}$ sea una función continua y creciente, el generador $\bar{Q}^{(n)}(\cdot)$ del proceso escalado aplicado a ella dará

$$Q_1^{(n)}(f(\bar{U}_t^{(n)})) \leq \bar{Q}^{(n)}(f(\bar{U}_t^{(n)})) \leq Q_2^{(n)}(f(\bar{U}_t^{(n)})), \quad (6.20)$$

donde $Q_1^{(n)}(f(\bar{U}_t^{(n)})) := n \sum_{i=1}^{\infty} \bar{p}_i^{(n)}(\bar{U}_t^{(n)}) (f(\bar{U}_t^{(n)} - i/n) - f(\bar{U}_t^{(n)}))$ y $Q_2^{(n)}(f(\bar{U}_t^{(n)})) := n \sum_{i=1}^{\infty} \bar{\bar{p}}_i^{(n)}(\bar{U}_t^{(n)}) (f(\bar{U}_t^{(n)} - i/n) - f(\bar{U}_t^{(n)}))$.

Como

$$Q_1(\bar{U}_t^{(n)}) = -\lambda^{(n)} \bar{U}_t^{(n)} \text{ y}$$

$$Q_2(\bar{U}_t^{(n)}) = -\lambda^{(n)} \bar{U}_t^{(n)} + \frac{(\sigma^{(n)})^2 + (\lambda^{(n)})^2}{n},$$

tendremos que las derivas de $\bar{U}_t^{(n)}$ son $\delta^{(n)}(x) := -\lambda^{(n)}x + o(1)$ y por lo tanto tienden al límite $\delta(x) := -\lambda x$. Entonces, podemos probar el siguiente límite de escala.

Proposición 6.4.2. *Sea $\bar{U}_t^{(n)}$ como arriba. Entonces, para todo $T \geq 0$, $\bar{U}_t^{(n)} \xrightarrow{L^1} e^{-\lambda t}$ uniformemente en $[0, T]$.*

Demostración. La prueba se deduce de la Proposición 3.2.1. En efecto, la hipótesis (i) se cumple ya que $\bar{U}_0^{(n)} = 1$ para todo $n \geq 1$. Además, debido a que $\lambda^{(n)} \xrightarrow{n \rightarrow \infty} \lambda$ y todas las derivas son casi lineales, también se cumple la condición (ii). Sólo nos queda acotar la variación cuadrática de los procesos $U_t^{(n)}$. Esta vendrá dada por

$$\mathbb{E} \left(\int_0^{nt} \sum_{i=0}^{\infty} \bar{p}(U_s^{(n)}/n) \left(U_s^{(n)} - k - U_s^{(n)} \right)^2 ds \right) \leq n(\sigma^{(n)2} + \lambda^{(n)2})t,$$

que es $o(n^2)$. Entonces, la conclusión se sigue de la Proposición 3.2.1. \square

La siguiente proposición da una estimación de grandes desvíos de las fluctuaciones del proceso alrededor de su límite de escala.

Proposición 6.4.3. *Si $A(n) = n$, entonces para todo $\gamma \in (0, 1/13)$*

$$\mathbb{P} \left(\sup_{t \leq n^\gamma} |\bar{U}_t^{(n)} - e^{-\lambda t}| > n^{-\gamma/2} \right) \leq e^{-n^{\gamma/2}}.$$

Demostración. Si definimos $\epsilon_t^{(n)} := \bar{U}_t^{(n)} - 1 + \lambda \int_0^t \bar{U}_s^{(n)} ds$ tenemos que $\epsilon_t^{(n)} = \bar{U}_t^{(n)} - e^{-\lambda t} + \lambda \int_0^t \bar{U}_s^{(n)} - e^{-\lambda s} ds$. Por otro lado, para todo par de funciones $f_1, f_2 : \mathbb{R} \rightarrow \mathbb{R}$ tales que para algún $c \geq 0$

$$f_1(t) = f_2(t) + c \int_0^t f_2(s) ds,$$

vale que si $\sup_{t \leq T} |f_1(t)| \leq \epsilon/2$ entonces $\sup_{t \leq T} |f_2(t)| \leq \epsilon$. Con lo cual tenemos que

$$\left\{ \sup_{t \leq T} |\bar{U}_t^{(n)} - e^{-\lambda t}| \geq \epsilon \right\} \subseteq \left\{ \sup_{t \leq T} |\epsilon_t^{(n)}| \geq \epsilon/2 \right\} = \left\{ \sup_{t \leq T} \epsilon_t^{(n)} \geq \epsilon/2 \right\} \cup \left\{ \inf_{t \leq T} \epsilon_t^{(n)} \leq -\epsilon/2 \right\}.$$

Alcanza entonces con acotar las probabilidades de estos últimos dos eventos. Ahora, si definimos (para $\rho > 0$) el proceso

$$R_t(\rho) := \exp \left(\rho \bar{U}_t^{(n)} - \rho - \int_0^t e^{-\rho \bar{U}_s^{(n)}} \bar{Q}(e^{\rho \bar{U}_s^{(n)}}) ds \right),$$

por el hecho de que $\bar{U}_t^{(n)} \leq 1$ para todo $t \geq 0$, tenemos que $e^{\rho \bar{U}_t^{(n)} - \rho}$ es integrable y por el Lema 3.1.3 es una martingala. De acuerdo con la ecuación (6.20),

$$\begin{aligned} \bar{Q} \left(e^{\rho \bar{U}_t^{(n)}} \right) &\leq Q_2 \left(e^{\rho \bar{U}_t^{(n)}} \right) = n e^{\rho \bar{U}_t^{(n)}} \sum_{i=1}^{\infty} \bar{p}_i^{(n)}(\bar{U}_t^{(n)}) \left(e^{-\rho i/n} - 1 \right) \\ &\leq -e^{\rho \bar{U}_t^{(n)}} \left[\rho \sum_{i=1}^K i \bar{p}_i^{(n)}(\bar{U}_t^{(n)}) + \mathcal{O}(\rho^2 K^3/n) - \rho \sum_{j=K+1}^{\infty} j \bar{p}_j^{(n)} \right] \\ &= -\lambda^{(n)} e^{\rho \bar{U}_t^{(n)}} \left[\rho \bar{U}_t^{(n)} + \mathcal{O}(\rho/n) + \mathcal{O}(\rho^2 K^3/n) + \mathcal{O}(\rho/K) \right], \end{aligned}$$

donde en la segunda desigualdad se utilizó la expansión de Taylor de $e^{-x} - 1$, que $\sum_{i \geq 1} i \bar{p}_i^{(n)}(\bar{U}_t^{(n)}) = \lambda^{(n)} \bar{U}_t^{(n)} + \mathcal{O}(1/n)$ y que (para todo $x \geq 0$) $e^{-x} - 1 \leq x$. Y para la

