



UNIVERSIDAD DE CONCEPCIÓN
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

Isomorfismos entre Grafos Aleatorios Densos No Homogéneos

POR

Constanza Valentina Gacitúa Fuentes

Tesis presentada a la Facultad de Ciencias Físicas y Matemáticas de la
Universidad de Concepción para optar al título profesional de
Ingeniera Civil Matemática

Profesores Guías: Dr. Nicolás Sanhueza Matamala y Dr. Matías Pavez Signé.

28 de marzo de 2025,
Concepción, Chile.

© 2025 Constanza Gacitúa Fuentes

Se autoriza la reproducción total o parcial, con fines académicos, por cualquier medio o procedimiento, incluyendo la cita bibliográfica del documento.

Título

COMISIÓN EVALUADORA

Dr. Nicolás Sanhueza Matamala [Profesor guía]

Departamento de Ingeniería Matemática, Universidad de Concepción, Chile.

Dr. Matías Pavez Signé [Profesor guía]

Departamento de Ingeniería Matemática, Universidad de Chile, Chile.

Dr. Christopher Thraves

Universidad de Concepción, Chile.

Dr. Amitai Linker

Universidad Andrés Bello, Chile.

FECHA DE DEFENSA: 28 de marzo de 2025.

Agradecimientos

Para comenzar, quiero agradecer a Raúl por ser la compañía más atenta, cercana, paciente y amorosa en estos últimos años, especialmente por escucharme hablar cada tontera (incluso mientras duermo). El espacio que formamos juntos ha sido un gran pilar para mi supervivencia.

Gracias a mi mamá, que ha estado para tirarme para arriba y darme un abrazo apretado cuando más lo he necesitado, además de ser un referente de admiración enorme para mí.

Gracias a mi familia, papá, hermano, Nilda, abuela Marta, abuela Jeannette, tata Dario y abuelo Pancho. Ellos han tenido una tremenda responsabilidad en permitir que yo haya podido llegar a esta altura y estoy muy feliz de poder compartir este logro con ustedes.

Agradezco a Luna, Scott, Emma, Bambú, Moka y Lisa, por darme su cálido amor perruno, por recibir mi amor hacia ellos también y por hacerme reír muchísimo. No solo los mantengo siempre en mi memoria, sino también en las teclas de mi PC que están llenas de sus pelos.

Igualmente agradezco a mis amigos y amigas del equipo Voley Salud, a mi amiga Josefa y mis amigos los ambrosolis. Tantas personas increíbles que han formado elementalmente quien soy y han hecho mi vida más amena. Gracias por admitir en sus vidas mi intensidad, les quiero a más no poder.

En la universidad también hay un montón de personas a quienes quiero agradecer profundamente, pero no podría mencionarlas a todas porque alargaría demasiado este documento (y porque mi memoria es lo suficientemente frágil como para que me olvide de mencionarlas también). Agradezco a todas y todos los miembros del grupo de whatsapp cuyo nombre se asemeja (vulgarmente) a “cosas serias y raras”. Atesoro todos los resúmenes y momentos con ustedes que me permitieron soportar los ramos que odiamos colectivamente, y disfrutar los ramos más entretenidos. Especialmente agradezco al Niko, que después de todo lo que compartimos juntos terminó siendo como mi hermano, a la Mile por ser una tremenda referente para mi vida y con quien espero podamos seguir recorriendo lo que nos traerá esta carrera juntas. Gracias al Feña y a la Danae por ser mis primeros amigos en la carrera, y al Vicho Daza, por su acogedor cariño. También al espacio de la Asamblea de Mujeres y Disidencias de Género de Ing. Civil

Matemática, a ellas les agradezco su sororidad y su ayuda a poder sentirme realmente parte de la comunidad ICM desde la disidencia.

Agradezco a las y los profes del DIM, especialmente a profes Mónica, Christopher, Anahí, Julio y Manuel, por enseñarme matemáticas y también orientarme en mi desarrollo integral dentro de la carrera. A pesar de mi alta dispersión, ustedes me ayudaron a direccionar mis decisiones.

Casi terminando, quiero darle un agradecimiento ultra especial a los jóvenes Matías y Nicolás, mis profes guías, por permitirme trabajar con académicos de tan alto calibre, por enseñarme a través de una paciencia no numerable, por contagiarme su pasión y diversión en el desarrollo de la investigación, por escucharme hablar de memes que no entienden o no conocen, por respetar mi sensibilidad y por hacer extremadamente ameno este proceso. Anhele que, mientras ustedes también quieran, encontremos más problemas para trabajar y sentirnos estúpidos juntos.

Finalmente, quiero agradecer el financiamiento dado por el proyecto Fondecyt regular 1241398 de Matías Pavez Signé, para realizar esta tesis.

Índice general

Agradecimientos	IV
Contenidos	VI
Índice de figuras	VIII
Resumen	IX
1. Introducción	1
1.1. Grafos aleatorios	2
1.2. Grafos aleatorios no homogéneos	4
1.3. Estructura del documento	5
2. Preliminares	6
2.1. Notaciones y desigualdades básicas	6
2.2. Conceptos básicos de probabilidades	8
2.3. Grafos	12

2.3.1. Grafos aleatorios	14
2.4. Estructura de las demostraciones	15
2.5. Grafones	17
3. Tamaño del clique más grande	19
3.1. Caso $G(n, 1/2)$	19
3.2. Caso $G(n, p)$	24
3.3. Caso $G(n, W)$	28
4. Tamaño del subgrafo inducido en común más grande entre grafos aleatorios homogéneos	32
4.1. Grafo auxiliar	33
4.2. Caso $G_1 \sim G(n, 1/2)$ y $G_2 \sim G(n, 1/2)$	36
4.3. Caso $G_1, G_2 \sim G(n, p)$	43
4.4. Caso $G_1 \sim G(n, p_1)$ y $G_2 \sim G(n, p_2)$	49
5. Tamaño del subgrafo inducido en común entre grafos aleatorios no homogéneos	54
5.1. Caso $G_1 \sim G(n, 1/2)$ y $G_2 \sim G(n, W)$	54
6. Conclusiones y trabajo futuro	65
Bibliografía	72

Índice de figuras

1.1. Esquema ilustrativo de como se relacionan los problemas estudiados a lo largo de este trabajo.	5
2.1. De izquierda a derecha: subgrafo inducido por $\{2, 3, 4\}$ (en magenta) que corresponde a un triángulo (clique de tamaño de 3) y un subgrafo (en naranja).	13
4.1. Ejemplo del grafo auxiliar Γ con $S = \{2, 3, 6\}$, $T = \{3, 4, 6\}$, $(\phi(2), \phi(3), \phi(6)) = (6, 4, 7)$ y $(\varphi(3), \varphi(4), \varphi(6)) = (4, 3, 2)$	34
4.2. Ejemplo de Γ cuando $ S \cap T , \phi(S) \cap \varphi(T) \leq 1$	35
4.3. Región de los valores $p_1, p_2 \in [0, 1]^2$ que satisfacen la desigualdad (4.19).	50
4.5. Región de los valores $p_1, p_2 \in (0, 1)$ que satisfacen la restricción (4.22).	52
4.6. Gráfica de ambas regiones definidas por las restricciones (4.19) (en rosado) y (4.22) (en morado).	53
5.1. Región de los valores $p_1, p_2 \in (0, 1)$ junto a las coordenadas $(\frac{1}{2}, \alpha)$ y $(\frac{1}{2}, 1 - \alpha)$	64

Resumen

En este trabajo estudiamos el problema de encontrar el tamaño del subgrafo inducido en común más grande entre dos grafos aleatorios G_1 y G_2 , que denotaremos por $L(G_1, G_2)$. Este problema tiene relevancia en distintas aplicaciones como en reconocimiento de patrones, bioquímica y ciencia molecular. Recientemente se ha estudiado el valor típico de $L(G_1, G_2)$ entre dos grafos aleatorios generados por el clásico modelo homogéneo de Erdős–Rényi.

La intención de este estudio es generalizar dicho resultado a otros modelos de grafos aleatorios. Un *grafón* $W : [0, 1]^2 \rightarrow [0, 1]$ es una función simétrica y medible. Los grafones son comúnmente estudiados en la teoría de límites de grafos, y a partir de un grafón W se puede definir un modelo no homogéneo de grafos aleatorios, que se denota por $G(n, W)$, y coincide con el modelo homogéneo cuando el grafón es una función constante. En esta tesis, revisamos resultados conocidos en grafos aleatorios homogéneos, como el estudio del tamaño del clique más grande y el tamaño del subgrafo inducido en común más grande entre dos grafos. Además, extendemos el estudio del tamaño del subgrafo inducido en común más grande entre grafos no necesariamente homogéneos. En particular, obtenemos un resultado que no existe en la literatura actualmente, que corresponde al valor de $L(G_1, G_2)$ cuando $G_1 \sim G(n, 1/2)$ y $G_2 \sim G(n, W)$, donde W es un grafón adecuado.

Introducción

Un problema ampliamente estudiado en matemáticas discretas es el de determinar cuándo dos estructuras son similares. Un ejemplo clásico es el problema de encontrar la subsecuencia en común más larga entre dos secuencias escogidas al azar, el que fue abordado por Chvátal y Sankoff [8] en 1975 y tiene relevancia en aplicaciones como el estudio de la similitud entre secuencias de ADN [31]. Otro problema interesante es el de encontrar subestructuras grandes, como por ejemplo, el estudio del largo de la subsecuencia creciente más larga posible en una permutación escogida al azar [29]. Este problema tiene un gran valor en la matemática dada la diversidad de áreas que se deben tocar para abordarlo, incluyendo el área de combinatoria, probabilidades, análisis y ecuaciones diferenciales, por mencionar algunas. Por otro lado, es también relevante estudiar esta clase de problemas cuando las estructuras son grafos. En lo que sigue, enunciamos el problema que será el mayor foco de estudio para esta tesis.

Problema 1 (Tamaño del subgrafo inducido en común más grande). *Dados dos grafos G_1 y G_2 , encontrar $L(G_1, G_2)$, que denota el mayor tamaño posible de un subgrafo inducido en G_1 que sea isomorfo a un subgrafo inducido de G_2 .*

Este problema tiene diversas aplicaciones en química y biología, como cuando se quiere predecir o conocer un grado de similitud que existe entre dos estructuras que se modelen a través de grafos. En general, modelar estructuras moleculares a través de grafos puede ser bastante intuitivo y tiene la ventaja de que provee de una buena base teórica para su implementación computacional. Sin embargo, encontrar el tamaño del subgrafo inducido en común más grande entre dos grafos es un problema NP-difícil, por ende también existe un gran interés en entender cómo se puede abordar de manera algorítmica. Ehrlich y Rarey [15] realizaron una revisión de distintos problemas aplicados de biología y química que son modelados a través de encontrar el subgrafo inducido en común más grande entre dos grafos.

En esta tesis, estudiaremos el valor típico de $L(G_1, G_2)$ cuando G_1 y G_2 son grafos generados a partir de un proceso aleatorio.

1.1. Grafos aleatorios

El modelo de *grafos aleatorios binomiales* $G(n, p)$ es una distribución sobre grafos con n vértices, donde cada par de vértices conforma una arista de manera idéntica e independiente con probabilidad p . Este modelo fue introducido por Gilbert [19] alrededor del año 1960, y es también conocido como el también conocido como el *modelo de Erdős-Rényi* de grafos aleatorios. En particular, debido a que el modelo considera la misma probabilidad de que exista una arista para cada par de vértices, nos referiremos a este modelo como el modelo *homogéneo*.

Muy recientemente, desde el año 2023, se han realizado distintos trabajos dedicados a estudiar $L(G_1, G_2)$ cuando G_1 y G_2 son grafos aleatorios homogéneos, motivados a partir del siguiente fenómeno bastante curioso. Un modelo muy conocido en la teoría de grafos es el *grafo de Rado* $G(\infty, 1/2)$, que se define como el grafo aleatorio cuyo conjunto de vértices corresponde a todos los enteros no negativos y cada par de vértices conforma una arista de manera independiente y con probabilidad $1/2$ (para más detalle ver [6]). Erdős y Rényi [17] probaron que dos grafos independientes generados por el grafo de Rado son isomorfos con probabilidad 1. Sin embargo, la probabilidad de que dos grafos independientes $G(n, 1/2)$ sean isomorfos es a lo más $n!2^{-\binom{n}{2}} = o(1)$. Esto nos muestra una diferencia drástica entre el estudio de $L(G_1, G_2)$

dependiendo de si estamos en el caso finito o infinito. Para comprender más precisamente este fenómeno, Chatterjee y Diaconis [7] propusieron estudiar cuál es el tamaño del subgrafo inducido en común más grande entre dos grafos $G_1, G_2 \sim G(n, 1/2)$, como una manera de describir la similitud entre ambos grafos. Finalmente, probaron que con alta probabilidad $L(G_1, G_2)$ se concentra en dos puntos alrededor de $4 \cdot \log_2 n$.

Posteriormente, Surya, Warnke y Zhu [30] estudiaron $L(G_1, G_2)$ para dos grafos aleatorios homogéneos $G_1 \sim G(n, p_1)$ y $G_2 \sim G(n, p_2)$, con $p_1, p_2 \in (0, 1)$ fijos, mostrando que con alta probabilidad $L(G_1, G_2) = (1 + o(1)) \cdot \beta(p_1, p_2) \cdot \log n$, donde $\beta(p_1, p_2)$ es una constante que depende solo de p_1 y p_2 y por \log nos referiremos al logaritmo natural. Diamantidis, Konstantopoulos y Yuan [13] también estudiaron el tamaño del subgrafo común más grande entre dos grafos $G_1 \sim G(n, p_1)$ y $G_2 \sim G(n, p_2)$, pero solo para (p_1, p_2) dentro de cierta región $\mathbb{Y} \subset [0, 1]^2$. Este problema también ha sido estudiado por Lenoir [25] para el caso de hipergrafos aleatorios uniformes.

Un caso particular del Problema 1 es cuando tomamos $G_1 \sim G(n, p)$ y $G_2 \sim G(n, 1)$. En este caso, cualquier subgrafo inducido en común entre G_1 y G_2 necesariamente será un clique, lo que nos lleva a otro problema que clásico en la literatura sobre grafos aleatorios

Problema 2 (Tamaño del clique más grande). *Dados $p \in (0, 1)$ y $G \sim G(n, p)$, encontrar $\omega(G)$, esto es; encontrar el tamaño del clique más grande en G .*

Determinar el tamaño del clique más grande en grafos aleatorios homogéneos ha sido ampliamente estudiado desde la década de 1970, donde destacan los trabajos de Bollobás y Erdős [4], Grimmett y McDiarmid [20] y Matula [28]. Ellos probaron que, con alta probabilidad, el tamaño del clique más grande en un grafo aleatorio $G(n, p)$, con $p \in (0, 1)$ una constante fija, es $(2 + o(1)) \cdot \log_{1/p} n$.

Nuestro objetivo es estudiar estos problemas cuando los grafos aleatorios no necesariamente sean homogéneos, donde por ejemplo la probabilidad asociada a cada par de vértices no sea necesariamente una misma constante.

1.2. Grafos aleatorios no homogéneos

Un *grafón* es una función $W : [0, 1]^2 \rightarrow [0, 1]$ simétrica y medible. El concepto de grafón fue introducido por Lovász y Szegedy [27] el 2006, el que aparece como objeto crucial en la teoría de límites de secuencias de grafos densos y probaron que, bajo cierta medida de convergencia, si una secuencia de grafos $(G_n)_n$ es convergente, entonces converge a un grafón. Recíprocamente, probaron que para todo grafón W , existe una secuencia de grafos $(G_n)_n$ que converge a W . Para probar esto último, es relevante definir un modelo *no homogéneo* de grafos aleatorios. A partir de un grafón W , podemos definir el modelo de grafos aleatorios no homogéneos $G \sim G(n, W)$ como sigue. Primero, hacemos un muestreo de n puntos independientes $x_1, x_2, \dots, x_n \in [0, 1]$, de modo que cada vértice $i \in V(G)$ es representado por $x_i \in [0, 1]$. Luego, cada par de vértices distintos $i, j \in V(G)$ conformará una arista de manera independiente con probabilidad $W(x_i, x_j)$. Lovász y Szegedy probaron que la secuencia de grafos $(G(n, W))_n$ converge al grafón W con probabilidad 1.

Un tipo de modelo no homogéneo es el *modelo bloque estocástico*, que equivale al modelo $G(n, W)$ cuando W es un *grafón escalonado*. Este modelo fue introducido en la década del 1980 (antes que el modelo más general $G(n, W)$ para cualquier grafón W), en el contexto del estudio de redes sociales [21].

Distintos problemas se han abordado respecto al modelo $G(n, W)$. Por ejemplo, se ha estudiado el problema de conectividad, es decir, determinar cuándo un grafo $G(n, W)$ es conexo [12], y también se ha estudiado el diámetro de un grafo no homogéneo [18]. Más aún, el 2017 Doležal, Hladký y Máthé [14] estudiaron el tamaño del clique más grande en un grafo $G(n, W)$. Específicamente, vieron que si W es un grafón que satisface ciertas propiedades técnicas, entonces con alta probabilidad el tamaño típico del clique más grande en $G(n, W)$ es $(1 + o(1)) \cdot \kappa(W) \cdot \log n$, e inclusive lograron precisar el valor de la constante $\kappa(W)$ que depende solo de W .

Entonces, una pregunta natural que nos podemos hacer es *¿cuál es tamaño del subgrafo inducido común más grande entre dos grafos aleatorios no homogéneos?* Dicha pregunta no se ha visto resuelta en la literatura y el objetivo de esta tesis es precisamente abordarla.

1.3. Estructura del documento

En el capítulo 2 veremos notaciones y conceptos básicos de probabilidades, teoría de grafos, grafos aleatorios y grafones. Además, enunciaremos resultados preliminares necesarios para el desarrollo de las demostraciones, incluyendo la explicación del método de primer y segundo momento, y cotas para coeficientes binomiales.

En el capítulo 3 estudiaremos el tamaño del clique más grande en grafos aleatorios homogéneos, y finalmente enunciaremos y explicaremos brevemente el resultado de Doležal, Hladký y Máthé, donde entregan una fórmula explícita para el tamaño del clique más grande en un grafo no homogéneo.

En el capítulo 4 estudiaremos el tamaño del subgrafo inducido en común más grande entre grafos homogéneos y compararemos algunos de los resultados obtenidos por Surya, Warnke y Zhu, y Diamantidis, Konstantopoulus y Yuan.

En el capítulo 5 mostramos que el tamaño del subgrafo inducido en común más grande entre $G(n, 1/2)$ y $G(n, W)$, para cierto grafón W , es con alta probabilidad igual a $(4 + o(1)) \cdot \log_2 n$. Recordemos que, hasta ahora, no se han visto estudios en la literatura sobre el tamaño del subgrafo inducido en común más grande entre grafos aleatorios no homogéneos.

Finalmente, en el capítulo 6 daremos las conclusiones del nuestro trabajo, mencionaremos algunos problemas que quedan por resolver y nuestra conjetura para el tamaño del subgrafo inducido en común más grande entre dos grafos aleatorios no homogéneos $G(n, W_1)$ y $G(n, W_2)$, con W_1 y W_2 grafones adecuados.

La idea es que mientras avancemos en el documento podamos ir apreciando cómo van variando las demostraciones a medida que los problemas se hacen cada vez más generales. Todas las demostraciones se realizan aplicando el método de primer y segundo momento.

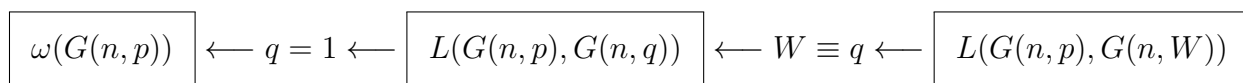


Figura 1.1: Esquema ilustrativo de como se relacionan los problemas estudiados a lo largo de este trabajo.

Preliminares

2.1. Notaciones y desigualdades básicas

Para $n \geq 1$ entero, definimos el conjunto $[n] := \{1, \dots, n\}$ y $[n]_0 := [n] \cup \{0\}$. Dado un conjunto V finito y $k \leq |V|$, definimos el conjunto $V^{(k)} := \{S \subseteq V : |S| = k\}$.

Usaremos la notación $x \ll y$ para decir que estamos tomando un valor x significativamente más pequeño que y .

El siguiente lema es una cota que usaremos frecuentemente en nuestro desarrollo.

Lema 2.1.1. Sean $0 \leq \alpha \leq \frac{1}{2}$ y $k > 0$. Entonces,

$$\sum_{i \geq k} \alpha^i \leq 2\alpha^k.$$

Demostración. Usando que la serie geométrica $\sum_{i \geq 0} \alpha^i$ converge a $\frac{1}{1-\alpha}$ y siguiente igualdad para

la suma parcial

$$\sum_{0 \leq i \leq k} \alpha^i = \frac{1 - \alpha^k}{1 - \alpha}$$

tenemos que

$$\begin{aligned} \sum_{i \geq k} \alpha^i &= \sum_{i \geq 0} \alpha^i - \sum_{0 \leq i \leq k} \alpha^i \\ &= \frac{1}{1 - \alpha} - \frac{1 - \alpha^k}{1 - \alpha} \\ &= \frac{\alpha^k}{1 - \alpha} \\ \boxed{\alpha \leq 1/2} &\leq 2\alpha^k, \end{aligned}$$

que era lo que queríamos. □

Como es usual, denotaremos por $\binom{n}{k}$ al coeficiente binomial dado por $\frac{n!}{k!(n-k)!}$, para todo $n, k \geq 0$ enteros. A menudo será necesario acotar los coeficientes binomiales y para esto damos las siguientes cotas conocidas para los coeficientes binomiales [11, ec. (2)].

Lema 2.1.2. *Sea n entero positivo. Luego, para todo $1 \leq k \leq n$ se tiene*

$$\left(\frac{n}{k}\right)^k \leq \binom{n}{k} \leq \left(\frac{en}{k}\right)^k.$$

Además, vamos a recurrir bastante a la aproximación de Stirling [11, ec. (4) y (6)].

Lema 2.1.3 (Aproximación de Stirling). *Sea n entero positivo.*

(I) *Para $\omega(1) = k = o(\sqrt{n})$*

$$\binom{n}{k} = \frac{1 + o(1)}{\sqrt{2\pi k}} \left(\frac{en}{k}\right)^k.$$

(II) *Para $k = o(n)$*

$$\binom{n}{k} = \exp \left\{ (1 + o(1))k \log \left(\frac{n}{k}\right) \right\}.$$

2.2. Conceptos básicos de probabilidades

A continuación, realizamos una breve revisión de algunos conceptos básicos de probabilidades desde la teoría de la medida. Veremos la definición de un espacio de probabilidad, variable aleatoria, valor esperado, varianza, covarianza y sus propiedades (para más detalle ver [1] [2]).

Sean Ω un conjunto no vacío, \mathcal{F} una σ -álgebra de subconjuntos de Ω . Decimos que \mathbb{P} es una *función de probabilidad*, o simplemente una *probabilidad*, definida en \mathcal{F} si satisface las siguientes propiedades:

- (I) $\mathbb{P}[A] \geq 0$, para todo $A \in \mathcal{F}$,
- (II) $\mathbb{P}[\Omega] = 1$, y
- (III) dada $\{A_n\}_{n=1}^{\infty}$ una secuencia de elementos disjuntos de \mathcal{F} , se tiene que

$$\mathbb{P}\left[\bigcup_{i=1}^{\infty} A_i\right] = \sum_{i=1}^{\infty} \mathbb{P}[A_i].$$

Luego, a la tupla $(\Omega, \mathcal{F}, \mathbb{P})$ se le denomina *espacio de probabilidad*. Aquí, Ω es el *espacio muestral*, los elementos de \mathcal{F} son *eventos* y para todo $A \in \mathcal{F}$, $\mathbb{P}[A]$ es la *probabilidad* de que ocurra A .

Dado un espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$, decimos que la función $X : \Omega \rightarrow \mathbb{R}$ es una *variable aleatoria* si es \mathcal{F} -medible. En otras palabras, X es una variable aleatoria si para todo conjunto medible $A \in \mathcal{B}(\mathbb{R})$, se tiene que $X^{-1}(A) \in \mathcal{F}$.

Sea $\mathcal{B}(\mathbb{R})$ la σ -álgebra de Borel en \mathbb{R} , esto es, la σ -álgebra generada por los intervalos abiertos en \mathbb{R} . Para todo $A \in \mathcal{B}(\mathbb{R})$, definimos

$$\mathbb{P}[X \in A] := \mathbb{P}[\{\omega \in \Omega : X(\omega) \in A\}] = \mathbb{P}[X^{-1}(A)].$$

Luego, la función \mathbb{P}_X a valores en $\mathcal{B}(\mathbb{R})$ definida por

$$\mathbb{P}_X[A] := \mathbb{P}[X \in A]$$

es una medida de probabilidad en el espacio $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Aquí, \mathbb{P}_X se denomina como la *distribución*

bución de X .

Dado $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad, decimos que el conjunto de eventos $\{A_1, A_2, \dots, A_n\} \subseteq \mathcal{F}$ es *independiente* si para todo subconjunto $\{A_{i_1}, A_{i_2}, \dots, A_{i_k}\}$ se satisface que

$$\mathbb{P} \left[\bigcap_{j \in [k]} A_{i_j} \right] = \prod_{j \in [k]} \mathbb{P}[A_{i_j}].$$

Análogamente, decimos que un conjunto de variables aleatorias $\{X_1, X_2, \dots, X_n\}$ es *independiente* si para todo subconjunto $\{X_{i_1}, X_{i_2}, \dots, X_{i_k}\}$ y para todo conjunto de eventos $\{A_{i_1}, A_{i_2}, \dots, A_{i_k}\}$ se tiene que

$$\mathbb{P} \left[\bigcap_{j \in [k]} \{X_{i_j} \in A_{i_j}\} \right] = \prod_{j \in [k]} \mathbb{P}[X_{i_j} \in A_{i_j}].$$

Sea X una variable aleatoria del espacio de probabilidad $(\Omega, \mathcal{F}, \mathbb{P})$. Se define el *valor esperado* o la *esperanza* de X , que denotamos por $\mathbb{E}[X]$, como

$$\mathbb{E}[X] = \int_{\Omega} X(\omega) d\mathbb{P}(\omega).$$

La mayor parte del tiempo trabajaremos con variables aleatorias cuyo espacio muestral es numerable. Si $(\Omega, \mathcal{F}, \mathbb{P})$ es un espacio de probabilidad con Ω un conjunto finito o infinito numerable, entonces definimos el evento $X = x$ como el conjunto $\{\omega \in \Omega : X(\omega) = x\}$ y

$$\mathbb{P}[X = x] := \sum_{\omega \in \Omega: X(\omega)=x} \mathbb{P}[\omega].$$

Así, el valor de la esperanza está dado por

$$\mathbb{E}[X] = \sum_x x \mathbb{P}[X = x],$$

donde sumamos sobre los posible valores x que puede tomar X . Aquí, $\mathbb{E}[X]$ es acotado si $\sum_x |x| \mathbb{P}[X = x]$ converge.

Dados $(\Omega, \mathcal{F}, \mathbb{P})$ un espacio de probabilidad y X una variable aleatoria, algunas propiedades de la esperanza se dan a continuación.

(I) Para todo conjunto de variables aleatorias $\{X_1, \dots, X_n\}$ con esperanza finita, se tiene que

$$\mathbb{E} \left[\sum_{i \in [n]} X_i \right] = \sum_{i \in [n]} \mathbb{E}[X_i].$$

(II) Para todo $a > 0$ real se tiene que $\mathbb{E}[a \cdot X] = a \cdot \mathbb{E}[X]$.

(III) Si $X \geq 0$ en casi todo punto (c.t.p.), entonces $\mathbb{E}[X] \geq 0$.

(IV) $|\mathbb{E}[X]| \leq \mathbb{E}[|X|]$.

Se define la *varianza* de una variable aleatoria X , que denotamos por $\text{Var}[X]$, como

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2.$$

Algunas propiedades de la varianza se enuncian a continuación.

(I) $\text{Var}[X] \geq 0$.

(II) Para todo $a > 0$ constante real, se tiene que $\text{Var}[a] = 0$ y $\text{Var}[a \cdot X] = a^2 \cdot \text{Var}[X]$.

(III) Si X_1, X_2, \dots, X_n son variables aleatorias, entonces

$$\text{Var} \left[\sum_{i \in [n]} X_i \right] = \sum_{i \in [n]} \text{Var}[X_i] + \sum_{i \neq j} \text{Cov}(X_i, X_j). \quad (2.1)$$

Dadas dos variables aleatorias X e Y , se define la *covarianza* entre X e Y , que denotamos por $\text{Cov}(X, Y)$, como

$$\text{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])].$$

A partir de lo anterior, no es difícil ver que $\text{Var}[X] = \text{Cov}(X, X)$.

Si dos variables aleatorias X e Y son independientes, entonces

- $\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$,
- $\text{Cov}(X, Y) = 0$, y

$$\blacksquare \text{ Var}[X + Y] = \text{Var}[X] + \text{Var}[Y].$$

A continuación, enunciemos un resultado simple pero que nos será de ayuda más adelante.

Lema 2.2.1. *Dada una variable aleatoria $X \in \{0, 1\}$, entonces $\text{Var}[X] \leq \mathbb{E}[X]$.*

Demostración. Notemos que, por definición de varianza de una variable aleatoria

$$\text{Var}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2 \leq \mathbb{E}[X^2].$$

Como $X \in \{0, 1\}$, se sigue que $X = X^2$. Luego,

$$\text{Var}[X] \leq \mathbb{E}[X^2] = \mathbb{P}[X^2 = 1] = \mathbb{P}[X = 1] = \mathbb{E}[X],$$

que era lo que queríamos probar. □

A partir de lo anterior, se obtiene el siguiente resultado.

Teorema 2.2.2. *Sea X una variable aleatoria definida como suma de variables aleatorias Y_1, \dots, Y_n que toman valores 0 o 1. Luego,*

$$\text{Var}[X] \leq \mathbb{E}[X] + \sum_{i \neq j} \text{Cov}(Y_i, Y_j).$$

Demostración. La demostración se sigue directamente de aplicar el Lema 2.2.1 en la ecuación (2.1). □

Algunas desigualdades conocidas y frecuentemente utilizadas para desarrollar argumentos probabilistas son la desigualdad de Markov y la desigualdad de Chebyshev, que son enunciadas a continuación.

Lema 2.2.3 (Desigualdad de Markov). *Dadas X una variable aleatoria positiva y $a > 0$ una constante real, se tiene*

$$\mathbb{P}[X \geq a] \leq \frac{\mathbb{E}[X]}{a}.$$

Lema 2.2.4 (Desigualdad de Chebyshev). *Sea X una variable aleatoria con esperanza y varianza están bien definidas, y $a > 0$ una constante real. Entonces,*

$$\mathbb{P}[|X - \mathbb{E}[X]| \geq a] \leq \frac{\text{Var}[X]}{a^2}.$$

2.3. Grafos

Un *grafo* es un par ordenado $G = (V, E)$, donde $E \subseteq V^{(2)}$ y V se denomina conjunto de *vértices* y E , conjunto de *aristas*. Dado un grafo G , usaremos la notación $V(G)$ para denotar a su conjunto de vértices y $E(G)$, a su conjunto de aristas. Dada la forma en la que definimos grafos, trabajaremos con grafos simples, es decir, que para cada par de vértices existe a lo más una arista entre ellos, y sin aristas del tipo $\{x, x\}$ (loops).

Una arista $\{x, y\}$ en un grafo G representa una conexión entre dos vértices distintos $x, y \in V(G)$. Para simplificar la notación, también denotaremos por xy o yx a la arista $\{x, y\}$. Si $xy \in E(G)$, diremos que x e y son *vecinos* en G . Denotamos por $N_G(x)$ el conjunto de vecinos de x en G . Siempre y cuando no estemos hablando de otros grafos, simplemente usaremos la notación $N(x)$ en vez de $N_G(x)$. El *grado* de un vértice x corresponde a $|N(x)|$ y lo denotaremos por $d(x)$. El grado máximo de un G está dado por $\max_{x \in V(G)} d(x)$ y lo denotaremos por $\Delta(G)$. Análogamente, el grado mínimo de G está dado por $\min_{x \in V(G)} d(x)$ y lo denotaremos por $\delta(G)$. Consideraremos el *tamaño de un grafo* como la cardinalidad de su conjunto de vértices.

Dados dos grafos G_1 y G_2 , decimos que G_1 y G_2 son *isomorfos* si es que existe una biyección $\varphi : V(G_1) \rightarrow V(G_2)$ tal que para todo par de vértices $i, j \in V(G_1)$ se tiene que $ij \in E(G_1)$ en G_1 si y solo si $\varphi(i)\varphi(j) \in E(G_2)$. Esta relación entre ambos grafos se denota por $G_1 \cong G_2$. Específicamente, si queremos indicar que G_1 es isomorfo a G_2 a través de la función φ , usaremos la notación $G_1 \cong_{\varphi} G_2$.

Si G es un grafo, $\omega(G)$ es el tamaño del clique más grande en G . Si G_1 y G_2 son grafos, denotaremos por $L(G_1, G_2)$ al tamaño del subgrafo inducido en común, vía isomorfismos, más grande entre G_1 y G_2 .

Un subconjunto de vértices $S \subseteq V(G)$ se denomina *conjunto independiente* si no existen

aristas entre sus vértices.

Una manera de representar un grafo G con $V(G) = [n]$ y m aristas es a través de una *matriz de adyacencia* $A = (a_{ij})_{i,j \in [n]}$, donde

$$a_{i,j} = \begin{cases} 1 & \text{si } ij \in E(G), \\ 0 & \text{si no,} \end{cases}$$

para todo $i, j \in [n]$.

Un grafo G se denomina *bipartito* si es que existe una partición V_1, V_2 de sus vértices, tales que V_1 y V_2 son conjuntos independientes de G . En particular, si todo vértice en V_1 es vecino de todos los vértices en V_2 , G se denomina *grafo bipartito completo*, y se denota por $K_{|V_1|, |V_2|}$. Un *grafo completo* es un grafo que contiene todas las aristas posibles entre todos los sus vértices, al grafo completo con n vértices lo denotaremos por K_n .

Un grafo H es *subgrafo* de G si se cumple que $V(H) \subseteq V(G)$ y $E(H) \subseteq E(G)$, y esto se denota por $H \subseteq G$. En particular, si $E(H)$ contiene **todas** las aristas en $E(G)$ contenidas en $V(H)$, entonces H es un *subgrafo inducido*. Otra manera de definir un subgrafo inducido es a partir de un subconjunto de vértices. Dado $U \subseteq V(G)$, el subgrafo inducido cuyo conjunto de vértices es U se denota por $G[U]$, en dicho caso diremos que $G[U]$ es el *subgrafo inducido por U en G* , o que U induce el subgrafo $G[U]$ en G . Si además H es un grafo completo, entonces diremos que H es un *clique* en G de tamaño $|V(H)|$ o un $|V(H)|$ -*clique*.

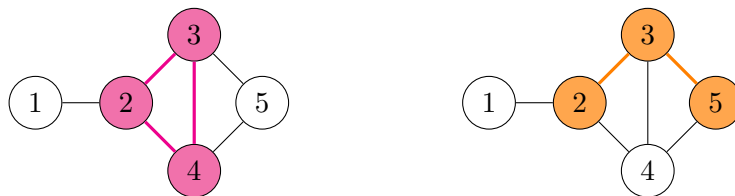


Figura 2.1: De izquierda a derecha: subgrafo inducido por $\{2, 3, 4\}$ (en magenta) que corresponde a un triángulo (clique de tamaño de 3) y un subgrafo (en naranja).

Un *emparejamiento* o *matching* en un grafo G corresponde a un subconjunto de aristas $M \subseteq E(G)$ que no poseen ningún vértice en común. Si un vértice pertenece a una arista en M , decimos que está *saturado* por M . Si M satura a todos los vértices de G , entonces M se

denomina *matching perfecto* en G .

Dado un grafo $G = (V, E)$, un *paseo* \mathcal{P} es una secuencia de vértices $\mathcal{P} = x_0x_1 \cdots x_n$, donde para todo $i \in \{0, \dots, n-1\}$ se tiene que $\{x_i, x_{i+1}\} \in E$. Si \mathcal{P} no repite vértices, es un *camino*.