última igualdad se usó que $(jp_j^{(n)})_{j \geq 1}$ es uniformemente sumable (porque satisface la (CC)). Entonces tenemos que

$$R_t(\rho) = \exp \left(\rho \epsilon_t^{(n)} + \int_0^t \left(\mathcal{O}(\rho/n) + \mathcal{O}(\rho^2 K^3/n) + \mathcal{O}(\rho/K) \right) ds \right).$$

Fijando $\rho(n) := n^\gamma$, $K(n) := n^{3\gamma}$ y $T(n) := n^\gamma$, llegamos a que para todo $t \leq T(n)$, $R_t(\rho) = \exp \left(\rho \epsilon_t^{(n)} + o(1) \right)$. Entonces,

$$\mathbb{P} \left(\sup_{t \leq n^\gamma} \epsilon_t^{(n)} > \epsilon/2 \right) = \mathbb{P} \left(\sup_{t \leq n^\gamma} R_t(\rho) > e^{\rho\epsilon/2+o(1)} \right) \leq e^{-n^\gamma\epsilon/2+o(1)},$$

donde se usó la desigualdad de Doob. De forma análoga, podemos probar que

$$\mathbb{P} \left(\inf_{t \leq n^\gamma} \epsilon_t^{(n)} < -\epsilon/2 \right) \leq e^{-n^\gamma\epsilon/2+o(1)}.$$

La conclusión se sigue fijando $\epsilon(n) = n^{-\gamma/2}$. □

Usando esta cota, podemos probar el resultado principal para este régimen.

Prueba de la Proposición 6.2.1. Por la Proposición 6.4.1, para todo $\epsilon > 0$ el evento $\{\mathcal{V}^{(n)} > n^\epsilon\}$ tiene una probabilidad exponencialmente chica de suceder. Por lo tanto, tenemos que $\mathcal{V}^{(n)} \leq n^\epsilon$ eventualmente. Entonces, como $\mathcal{S}^{(n)} \leq A(n)\mathcal{V}^{(n)}$ c.s., tenemos que $\mathcal{S}^{(n)}/n \leq n^\epsilon$ eventualmente.

Por la Proposición 6.4.3, $\bar{U}_t^{(n)} \xrightarrow{n \rightarrow \infty} e^{-\lambda t}$ casi seguro y uniformemente para $t(n) \leq n^{2\epsilon}$, para ϵ lo suficientemente chico. Definamos $t^*(n)$ como el tiempo del $\mathcal{S}^{(n)}$ -ésimo salto de $U_t^{(n)}$. Porque $\mathcal{S}^{(n)}$ es eventualmente más chico que $n^{1+\epsilon}$, por grandes desvíos de procesos de Poisson tenemos que eventualmente $t^*(n) \leq n^{1+2\epsilon}$. Entonces, $\bar{U}_{\mathcal{S}^{(n)}/n}^{(n)} = e^{-\lambda(\mathcal{S}^{(n)}/n)} + o(1)$. Y como $\mathcal{E}^{(n)}$ es el número total de vértices explorados, $\mathcal{E}^{(n)}/n = 1 - \bar{U}_{\mathcal{S}^{(n)}/n}^{(n)}$. Entonces, tenemos que $\mathcal{E}^{(n)}/n = 1 - e^{-\lambda(\mathcal{S}^{(n)}/n)} + o(1)$. La conclusión se obtiene invirtiendo esta expresión. □

6.4.4. Prueba del Teorema 6.2.2

El caso supercrítico es el más sencillo de analizar de los tres, ya que en él el grafo global $G_g^{(n)}$ es completo (excepto en un evento de probabilidad exponencialmente chica).

Lema 6.4.2. *Sea C el evento bajo el cual $G_g^{(n)}$ es completo. Entonces, si $A(n) = \lceil n^\alpha \rceil$ con $\alpha > 1$, (para algún $\gamma > 0$) $\mathbb{P}(C^c) = \mathcal{O}(e^{-\gamma n})$.*

Demostración. Para ver esto, damos una cota exponencial al evento en el que cierto par de vértices $v, w \in [n]$ no está conectado

$$\mathbb{P}(v \not\leftrightarrow w) \leq \left(1 - \frac{1}{n}\right)^{A(n)} = \mathcal{O}\left(e^{-n^{\alpha-1}}\right).$$

Entonces, la probabilidad de que haya algún par de vértices no conectado será a lo sumo $\mathcal{O}\left(n^2 e^{-n^{\alpha-1}}\right)$. □

Prueba del Teorema 6.2.2. Notar que la variable $\mathcal{S}^{(n)} - A(n)$ da el número de pasos antes de encontrar un máximo. Entonces, condicionalmente a C tenemos que

$$\mathbb{P}(\mathcal{S}^{(n)} - A(n) > k | D_1^{(n)}, \dots, D_k^{(n)}, C) = \left(1 - \frac{D_1^{(n)}}{n}\right) \cdots \left(1 - \frac{D_k^{(n)}}{n}\right),$$

como en cada paso, cada vértice en la vecindad del vértice visitado tiene una probabilidad uniforme de ser el máximo global. Luego, en el caso en que $G_g^{(n)}$ sea completo, tenemos ese $\mathcal{S}^{(n)} - A(n) \sim \text{Geom}(\lambda^{(n)}/n)$. La conclusión se deduce del hecho de que este último evento tiene una probabilidad exponencialmente pequeña de no suceder. \square

Apéndice A

Procesos de ramificación

Los *procesos de ramificación* son modelos de poblaciones simples propuestos por Francis Galton¹ en 1873 para estudiar la extinción de los apellidos a través de la desaparición de líneas masculinas. Este origen histórico impulsa muchos de los términos utilizados. El modelo describe los tamaños de diferentes generaciones de una familia de individuos, cada uno de los cuales tiene una progenie aleatoria independiente de los demás y dada por una distribución fija $(p_i)_{i \geq 0}$, llamada *distribución de descendencia*. Cada uno de los *hijos* de un individuo de la generación $k \geq 0$ es parte de la generación $k + 1$. En la generación 0, sólo existe un individuo.

Más específicamente, se definirá el proceso $(Z_i)_{i \geq 1}$ que describe los tamaños de las generaciones de un proceso de ramificación, para $k \geq 1$, de acuerdo con la siguiente recursión:

$$Z_{k+1} = \sum_{j=1}^{Z_k} X_j^{(k)}, \quad (\text{A.1})$$

donde $X_j^{(k)}$ son variables aleatorias i.i.d. de distribución $(p_i)_{i \geq 0}$ y Z_0 se toma igual a 1. Los procesos de ramificación se pueden representar como árboles con raíz², donde cada arista representa una relación *padre-hijo* entre dos individuos. Por lo general, se los denominan *árboles de Galton-Watson*.