Si \mathcal{P} es un (x, y) -camino pero $x = y$, \mathcal{P} es un *ciclo*. A un ciclo de largo 3 se le suele llamar *triángulo*. Si un ciclo pasa por todos vértices de un grafo se le denomina *ciclo hamiltoniano*. Si un grafo G posee un ciclo hamiltoniano, decimos que G es un *grafo hamiltoniano*.

Un grafo G es *conexo* si para todo par de vértices $x, y \in V(G)$ existe un (x, y) -camino que los conecte.

2.3.1. Grafos aleatorios

Dados $p \in (0, 1)$ y n un entero positivo, el modelo $G \sim G(n, p)$ define el espacio de probabilidades sobre todos los grafos de n vértices (en general, supondremos siempre que el conjunto de vértices será $[n]$). Las aristas de $G(n, p)$ se definen tal que, para todo par de vértices distintos i, j incluimos la arista ij en $G(n, p)$ con probabilidad p y de manera independiente del resto de las elecciones. A la probabilidad p se le denomina la *intensidad de aristas*.

Algunas propiedades básicas del grafo aleatorio $G \sim G(n, p)$ son:

- (I) $\mathbb{E}[|E(G)|] = \binom{n}{2}p$,
- (II) dado $S \subseteq V(G)$, $\mathbb{E}[|E(G[S])|] = \binom{|S|}{2}p$,
- (III) para todo $v \in V(G)$, $\mathbb{E}[d(v)] = (n-1)p$.

Una *propiedad en grafos* consiste en una familia de grafos que es invariante bajo isomorfismos, es decir, es independiente del etiquetado de sus vértices. Por ejemplo, si definimos la propiedad \mathcal{H} como la propiedad de ser conexo y P_n es un camino de largo n , entonces $P_n \in \mathcal{H}$, es decir, P_n satisface la propiedad \mathcal{H} .

Una propiedad \mathcal{H} se dice *monótona* si para un grafo $G \in \mathcal{H}$ se cumple que dada una arista $e \notin E(G)$, se tiene que $G + e \in \mathcal{H}$. Por ejemplo, la propiedad de ser bipartito no es monótona, pero la propiedad de ser hamiltoniano lo es. No es difícil ver que la propiedad de poseer un k -clique es monótona.

Es posible estudiar propiedades de un grafo $G(n, p)$ dependiendo del valor de p . Específicamente, veremos que hay algunas propiedades en $G(n, p)$ que se satisfacen con alta probabilidad dependiendo de si p está por debajo de cierto valor *umbral*, y que no se satisfacen con alta probabilidad cuando p supera dicho umbral. Formalmente, se define una función f como *umbral* para una propiedad \mathcal{H} , tal que

(I) si $p = p(n) = o(f(n))$, entonces $\lim_{n \rightarrow \infty} \mathbb{P}[G(n, p) \in \mathcal{H}] = 0$,

(II) si $p = p(n) = \omega(f(n))$, entonces $\lim_{n \rightarrow \infty} \mathbb{P}[G(n, p) \in \mathcal{H}] = 1$.

A partir del siguiente teorema [3, Theorem 4] veremos que las funciones umbrales cumplen un rol importante a la hora de estudiar propiedades monótonas.

Teorema 2.3.1 (Bollobás y Thomasson). *Toda propiedad monótona posee una función umbral.*

Algunos ejemplos de funciones umbrales para distintas propiedades monótonas de grafos son:

- $f(n) = \log n/n$, para la propiedad de “ser conexo”,
- $f(n) = 1/n$, para la propiedad de “tener triángulos”.

Decimos que una secuencia de eventos, que dependen de un cierto n , ocurre *con alta probabilidad* si es que ocurre con probabilidad que tiende a 1 cuando n tiende a infinito.

2.4. Estructura de las demostraciones

Una herramienta fundamental para el desarrollo de toda esta tesis es el Método del segundo momento. Este método nos permitirá encontrar funciones umbrales para propiedades en grafos. Específicamente, buscaremos funciones umbrales para determinar el tamaño más grande posible de ciertas subestructuras en grafos, como un clique o subgrafo inducido.

El procedimiento usual será el siguiente. Supongamos que estamos estudiando sobre la secuencia de grafos $(G(n, p))_n$ la propiedad de “contener cierta subestructura de tamaño $k =$

$k(n)$ ”, que denotaremos por \mathcal{H}_k . Para esto, definiremos una secuencia de variables aleatorias X_k , para todo $k \geq 1$, tal que X_k sea igual a la cantidad de subestructuras de tamaño k que estemos buscando. De esta forma, necesitamos encontrar una función $f : \mathbb{N} \rightarrow \mathbb{R}$ tal que para todo $\varepsilon > 0$

(I) si $k = (1 + \varepsilon) \cdot f(n)$, entonces $\lim_{n \rightarrow \infty} \mathbb{P}[G(n, p) \in \mathcal{H}_k] = 0$,

(II) si $k = (1 - \varepsilon) \cdot f(n)$, entonces $\lim_{n \rightarrow \infty} \mathbb{P}[G(n, p) \in \mathcal{H}_k] = 1$.

Intuitivamente, la parte (I) nos indica que si el tamaño de la subestructura es mayor a $f(n)$, entonces con alta probabilidad no existirá dicha subestructura en el grafo o equivalentemente, con alta probabilidad el tamaño de la subestructura más grande no supera al valor de $f(n)$, es decir, encontramos una cota superior para k . Para demostrar esto, necesitamos probar que

$$\mathbb{P}[X_k \geq 1] = o(1).$$

Lo anterior podemos conseguirlo a través de la desigualdad de Markov (Lema 2.2.3). Por lo tanto, necesitamos que cuando $k = k(n) \gg f(n)$, tengamos que $\mathbb{E}[X_k] = o(1)$, y así por la desigualdad de Markov (Lema 2.2.3) obtendremos que

$$\mathbb{P}[X_k \geq 1] \leq \mathbb{E}[X_k] = o(1),$$

que era lo que queríamos.

Para la parte (II) necesitamos hacer un proceso análogo al anterior, pero ahora queremos mostrar que si el tamaño de la subestructura que buscamos es mucho menor a $f(n)$, entonces con alta probabilidad existirá dicha subestructura en $G(n, p)$. Equivalentemente, queremos mostrar que si $k \ll f(n)$, entonces

$$\mathbb{P}[X_k = 0] = o(1).$$

Para conseguir esto, necesitamos controlar la varianza de X_k , de manera que al aplicar la siguiente desigualdad podamos concluir lo que queremos.

Lema 2.4.1. *Dada X una variable aleatoria que toma valores enteros positivos, se tiene*

$$\mathbb{P}[X = 0] \leq \frac{\text{Var}[X]}{\mathbb{E}[X]^2}.$$

Demostración. Notemos que, como $X > 0$

$$\mathbb{P}[X = 0] = \mathbb{P}[X \leq 0] = \mathbb{P}[X - \mathbb{E}[X] \leq -\mathbb{E}[X]] \leq \mathbb{P}[|X - \mathbb{E}[X]| \geq \mathbb{E}[X]] \leq \frac{\text{Var}[X]}{\mathbb{E}[X]^2},$$

donde la última cota se obtiene usando la desigualdad de Chebyshev (Lema 2.2.4). \square

Diremos que el método del segundo momento “funciona” cuando tenemos que $k \ll f$ y $\mathbb{E}[X_k] \xrightarrow[n]{} \infty$ implican que $\mathbb{P}[X_k = 0] = o(1)$. Si bien, a lo largo de este documento estudiaremos casos en los que este método funciona, esto no siempre es así cuando no es posible acotar adecuadamente la varianza. Por ejemplo, si tuviésemos una variable aleatoria del tipo $X_n = nY$, donde Y es una variable aleatoria que toma valor 1 con probabilidad 1/2 y toma valor 0 con probabilidad 1/2. Entonces, $\mathbb{E}[X_n] = \frac{1}{2} \cdot n \xrightarrow[n]{} \infty$, pero $\mathbb{P}[X_n = 0] = 1/2 > 0$.

2.5. Grafones

En esta sección daremos algunas definiciones elementales para nuestro trabajo, en torno a materia de grafones (para más detalle ver [26]).

Un *grafón* $W : [0, 1]^2 \rightarrow [0, 1]$ es una función simétrica, es decir, para todo $(x, y) \in [0, 1]^2$, $W(x, y) = W(y, x)$, y medible.

Definición 1. *Sea $\{\Omega_1, \dots, \Omega_k\}$ una partición de $[0, 1]$ y para todo $i, j \in [k]$, $p_{ij} \in [0, 1]$ es una constante. Un *grafón escalonado* es un grafón $W : [0, 1]^2 \rightarrow [0, 1]$ definido tal que para todo $i, j \in [k]$ y para todo $(x, y) \in (\Omega_i \times \Omega_j) \cup (\Omega_j \times \Omega_i)$, $W(x, y) = p_{ij}$.*

Dado un grafón W podemos definir un modelo de grafos aleatorios más general que el modelo binomial. El modelo de grafos aleatorios *no homogéneo* $G(n, W)$ se construye de la siguiente forma. Supondremos siempre que su conjunto de vértices es $[n]$. Tomamos una muestra de

n elementos independientes $x_1, \dots, x_n \in [0, 1]$. Luego, todo par de vértices i, j conforma una arista con probabilidad $W(x_i, x_j)$ y de manera independiente del resto de las elecciones.

Una manera equivalente de generar un grafo $G(n, W)$ es a través del grafo aleatorio $H \sim \mathbb{H}(n, W)$, el cual se obtiene de la siguiente manera. Se extrae una muestra de n elementos independientes $x_1, x_2, \dots, x_n \in [0, 1]$. Se tiene el conjunto de vértices $V(H) = [n]$. Luego, H será un grafo completo con pesos, tal que cada arista ij en H tendrá peso $W(x_i, x_j)$. De esta manera, podemos obtener el grafo $G \sim G(n, W)$ tomando todos los vértices de H e incluyendo en G cada arista ij de H con probabilidad $W(x_i, x_j)$, correspondiente a su peso en H .

Dado $X \subseteq \Omega$, decimos que $U : X^2 \rightarrow [0, 1]$ es un *subgrafón de W obtenido por X* si $U \equiv W|_{X^2}$. Respecto a esto, supondremos que X como espacio de medida estará dotado de la medida $\lambda_X := \frac{\lambda(Y)}{\lambda(X)}$, para todo $Y \subseteq X$, donde λ denota la medida de Lebesgue.

Dado H un grafo con d vértices, podemos definir el *grafón representante de H* , que denotamos por W_H , como sigue. Tomamos una partición $\{X_1, X_2, \dots, X_d\}$ de Ω , tal que $\lambda(X_i) = 1/d$, para todo $i \in [d]$. Luego, para todo $x_i, x_j \in X_i \times X_j$

$$W_H(x_i, x_j) = \begin{cases} 1, & \text{si } ij \in E(H) \\ 0, & \text{si no.} \end{cases}$$

Aquí, W_H es un grafón escalonado. Notemos que no existe un único grafón representante para H , ya que depende de la elección de los X_i .

Para una función medible $f : X \rightarrow \mathbb{R}$ se define su *supremo esencial*, que denotamos por $\text{ess sup } f$, como la menor cota superior α tal que la medida del conjunto $\{x \in X : f(x) > \alpha\}$ sea nula. El *ínfimo esencial* de f , que denotamos por $\text{ess inf } f$, se define de manera análoga.

Tamaño del clique más grande

Nuestro objetivo en este capítulo es estudiar cuál es el tamaño típico del clique más grande en un grafo aleatorio binomial. Lo que nos motiva a realizar este estudio es que determinar el tamaño típico del clique más grande en un grafo aleatorio binomial es un caso particular del problema de encontrar el tamaño del subgrafo inducido en común más grande entre dos grafos aleatorios, que es nuestro problema foco. De esta manera podremos identificar detalladamente cuáles son las herramientas necesarias para ir obteniendo resultados cada vez más generales.

3.1. Caso $G(n, 1/2)$

El problema más elemental que estudiaremos será encontrar el tamaño del clique más grande en un grafo $G(n, 1/2)$, que denotaremos por $\omega(G(n, 1/2))$. Este problema corresponde a un caso particular del estudio de Matula [28], Grimmett y McDiarmid [20], y Bollobás y Erdős [4].

Teorema 3.1.1. *Con alta probabilidad, $\omega(G(n, 1/2)) = (2 + o(1)) \cdot \log_2 n$.*

Demostración. Sea $G \sim G(n, 1/2)$. Para cada entero positivo k , definimos la variable aleatoria X_k correspondiente a la cantidad de k -cliques en G . Asimismo, dado $S \subseteq V(G)$, definimos Y_S como la variable aleatoria que toma valor 1 si $G[S]$ es un clique y 0, si no. De esta forma,

$$X_k = \sum_{S \in V(G)^{(k)}} Y_S. \quad (3.1)$$

Así, se obtiene la esperanza de X_k como

$$\mathbb{E}[X_k] = \sum_{S \in V(G)^{(k)}} \mathbb{P}[Y_S = 1] = \sum_{S \in V(G)^{(k)}} 2^{-\binom{k}{2}} = \binom{n}{k} 2^{-\binom{k}{2}}. \quad (3.2)$$

La demostración estará estructurada de la siguiente manera: primero veremos una cota superior para $\omega(G)$ a través de un argumento de primer momento, éste nos permitirá mostrar que cuando k supera cierto umbral, la esperanza de X_k convergerá a 0 con n y que por tanto, la probabilidad de que exista un k -clique tenderá a 0 cuando n tiende a infinito. Luego, veremos una cota inferior para $\omega(G)$ a través de un argumento de segundo momento, el cual nos permitirá concluir que, cuando la esperanza de X_l diverge con n y si es que podemos controlar la varianza de X_l , entonces con alta probabilidad existirá un l -clique en G . Equivalentemente, mostraremos que la probabilidad de que X_l sea menor que 1 converge a 0 cuando n tiende a infinito.

COTA SUPERIOR: Dado $\varepsilon > 0$, fijemos $k = (2 + \varepsilon) \cdot \log_2 n$. Usando el Lema 2.1.2 se tiene la siguiente cota superior para la esperanza,

$$\mathbb{E}[X_k] = \binom{n}{k} 2^{-\binom{k}{2}} \leq \left(\frac{en}{k}\right)^k 2^{-\frac{k(k-1)}{2}} = \left[\frac{en}{k} \cdot 2^{-\frac{(k-1)}{2}}\right]^k.$$

Aquí, notamos que $2^{-\frac{(k-1)}{2}} = \sqrt{2} \cdot n^{-1-\frac{\varepsilon}{2}}$ y además, $k = (2 + \varepsilon) \cdot \log_2 n \xrightarrow[n]{\rightarrow} \infty$. Por ende,

$$\mathbb{E}[X_k] \leq \underbrace{\left[\frac{\sqrt{2}e}{k} \cdot n^{-\frac{\varepsilon}{2}}\right]}_{<1}^k = o(1).$$

Notemos que lo anterior se puede interpretar como que en promedio no hay cliques de tamaño mayor a $(2 + \varepsilon) \cdot \log_2 n$, cuando $n \rightarrow +\infty$. Esto último, junto a la desigualdad de Markov

(Lema 2.2.3), implica que

$$\mathbb{P}[X_k \geq 1] \leq \mathbb{E}[X_k] = o(1).$$

Por lo tanto, con alta probabilidad no existe un clique en G de tamaño mayor a $(2 + \varepsilon) \cdot \log_2 n$, que era lo que queríamos mostrar.

COTA INFERIOR: Dado $\varepsilon > 0$, fijemos $k = (2 - \varepsilon) \cdot \log_2 n$. Usando el Lema 2.1.2 obtenemos la siguiente cota inferior para la esperanza

$$\mathbb{E}[X_k] = \binom{n}{k} 2^{-\binom{k}{2}} \geq \left(\frac{n}{k}\right)^k 2^{-\frac{k(k-1)}{2}} = \left[\frac{n}{k} \cdot 2^{-\frac{(k-1)}{2}}\right]^k.$$

Aquí, notamos que $2^{-\frac{(k-1)}{2}} = \sqrt{2} \cdot n^{-1+\frac{\varepsilon}{2}}$, y como además $k = o(n)$, obtenemos que

$$\mathbb{E}[X_k] \geq \left[\frac{\sqrt{2}}{k} \cdot n^{\frac{\varepsilon}{2}}\right]^k \xrightarrow{n} \infty.$$

Observemos que lo anterior se puede interpretar como que en promedio hay una gran cantidad de cliques de tamaño menor a $(2 - \varepsilon) \cdot \log_2 n$, pero esto no siempre nos garantiza que con alta probabilidad exista un clique de tamaño menor a $(2 - \varepsilon) \cdot \log_2 n$. Recordemos que el Lema 2.4.1 nos dice que

$$\mathbb{P}[X_k = 0] \leq \frac{\text{Var}[X_k]}{\mathbb{E}[X_k]^2}. \quad (3.3)$$

Por ende, para concluir lo que buscamos, es necesario estudiar la varianza de X_k . Recordemos que, para $S \in V(G)^{(k)}$, Y_S corresponde a la variable aleatoria que toma valor 1 si $G[S]$ es un clique, y 0 si no. Luego, por (3.1) y por Teorema 2.2.2 se sigue que

$$\text{Var}[X_k] \leq \mathbb{E}[X_k] + \underbrace{\sum_{\substack{S, T \in V(G)^{(k)} \\ 1 \leq |S \cap T| \leq k-1}} \text{Cov}(Y_S, Y_T)}_{(\diamond)}. \quad (3.4)$$

En lo que sigue, acotaremos el término (\diamond) en la desigualdad (3.4). Notemos que si S y T no se intersecan en al menos un par de vértices, esto es, si $|S \cap T| \leq 1$, entonces las variables Y_S e Y_T son independientes. Lo anterior se debe a que si S y T no poseen al menos un par de vértices en común, entonces no conforman ninguna posible arista en común en G . Luego, la covarianza

entre ambas variables aleatorias es 0. Por ello, en (\blacklozenge) solo consideraremos las covarianzas sobre los S y T tales que $|S \cap T| \geq 2$. Ahora, nos concentramos en acotar las covarianzas en (\blacklozenge) . Como Y_S e Y_T son no negativas, entonces

$$\text{Cov}(Y_S, Y_T) = \mathbb{E}[Y_S Y_T] - \mathbb{E}[Y_S]\mathbb{E}[Y_T] \leq \mathbb{E}[Y_S Y_T] = \mathbb{P}[Y_S = 1, Y_T = 1] = 2^{-2\binom{k}{2} + \binom{|S \cap T|}{2}}.$$

Notemos que la última igualdad se obtiene porque el total de pares de vértices que potencialmente conforman una arista en $G[S]$ es $\binom{k}{2}$. Cada uno conforma una arista con probabilidad independiente y uniforme igual a $\frac{1}{2}$. Luego, se tiene exactamente el mismo conteo para los pares de vértices en $G[T]$, descontando los $\binom{|S \cap T|}{2}$ pares de vértices que hay en común con $G[S]$. Usando lo anterior, se tiene que

$$\begin{aligned} (\blacklozenge) &= \sum_{\substack{S, T \in V(G)^{(k)}: \\ 2 \leq |S \cap T| \leq k-1}} \text{Cov}(Y_S, Y_T) \leq \sum_{\substack{S, T \in V(G)^{(k)}: \\ 2 \leq |S \cap T| \leq k-1}} 2^{-2\binom{k}{2} + \binom{|S \cap T|}{2}} \\ &= \sum_{2 \leq i \leq k-1} \sum_{\substack{S, T \in V(G)^{(k)}: \\ |S \cap T| = i}} 2^{-2\binom{k}{2} + \binom{|S \cap T|}{2}} \\ &= \sum_{2 \leq i \leq k-1} \binom{n}{k} \binom{k}{i} \binom{n-k}{k-i} 2^{-2\binom{k}{2} + \binom{i}{2}}. \end{aligned} \quad (3.5)$$

Notar que en la última ecuación (3.5), $\binom{n}{k}$ cuenta las posibles maneras de elegir el conjunto S , $\binom{k}{i}$ cuenta las posibles maneras de escoger $S \cap T$ y $\binom{n-k}{k-i}$ cuenta las posibles maneras de elegir $T \setminus S$. De esta forma, aplicando la igualdad vista en (3.5) en la desigualdad (3.4) obtenemos

$$\text{Var}[X_k] \leq \mathbb{E}[X_k] + \sum_{2 \leq i \leq k-1} \binom{n}{k} \binom{k}{i} \binom{n-k}{k-i} 2^{-2\binom{k}{2} + \binom{i}{2}}. \quad (3.6)$$

Aplicando la última desigualdad (3.6) en la cota del Segundo Momento (3.3) se tiene que

$$\begin{aligned} \mathbb{P}[X_k = 0] &\leq \frac{1}{\mathbb{E}[X_k]} + \sum_{2 \leq i \leq k-1} \frac{\binom{n}{k} \binom{k}{i} \binom{n-k}{k-i} 2^{-2\binom{k}{2} + \binom{i}{2}}}{\binom{n}{k}^2 \cdot 2^{-2\binom{k}{2}}} \\ &= \frac{1}{\mathbb{E}[X_k]} + \sum_{2 \leq i \leq k-1} \frac{\binom{k}{i} \binom{n-k}{k-i} 2^{\binom{i}{2}}}{\binom{n}{k}} \end{aligned}$$

$$= \frac{1}{\mathbb{E}[X_k]} + \sum_{2 \leq i \leq k-1} a_i, \quad (3.7)$$

donde en la ecuación anterior definimos, para todo $i \in \{2, \dots, k-1\}$,

$$a_i := \frac{\binom{k}{i} \binom{n-k}{k-i} 2^{\binom{i}{2}}}{\binom{n}{k}}.$$

Para continuar acotando el último término de la desigualdad (3.7) veremos que, para n suficientemente grande, a_i se maximiza en $i = 2$. En efecto,

$$\begin{aligned} \frac{a_i}{a_2} &= \frac{\binom{k}{i} \binom{n-k}{k-i} 2^{\binom{i}{2}}}{\binom{k}{2} \binom{n-k}{k-2} \cdot 2} \\ &= \frac{2}{i!} \cdot \frac{(n-2k+2)!}{(n-2k+i)!} \cdot \left[\frac{(k-2)!}{(k-i)!} \right]^2 \cdot 2^{\frac{i(i-1)}{2}-1} \\ &\leq \frac{1}{(n-2k)^{i-2}} \cdot k^{2(i-2)} \cdot 2^{\frac{i(i-1)}{2}-1} \\ &= \left[\frac{k^2 \cdot 2^{\frac{i+1}{2}}}{n-2k} \right]^{i-2} \\ &\leq \left[\frac{k^2 \cdot 2^{\frac{k+1}{2}}}{n-2k} \right]^{i-2}. \end{aligned}$$

Como $k = (2 - \varepsilon) \cdot \log_2 n$, entonces $2^{\frac{k+1}{2}} = \sqrt{2} n^{1-\frac{\varepsilon}{2}}$. Además, para n suficientemente grande se tiene que $k \leq \frac{n}{4}$, entonces

$$\frac{a_i}{a_2} \leq \left[\underbrace{\frac{2\sqrt{2} \cdot n^{1-\frac{\varepsilon}{2}}}{n}}_{<1} \right]^{i-2} < 1. \quad (3.8)$$

Ahora, podemos aplicar la propiedad anterior (3.8) en la desigualdad (3.7), de donde se obtiene, para n suficientemente grande

$$\mathbb{P}[X_k = 0] \leq \frac{1}{\mathbb{E}[X_k]} + \sum_{2 \leq i \leq k-1} a_i$$

$$\leq \frac{1}{\mathbb{E}[X_k]} + k \cdot \frac{\binom{k}{2} \binom{n-k}{k-2} \cdot 2}{\binom{n}{k}} \quad (3.9)$$

$$\leq \frac{1}{\mathbb{E}[X_k]} + 2k \cdot \frac{\binom{k}{2} \binom{n-k}{k-2}}{\binom{n-k}{k}}$$

$$= \frac{1}{\mathbb{E}[X_k]} + 2k \cdot \frac{\binom{k}{2} \binom{n-k}{k-2}}{\binom{n-k}{k-2} \cdot \frac{(n-2k+1)(n-2k+2)}{k(k-1)}}$$

$$= \frac{1}{\mathbb{E}[X_k]} + 2k \cdot \frac{k(k-1) \binom{k}{2}}{(n-2k+1)(n-2k+2)}$$

$$\leq \frac{1}{\mathbb{E}[X_k]} + 2k \cdot \frac{k(k-1) \binom{k}{2}}{(n-2k)^2}$$

$$\boxed{\text{para } n \text{ suficientemente grande}} \leq \frac{1}{\mathbb{E}[X_k]} + 8k \cdot \frac{k(k-1) \binom{k}{2}}{n^2}$$

$$\leq \frac{1}{\mathbb{E}[X_k]} + 8 \frac{k^5}{n^2}. \quad (3.10)$$

Luego, para todo $\varepsilon > 0$ y $k = (2 - \varepsilon) \cdot \log_2 n$, se tiene que $\mathbb{E}[X_k] \xrightarrow[n]{\infty} \infty$ y $k^5 n^{-2} = o(1)$. Por lo cual, obtenemos que con alta probabilidad existe al menos un clique de tamaño menor a $(2 - \varepsilon) \cdot \log_2 n$, lo que concluye la demostración. \square

3.2. Caso $G(n, p)$

Ahora, veremos un resultado más general que el Teorema 3.1.1. Debido a que las demostraciones son muy similares, simplificaremos varios argumentos. En la década del 1970, Matula [28], Grimmett y McDiarmid [20], y Bollobás y Erdős [4] dieron distintas demostraciones para probar el siguiente resultado.

Teorema 3.2.1. *Sea $p \in (0, 1)$. Luego, con alta probabilidad $\omega(G(n, p)) = (2 + o(1)) \cdot \log_{1/p} n$.*

Demostración. Sean $p \in (0, 1)$ y $G \sim G(n, p)$. Para cada entero positivo k , definimos la variable aleatoria X_k como la cantidad de k -cliques en G y para todo $S \subseteq V(G)$, Y_S es la variable aleatoria que toma valor 1 si $G[S]$ es un clique y 0, si no. Así, si calculamos la esperanza de X_k

de manera análoga a la vista en la igualdad (3.2), obtenemos que

$$\mathbb{E}[X_k] = \binom{n}{k} p^{\binom{k}{2}}. \quad (3.11)$$

La demostración estará estructurada de la manera vista en la demostración de Teorema 3.1.1. Comenzaremos desarrollando un argumento de primer momento.

COTA SUPERIOR: Dado $\varepsilon > 0$, fijamos $k = (2 + \varepsilon) \cdot \log_b n$, donde $b := p^{-1}$. Usando Lema 2.1.2, se tiene la siguiente cota superior para la esperanza,

$$\mathbb{E}[X_k] = \binom{n}{k} p^{\binom{k}{2}} \leq \left(\frac{en}{k}\right)^k p^{\frac{k(k-1)}{2}} = \left[\frac{en}{k} \cdot p^{\frac{k-1}{2}}\right]^k.$$

Y aquí, notamos que $p^{\frac{k-1}{2}} = \sqrt{b} n^{-\frac{(2+\varepsilon)}{2}}$. Así,

$$\mathbb{E}[X_k] \leq \left[\frac{e\sqrt{b}}{k} \cdot n^{-\frac{\varepsilon}{2}}\right]^k = o(1).$$

Luego, para todo $\varepsilon > 0$ y $k \geq (2 + \varepsilon) \cdot \log_b n$ se tiene que $\mathbb{E}[X_k] = o(1)$. Por la desigualdad de Markov (Lema 2.2.3), se obtiene que

$$\mathbb{P}[X_k \geq 1] \leq \mathbb{E}[X_k] = o(1).$$

Por ende, con alta probabilidad no existe clique de tamaño mayor a $(2 + \varepsilon) \cdot \log_b n$, que era lo que queríamos probar.

COTA INFERIOR: Veremos ahora el argumento de segundo momento, de la misma manera que se vio en la demostración de Teorema 3.1.1.

Dado $\varepsilon > 0$, fijamos $k = (2 - \varepsilon) \cdot \log_b n$. Usando Lema 2.1.2 acotamos inferiormente la cantidad esperada de k -cliques como sigue

$$\mathbb{E}[X_k] = \binom{n}{k} p^{\binom{k}{2}} \geq \left(\frac{n}{k}\right)^k p^{\frac{k(k-1)}{2}} = \left[\frac{n}{k} \cdot p^{\frac{k-1}{2}}\right]^k.$$

Aquí, notamos que $p^{\frac{k-1}{2}} = \sqrt{b}n^{-1+\frac{\varepsilon}{2}}$ y por ende

$$\mathbb{E}[X_k] \geq \left[\frac{\sqrt{b}}{k} \cdot n^{\frac{\varepsilon}{2}} \right]^k \xrightarrow{n} +\infty.$$

Ahora, queremos estudiar la probabilidad de que existan cliques de tamaño menor a $(2 - \varepsilon) \cdot \log_b n$. Equivalentemente, queremos ver la probabilidad de que no existan cliques de tamaño a lo más $(2 - \varepsilon) \cdot \log_b n$. Por Lema 2.4.1, tenemos que

$$\mathbb{P}[X_k = 0] \leq \frac{\text{Var}[X_k]}{\mathbb{E}[X_k]^2}. \quad (3.12)$$

Recordemos que, para $S \in V(G)^{\binom{k}{k}}$, Y_S corresponde a la variable aleatoria que toma valor 1 si $G[S]$ es un clique, y 0 si no. Luego, repitiendo de manera análoga los cálculos y argumentos vistos para acotar los términos de las desigualdades (3.4), (3.5) y (3.6) tenemos que

$$\text{Var}[X_k] \leq \mathbb{E}[X_k] + \sum_{2 \leq i \leq k-1} \binom{n}{k} \binom{k}{i} \binom{n-k}{k-i} p^{2^{\binom{k}{2}} - \binom{i}{2}}. \quad (3.13)$$

De esta manera, aplicando la igualdad (3.11) y la desigualdad (3.13) en la cota del Segundo Momento (3.12) se tiene que

$$\begin{aligned} \mathbb{P}[X_k = 0] &\leq \frac{1}{\mathbb{E}[X_k]} + \sum_{2 \leq i \leq k-1} \frac{\binom{n}{k} \binom{k}{i} \binom{n-k}{k-i} p^{2^{\binom{k}{2}} - \binom{i}{2}}}{\binom{n}{k}^2 \cdot p^{2^{\binom{k}{2}}}} \\ &= \frac{1}{\mathbb{E}[X_k]} + \sum_{2 \leq i \leq k-1} \frac{\binom{k}{i} \binom{n-k}{k-i} b^{\binom{i}{2}}}{\binom{n}{k}} \\ &= \frac{1}{\mathbb{E}[X_k]} + \sum_{2 \leq i \leq k-1} d_i. \end{aligned} \quad (3.14)$$

En la ecuación anterior definimos, para todo $i \in \{2, \dots, k-1\}$,

$$d_i := \frac{\binom{k}{i} \binom{n-k}{k-i} b^{\binom{i}{2}}}{\binom{n}{k}}.$$

Para continuar acotando el último término de la desigualdad (3.14) veremos que d_i se maximiza

en $i = 2$. En efecto, repitiendo de manera análoga los cálculos vistos para a_i en la demostración de Teorema 3.1.1 se obtiene que

$$\frac{d_i}{d_2} \leq \left[\underbrace{\frac{2\sqrt{b} \cdot n^{1-\frac{\varepsilon}{2}}}{n}}_{<1} \right]^{i-2} < 1.$$

Continuando con la desigualdad (3.14), a partir de la maximización de d_i , podemos repetir de manera análoga los cálculos vistos en la desigualdad vista anteriormente (3.10). Así, se tiene

$$\begin{aligned} \mathbb{P}[X_k = 0] &\leq \frac{1}{\mathbb{E}[X_k]} + \sum_{2 \leq i \leq k-1} d_2 \\ &\leq \frac{1}{\mathbb{E}[X_k]} + k \cdot \frac{\binom{k}{2} \binom{n-k}{k-2} \cdot b}{\binom{n}{k}} \\ &= \frac{1}{\mathbb{E}[X_k]} + bk \cdot \frac{\binom{k}{2} \binom{n-k}{k-2}}{\binom{n-k}{k}} \\ &= \frac{1}{\mathbb{E}[X_k]} + bk \cdot \frac{\binom{k}{2} \binom{n-k}{k-2}}{\binom{n-k}{k-2} \cdot \frac{(n-2k+1)(n-2k+2)}{k(k-1)}} \\ &= \frac{1}{\mathbb{E}[X_k]} + bk \cdot \frac{k(k-1) \binom{k}{2}}{(n-2k+1)(n-2k+2)} \\ &\leq \frac{1}{\mathbb{E}[X_k]} + bk \cdot \frac{k(k-1) \binom{k}{2}}{(n-2k)^2} \\ &\stackrel{\boxed{\text{para } n \text{ suficientemente grande}}}{\leq} \frac{1}{\mathbb{E}[X_k]} + 4bk \cdot \frac{k(k-1) \binom{k}{2}}{n^2} \\ &\leq \frac{1}{\mathbb{E}[X_k]} + 4b \frac{k^5}{n^2}. \end{aligned} \tag{3.15}$$

Luego, para todo $\varepsilon > 0$ y $k = (2 - \varepsilon) \cdot \log_b n$ se tiene que $\mathbb{E}[X_k] \xrightarrow[n]{\rightarrow} \infty$ y $k^5 n^{-2} = o(1)$. Por lo tanto, con alta probabilidad existe al menos un clique de tamaño menor a $(2 - \varepsilon) \cdot \log_b n$, lo que concluye la demostración. \square

3.3. Caso $G(n, W)$

A continuación veremos el tamaño del clique más grande en un grafo aleatorio no homogéneo, el cual generaliza los resultados vistos en el Teorema 3.1.1 y el Teorema 3.2.1.

Sea $W : [0, 1]^2 \rightarrow [0, 1]$ un grafón con ínfimo esencial positivo. Definimos

$$\kappa(W) := \sup \left\{ \frac{2\|h\|_1^2}{\int_{(x,y) \in [0,1]^2} h(x)h(y) \log(1/W(x,y)) d(\lambda^2)} : h \text{ es una función } L^1 \text{ no negativa en } [0, 1] \right\}.$$

Considerando que $\frac{0}{0} = 0$ y que para todo $a \in \mathbb{R} - \{0\}$, $\frac{a}{0} = +\infty$, el siguiente teorema fue demostrado en el año 2017 por Doležal, Hladký y Máthé [14, Theorem 2.4].

Teorema 3.3.1. *Sea W un grafón con ínfimo esencial positivo. Luego, con alta probabilidad se tiene que*

- si $\kappa(W) < +\infty$, $\omega(G(n, W)) = (1 + o(1)) \cdot \kappa(W) \cdot \log n$, y
- si $\kappa(W) = +\infty$, $\omega(G(n, W)) \gg \log n$.

Es claro que podemos recuperar el Teorema 3.2.1 a partir del Teorema 3.3.1, ya que si tomamos $W \equiv p$ se tiene que $\text{ess inf } W = p > 0$ y

$$\frac{2\|h\|_1^2}{\int_{(x,y) \in [0,1]^2} h(x)h(y) \log(1/W(x,y)) d(\lambda^2)} = \frac{2\|h\|_1^2}{\int_{(x,y) \in [0,1]^2} h(x)h(y) \log \frac{1}{p} d(\lambda^2)} = \frac{2}{\log \frac{1}{p}}, \quad (3.16)$$

para toda función h . Por lo tanto, $\kappa(W) = \frac{2}{\log \frac{1}{p}}$ y por ende, del Teorema 3.3.1 se sigue que $\omega(G(n, p)) = (1 + o(1)) \cdot \frac{2}{\log \frac{1}{p}} \cdot \log n = (1 + o(1)) \cdot 2 \cdot \log_{1/p} n$, que es justamente el resultado dado por el Teorema 3.2.1.