A.1. Supervivencia y extinción

Si el proceso posee una generación $m \geq 1$ tal que $Z_m = 0$, entonces permanece en ese estado para siempre y se dice que se ha *extinguido*. En otras palabras, 0 es un estado absorbente, y el evento de extinción es el evento de visitarlo. Si en cada generación hay al menos un individuo, se dice que el proceso ha *sobrevivido*. Aquí denotaremos por E el evento de extinción dado por $\{\exists n \geq 1 : Z_n = 0\}$.

Como, para todo $m \geq 1$, la progenie de un individuo particular en la m -ésima generación es independiente de Z_m , por la definición del proceso es fácil ver que

$$\mathbb{E}(Z_m) = \mu^m,$$

¹El mismo Galton conocido infamemente por ser creador y promotor de la eugenesia.

²Árboles que tienen un vértice especial llamado *raíz*.

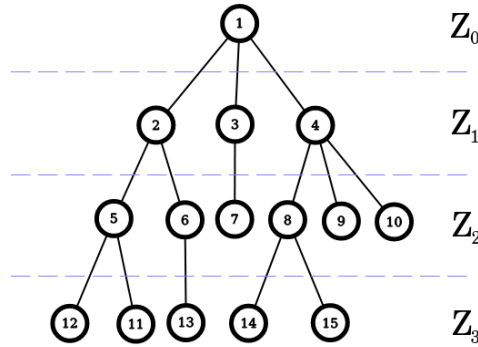


Figura A.1: Ejemplo de un árbol de Galton-Watson. En este árbol, el individuo 1 es la raíz y los otros se enumeran según el orden de una exploración en anchura. El proceso del tamaño de las generaciones viene dado por $Z_1 = 3$, $Z_2 = 6$, $Z_3 = 5$ y $Z_4 = 0$. El número de hijos del individuo 8 es dos y la progenie total del proceso es $T = 15$. La extinción tiene lugar después de tres generaciones.

donde $\mu > 0$ es la media de la distribución de descendencia. Esto implica, por la desigualdad de Markov, que la probabilidad del evento $\{Z_m \geq 1\}$ es tal que

$$\mathbb{P}(Z_m \geq 1) \leq \mu^m. \tag{A.2}$$

Esto, debido a que el evento de supervivencia viene dado por $E^c = \bigcup_{i \geq 1} \{Z_i \geq 1\}$, implica por Borel-Cantelli que si la distribución de la descendencia tiene un valor medio $\mu < 1$, entonces el proceso casi seguramente se extinguirá en algún momento. Utilizando la naturaleza recursiva del proceso de ramificación y analizando la función generadora de probabilidad de la distribución de descendencia, se puede probar la siguiente extensión [VDH16] de este resultado:

Teorema A.1.1. *Dado un proceso de ramificación $(Z_n)_{n \geq 1}$ con una distribución de descendencia de media μ y tal que $p_0 > 0$, $\mathbb{P}(E) = 1$ si $\mu \leq 1$.*

En otras palabras, en el caso no trivial, la extinción está casi seguro garantizada si y sólo si el número promedio de hijos que cada individuo tiene es 1 o menos. Por eso, se dice que un proceso de ramificación es:

- *Subcrítico* si $\mu < 1$, en cuyo caso por (A.2) la probabilidad de que la extinción no haya aún sucedido decrece exponencialmente con el número de generación.
- *Crítico* si $\mu = 1$.
- Y *supercrítico* si $\mu > 1$; y por lo tanto, hay una probabilidad positiva de supervivencia.

A.2. Exploración de árboles de Galton-Watson

En esta sección daremos una descripción de la *exploración en anchura* de un árbol Galton-Watson. Por exploración en anchura, nos referimos a la exploración secuencial de los individuos en cada generación. El proceso de exploración se llevará a cabo de la siguiente manera: en cada paso, un individuo de la generación actual se declara *explorado*, se revela su número de hijos y, finalmente, cada uno de sus hijos se declara *inexplorado*. Esto se hace hasta que se

haya explorado toda la generación, en cuyo momento comienza la exploración de la próxima generación. Estudiaremos el proceso $(U_i)_{i \geq 1}$ asociado a la cantidad de vértices inexplorados en cada paso, donde U_0 se toma como 0.

Es fácil ver que si la variable aleatoria X_i (distribuida de acuerdo con $(p_k)_{k \geq 0}$, como antes) da el número de hijos del i -ésimo individuo explorado, entonces en el paso i de la exploración, la variación del número de individuos inexplorados viene dada por $X_i - 1$. Esto es así porque en cada paso un individuo inexplorado deja de ser inexplorado (el explorado) y todos sus hijos se agregan al conjunto de individuos inexplorados. Esto significa que el único momento en que U_i disminuye es cuando se explora un individuo sin descendencia; es decir, cuando se visita una hoja del árbol de Galton-Watson. Luego, el proceso de exploración U_i de un árbol de Galton-Watson puede describirse como un paseo aleatorio simple de tamaño de paso dado por $X_i - 1$ y con 0 como estado absorbente. Esto permite la descripción del proceso de exploración como una suma de variables aleatorias i.i.d.:

$$U_i = \sum_{j=1}^i (X_j - 1).$$

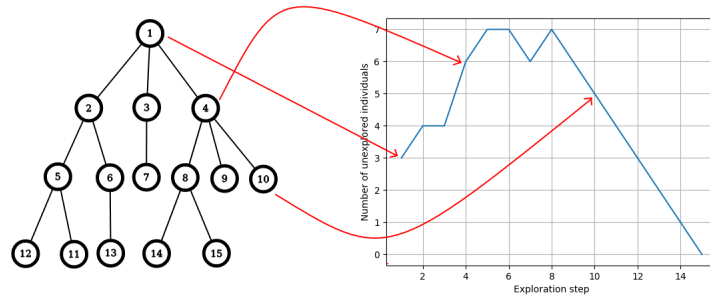


Figura A.2: Ejemplo de exploración de un árbol de Galton-Watson. Como se puede ver en la figura, el proceso comienza con altura tres, ya que el número de individuos inexplorados descubiertos durante la exploración del individuo 1 es tres. Del mismo modo, la altura en el paso 4 aumenta en dos ya que en ese momento un individuo pasa a estar explorado (el individuo 4) y se descubren tres nuevos individuos inexplorados (los hijos de 4).

Bajo esta descripción, la *progenie total* $T := \sum_{n=1}^{\infty} Z_n$ de un proceso de ramificación se puede describir en términos de este paseo aleatorio como un tiempo de parada: $T \sim \inf\{i \geq 0 : U_i = 0\}$. En caso de supervivencia, este tiempo de parada tendrá un valor infinito.

El siguiente teorema presente en [NSS04], es útil para mostrar la finitud de momentos exponenciales de la progenie total.

Teorema A.2.1. *Dado un proceso de ramificación con progenie distribuida como copias de la variable X . Entonces, existe $\gamma_1 > 0$ tal que $\mathbb{E}(e^{\gamma_1 T}) < \infty$ sii existe algún otro $\gamma_2 > 0$ tal que $\mathbb{E}(e^{\gamma_2 X}) < \infty$.*

Apéndice B

Medidas puntuales de Poisson

B.1. Definición y resultados básicos

Sea (E, \mathcal{M}) un espacio medible (generalmente $\mathbb{R}_{\geq 0}$ con los conjuntos de Borel \mathcal{B}), las medidas puntuales de Poisson serán medidas aleatorias σ -finitas y atómicas. Es decir, medidas aleatorias que dan masa unitaria a un cierto subconjunto de puntos numerable en E .

Definición B.1.1. Sea $\lambda(\cdot)$ una medida en (E, \mathcal{M}) . Una *medida puntual de Poisson* $\mu(\cdot)$ con intensidad $\lambda(\cdot)$ es una medida atómica σ -finita en (E, \mathcal{M}) tal que para toda sucesión de conjuntos disjuntos de medida λ finita $(A_i)_{i \geq 1} \subseteq E$, las variables aleatorias $\mu(A_i)$ son variables de Poisson independientes de media $\lambda(A_i)$.

Una prueba de la existencia de tales medidas se puede encontrar en [Mie06], esta se basa en una construcción explícita para las medidas. Si una medida puntual de Poisson se define en $(\mathbb{R}, \mathcal{B})$, se dice que la medida es *homogénea* si $\lambda(\cdot)$ es proporcional a la medida de Lebesgue e *inhomogénea* si no lo es. En el caso homogéneo, la intensidad puede describirse alternatively mediante un número real $\lambda > 0$, que proporciona el factor de proporcionalidad entre la medida de intensidad y la medida de Lebesgue.

Definición B.1.2. Dada una medida puntual de Poisson $\mu(\cdot)$ en (E, \mathcal{M}) , el conjunto asociado de puntos de medida positiva define lo que se llama *proceso puntual de Poisson*.

Cuando no nos referimos a una medida puntual de Poisson específica, diremos que un conjunto de puntos \mathcal{P} es un proceso puntual de Poisson de intensidad $\lambda(\cdot)$ si se distribuye como el proceso puntual de Poisson de una medida puntual de Poisson con esta intensidad.

Definición B.1.3. Dada una medida puntual de Poisson $\mu(\cdot)$ en $(\mathbb{R}_{\geq 0}, \mathcal{B})$, el proceso estocástico a tiempo continuo asociado $(N_t)_{t \geq 0}$ definido (para $t \geq 0$) como $N_t = \mu([0, t])$ se llama *proceso de Poisson*.

De manera análoga, cuando no nos referimos a una medida puntual de Poisson específica, diremos que el proceso $(N_t)_{t \geq 0}$ es un proceso de Poisson de intensidad $\lambda(\cdot)$ si se distribuye como el Poisson proceso de una medida puntual de Poisson con esta intensidad.

Dado una medida puntual de Poisson $\mu(\cdot)$ en (E, \mathcal{M}) y una función medible $f : E \rightarrow \mathbb{R}$, definiremos la integral de f con respecto a esta medida aleatoria de acuerdo a $\mu(f) := \int_E f(x) \mu(dx)$.

Proposición B.1.1. *Sea $\mu(\cdot)$ una medida puntual de Poisson de intensidad $\lambda(\cdot)$ en (E, \mathcal{M}) y $f : E \rightarrow \mathbb{R}$ una función medible. Entonces, si $f \in L^1(\lambda)$, tendremos que $f \in L^1(\mu)$ casi seguro, $\mu(f)$ define una variable aleatoria y*

$$\mathbb{E}(\exp(i\mu(f))) = \exp\left(\int_E (e^{if(x)} - 1)\lambda(dx)\right)$$

La prueba de esta proposición se puede encontrar en [Mie06]. Aplicando la última igualdad de la proposición a af , diferenciando con respecto a a un número adecuado de veces y tomando el límite $a \searrow 0$, obtenemos bajo las hipótesis de la proposición que

$$\mathbb{E}(\mu(f)) = \int_E f(x)\lambda(dx) \quad \text{y que} \tag{B.1}$$

$$\text{Var}(\mu(f)) = \int_E f^2(x)\lambda(dx). \tag{B.2}$$

Las relaciones (B.1) y (B.2) son normalmente conocidas como *fórmulas de Campbell*.

B.2. Adelgazamientos de medidas puntuales de Poisson

Definición B.2.1. Dado un proceso puntual de Poisson \mathcal{P} en $(\mathbb{R}_{\geq 0}, \mathcal{B})$ y una función integrable $f : \mathbb{R}_{\geq 0} \rightarrow [0, 1]$, un *f-adelgazamiento* de \mathcal{P} es un nuevo proceso puntual definido a partir de quedarse con cada punto $x \in \mathcal{P}$ con probabilidad $f(x)$ e independientemente de los otros.

Es fácil ver que, dada una función f como en la definición, un adelgazamiento de un proceso puntual de Poisson $\mu(\cdot)$ de intensidad $\lambda(\cdot)$ es un nuevo proceso puntual de Poisson con medida de intensidad $\int_{(\cdot)} f\lambda(dx)$. Es por eso que podemos, de manera análoga, hablar de adelgazamiento de las medidas puntuales de Poisson.

Proposición B.2.1. *Dada una medida puntual de Poisson en $(\mathbb{R}_{\geq 0}, \mathcal{B})$ con intensidad $\lambda(\cdot)$ y una familia numerable de funciones integrables $f_1, \dots, f_k, \dots : \mathbb{R}_{\geq 0} \rightarrow [0, 1]$ tal que por cada $x \in \mathbb{R}_{\geq 0}$ $\sum_{i \geq 1} f_i(x) = 1$, definamos una nueva familia de medidas puntuales de Poisson en $(\mathbb{R}_{\geq 0}, \mathcal{B})$ de intensidades $\lambda_i(\cdot) := \int_{(\cdot)} f_i\lambda(dx)$ como $\mu_i(\cdot)$. Estas medidas se pueden acoplar de manera que sean mutuamente independientes y $\mu(\cdot) = \sum_{i \geq 1} \mu_i(\cdot)$.*