Doležal, Hladký y Máthé [14, Fact 2.6] probaron la siguiente fórmula equivalente para calcular $\kappa(W)$.

$$\kappa(W) = \sup \left\{ \|h\|_1 : h \text{ es una función } L^1 \text{ no negativa en } [0, 1], \Gamma(f, W) \geq 0 \right\}, \quad (3.17)$$

donde

$$\Gamma(h, W) = \int_{x \in [0,1]} h(x) d(\lambda) + \frac{1}{2} \int_{(x,y) \in [0,1]^2} h(x)h(y) \log W(x,y) d(\lambda^2). \quad (3.18)$$

No enseñaremos la demostración del Teorema 3.3.1, sin embargo, daremos algunas ideas generales que nos ayudarán a entender mejor este problema.

Sean $p_1, p_2 \in (0, 1)$ tales que $p_1 \leq p_2$. Sea además $W : [0, 1]^2 \rightarrow [0, 1]$ un grafón tal que $p_1 = \text{ess inf } W \leq \text{ess sup } W = p_2$. Tomamos ahora de manera uniforme e independiente, los elementos $x_i \in [0, 1]$, para todo $i \in [n]$ y los elementos $y_{jk} \in [0, 1]$, para todo $jk \in [n]^{(2)}$. A partir de esto, generamos los grafos aleatorios G_1 , G_2 y G_W como sigue:

- (I) si $y_{jk} \leq p_1$, jk conforma una arista en G_1 ,
- (II) si $y_{jk} \leq p_2$, jk conforma una arista en G_2 ,
- (III) si $y_{jk} \leq W(x_i, x_j)$, jk conforma una arista en G_W .

De lo anterior, se sigue que $E(G_1) \subseteq E(G_W) \subseteq E(G_2)$, y por ende $G_1 \subseteq G_W \subseteq G_2$. Lo anterior implica que

$$\omega(G_1) \leq \omega(G_W) \leq \omega(G_2). \quad (3.19)$$

Por otra parte, los eventos definidos por (I), (II) e (III) ocurren con probabilidad p_1 , p_2 y $W(x_i, x_j)$, respectivamente. Entonces, $G_1 \sim G(n, p_1)$, $G_2 \sim G(n, p_2)$ y $G_W \sim G(n, W)$, por ello podemos aplicar el Teorema 3.2.1. Así, obtenemos que con alta probabilidad

- $\omega(G_1) = (2 + o(1)) \cdot \log_{1/p_1} n = \frac{2}{\log \frac{1}{p_1}} \cdot (1 + o(1)) \cdot \log n$, y
- $\omega(G_2) = (2 + o(1)) \cdot \log_{1/p_2} n = \frac{2}{\log \frac{1}{p_2}} \cdot (1 + o(1)) \cdot \log n$.

Usando lo anterior en las desigualdades (3.19), obtenemos que con alta probabilidad

$$\frac{\omega(G_1)}{\log n} = (1 + o(1)) \cdot \frac{2}{\log \frac{1}{p_1}} \leq \frac{\omega(G)}{\log n} \leq (1 + o(1)) \cdot \frac{2}{\log \frac{1}{p_2}} = \frac{\omega(G_2)}{\log n}.$$

Finalmente, mostramos que el tamaño del clique más grande en $G(n, W)$ crece en orden logarítmico. La dificultad principal de demostrar el Teorema 3.3.1 está en encontrar la constante $\kappa(W)$.

El esquema de la demostración del Teorema 3.3.1 es similar a los que hemos visto hasta ahora, en el sentido de que también es a través de un argumento de primer y segundo momento.

Para ilustrar esto, veamos el caso para un grafón escalonado. Sea $W : [0, 1]^2 \rightarrow [0, 1]$ tal que, para todo $(x, y) \in [0, 1]^2$

$$W(x, y) = \begin{cases} p_{11}, & \text{si } (x, y) \in \Omega_1 \times \Omega_1, \\ p_{22}, & \text{si } (x, y) \in \Omega_2 \times \Omega_2, \\ p_{12}, & \text{en otro caso,} \end{cases}$$

donde $p_{11}, p_{22}, p_{12} \in (0, 1)$ y $\{\Omega_1, \Omega_2\}$ es una partición de $[0, 1]$ tal que Ω_1 y Ω_2 son conjuntos de medida positiva β_1 y β_2 , respectivamente.

El objetivo, al igual que en las otras demostraciones que hemos visto, es determinar para qué constante real $c > 0$ se tiene que con alta probabilidad existe un clique de tamaño $c \log n$, y para qué constante c no existe con alta probabilidad un clique de tamaño $c \log n$. Para este caso, como estamos trabajando sobre una partición en 2 conjuntos de $[0, 1]$, nos interesa contar la cantidad de cliques que poseen c_1 y c_2 vértices representados en Ω_1 y Ω_2 , respectivamente, de manera que maximicemos el valor de $c_1 + c_2 = c$. De lo anterior, podemos contar la cantidad de aristas en un clique que posee c_1 y c_2 vértices representados en Ω_1 y Ω_2 , respectivamente, como sigue

$$\underbrace{\binom{c_1 \log n}{2}}_{\text{aristas con ambos vértices en } \Omega_1} + \underbrace{\binom{c_2 \log n}{2}}_{\text{aristas con ambos vértices en } \Omega_2} + \underbrace{c_1 c_2 \log^2 n}_{\text{aristas con un vértice en } \Omega_1 \text{ y un vértice en } \Omega_2}.$$

Ahora, definimos X_{c_1, c_2} como la variable aleatoria que enumera la cantidad de cliques en G con c_1 y c_2 vértices representados en Ω_1 y Ω_2 , respectivamente. Realizando cálculos análogos a los vistos en las igualdades (3.2) y en (3.11), el valor de la esperanza de X_k es

$$\begin{aligned} \mathbb{E}[X_{c_1, c_2}] &\approx \binom{\beta_1 n}{c_1 \log n} \cdot p_{11}^{\binom{c_1 \log n}{2}} \cdot \binom{\beta_2 n}{c_2 \log n} \cdot p_{22}^{\binom{c_2 \log n}{2}} \cdot p_{12}^{c_1 c_2 \log^2 n} \\ &= \exp \left\{ (1 + o(1)) \log^2 n \left[c_1 + c_2 + \frac{c_1^2}{2} \log p_{11} + \frac{c_2^2}{2} \log p_{22} + c_1 c_2 \log p_{12} \right] \right\}, \end{aligned}$$

donde la aproximación se debe a que estamos aproximando la cantidad de vértices en G que

están representados en Ω_1 y Ω_2 como $\beta_1 n$ y $\beta_2 n$, respectivamente, y la última igualdad se obtiene ocupando aproximación de Stirling (caso (II) Lema 2.1.3) en los coeficientes binomiales. Observemos que, a pesar de lo anterior, finalmente obtuvimos una expresión de $\mathbb{E}[X_{c_1, c_2}]$ que no depende de β_1 ni de β_2 . Así, se sigue que podemos controlar el valor de la esperanza de X_{c_1, c_2} dependiendo del signo del siguiente valor

$$\theta = \theta(c_1, c_2) := c_1 + c_2 + \frac{c_1^2}{2} \log p_{11} + \frac{c_2^2}{2} \log p_{22} + c_1 c_2 \log p_{12}.$$

De esta manera, tenemos la siguiente relación entre θ y $\mathbb{E}[X_{c_1, c_2}]$

$$\mathbb{E}[X_{c_1, c_2}] \xrightarrow[n]{} 0 \quad \text{si y solo si } \theta < 0, \quad \text{y} \quad \mathbb{E}[X_{c_1, c_2}] \xrightarrow[n]{} \infty \quad \text{si y solo si } \theta > 0. \quad (3.20)$$

A partir de esto, por un argumento de primer y segundo momento (que no es trivial de desarrollar, pero debido a los resultados vistos anteriormente podríamos esperar que funcionaría), obtendríamos que con alta probabilidad existe un clique en G de tamaño $c \log n$ si y solo si existen $c_1, c_2 > 0$ tales que $c_1 + c_2 = c$ y $\theta > 0$, donde esto último equivale a resolver el siguiente problema de optimización

$$\begin{aligned} \text{máx. } \quad & \theta(c_1, c_2) := c_1 + c_2 + \frac{c_1^2}{2} \log p_{11} + \frac{c_2^2}{2} \log p_{22} + c_1 c_2 \log p_{12} \\ \text{sujeto a } \quad & c_1 + c_2 = c. \end{aligned}$$

Notemos que podríamos haber tomado una partición de Ω con conjuntos de medida β_i de tamaño infinitesimal. Esto nos permite intuir una relación directa entre el valor de θ y la fórmula de $\Gamma(h, W)$ en la identidad (3.18). Más aún, la dependencia entre la convergencia de $\mathbb{E}[X_k]$ y el signo de θ vista en (3.20) se relaciona directamente con la condición $\Gamma(h, W) \geq 0$ en la fórmula de $\kappa(W)$ (3.17).

Tamaño del subgrafo inducido en común más grande entre grafos aleatorios homogéneos

Dados dos grafos aleatorios homogéneos independientes G_1 y G_2 , nuestro objetivo ahora consiste en determinar el tamaño típico del subgrafo inducido en común más grande entre ambos grafos, el cual denotamos por $L(G_1, G_2)$. Recordemos que el tamaño de un grafo es la cantidad de vértices que posee. En esta etapa del documento, estamos estudiando un problema que generaliza el problema de encontrar el tamaño del clique más grande en un grafo aleatorio, que estudiamos en el capítulo anterior. Esto se debe a que si fijamos a G_1 como un grafo completo y G_2 es un grafo aleatorio binomial, entonces cualquier subgrafo inducido que tengan en común será un clique, en particular, el subgrafo inducido en común más grande que posean será el clique más grande en G_2 .

Al igual que en el capítulo anterior, iremos avanzando de un resultado a otro en la medida que este último que sea más general que el anterior.

4.1. Grafo auxiliar

Una herramienta que utilizaremos a lo largo del estudio del tamaño del subgrafo inducido en común más grande entre dos grafos aleatorios es un grafo auxiliar, que definiremos a continuación. Sean $G_1 \sim G(n, p_1)$ y $G_2 \sim G(n, p_2)$ dos grafos aleatorios independientes, cuyos conjuntos de vértices denotaremos por V_1 y V_2 , respectivamente. Sea $k \geq 1$ un entero y sean $S, T \in V_1^{(k)}$, $\phi : S \rightarrow V_2$ y $\varphi : T \rightarrow V_2$ funciones inyectivas. A medida de que desarrollemos las demostraciones, nos va a interesar calcular el valor de

$$\mathbb{P}[G_1[S] \cong_\phi G_2[\phi(S)] \text{ y } G_1[T] \cong_\varphi G_2[\varphi(T)]],$$

el cual evidentemente sería sencillo de calcular si es que los eventos $\{G_1[S] \cong_\phi G_2[\phi(S)]\}$ y $\{G_1[T] \cong_\varphi G_2[\varphi(T)]\}$ fuesen independientes. Sin embargo, la mayoría de la veces este no será el caso. Para abordar esta dificultad, definimos el siguiente grafo auxiliar $\Gamma = \Gamma(S, T, \phi, \varphi)$

$$V(\Gamma) := A \cup B \text{ y } E(\Gamma) := E_\phi \cup E_\varphi, \quad (4.1)$$

donde

$$\begin{aligned} A &:= S^{(2)} \cup T^{(2)}, \\ B &:= \phi(S)^{(2)} \cup \varphi(T)^{(2)}, \\ E_\phi &:= \{ab, \phi(a)\phi(b) : ab \in S^{(2)}\}, \text{ y} \\ E_\varphi &:= \{xy, \varphi(x)\varphi(y) : xy \in T^{(2)}\}, \end{aligned}$$

La idea es poder representar, a través de cada vértice en Γ , pares de vértices en G_1 y G_2 . Específicamente, los elementos de A representan todos los pares de vértices en G_1 y los elementos de B , a todos los pares de vértices en G_2 . De esta forma, una arista en Γ representará la conexión entre un par de vértices en G_1 con su par de vértices en G_2 asociado a través de ϕ o a través de φ , según corresponda. Es fácil ver que Γ es bipartito, con A, B su bipartición de vértices. Entonces, dados $ij, st \in V(\Gamma)$, la arista $\{ij, st\} \in E(\Gamma)$ nos indica que ij conforma una arista en

G_1 si y solo si st conforma una arista en G_2 . Por ende, cada componente conexa en Γ representa un grupo de pares de vértices en G_1 y G_2 que debe actuar de manera sincronizada en cuanto a si conforman o no una arista.

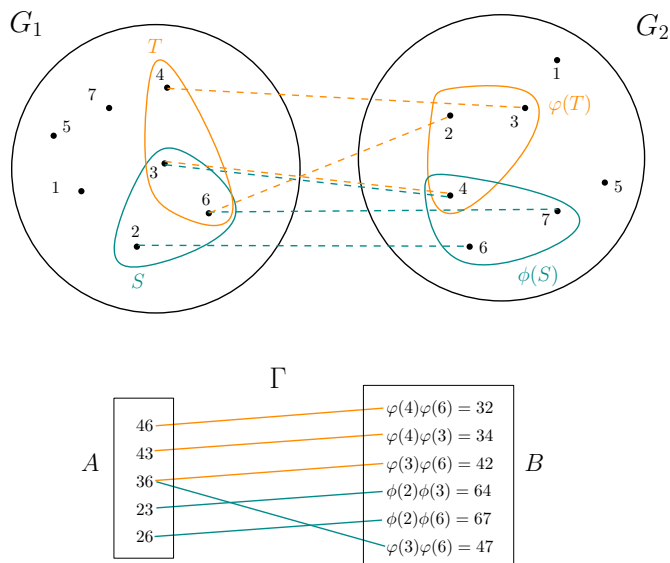


Figura 4.1: Ejemplo del grafo auxiliar Γ con $S = \{2, 3, 6\}$, $T = \{3, 4, 6\}$, $(\phi(2), \phi(3), \phi(6)) = (6, 4, 7)$ y $(\varphi(3), \varphi(4), \varphi(6)) = (4, 3, 2)$.

Más aún, podemos detallar la forma que tienen las componentes conexas de Γ . Para esto, notemos que Γ corresponde a la unión de los grafos $\Gamma_1 := (A \cup B, E_\phi)$ y $\Gamma_2 := (A \cup B, E_\varphi)$. Es claro que Γ_1 y Γ_2 son bipartitos, además por la inyectividad de ϕ y φ , se tiene que todos sus vértices poseen grado exactamente 1. Por ende, E_ϕ y E_φ son matchings perfectos de Γ_1 y Γ_2 , respectivamente. Luego, $\Delta(\Gamma) \leq 2$ y por ende las componentes conexas de Γ son caminos o ciclos de largo par.

Debido a que nuestro interés es calcular una probabilidad conjunta, queremos entender, a través de nuestro grafo Γ , cómo es la relación de dependencia entre los eventos $\{G_1[S] \cong_\phi G_2[\phi(S)]\}$ y $\{G_1[T] \cong_\varphi G_2[\varphi(T)]\}$. El caso más sencillo de estudiar es cuando el tamaño de la intersección entre S y T o entre $\phi(S)$ y $\varphi(T)$ es a lo más 1. En tal caso, el evento donde $G_1[S] \cong_\phi G_2[\phi(S)]$ es independiente de si $G_1[T] \cong_\varphi G_2[\varphi(T)]$, ya que no hay pares de vértices en común entre $G_1[S]$ y $G_1[T]$, ni tampoco entre $G_2[\phi(S)]$ y $G_2[\varphi(T)]$. En el grafo auxiliar, esto se traduce a que no hay ningún par de aristas adyacentes en Γ .

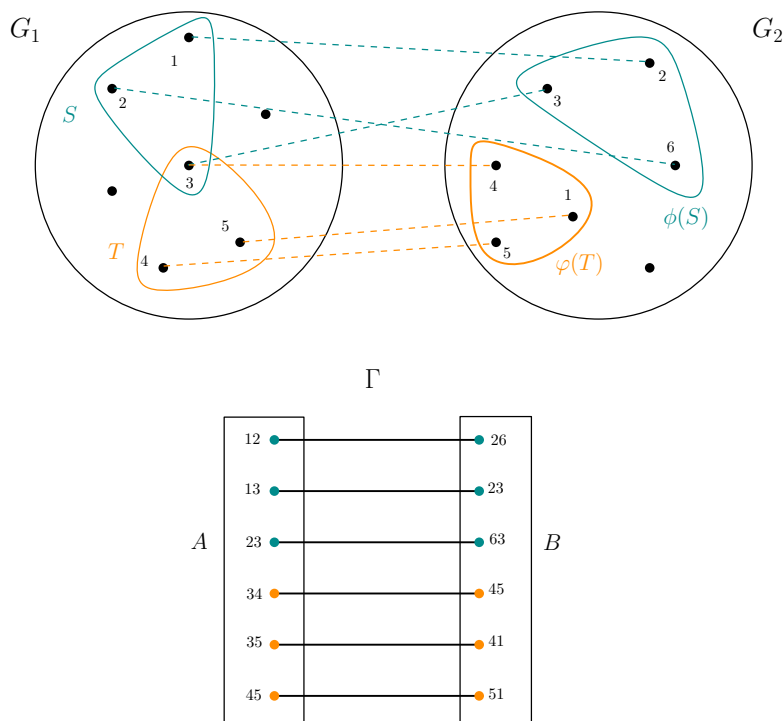


Figura 4.2: Ejemplo de Γ cuando $|S \cap T|, |\phi(S) \cap \varphi(T)| \leq 1$.

Por otro lado, fijemos ahora $j_1 = |S \cap T|$ y $j_2 = |\phi(S) \cap \varphi(T)|$ tal que $\max\{j_1, j_2\} \geq 2$. Calcularemos el valor de $\mathbb{P}[G_1[S] \cong_\phi G_2[\phi(S)] \text{ y } G_1[T] \cong_\varphi G_2[\varphi(T)]]$ dependiendo de las componentes conexas de Γ . Sea $\mathcal{C} = \mathcal{C}(\Gamma)$ el conjunto de componentes conexas de Γ . Dada una componente conexa $C \in \mathcal{C}$, necesitamos que cada par de vértices en G_1 y G_2 , representados por vértices en C , conformen todos simultáneamente una arista, en G_1 y G_2 , o que ninguno conforme una arista. Entonces, como cada par de vértices conforma una arista en G_1 (respect. G_2) de manera independiente y con probabilidad p_1 (respect. p_2), el evento donde todos los pares de vértices en G_1 y G_2 , representados en una misma componente conexa C en Γ , conforman o no una arista en G_1 y G_2 ocurre con probabilidad

$$p_1^{|C^A|} p_2^{|C^B|} + (1 - p_1)^{|C^A|} (1 - p_2)^{|C^B|},$$

donde $C^A := A \cap C$ y $C^B := B \cap C$. De lo anterior, obtenemos finalmente que

$$\mathbb{P}[G_1[S] \cong_\phi G_2[\phi(S)] \text{ y } G_1[T] \cong_\varphi G_2[\varphi(T)]] = \prod_{C \in \mathcal{C}} \left[p_1^{|C^A|} p_2^{|C^B|} + (1 - p_1)^{|C^A|} (1 - p_2)^{|C^B|} \right] \quad (4.2)$$

que era lo que necesitábamos.