Demostración. Para probar esto, construyamos un proceso puntual $\bar{\mathcal{P}}$ en la banda $\mathbb{R}_{\geq 0} \times [0, 1]$ colocando cada punto x_i en el proceso de puntos \mathcal{P} de $\mu(\cdot)$ en $x_i \times U_i$, donde U_i son variables uniformes independientes en $[0, 1]$. Es fácil comprobar que $\bar{\mathcal{P}}$ es un proceso puntual de Poisson de intensidad $\lambda(\cdot) \times \mathcal{L}(\cdot)$, donde $\mathcal{L}(\cdot)$ es la medida de Lebesgue en $[0, 1]$. Ahora, si definimos, para $i \geq 1$, los conjuntos $A_i := \{(x, y) \in \mathbb{R}_{\geq 0} \times [0, 1] : \sum_{j=1}^{i-1} f_j(x) \leq y < \sum_{j=1}^i f_j(x)\}$; estos constituyen una partición de la banda. Además, si definimos (para $i \geq 1$) el proceso puntual $\mathcal{P}_i := \{x \in \mathbb{R}_{\geq 0} : (x, y) \in \bar{\mathcal{P}} \cap A_i\}$, es fácil ver que se distribuyen como procesos puntuales de Poisson de intensidad $\int_{(\cdot)} f_i(x)\lambda(dx)$. Finalmente, estos procesos son independientes ya que se corresponden a puntos de un proceso puntual de Poisson que caen en conjuntos disjuntos. \square

En la Figura B.1 presentamos una visualización de la construcción utilizada en la prueba de la Proposición B.2.1.

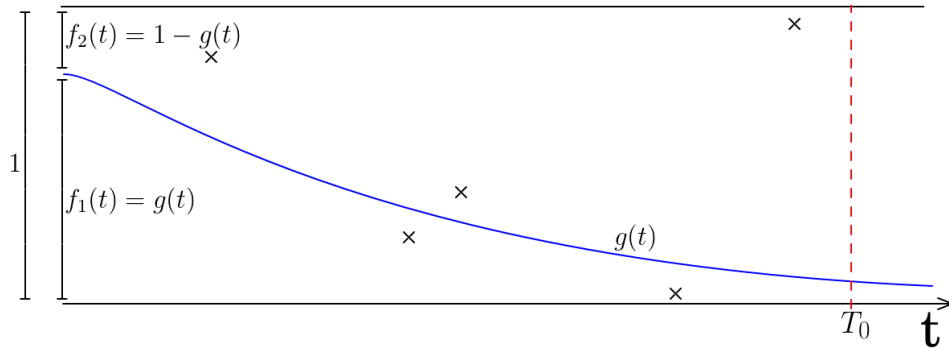


Figura B.1: Representación de la construcción que da lugar a los adelgazamientos independientes de un proceso puntual de Poisson en la línea. Aquí, una medida puntual de Poisson homogénea de intensidad $\lambda = 2$ se adelgaza con las funciones $f_1(t) = g(t)$ y $f_2(t) = 1 - g(t)$; resultando en dos procesos puntuales de Poisson inhomogéneos independientes de intensidad $\lambda_1(\cdot) = 2 \int_{(\cdot)} g(t) dt$ y $\lambda_2(\cdot) = 2 \int_{(\cdot)} 1 - g(t) dt$. En este caso, $\mu([0, T_0]) = 5$, $\mu_1([0, T_0]) = 2$ y $\mu_2([0, T_0]) = 3$.

Apéndice C

Topología de Skorohod y convergencia débil

C.1. Funciones càdlàg y topología de Skorohod

Los procesos estocásticos pueden interpretarse como elementos aleatorios de un espacio de funciones. Aquí discutiremos cuál es el conjunto natural para describir los procesos de salto markovianos y cómo dotarlo de una topología que le otorgue propiedades razonables y que sea útil para probar resultados de convergencia. Una buena manera de ver este problema es a través de las diferencias y similitudes con el caso canónico del conjunto de funciones continuas $\mathcal{C}[0, T]$ con la métrica inducida por la norma uniforme.

Primero analicemos el caso en el que los procesos se definen en un conjunto compacto $[0, T]$ en lugar de en la semilínea $\mathbb{R}_{\geq 0}$. El conjunto de funciones a considerar es el conjunto conocido como càdlàg (acrónimo en francés para “límite izquierdo continua a derecha”, y denotado por $D[0, T]$), definido por las funciones $f : [0, T] \rightarrow [0, 1]$ tales que:

(I) $\forall t \in [0, T]$, el límite $\lim_{s \rightarrow t^+} f(s)$ existe y es igual a $f(t)$,

(II) y $\forall t \in (0, T]$, el límite $\lim_{s \rightarrow t^-} f(s)$ existe.

Estas son funciones que son continuas excepto, posiblemente, por un cierto número de saltos. Tengamos en cuenta que si se le diera a este conjunto la métrica inducida por la norma uniforme, el resultado no sería del todo bueno: la familia no numerable de funciones $(\mathbb{I}_{[0, \alpha]})_{\alpha \in (0, 1)}$ estaría compuesta por elementos a distancia 1 entre sí, lo que haría que el espacio no sea separable. Además, para $\epsilon > 0$ arbitrariamente pequeño, las funciones $\mathbb{I}_{[0, T/2]}$ y $\mathbb{I}_{[0, T/2 - \epsilon]}$ estarían a distancia 1; lo que significaría que dos realizaciones de procesos que pueden parecer intuitivamente muy similares, sin embargo, podrían estar lejos. Ambos problemas se pueden resolver dotando al conjunto de otra topología que tenga en cuenta estas cuestiones.

Definición C.1.1. Sea Λ el conjunto de asignaciones continuas y estrictamente crecientes de $[0, T]$ en $[0, T]$ y $d(f, g) := \inf_{\lambda \in \Lambda} \|\lambda - id\|_{\infty} \vee \|f - g \circ \lambda\|_{\infty}$ sea una métrica en $D[0, T]$. La topología inducida por esta métrica se conoce como *topología de Skorohod*.

Es decir, dos funciones estarán cerca en esta métrica si existe un cambio de tiempo cercano a la identidad que hace que las funciones se encuentren a una distancia pequeña en la norma

uniforme. Como se muestra en [Bil13], $D[0, T]$ con esta topología es separable. Es fácil ver que si el límite es continuo, la convergencia en la topología de Skorohod es equivalente a la convergencia uniforme.

Si definimos la norma $\|\lambda\|^\circ := \sup_{s \neq t \in [0, T]} \left\| \log \left(\frac{\lambda(t) - \lambda(s)}{t - s} \right) \right\|$ sobre Λ , la métrica resultante $d^\circ(f, g) := \inf_{\lambda \in \Lambda} \|\lambda - id\|^\circ \vee \|f - g \circ \lambda\|_\infty$ en $D[0, T]$ es equivalente a $d(\cdot, \cdot)$. Esto significa que esta métrica también genera a la topología de Skorohod, por lo cual hace al espacio separable; pero esta también lo hace completo [Bil13].