4.2. Caso $G_1 \sim G(n, 1/2)$ y $G_2 \sim G(n, 1/2)$

El caso más elemental que estudiaremos será cuando ambos grafos son generados por el modelo homogéneo, ambos con intensidad de aristas igual a $1/2$. Este caso fue estudiado por Chatterjee y Diaconis [7] el 2023.

Teorema 4.2.1. *Sean $G_1, G_2 \sim G(n, 1/2)$ independientes. Luego, con alta probabilidad $L(G_1, G_2) = (4 + o(1)) \cdot \log_2 n$.*

Demostración. Tomemos $V_1 := V(G_1)$ y $V_2 := V(G_2)$. Además, para cada k entero positivo, definimos la variable aleatoria $X_k := |\mathcal{X}|$, donde

$$\mathcal{X} := \left\{ (S, \phi) : S \in V_1^{(k)}, \phi : S \rightarrow V_2 \text{ tal que } \phi \text{ es inyectiva y } G_1[S] \cong_{\phi} G_2[\phi(S)] \right\}. \quad (4.3)$$

Dados $S \in V_1^{(k)}$ y $\phi : S \rightarrow V_2$, definimos $Y_{S,\phi}$ como la variable indicatriz de si $(S, \phi) \in \mathcal{X}$. De esta manera, obtenemos la siguiente expresión para X_k

$$X_k = \sum_{S \in V_1^{(k)}} \sum_{\substack{\phi : S \rightarrow V_2 \\ \phi \text{ inyectiva}}} Y_{S,\phi}. \quad (4.4)$$

Por otro lado, para todo $S \in V_1^{(k)}$ y $\phi : S \rightarrow V_2$ inyectiva, se tiene que

$$\mathbb{E}[Y_{S,\phi}] = \mathbb{P}[Y_{S,\phi} = 1] = 2^{-\binom{k}{2}}.$$

El cálculo anterior resulta de que G_1 y G_2 son independientes, y cada par de vértices conforma una arista de manera uniforme e independiente con probabilidad $\frac{1}{2}$. Luego, el evento donde un par $a, b \in V_1$ conforma una arista en G_1 si y solo si $\phi(a)\phi(b) \in V_2$ conforma una arista en G_2 ,

ocurre con probabilidad $\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}$. Luego,

$$\begin{aligned}
 \mathbb{E}[X_k] &= \sum_{S \in V_1^{(k)}} \sum_{\substack{\phi: S \rightarrow V_2 \\ \text{inyectiva}}} \mathbb{E}[Y_{S,\phi}] \\
 &= \sum_{S \in V_1^{(k)}} \sum_{\substack{\phi: S \rightarrow V_2 \\ \text{inyectiva}}} 2^{-\binom{k}{2}} \\
 &= \binom{n}{k} \cdot \binom{n}{k} \cdot k! \cdot 2^{-\binom{k}{2}} \\
 &= \binom{n}{k}^2 \cdot k! \cdot 2^{-\binom{k}{2}}. \tag{4.5}
 \end{aligned}$$

Ahora, desarrollaremos el argumento de primer momento, ya visto anteriormente, para encontrar una cota superior de $L(G_1, G_2)$.

COTA SUPERIOR: Dado $\varepsilon > 0$, fijamos $k = (4 + \varepsilon) \cdot \log_2 n$. Usando Lema 2.1.2, se tiene la siguiente cota superior para la esperanza

$$\mathbb{E}[X_k] = \binom{n}{k}^2 \cdot k! \cdot 2^{-\binom{k}{2}} \leq \left(\frac{en}{k}\right)^{2k} \cdot k^k \cdot 2^{-\frac{k(k-1)}{2}} = \left[\frac{e^2 n^2}{k} \cdot 2^{-\frac{(k-1)}{2}}\right]^k.$$

Aquí, notamos que $2^{-\frac{(k-1)}{2}} = \sqrt{2} \cdot n^{-2-\frac{\varepsilon}{2}}$ y por ende

$$\mathbb{E}[X_k] \leq \left[\frac{e^2 \sqrt{2}}{k} \cdot n^{-\frac{\varepsilon}{2}}\right]^k = o(1).$$

Luego, por desigualdad de Markov (Lema 2.2.3), obtenemos que

$$\mathbb{P}[X_k \geq 1] \leq \mathbb{E}[X_k] = o(1).$$

Por lo tanto, con alta probabilidad no existe un subgrafo inducido en común entre G_1 y G_2 de tamaño mayor que $(4 + \varepsilon) \cdot \log_2 n$. Equivalentemente, el tamaño del subgrafo inducido en común más grande entre G_1 y G_2 es a lo más $(4 + \varepsilon) \cdot \log_2 n$.

COTA INFERIOR: Ahora, desarrollaremos el argumento de segundo momento, que fue visto y explicado también anteriormente, para encontrar una cota inferior de $L(G_1, G_2)$.

Dado $\varepsilon > 0$, fijamos $k = (4 - \varepsilon) \cdot \log_2 n$. Por Lema 2.1.2 y usando la cota $k! \geq \left(\frac{k}{e}\right)^k$ se tiene que

$$\mathbb{E}[X_k] = \binom{n}{k}^2 \cdot k! \cdot 2^{-\binom{k}{2}} \geq \left(\frac{n}{k}\right)^{2k} \cdot \left(\frac{k}{e}\right)^k \cdot 2^{-\frac{k(k-1)}{2}} = \left[\frac{n^2}{ek} \cdot 2^{-\frac{(k-1)}{2}}\right]^k. \quad (4.6)$$

Aquí, notamos que $2^{-\frac{(k-1)}{2}} = n^{-2+\frac{\varepsilon}{2}}$. Luego, para todo $\varepsilon > 0$ y $k = (4 - \varepsilon) \cdot \log_2 n$ se tiene que

$$\mathbb{E}[X_k] \geq \left[\frac{\sqrt{2}}{k} \cdot n^{\frac{\varepsilon}{2}}\right]^k \xrightarrow{n} +\infty.$$

Ahora, queremos estudiar la probabilidad de que no existan subgrafos inducidos en común entre G_1 y G_2 de tamaño a lo más $(4 - \varepsilon) \cdot \log_2 n$. Por Lema 2.4.1, tenemos que

$$\mathbb{P}[X_k = 0] \leq \frac{\text{Var}[X_k]}{\mathbb{E}[X_k]^2}. \quad (4.7)$$

Además, por Teorema 2.2.2 tenemos que

$$\text{Var}[X_k] \leq \mathbb{E}[X_k] + \sum_{\substack{(S,\phi),(T,\varphi) \\ \in \mathcal{X}}} \text{Cov}(Y_{S,\phi}, Y_{T,\varphi}). \quad (4.8)$$

De esta forma, aplicando la cota de la varianza (4.8) a la desigualdad del Segundo Momento (4.7) resulta

$$\mathbb{P}[X_k = 0] \leq \frac{1}{\mathbb{E}[X_k]} + \frac{1}{\mathbb{E}[X_k]^2} \sum_{\substack{(S,\phi),(T,\varphi) \\ \in \mathcal{X}}} \text{Cov}(Y_{S,\phi}, Y_{T,\varphi}). \quad (4.9)$$

Para acotar la suma de la covarianzas en la desigualdad (4.9), recurrimos al grafo auxiliar Γ , cuya definición recordamos a continuación:

$$V(\Gamma) := A \cup B \text{ y } E(\Gamma) := E_\phi \cup E_\varphi,$$

donde

$$A := S^{(2)} \cup T^{(2)},$$

$$\begin{aligned} B &:= \phi(S)^{(2)} \cup \varphi(T)^{(2)}, \\ E_\phi &:= \{ab, \phi(a)\phi(b) : ab \in S^{(2)}\}, \text{ y} \\ E_\varphi &:= \{xy, \varphi(x)\varphi(y) : xy \in T^{(2)}\}, \end{aligned}$$

Ahora, definimos el siguiente conjunto

$$B_{j_1, j_2} := \{((S, \phi), (T, \varphi)) \in \mathcal{X} \times \mathcal{X} : |S \cap T| = j_1 \text{ y } |\phi(S) \cap \varphi(T)| = j_2\}, \quad (4.10)$$

para todo $j_1, j_2 \in [k]_0$. Estudiaremos para qué pares $(S, \phi), (T, \varphi)$ el término de la covarianza en la desigualdad (4.9) se anula.

Afirmación 1. *Para todo $((S, \phi), (T, \varphi)) \in B_{j_1, j_2}$ tal que $\min\{j_1, j_2\} \leq 1$, se tiene que $\text{Cov}(Y_{S, \phi}, Y_{T, \varphi}) = 0$.*

Demostración. Ya vimos anteriormente en Sección 4.1 que cuando $j_1 = |S \cap T|$ y $j_2 = |\phi(S) \cap \varphi(T)| \leq 1$, los eventos $\{G_1[S] \cong_\phi G_2[\phi(S)]\}$, $\{G_1[T] \cong_\varphi G_2[\varphi(T)]\}$ son independientes. Luego, en este caso $\text{Cov}(Y_{S, \phi}, Y_{T, \varphi}) = 0$.

Sin pérdida de generalidad, supongamos ahora que $|S \cap T| \leq 1$ y $|\phi(S) \cap \varphi(T)| \geq 2$ (para ver el caso donde $|S \cap T| \geq 2$ y $|\phi(S) \cap \varphi(T)| \leq 1$ se intercambian las propiedades de A y $S^{(2)} \cup T^{(2)}$ por B y $\phi(S) \cup \varphi(T)$, respectivamente). Queremos mostrar que

$$\text{Cov}(Y_{S, \phi}, Y_{T, \varphi}) = \mathbb{E}[Y_{S, \phi} \cdot Y_{T, \varphi}] - \mathbb{E}[Y_{S, \phi}] \cdot \mathbb{E}[Y_{T, \varphi}] = 0,$$

donde $\mathbb{E}[Y_{S, \phi} \cdot Y_{T, \varphi}] = \mathbb{P}[Y_{S, \phi} = 1 \text{ y } Y_{T, \varphi} = 1]$ corresponde a la probabilidad conjunta de que $G_1[S] \cong_\phi G_2[\phi(S)]$ y $G_1[T] \cong_\varphi G_2[\varphi(T)]$. Por lo tanto, ocupando la fórmula vista para la probabilidad conjunta (4.2) obtenemos que

$$\begin{aligned} \mathbb{P}[G_1[S] \cong_\phi G_2[\phi(S)] \text{ y } G_1[T] \cong_\varphi G_2[\varphi(T)]] &= \prod_{C \in \mathcal{E}} \left[\frac{1}{2}^{|C|} + \left(1 - \frac{1}{2}\right)^{|C|} \right] \\ &= \prod_{C \in \mathcal{E}} 2 \cdot 2^{-|C|} \\ &= 2^{|\mathcal{E}|} \cdot 2^{-|A|-|B|} \end{aligned}$$

donde \mathcal{C} es el conjunto de componentes conexas de Γ . Ahora, para contar la cantidad de componentes conexas en Γ recordemos que estamos bajo la suposición de que $|S \cap T| \leq 1$ y $|\phi(S) \cap \varphi(T)| \geq 2$. Luego, los vértices en A poseen grado igual a 1 y los vértices en B poseen grado a lo más 2. Por lo tanto, hay tantas componentes conexas en Γ como elementos en B . De esta forma,

$$\mathbb{P}[G_1[S] \cong G_2[\phi(S)] \text{ y } G_1[T] \cong G_2[\varphi(T)]] = 2^{-|A|}.$$

Por otro lado, como no hay pares de vértices en común entre S y T (porque su intersección es a lo más 1), se sigue que $|S^{(2)} \cup T^{(2)}| = |S^{(2)}| + |T^{(2)}|$. Por ende,

$$\begin{aligned} \mathbb{E}[Y_{S,\phi}] \cdot \mathbb{E}[Y_{T,\varphi}] &= \mathbb{P}[Y_{S,\phi} = 1] \cdot \mathbb{P}[Y_{T,\varphi} = 1] \\ &= 2^{-|S^{(2)}| - |T^{(2)}|} \\ &= 2^{-|S^{(2)} \cup T^{(2)}|} \\ &= 2^{-|A|}. \end{aligned}$$

Y de todo lo anterior, obtenemos que

$$\text{Cov}(Y_{S,\phi}, Y_{T,\varphi}) = \mathbb{E}[Y_{S,\phi} \cdot Y_{T,\varphi}] - \mathbb{E}[Y_{S,\phi}] \cdot \mathbb{E}[Y_{T,\varphi}] = 2^{-|A|} - 2^{-|A|} = 0, \quad (4.11)$$

que era lo que queríamos ver. \square

Por otro lado, si $j_1 = |S \cap T| \geq 2$ y $j_2 = |\phi(S) \cap \varphi(T)| \geq 2$, entonces tenemos que

$$\text{Cov}(Y_{S,\phi}, Y_{T,\varphi}) \leq \mathbb{E}[Y_{S,\phi} \cdot Y_{T,\varphi}] = 2^{|\mathcal{C}|} \cdot 2^{-|A| - |B|} \leq 2^{-|A| - |B| + \min\{|A|, |B|\}}, \quad (4.12)$$

donde

$$|A| = 2 \binom{k}{2} - \binom{j_1}{2} \quad \text{y} \quad |B| = 2 \binom{k}{2} - \binom{j_2}{2}.$$

Ahora, volviendo a la desigualdad (4.9) obtenemos que

$$\mathbb{P}[X_k = 0] \leq \frac{1}{\mathbb{E}[X_k]} + \frac{1}{\mathbb{E}[X_k]^2} \sum_{\substack{(S,\phi),(T,\varphi) \\ \in \mathcal{X} \times \mathcal{X}}} \text{Cov}(Y_{S,\phi}, Y_{T,\varphi}) \quad (4.13)$$

$$\begin{aligned}
 & \stackrel{(4.11)}{=} \frac{1}{\mathbb{E}[X_k]} + \frac{1}{\mathbb{E}[X_k]^2} \sum_{2 \leq j_1, j_2 \leq k} \sum_{\substack{(S, \phi), (T, \varphi) \\ \in B_{j_1, j_2}}} \text{Cov}(Y_{S, \phi}, Y_{T, \varphi}) \\
 & \stackrel{(4.12)}{\leq} \frac{1}{\mathbb{E}[X_k]} + \frac{1}{\mathbb{E}[X_k]^2} \sum_{2 \leq j_1, j_2 \leq k} \sum_{\in B_{j_1, j_2}} 2^{-|A| - |B| + \min\{|A|, |B|\}} \\
 & = \frac{1}{\mathbb{E}[X_k]} + \frac{1}{\mathbb{E}[X_k]^2} \sum_{2 \leq j_1, j_2 \leq k} \sum_{\in B_{j_1, j_2}} 2^{-2\binom{k}{2} + \min\{\binom{j_1}{2}, \binom{j_2}{2}\}} \\
 & = \frac{1}{\mathbb{E}[X_k]} + \underbrace{\sum_{2 \leq j_1, j_2 \leq k} |B_{j_1, j_2}| \frac{2^{\min\{\binom{j_1}{2}, \binom{j_2}{2}\}}}{\binom{n}{k}^4 \cdot k!^2}}_{(\spadesuit)}.
 \end{aligned}$$

Como $\mathbb{E}[X_k]^{-1} = o(1)$, solo nos resta mostrar que el término (\spadesuit) también converge a 0 con n . Para conseguir esto, mostraremos a continuación una cota superior para $\frac{|B_{j_1, j_2}|}{\binom{n}{k}^4 \cdot k!^2}$.