Usando esta construcción para las funciones càdlàg en conjuntos compactos, podemos extender estos resultados a $D[0, \infty)$. Para esto, definamos la métrica $d^\circ(f, g) := \sum_{m=1}^\infty 2^{-m} 1 \wedge \|f - g\|_m^\circ$ en $D[0, \infty)$, donde $\|\cdot\|_m^\circ$ es la norma en $D[0, m]$ definida anteriormente. El espacio topológico resultante hereda las propiedades de separabilidad y completitud de la construcción anterior. Además, las convergencias en $D[0, \infty)$ y $D[0, T]$ están conectadas por el siguiente teorema [Bil13]:

Teorema C.1.1. *Sea $(f_n)_{n \geq 1}$, $f \in D[0, \infty)$ y (para $T > 0$) sea $R_T : D[0, \infty) \rightarrow D[0, T]$ el mapa que actúa según $R_T(g) = g|_{[0, T]}$. Entonces, $d^\circ(f_n, f) \rightarrow 0$ en $D[0, \infty)$, si y sólo si $d^\circ(R_T f_n, R_T f) \rightarrow 0$ en $D[0, T]$ para todo punto de continuidad T de f .*

C.2. Convergencia débil de procesos estocásticos

En los cursos introductorios de probabilidad, generalmente se estudian tres tipos de convergencia para variables aleatorias: casi segura, en probabilidad y en distribución. Las dos primeras son convergencias para variables aleatorias definidas en el mismo espacio de probabilidad, mientras que la convergencia en distribución se define más ampliamente para variables aleatorias que pueden pertenecer a diferentes espacios. Aquí presentaremos una generalización de la convergencia en distribución para procesos estocásticos, junto con un teorema de acoplamiento para sucesiones de procesos que convergen de esta manera. Para esto, primero definiremos la convergencia débil para medidas de probabilidad.

Definición C.2.1. Dado un espacio topológico (X, τ) y una sucesión de medidas de probabilidad $(\mathbb{P}_n(\cdot))_{n \geq 1}$ y otra medida de probabilidad $\mathbb{P}(\cdot)$ (todas sobre (X, \mathcal{B})), diremos que $\mathbb{P}_n(\cdot)$ converge débilmente a $\mathbb{P}(\cdot)$ si para toda función $f \in \mathcal{C}_b(X)$ $\mathbb{E}_n(f) \xrightarrow{n \rightarrow \infty} \mathbb{E}(f)$. A esto se lo va a denotar como $\mathbb{P}_n \Rightarrow \mathbb{P}$.

Donde $\mathcal{C}_b(X)$ es el conjunto de funciones reales, continuas y acotadas en X . Es decir, una sucesión de medidas de probabilidad converge débilmente a otra probabilidad si, para cada función de prueba, sus respectivas medias convergen a la media límite. Tengamos en cuenta que la noción de convergencia débil depende directamente de la topología de X , ya que esta determina qué funciones son continuas y cuáles no (definiendo así el conjunto de funciones de prueba). Aunque la convergencia débil se definió para espacios topológicos, en la mayoría de las situaciones prácticas los espacios en cuestión serán de hecho métricos. Una serie de equivalencias útiles para esta convergencia se dan en el Teorema de Portmanteau [Mie06].

Entonces, la convergencia débil de una sucesión de variables aleatorias a una variable límite puede definirse en términos de la convergencia débil de las medidas imagen de ellas. Esta definición es equivalente a decir que una sucesión de variables aleatorias $(X_n)_{n \geq 1}$ converge a X si, por cada $f \in \mathcal{C}_b(\mathbb{R})$, tenemos que $\mathbb{E}(f(X_n)) \xrightarrow{n \rightarrow \infty} \mathbb{E}(f(X))$. Definiremos la convergencia débil de procesos estocásticos en analogía a esto.

Definición C.2.2. Dada una sucesión de procesos estocásticos $(X_t^{(n)})_{n \geq 1} \subseteq D[0, \infty)$ y otro proceso $(X_t)_{t \geq 0} \in D[0, \infty)$ diremos que $X_t^{(n)}$ converge débilmente a X_t ($X_t^{(n)} \Rightarrow X_t$) si, para toda $F : D[0, \infty) \rightarrow \mathbb{R}$ continua en el sentido de Skorohod y acotada, $\mathbb{E}(F(X_t^{(n)})) \xrightarrow{n \rightarrow \infty} \mathbb{E}(F(X_t))$.

La secuencia habitual de pasos para demostrar una convergencia débil para procesos estocásticos a tiempo continuo es primero probar que es apretada y luego mostrar la convergencia de las medias de las proyecciones de los procesos en ciertos subconjuntos de tiempos finitos. La sutileza en esto es que estas proyecciones no siempre son continuas en la topología de Skorohod, problema que puede ser eludido por una serie de teoremas (consulte la Sección 13 de [Bil13]).

Como comentamos anteriormente, la convergencia débil se aplica a procesos estocásticos no necesariamente definidos en el mismo espacio de probabilidad. Sin embargo, como muestra el siguiente teorema, si uno tiene convergencia débil de una sucesión de procesos estocásticos a tiempo continuo, existe un espacio de probabilidad común donde todos los procesos y el límite pueden definirse juntos y para el cual la convergencia es casi segura¹. En muchas situaciones, esto simplifica enormemente las pruebas, ya que muchos de los teoremas de convergencia habituales del análisis real se pueden usar una vez que la convergencia se establece como casi segura.

Teorema C.2.1 (Representación de Skorohod [Bil13]). *Dada una sucesión de procesos estocásticos $(X_t^{(n)})_{n \geq 1} \subseteq D[0, T]$ y $X_t \in D[0, T]$, con $X_t^{(n)} \Rightarrow X_t$. Entonces, existe un espacio de probabilidad común $(D[0, T], \mathcal{F}, \mathbb{P})$ y procesos estocásticos $\tilde{X}_t, (\tilde{X}_t^{(n)})_{n \geq 1}$ en este espacio tales que $\tilde{X}_t \sim X_t$, (para todo $n \geq 1$) $\tilde{X}_t^{(n)} \sim X_t^{(n)}$ y $\tilde{X}_t^{(n)} \xrightarrow{a.s.} \tilde{X}_t$ para todo $t \in [0, T]$.*

Como una aplicación simple de este teorema, se puede dar una breve prueba del teorema del mapeo continuo.

¹De hecho, incluso puede tomarse como *convergencia segura*. Pero, en cualquier caso, para todos los asuntos prácticos, ambas convergencias son iguales.

Bibliografía

- [ACG⁺12] G. Ausiello, P. Crescenzi, G. Gambosi, V. Kann, A. Marchetti-Spaccamela, and M. Protasi. *Complexity and approximation: Combinatorial optimization problems and their approximability properties*. Springer Science & Business Media, 2012.
- [AFP98] J. Aronson, A.M. Frieze, and B.G. Pittel. Maximum matchings in sparse random graphs: Karp–sipser revisited. *Random Structures & Algorithms*, 12(2):111–177, 1998.
- [AHK07] E. Altman, Y. Hayel, and H. Kameda. Evolutionary dynamics and potential games in non-cooperative routing. In *WiOpt*, pages 1–5, 2007.
- [ARV08] H. Ackermann, H. Röglin, and B. Vöcking. On the impact of combinatorial structure on congestion games. *Journal of the ACM (JACM)*, 55(6):25, 2008.
- [Bar36] D.F. Barrow. Infinite exponentials. *The American Mathematical Monthly*, 43(3):150–160, 1936.
- [BE76] B. Bollobás and P. Erdős. Cliques in random graphs. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 80 - 3, pages 419–427. Cambridge University Press, 1976.
- [BGT10] M. Bayati, D. Gamarnik, and P. Tetali. Combinatorial approach to the interpolation method and scaling limits in sparse random graphs. In *Proceedings of the forty-second ACM symposium on Theory of computing*, pages 105–114. ACM, 2010.
- [Bia00] G. Bianchi. Performance analysis of the ieee 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 18(3):535–547, March 2000.
- [Bil13] P. Billingsley. *Convergence of probability measures*. John Wiley & Sons, 2013.
- [BJL16] G. Brightwell, S. Janson, and M. Luczak. The greedy independent set in a random graph with given degrees. *Random Structures & Algorithms*, 2016.
- [BJLM14] P. Bermolen, M. Jonckheere, F. Larroca, and P. Moyal. Estimating the spatial reuse with configuration models. *arXiv preprint arXiv:1411.0143*, 2014.

- [BJLM16] P. Bermolen, M. Jonckheere, F. Larroca, and P. Moyal. Estimating the transmission probability in wireless networks with configuration models. *ACM Transactions on Modeling and Performance Evaluation of Computing Systems*, 1(2):9, 2016.
- [BJLS19] P. Bermolen, M. Jonckheere, F. Larroca, and M. Sáenz. Degree-greedy algorithms on large random graphs. *ACM SIGMETRICS Performance Evaluation Review*, 46(3):27–32, 2019.
- [BJM17] P. Bermolen, M. Jonckheere, and P. Moyal. The jamming constant of uniform random graphs. *Stochastic Processes and their Applications*, 127(7):2138–2178, 2017.
- [BJS17] P. Bermolen, M. Jonckheere, and J. Sanders. Scaling limits and generic bounds for exploration processes. *Journal of Statistical Physics*, 169(5):989–1018, 2017.
- [BM15] R. Bardenet and O.A. Maillard. Concentration inequalities for sampling without replacement. *Bernoulli*, 21(3):1361–1385, 2015.
- [Bol81] B. Bollobás. The independence ratio of regular graphs. *Proceedings of the American Mathematical Society*, pages 433–436, 1981.
- [Bol01] B. Bollobás. *Random graphs*. Number 73. Cambridge university press, 2001.
- [Bor16] C. Bordenave. Lecture notes on random graphs and probabilistic combinatorial optimization. <https://bit.ly/2ZACYTF>. *Lecture Notes*, 2016.
- [Boy18] K. Boyadzhiev. *Notes on the binomial transform*. World Scientific, 2018.
- [Bur91] A. Burns. Scheduling hard real-time systems: a review. *Software Engineering Journal*, 6(3):116–128, 1991.
- [CAP07] A. Cadilhe, N.A.M. Araújo, and V. Privman. Random sequential adsorption: from continuum to lattice and pre-patterned substrates. *Journal of Physics: Condensed Matter*, 19(6):065124, 2007.
- [Chi06] C. Chicone. *Ordinary differential equations with applications*, volume 34. Springer Science & Business Media, 2006.
- [CMOP10] U.O. Candogan, I. Menache, A. Ozdaglar, and P.A. Parrilo. Near-optimal power control in wireless networks: A potential game approach. In *2010 Proceedings IEEE INFOCOM*, pages 1–9. IEEE, 2010.
- [Dar02] R.W.R. Darling. Fluid limits of pure jump markov processes: a practical guide. *arXiv preprint math/0210109*, 2002.
- [DG16] S. Durand and B. Gaujal. Complexity and optimality of the best response algorithm in random potential games. In *International Symposium on Algorithmic Game Theory*, pages 40–51. Springer, 2016.
- [DN08] R.W.R. Darling and J.R. Norris. Differential equation approximations for markov chains. *Probability surveys*, 5:37–79, 2008.

-
- [DSS16] J. Ding, A. Sly, and N. Sun. Maximum independent sets on random regular graphs. *Acta Mathematica*, 217(2):263–340, 2016.
- [DvLM16] S. Dhara, J.S.H. van Leeuwen, and D. Mukherjee. Generalized random sequential adsorption on erdős–rényi random graphs. *Journal of Statistical Physics*, 164(5):1217–1232, 2016.
- [DZ10] A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer-Verlag Berlin Heidelberg, 2010.
- [ER59] P. Erdős and A. Rényi. On random graphs, i. *Publicationes Mathematicae (Debrecen)*, 6:290–297, 1959.
- [ER60] P. Erdős and A. Rényi. On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci.*, 5:17–61, 1960.
- [Erd59] P. Erdős. Graph theory and probability. *Canadian Journal of Mathematics*, 11:34–38, 1959.
- [Eva93] J.W. Evans. Random and cooperative sequential adsorption. *Reviews of modern physics*, 65(4):1281, 1993.
- [FL92] A.M. Frieze and T. Łuczak. On the independence and chromatic numbers of random regular graphs. *Journal of Combinatorial Theory, Series B*, 54(1):123–132, 1992.
- [FM97] A. Frieze and C. McDiarmid. Algorithmic theory of random graphs. *Random Structures & Algorithms*, 10(1-2):5–42, 1997.
- [Fri90] A.M. Frieze. On the independence number of random graphs. *Discrete Mathematics*, 81(2):171–175, 1990.
- [GG10] D. Gamarnik and D.A. Goldberg. Randomized greedy algorithms for independent sets and matchings in regular graphs: Exact results and finite girth corrections. *Combinatorics, Probability and Computing*, 19(1):61–85, 2010.
- [GH85] R.L. Graham and P. Hell. On the history of the minimum spanning tree problem. *Annals of the History of Computing*, 7(1):43–57, 1985.
- [Gil59] E.N. Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.
- [GM75] G.R. Grimmett and C.J.H. McDiarmid. On colouring random graphs. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 77-2, pages 313–324. Cambridge University Press, 1975.
- [Gri99] G. Grimmett. What is percolation? In *Percolation*, pages 1–31. Springer, 1999.
- [GT06] D. Galvin and P. Tetali. Slow mixing of glaucer dynamics for the hard-core model on regular bipartite graphs. *Random Structures & Algorithms*, 28(4):427–443, 2006.