Afirmación 2. *Para todo $j_1, j_2 > 0$ se tiene que*

$$\frac{|B_{j_1, j_2}|}{\binom{n}{k}^4 \cdot k!^2} \leq 4\pi^2 k^2 \left(\frac{ek}{n}\right)^{j_1 + j_2} \left(\frac{k}{j_1}\right)^{j_1} \left(\frac{k}{j_2}\right)^{j_2}. \quad (4.14)$$

Demostración. Primero, observemos que

$$|B_{j_1, j_2}| = (k!)^2 \binom{n}{j_1} \binom{n - j_1}{k - j_1} \binom{n - k}{k - j_1} \binom{n}{j_2} \binom{n - j_2}{k - j_2} \binom{n - k}{k - j_2}.$$

En efecto, en B_{j_1, j_2} viven los pares $(S, \phi), (T, \varphi)$ tales que $|S \cap T| = j_1$ y $|\phi(S) \cap \varphi(T)| = j_2$. Entonces, podemos contar las maneras de elegir a S y a T tales que $\binom{n}{j_1}$ corresponde a las formas de elegir los elementos de $S \cap T$, $\binom{n - j_1}{k - j_1}$ es la cantidad de elementos en $S \setminus T$ y $\binom{n - k}{k - j_1}$ es la cantidad de elementos en $T \setminus S$. Luego, repetimos el conteo para los elementos en $\phi(S) \cup \varphi(T)$, ponderando por todas las $k!^2$ posibles permutaciones de ϕ y φ . Así,

$$\begin{aligned}
 \frac{|B_{j_1, j_2}|}{\binom{n}{k}^4 \cdot k!^2} &= \frac{k!^2 \binom{n}{j_1} \binom{n - j_1}{k - j_1} \binom{n - k}{k - j_1} \binom{n}{j_2} \binom{n - j_2}{k - j_2} \binom{n - k}{k - j_2}}{\binom{n}{k}^4 \cdot k!^2} \\
 & \stackrel{\text{por Lema 2.1.3 en el denominador}}{=} \frac{\binom{n}{j_1} \binom{n - j_1}{k - j_1} \binom{n - k}{k - j_1} \binom{n}{j_2} \binom{n - j_2}{k - j_2} \binom{n - k}{k - j_2}}{\left[\frac{(1 + o(1))}{\sqrt{2\pi k}} \cdot \left(\frac{en}{k}\right)^k\right]^4}
 \end{aligned}$$

$$\begin{aligned}
 &\leq 4\pi^2 k^2 \cdot \frac{\binom{n}{j_1} \binom{n}{k-j_1}^2 \binom{n}{j_2} \binom{n}{k-j_2}^2}{\left(\frac{en}{k}\right)^{4k}} \\
 \text{por Lema 2.1.2} &\leq 4\pi^2 k^2 \cdot \frac{\left(\frac{en}{j_1}\right)^{j_1} \left(\frac{en}{k-j_1}\right)^{2(k-j_1)} \left(\frac{en}{j_2}\right)^{j_2} \left(\frac{en}{k-j_2}\right)^{2(k-j_2)}}{\left(\frac{en}{k}\right)^{4k}} \\
 &= 4\pi^2 k^2 \left(\frac{k}{en}\right)^{j_1+j_2} \left(\frac{k}{k-j_1}\right)^{2(k-j_1)} \left(\frac{k}{k-j_2}\right)^{2(k-j_2)} \left(\frac{k}{j_1}\right)^{j_1} \left(\frac{k}{j_2}\right)^{j_2}
 \end{aligned}$$

Notamos que, para $i \in \{1, 2\}$ y usando que para todo real x se tiene que $1+x \leq e^x$, obtenemos que

$$\frac{k}{k-j_i} = 1 + \frac{j_i}{k-j_i} \leq e^{\frac{j_i}{k-j_i}}$$

entonces,

$$\left(\frac{k}{k-j_i}\right)^{2(k-j_i)} \leq e^{2(k-j_i) \cdot \frac{j_i}{k-j_i}} = e^{2j_i} \quad (4.15)$$

y así tenemos que

$$\frac{|B_{j_1, j_2}|}{\binom{n}{k}^4 \cdot k!^2} \leq 4\pi^2 k^2 \cdot \left(\frac{ek}{n}\right)^{j_1+j_2} \cdot \left(\frac{k}{j_1}\right)^{j_1} \cdot \left(\frac{k}{j_2}\right)^{j_2}.$$

□

Ahora, tomando n suficientemente grande y ocupando la Afirmación 2 obtenemos

$$\begin{aligned}
 (\spadesuit) &= \sum_{2 \leq j_1, j_2 \leq k} |B_{j_1, j_2}| \frac{2^{\min\{\binom{j_1}{2}, \binom{j_2}{2}\}}}{\binom{n}{k}^4 \cdot k!^2} \leq 4\pi^2 k^2 \sum_{2 \leq j_1, j_2 \leq k} 2^{\min\{\binom{j_1}{2}, \binom{j_2}{2}\}} \cdot \left(\frac{ek}{n}\right)^{j_1+j_2} \cdot \left(\frac{k}{j_1}\right)^{j_1} \cdot \left(\frac{k}{j_2}\right)^{j_2} \\
 &\quad \text{por simetría entre } G_1 \text{ y } G_2 = 8\pi^2 k^2 \sum_{2 \leq j_1 \leq j_2 \leq k} 2^{\binom{j_1}{2}} \cdot \left(\frac{ek}{n}\right)^{j_1+j_2} \cdot \left(\frac{k}{j_1}\right)^{j_1} \cdot \left(\frac{k}{j_2}\right)^{j_2} \\
 &= 8\pi^2 k^2 \sum_{2 \leq j_1 \leq k} 2^{\binom{j_1}{2}} \cdot \left(\frac{ek^2}{nj_1}\right)^{j_1} \cdot \sum_{j_1 \leq j_2 \leq k} \left(\frac{ek^2}{nj_2}\right)^{j_2} \\
 \text{Lema 2.1.1} &\leq 8\pi^2 k^2 \sum_{2 \leq j_1 \leq k} 2^{\binom{j_1}{2}} \cdot \left(\frac{ek^2}{nj_1}\right)^{j_1}
 \end{aligned}$$

$$\begin{aligned}
 &\leq 8\pi^2 k^2 \sum_{2 \leq j_1 \leq k} \left(\frac{2^{\frac{k}{2}} e^2 k^4}{n^2 j_1^2} \right)^{j_1} \\
 &\leq 8\pi^2 k^2 \sum_{2 \leq j_1 \leq k} \left(\frac{e^2 k^4}{4n^{\frac{\varepsilon}{2}}} \right)^{j_1} \\
 \boxed{\text{Lema 2.1.1}} &\leq 16\pi^2 k^2 \left(\frac{e^2 k^4}{4n^{\frac{\varepsilon}{2}}} \right)^2 \\
 &= O(k^{10} \cdot n^{-\varepsilon}) \\
 \boxed{\text{ya que } k = O(\log n)} &= o(1).
 \end{aligned}$$

Luego, con alta probabilidad existe un subgrafo inducido en común entre G_1 y G_2 cuyo tamaño es a lo más $(4 - \varepsilon) \cdot \log_2 n$, lo que concluye la demostración. \square

4.3. Caso $G_1, G_2 \sim G(n, p)$

A continuación, estudiaremos el típico valor de $L(G_1, G_2)$ cuando $G_1, G_2 \sim G(n, p)$, para $p \in (0, 1)$ fijo. Este es un caso particular del resultado visto por Surya, Warnke y Zhu [30], y Diamantidis, Konstantopoulus y Yuan [13]. El desarrollo de la demostración que veremos a continuación nos permite ir entendiendo en detalle cómo se va generalizando el resultado a medida que estudiamos modelos de grafos aleatorios con intensidades de aristas diferentes.

Teorema 4.3.1. *Sean $p \in (0, 1)$ y $G_1, G_2 \sim G(n, p)$ independientes. Luego, con alta probabilidad $L(G_1, G_2) = (4 + o(1)) \cdot \log_{1/q} n$, donde $q := p^2 + (1 - p)^2$.*

Demostración. Sean $V_1 := V(G_1)$ y $V_2 := V(G_2)$. Para cada k entero positivo, definimos la variable aleatoria $X_k := |\mathcal{X}|$, donde \mathcal{X} se define como en (4.3), y dados $S \in V_1^{(k)}$ y $\phi : S \rightarrow V_2$, definimos $Y_{S, \phi}$ como la variable indicatriz de si $(S, \phi) \in \mathcal{X}$. De esta manera, se obtiene

$$X_k = \sum_{S \in V_1^{(k)}} \sum_{\substack{\phi: S \rightarrow V_2 \\ \phi \text{ inyectiva}}} Y_{S, \phi}.$$

Por otro lado, para todo $S \in V_1^{(k)}$ y $\phi : S \rightarrow V_2$ inyectiva, se tiene que

$$\mathbb{E}[Y_{S,\phi}] = \mathbb{P}[Y_{S,\phi} = 1] = [p^2 + (1-p)^2]^{\binom{k}{2}} = q^{\binom{k}{2}}.$$

El cálculo anterior resulta de que G_1 y G_2 son independientes y cada par de vértices conforma una arista de manera uniforme e independiente con probabilidad p . Luego, el evento donde un par $a, b \in V_1$ conforma una arista en G_1 si y solo si $\phi(a), \phi(b)$ conforma una arista en G_2 , ocurre con probabilidad $p^2 + (1-p)^2 = q$. Luego,

$$\begin{aligned} \mathbb{E}[X_k] &= \sum_{S \in V_1^{(k)}} \sum_{\substack{\phi: S \rightarrow V_2 \\ \text{inyectiva}}} \mathbb{E}[Y_{S,\phi}] \\ &= \sum_{S \in V_1^{(k)}} \sum_{\substack{\phi: S \rightarrow V_2 \\ \text{inyectiva}}} q^{\binom{k}{2}} \\ &= \binom{n}{k} \cdot \frac{n!}{(n-k)!} \cdot q^{\binom{k}{2}} \\ &= \binom{n}{k}^2 \cdot k! \cdot q^{\binom{k}{2}}. \end{aligned}$$

Desarrollaremos el argumento de primer momento, ya visto anteriormente, para estudiar la probabilidad de que X_k sea al menos 1.

COTA SUPERIOR: Dado $\varepsilon > 0$, fijaremos $k = (4 + \varepsilon) \cdot \log_b n$, donde $b := q^{-1}$. Usando Lema 2.1.2 se tiene la siguiente cota superior para la esperanza

$$\mathbb{E}[X_k] = \binom{n}{k}^2 \cdot k! \cdot q^{\binom{k}{2}} \leq \left(\frac{en}{k}\right)^{2k} \cdot k^k \cdot q^{\frac{k(k-1)}{2}} = \left[\frac{e^2 n^2}{k} \cdot q^{\frac{(k-1)}{2}}\right]^k.$$

Y aquí, notamos que $q^{\frac{(k-1)}{2}} = \sqrt{b} \cdot n^{-2-\frac{\varepsilon}{2}}$. Luego, para todo $\varepsilon > 0$ y $k > (4 + \varepsilon) \cdot \log_b n$, se tiene que

$$\mathbb{E}[X_k] \leq \left[\frac{e^2}{k} \cdot n^{-\frac{\varepsilon}{2}}\right]^k = o(1).$$

Así, por desigualdad de Markov (Lema 2.2.3)

$$\mathbb{P}[X_k \geq 1] \leq \mathbb{E}[X_k] = o(1).$$

Por ende, con alta probabilidad no existe un subgrafo inducido en común entre G_1 y G_2 de tamaño mayor a $(4 + \varepsilon) \cdot \log_b n$.

COTA INFERIOR: Ahora, veremos el argumento de Segundo Momento. Dado $\varepsilon > 0$, fijaremos $k = (4 - \varepsilon) \cdot \log_b n$. Usando Lema 2.1.2 se tiene la siguiente cota inferior para la esperanza,

$$\mathbb{E}[X_k] = \binom{n}{k}^2 \cdot k! \cdot q^{\binom{k}{2}} \geq \left(\frac{n}{k}\right)^{2k} \cdot \left(\frac{k}{e}\right)^k \cdot q^{\frac{k(k-1)}{2}} = \left[\frac{n^2}{ek} \cdot q^{\frac{(k-1)}{2}}\right]^k.$$

Aquí, notamos que $q^{\frac{(k-1)}{2}} = \sqrt{bn}^{-2+\frac{\varepsilon}{2}}$. Luego, para todo $\varepsilon > 0$ y $k = (4 - \varepsilon) \cdot \log_b n$ se tiene que

$$\mathbb{E}[X_k] \geq \left[\frac{\sqrt{b}}{ek} \cdot n^{-\frac{\varepsilon}{2}}\right]^k \xrightarrow{n} +\infty.$$

Ahora, queremos estudiar la probabilidad de que no existan subgrafos inducidos en común entre G_1 y G_2 de tamaño $(4 - \varepsilon) \cdot \log_2 n$. Por Lema 2.4.1, tenemos que

$$\mathbb{P}[X_k = 0] \leq \frac{\text{Var}[X_k]}{\mathbb{E}[X_k]^2}. \quad (4.16)$$

Por Teorema 2.2.2 y por cálculos análogos a los vistos en la desigualdad (4.9) tenemos que

$$\mathbb{P}[X_k = 0] \leq \frac{1}{\mathbb{E}[X_k]} + \underbrace{\frac{1}{\mathbb{E}[X_k]^2} \sum_{\substack{(S,\phi),(T,\varphi) \\ \in \mathcal{X}}} \text{Cov}(Y_{S,\phi}, Y_{T,\varphi})}_{(\star)}. \quad (4.17)$$

Como $\mathbb{E}[X_k]^{-1} = o(1)$, basta con acotar el término (\star) en la desigualdad (4.17). Para esto, recurriremos a nuestro grafo bipartito auxiliar Γ . Recordemos que Γ se define como sigue

$$V(H) := A \cup B, \text{ y}$$

$$E(H) := \{ab, \phi(a)\phi(b) : ab \in S^{(2)}\} \cup \{xy, \varphi(x)\varphi(y) : xy \in T^{(2)}\},$$

donde $A = S^{(2)} \cup T^{(2)}$ y $B = \phi(S)^{(2)} \cup \varphi(T)^{(2)}$. Dados $j_1, j_2 \in [k]_0$. Definimos también el conjunto B_{j_1, j_2} igual que en la ecuación (4.10). Aquí, sabemos que la covarianza en (\star) se anula en B_{j_1, j_2} cuando $j_1, j_2 \leq 1$, es decir, cuando el tamaño de la intersección entre S y T , y entre $\phi(S)$ y

$\varphi(T)$ son a lo más 1. Luego,

$$(\star) = \frac{1}{\mathbb{E}[X_k]^2} \sum_{\substack{0 \leq j_1, j_2 \leq k \\ \max\{j_1, j_2\} \geq 2}} \sum_{\substack{(S, \phi), (T, \varphi) \\ \in B_{j_1, j_2}}} \text{Cov}(Y_{S, \phi}, Y_{T, \varphi}).$$

Así, solo nos resta ver que la expresión en (\star) tiende a 0 cuando n tiende a infinito.

Afirmación 3.

$$\text{Cov}(Y_{S, \phi}, Y_{T, \varphi}) \leq q^{2\binom{k}{2} - \frac{1}{2}[(\binom{j_1}{2}) + (\binom{j_2}{2})]}.$$

Demostración. Recordemos que $Y_{S, \phi}$ es la variable indicatriz de si $(S, \phi) \in \mathcal{X}$, o equivalentemente, es la variable indicatriz de si $G_1[S] \cong_{\phi} G_2[\phi(S)]$. Es claro que

$$\text{Cov}(Y_{S, \phi}, Y_{T, \varphi}) \leq \mathbb{E}[Y_{S, \phi} \cdot Y_{T, \varphi}].$$

Luego, nos interesa calcular la probabilidad de que $G_1[S] \cong_{\phi} G_2[\phi(S)]$ y $G_1[T] \cong_{\varphi} G_2[\varphi(T)]$. Para esto, definimos \mathcal{C} como el conjunto de componentes conexas de Γ y podemos ocupar la fórmula vista para la probabilidad conjunta (4.2). Así, obtenemos que

$$\mathbb{E}[Y_{S, \phi} \cdot Y_{T, \varphi}] = \mathbb{P}[G_1[S] \cong_{\phi} G_2[\phi(S)] \text{ y } G_1[T] \cong_{\varphi} G_2[\varphi(T)]] = \prod_{C \in \mathcal{C}} [p^{|C|} + (1-p)^{|C|}].$$

Luego, usando que para todo $x, y > 0$ y $a \geq 1$ se cumple que $(x + y)^a \geq x^a + y^a$, se sigue que

$$\begin{aligned} \text{Cov}(Y_{S, \phi}, Y_{T, \varphi}) &\leq \prod_{C \in \mathcal{C}} [p^{|C|} + (1-p)^{|C|}] \\ &= \prod_{C \in \mathcal{C}} \left[(p^2)^{\frac{|C|}{2}} + ((1-p)^2)^{\frac{|C|}{2}} \right] \\ &\leq \prod_{C \in \mathcal{C}} [p^2 + (1-p)^2]^{\frac{|C|}{2}} \\ &= \prod_{C \in \mathcal{C}} q^{\frac{|C|}{2}} \\ &= q^{\frac{|A|+|B|}{2}} \\ &= q^{2\binom{k}{2} - \frac{1}{2}[(\binom{j_1}{2}) + (\binom{j_2}{2})]}. \end{aligned}$$

□

Así, para n suficientemente grande se sigue que

$$\begin{aligned}
 \sum_{\substack{0 \leq j_1 \leq j_2 \leq k \\ \max\{j_1, 2\} \leq j_2}} \sum_{\substack{(S, \phi), (T, \varphi) \\ \in B_{j_1, j_2}}} \frac{\text{Cov}(Y_{S, \phi}, Y_{T, \varphi})}{\mathbb{E}[X_k]^2} &\leq \sum_{\substack{0 \leq j_1, j_2 \leq k \\ \max\{j_1, j_2\} \geq 2}} \frac{|B_{j_1, j_2}| \cdot q^{2\binom{k}{2} - \frac{1}{2}[(\binom{j_1}{2}) + (\binom{j_2}{2})]}}{\binom{n}{k}^4 \cdot (k!)^2 \cdot q^{2\binom{k}{2}}} \\
 &= 2 \sum_{\substack{0 \leq j_1 \leq j_2 \leq k \\ \max\{j_1, 2\} \leq j_2}} \frac{|B_{j_1, j_2}| \cdot b^{\frac{1}{2}[(\binom{j_1}{2}) + (\binom{j_2}{2})]}}{\binom{n}{k}^4 \cdot k!^2} \\
 &= 2 \sum_{\substack{0 \leq j_1 \leq j_2 \leq k \\ \max\{j_1, 2\} \leq j_2}} f_{j_1, j_2}
 \end{aligned}$$

donde en la ecuación anterior definimos el término

$$f_{j_1, j_2} := \frac{|B_{j_1, j_2}| \cdot b^{\frac{1}{2}[(\binom{j_1}{2}) + (\binom{j_2}{2})]}}{\binom{n}{k}^4 \cdot k!^2}.$$

Restaría ver que $\sum_{j_1, j_2} f_{j_1, j_2}$ tiende a 0 cuando n tiende a infinito. Para ello, veremos por separado los casos cuando

(i) $j_1 = 0$ y $2 \leq j_2 \leq k$, y

(ii) $1 \leq j_1 \leq k$ y $\max\{2, j_1\} \leq j_2 \leq k$.