- [HR97] M.M. Halldórsson and J. Radhakrishnan. Greed is good: Approximating independent sets in sparse and bounded-degree graphs. *Algorithmica*, 18(1):145–163, 1997.
- [Jan08] S. Janson. The largest component in a subcritical random graph with a power law degree distribution. *The Annals of Applied Probability*, pages 1651–1668, 2008.
- [Jan09a] S. Janson. On percolation in random graphs with given vertex degrees. *Electronic Journal of Probability*, 14:86–118, 2009.
- [Jan09b] S. Janson. The probability that a random multigraph is simple. *Combinatorics, Probability and Computing*, 18(1-2):205–225, 2009.
- [JL07] S. Janson and M.J. Luczak. A simple solution to the k -core problem. *Random Structures & Algorithms*, 30(1-2):50–62, 2007.
- [JL09] S. Janson and M.J. Luczak. A new approach to the giant component problem. *Random Structures & Algorithms*, 34(2):197–216, 2009.
- [JS19a] M. Jonckheere and M. Sáenz. Asymptotic optimality and complexity of best response dynamics on random sparse distributed optimisation problems. *Under preparation*, 2019.
- [JS19b] M. Jonckheere and M. Sáenz. Asymptotic optimality of degree-greedy discovering of independent sets in configuration model graphs. *Under review in Stochastic Processes and their Applications*, 2019.
- [Kar76] R.M. Karp. The probabilistic analysis of some combinatorial search algorithms. *Algorithms and complexity: New directions and recent results*, 1:19, 1976.
- [Kim07] J.H. Kim. Poisson cloning model for random graphs. *Expositions of current mathematics*, 2007(Autumn-Meeting1):104–120, 2007.
- [KN05] M. Kodialam and T. Nandagopal. Characterizing the capacity region in multi-radio multi-channel wireless mesh networks. In *Proceedings of the 11th annual international conference on Mobile computing and networking*, pages 73–87. ACM, 2005.
- [KS81] R.M. Karp and M. Sipser. Maximum matching in sparse random graphs. In *Foundations of Computer Science, 1981. SFCS'81. 22nd Annual Symposium on*, pages 364–375. IEEE, 1981.
- [LCZW10] H. Li, Y. Cheng, C. Zhou, and P. Wan. Multi-dimensional conflict graph based computing for optimal capacity in mr-mc wireless networks. In *2010 IEEE 30th International Conference on Distributed Computing Systems*, pages 774–783. IEEE, 2010.
- [LK16] R. Laufer and L. Kleinrock. The capacity of wireless csma/ca networks. *IEEE/ACM Transactions on Networking*, 24(3):1518–1532, June 2016.

-
- [LKLW10] S.C. Liew, C.H. Kai, H.C. Leung, and P. Wong. Back-of-the-envelope computation of throughput distributions in csma wireless networks. *IEEE Transactions on Mobile Computing*, 9(9):1319–1331, Sept 2010.
- [Lor12] G.G. Lorentz. *Bernstein polynomials*. American Mathematical Soc., 2012.
- [LW07] J. Lauer and N.C. Wormald. Large independent sets in regular graphs of large girth. *Journal of Combinatorial Theory, Series B*, 97(6):999–1009, 2007.
- [McK87] B.D. McKay. Independent sets in regular graphs of high girth. *Ars Combinatoria*, 23:179–185, 1987.
- [Mie06] G. Miermont. Advanced probability. <https://bit.ly/2XtF62f>. *Lecture Notes*, 2006.
- [MM09] M. Mezard and A. Montanari. *Information, physics, and computation*. Oxford University Press, 2009.
- [MR98] M. Molloy and B. Reed. The size of the giant component of a random graph with a given degree sequence. *Combinatorics, probability and computing*, 7(3):295–305, 1998.
- [MS15] J.R. Marden and J.S. Shamma. Game theory and distributed control. In *Handbook of game theory with economic applications*, volume 4, pages 861–899. Elsevier, 2015.
- [MU17] M. Mitzenmacher and E. Upfal. *Probability and computing: randomization and probabilistic techniques in algorithms and data analysis*. Cambridge university press, 2017.
- [Mul68] J.P. Mullooly. A one dimensional random space-filling problem. *Journal of Applied Probability*, 5(2):427–435, 1968.
- [NSS04] M.K. Nakayama, P. Shahabuddin, and K. Sigman. On finite exponential moments for branching processes and busy periods for queues. *Journal of Applied Probability*, 41(A):273–280, 2004.
- [RLBB17] C. Rattaro, F. Larroca, P. Bermolen, and P. Belzarena. Estimating the medium access probability in large cognitive radio networks. *Ad Hoc Networks*, 63:1 – 13, 2017.
- [Ros73] R.W. Rosenthal. A class of games possessing pure-strategy nash equilibria. *International Journal of Game Theory*, 2(1):65–67, 1973.
- [SJK15] J. Sanders, M. Jonckheere, and S. Kokkelmans. Sub-poissonian statistics of jamming limits in ultracold rydberg gases. *Physical review letters*, 115(4):043002, 2015.
- [SW95] A. Shwartz and A. Weiss. *Large deviations for performance analysis: queues, communication and computing*, volume 5. CRC Press, 1995.

- [TV02] P. Toth and D. Vigo. *The vehicle routing problem*. SIAM, 2002.
- [Unw18] Unwiredlabs. OpenCellID, 2018.
- [VDH16] R. Van Der Hofstad. *Random graphs and complex networks*, volume 43. Cambridge university press, 2016.
- [Vig01] E. Vigoda. A note on the glauber dynamics for sampling independent sets. *The electronic journal of combinatorics*, 8(1):R8, 2001.
- [Wor95] N.C. Wormald. Differential equations for random processes and random graphs. *The annals of applied probability*, pages 1217–1235, 1995.
- [Wor99] N.C. Wormald. Models of random regular graphs. *London Mathematical Society Lecture Note Series*, pages 239–298, 1999.
- [YZH14] P. Yi, Y. Zhang, and Y. Hong. Potential game design for a class of distributed optimisation problems. *Journal of Control and Decision*, 1(2):166–179, 2014.