Para ver el caso (i), fijamos $j_1 = 0$. Así, tenemos que

$$\begin{aligned}
 \sum_{2 \leq j_2 \leq k} f_{0, j_2} &= \sum_{2 \leq j_2 \leq k} \frac{|B_{0, j_2}| \cdot b^{\frac{1}{2}\binom{j_2}{2}}}{\binom{n}{k}^4 \cdot k!^2} \\
 &= \sum_{2 \leq j_2 \leq k} \frac{b^{\frac{1}{2}\binom{j_2}{2}} \binom{n}{k} \binom{n-k}{k} \binom{n}{j_2} \binom{n-j_2}{k-j_2} \binom{n-k}{k-j_2}}{\binom{n}{k}^4} \\
 &\leq \sum_{2 \leq j_2 \leq k} \frac{b^{\frac{1}{2}\binom{j_2}{2}} \binom{n}{j_2} \binom{n}{k-j_2}^2}{\binom{n}{k}^2} \\
 &\stackrel{\text{Lema 2.1.2}}{\leq} \sum_{2 \leq j_2 \leq k} \frac{b^{\frac{1}{2}\binom{j_2}{2}} \left(\frac{en}{j_2}\right)^{j_2} \left(\frac{en}{k-j_2}\right)^{2(k-j_2)}}{\binom{n}{k}^2}
 \end{aligned}$$

$$\begin{aligned}
 \boxed{\text{aproximación de Stirling}} &\leq \sum_{2 \leq j_2 \leq k} \frac{b^{\frac{1}{2} \binom{j_2}{2}} \left(\frac{en}{j_2}\right)^{j_2} \left(\frac{en}{k-j_2}\right)^{2(k-j_2)}}{\left[\frac{1+o(1)}{\sqrt{2\pi k}} \left(\frac{en}{k}\right)^k\right]^2} \\
 &\leq 4\pi^2 k^2 \sum_{2 \leq j_2 \leq k} \frac{b^{\frac{1}{2} \binom{j_2}{2}} \left(\frac{en}{j_2}\right)^{j_2} \left(\frac{en}{k-j_2}\right)^{2(k-j_2)}}{\left(\frac{en}{k}\right)^{2k}} \\
 &= 4\pi^2 k^2 \sum_{2 \leq j_2 \leq k} b^{\frac{1}{2} \binom{j_2}{2}} \left(\frac{k^2}{j_2 en}\right)^{j_2} \left(\frac{k}{k-j_2}\right)^{2(k-j_2)} \\
 &\stackrel{(4.15)}{\leq} 4\pi^2 k^2 \sum_{2 \leq j_2 \leq k} b^{\frac{j_2^2}{4}} \left(\frac{k^2}{j_2 en}\right)^{j_2} e^{2j_2} \\
 &\leq 4\pi^2 k^2 \sum_{2 \leq j_2 \leq k} \left(\frac{b^{\frac{j_2}{4}} k^2 e}{j_2 n}\right)^{j_2} \\
 \boxed{b^{\frac{j_2}{4}} \leq n^{1-\frac{\epsilon}{4}}} &\leq 4\pi^2 k^2 \sum_{2 \leq j_2 \leq k} \left(\frac{k^2 en^{1-\frac{\epsilon}{4}}}{j_2 n}\right)^{j_2} \\
 \boxed{\text{Lema 2.1.1}} &\leq 4\pi^2 k^2 ne^2 \left(\frac{k^2 e}{2n^{\frac{\epsilon}{4}}}\right)^2 \\
 &= O(k^6 n^{-\frac{\epsilon}{2}}) \\
 &= o(1).
 \end{aligned}$$

Finalmente para el caso (ii), tenemos que por Afirmación 2

$$\begin{aligned}
 \sum_{\substack{1 \leq j_1 \leq j_2 \leq k \\ \max\{2, j_1\} \leq j_2}} f_{j_1, j_2} &\leq 8\pi^2 k^2 \sum_{\substack{1 \leq j_1 \leq j_2 \leq k \\ \max\{j_1, 2\} \leq j_2}} b^{\frac{1}{2}[\binom{j_1}{2} + \binom{j_2}{2}]} \cdot \left(\frac{ek}{n}\right)^{j_1+j_2} \cdot \left(\frac{k}{j_1}\right)^{j_1} \cdot \left(\frac{k}{j_2}\right)^{j_2} \\
 \boxed{\text{como } b > 1} &\leq 8\pi^2 k^2 \sum_{\substack{1 \leq j_1 \leq j_2 \leq k \\ \max\{j_1, 2\} \leq j_2}} b^{\frac{j_1^2}{4} + \frac{j_2^2}{4}} \cdot \left(\frac{ek}{n}\right)^{j_1+j_2} \cdot \left(\frac{k}{j_1}\right)^{j_1} \cdot \left(\frac{k}{j_2}\right)^{j_2} \\
 &= 8\pi^2 k^2 \sum_{1 \leq j_1 \leq k} \left(\frac{b^{\frac{j_1}{4}} ek^2}{nj_1}\right)^{j_1} \sum_{\max\{j_1, 2\} \leq j_2 \leq k-1} \left(\frac{b^{\frac{j_1}{4}} ek^2}{nj_2}\right)^{j_2} \\
 \boxed{\text{por Lema 2.1.1}} &\leq 16\pi^2 k^2 \sum_{1 \leq j_1 \leq k} \left(\frac{b^{\frac{j_1}{4}} ek^2}{nj_1}\right)^{j_1} \cdot \left(\frac{b^{\frac{j_1}{4}} ek^2}{nj_1}\right)^{\max\{j_1, 2\}}
 \end{aligned}$$

$$\begin{aligned}
 &\leq 16\pi^2 k^2 \sum_{1 \leq j_1 \leq k} \left(\frac{b^{\frac{k}{2}} e^2 k^4}{n^2 j_1^2} \right)^{j_1} \\
 \boxed{\text{Lema 2.1.1}} &\leq 32\pi^2 k^2 \left(\frac{e^2 k^4}{4n^{\frac{\varepsilon}{2}}} \right) \\
 &= O(k^{10} \cdot n^{-\varepsilon}).
 \end{aligned}$$

Luego, con alta probabilidad existe un subgrafo inducido en común entre G_1 y G_2 cuyo tamaño es menor a $(4 - \varepsilon) \cdot \log_b n$, lo que concluye la demostración. \square

4.4. Caso $G_1 \sim G(n, p_1)$ y $G_2 \sim G(n, p_2)$

El estudio del tamaño del subgrafo inducido en común más grande entre dos grafos aleatorios homogéneos con intensidades de aristas no necesariamente iguales, pero constantes fijas, fue abordado por dos grupos: Surya, Warnke y Zhu [30], y Diamantidis, Konstantopoulus y Yuan [13]. Ambos grupos tienen demostraciones diferentes para abordar el problema. En esta sección, compararemos brevemente ambos resultados, las distintas herramientas que implementan para probar sus resultados y cómo varía el valor del tamaño del subgrafo inducido en común más grande dependiendo de la región en $[0, 1]^2$ donde se escojan las intensidades de aristas.

El primer estudio fue realizado por Surya, Warnke y Zhu el 2023 [30]. Mostraron que para todo $p_1, p_2 \in (0, 1)$, dados dos grafos independientes $G_1 \sim G(n, p_1)$ y $G_2 \sim G(n, p_2)$, con alta probabilidad el tamaño del subgrafo inducido en común más grande entre G_1 y G_2 se concentra alrededor de

$$L_n = L_n(p_1, p_2) := \frac{4 \log n}{\min_{p \in [0, 1]} \max\{\log b_0(p), 2 \log b_1(p), 2 \log b_2(p)\}} \quad (4.18)$$

donde, para $i \in \{1, 2\}$

$$b_0 = b_0(p) := \left(\frac{p}{p_1 p_2} \right)^p \left(\frac{1-p}{(1-p_1)(1-p_2)} \right)^{1-p} \quad \text{y} \quad b_i = b_i(p) := \left(\frac{p}{p_i} \right)^p \left(\frac{1-p}{1-p_i} \right)^{1-p}.$$

Cabe destacar que el resultado de Surya, Warnke y Zhu es más fuerte, ya que probaron una

concentración en dos puntos para el valor de $L(G_1, G_2)$, sin embargo para efectos de este trabajo no entraremos en esos detalles para no salirnos de foco. A partir del término L_n en la fórmula (4.18), podemos distinguir distintas regiones para las cuales la fórmula de L_n cambia dependiendo del valor de p_1 y p_2 . Para esto, primero vamos a definir un parámetro que depende de p_1 y p_2

$$\hat{p} = \hat{p}(p_1, p_2) := \frac{p_1 p_2}{p_1 p_2 + (1 - p_1)(1 - p_2)}.$$

CASO 1: $p_1, p_2 \in (0, 1)$ satisfacen que

$$\log b_0(\hat{p}) > \max\{2 \log b_1(\hat{p}), 2 \log b_2(\hat{p})\}. \quad (4.19)$$

Entonces, para $G_1 \sim G(n, p_1)$ y $G_2 \sim G(n, p_2)$ independientes, se tiene que

$$L(G_1, G_2) = (1 + o(1)) \cdot 4 \log_{b_0(\hat{p})} n. \quad (4.20)$$

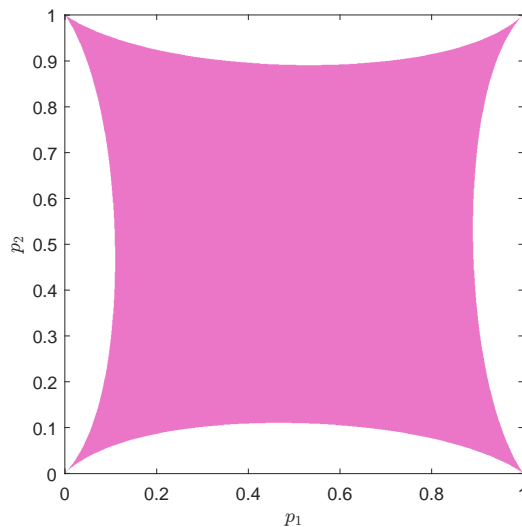


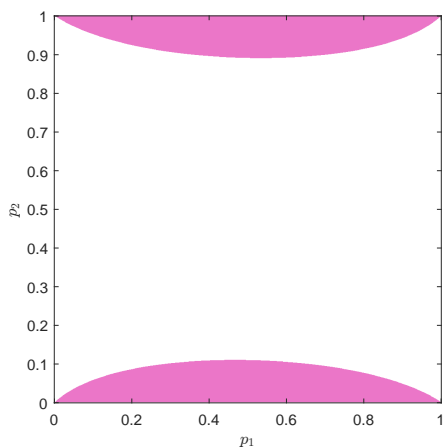
Figura 4.3: Región de los valores $p_1, p_2 \in [0, 1]^2$ que satisfacen la desigualdad (4.19).

CASO 2: $p_1, p_2 \in (0, 1)$ son tales que

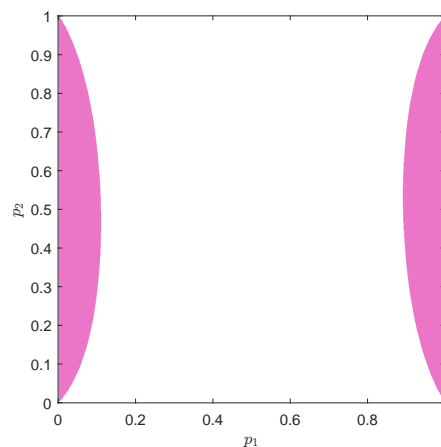
$$\log b_0(\hat{p}) \leq 2 \log b_i(\hat{p}), \quad (4.21)$$

para $i \in \{1, 2\}$. Sea $p_0 \in (p_i, \hat{p})$ definido como la única solución de la ecuación $\log b_0(p) = 2 \log b_i(p)$. Entonces, para n suficientemente grande, y dados $G_1 \sim G(n, p_1)$ y $G_2 \sim G(n, p_2)$ independientes se tiene que

$$L_n = (1 + o(1)) \cdot 2 \log_{b_i(p_0)} n.$$



(a) Región de los valores $p_1, p_2 \in (0, 1)$ que satisfacen la restricción (4.21) para $i = 1$.



(b) Región de los valores $p_1, p_2 \in (0, 1)$ que satisfacen la restricción (4.21) para $i = 2$.

En este trabajo, Surya, Warnke y Zhu realizan una estrategia de primer y segundo momento a partir de las siguientes variables aleatorias

$$X_{n,m} := \sum_{H \in \mathcal{G}_{n,m}} X_H^{(1)} X_H^{(2)},$$

donde $\mathcal{G}_{n,m}$ representa el conjunto de los grafos no etiquetados con n vértices y m aristas, y $X_H^{(i)}$ cuenta la cantidad de copias que induce H en el grafo $G(n, p_i)$, para $i \in \{1, 2\}$.

Por otra parte, en el estudio realizado por Diamantidis, Konstantopoulus y Yuan el 2024 [13], determinan también una concentración en dos puntos para $L(G_1, G_2)$, donde $G_1 \sim G(n, p_1)$ y $G_2 \sim G(n, p_2)$ son independientes. Para esto, definen el parámetro $\tau = \tau(p_1, p_2) := p_1 p_2 + (1 - p_1)(1 - p_2)$, y toman $p_1, p_2 \in (0, 1)$ tales que

$$\max\{p_1^2 p_2 + (1 - p_1)^2 (1 - p_2), p_1 p_2^2 + (1 - p_1)(1 - p_2)^2\} < \tau^{3/2}. \quad (4.22)$$

En este caso, mostraron que

$$L(G_1, G_2) = (1 + o(1)) \cdot 4 \log_{1/\tau} n. \quad (4.23)$$

Nuevamente, para simplificar el análisis, no detallaremos la expresión completa de la concentración en dos puntos que ellos determinan en [13].

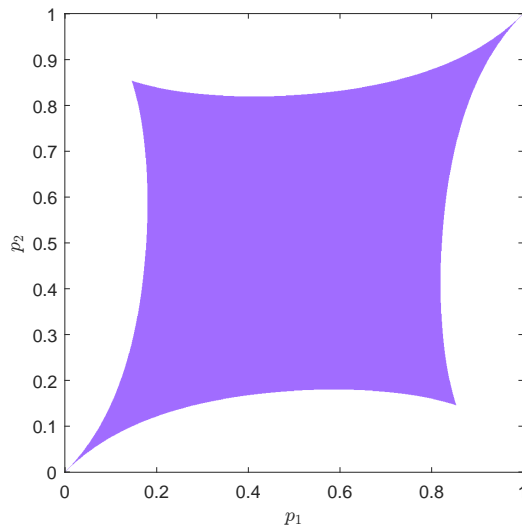


Figura 4.5: Región de los valores $p_1, p_2 \in (0, 1)$ que satisfacen la restricción (4.22).

Diamantidis, Konstantopoulos y Yuan también realizan una estrategia de primer y segundo momento. Para ello, definen la variable aleatoria N que cuenta la cantidad de m -isomorfismos entre G_1 y G_2 , donde un m -isomorfismo es una biyección entre algún par de subgrafos inducidos $H_1 \subseteq G_1$ y $H_2 \subseteq G_2$, ambos de tamaño m .

Para comparar ambos resultados, notemos que

$$\begin{aligned} b_0(\hat{p}) &= \left(\frac{\hat{p}}{p_1 p_2} \right)^{\hat{p}} \left(\frac{1 - \hat{p}}{(1 - p_1)(1 - p_2)} \right)^{1 - \hat{p}} \\ &= \left(\frac{1}{p_1 p_2 + (1 - p_1)(1 - p_2)} \right)^{\hat{p}} \left(\frac{1}{p_1 p_2 + (1 - p_1)(1 - p_2)} \right)^{1 - \hat{p}} \\ &= \frac{1}{\tau}. \end{aligned}$$

De esta forma, vemos que las ecuaciones (4.20) y (4.23) coinciden.

Para comparar esta similitud entre los resultados, en la siguiente gráfica tenemos las regiones sobre las cuales cada grupo define este valor para el tamaño del subgrafo inducido en común más grande. Desde la Figura 4.6, observamos que la región dada por la condición vista por Surya, Warnke y Zhu (4.20) contiene a la región dada por la condición de Diamantidis, Konstantopoulus y Yuan (4.22).

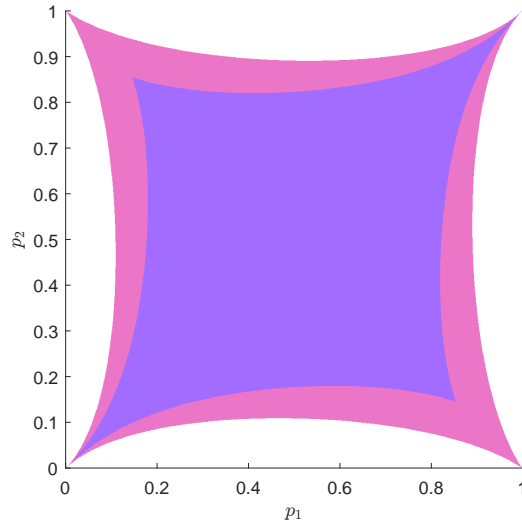


Figura 4.6: Gráfica de ambas regiones definidas por las restricciones (4.19) (en rosado) y (4.22) (en morado).

Tamaño del subgrafo inducido en común entre grafos aleatorios no homogéneos

5.1. Caso $G_1 \sim G(n, 1/2)$ y $G_2 \sim G(n, W)$

En esta sección estudiamos el tamaño típico del subgrafo inducido en común entre grafos aleatorios no necesariamente homogéneos. Recordemos que los resultados para determinar el valor de $L(G_1, G_2)$ para G_1 y G_2 grafos aleatorios homogéneos han sido (curiosamente) bastante recientes y para el caso en el que los grafos son generados por el modelo no homogéneo no ha sido visto en la literatura hasta ahora. Nosotros en esta sección realizamos un primer acercamiento a este último caso.

En lo que sigue, estudiamos $L(G_1, G_2)$ cuando $G_1 \sim G(n, 1/2)$ y $G_2 \sim G(n, W)$, donde W es un grafón con ínfimo esencial y supremo esencial acotados. Veremos también cómo se compara la restricción que imponemos sobre el grafón con respecto a las restricciones que definen Surya, Warnke y Zhu, y Diamantidis, Konstantopoulos y Yuan para el caso homogéneo.

Teorema 5.1.1. *Sea $\alpha > 0$ la menor raíz de la ecuación $x^2 + (1-x)^2 = \frac{1}{\sqrt{2}}$. Sean $G_1 \sim G(n, 1/2)$ y $G_2 \sim G(n, W)$ independientes, con W un grafón tal que $\alpha \leq \text{ess inf } W \leq \text{ess sup } W \leq 1 - \alpha$. Luego, con alta probabilidad se tiene que $L(G_1, G_2) = (4 + o(1)) \cdot \log_2 n$.*

Demostración. Sean $G_1 \sim G(n, 1/2)$ y $G_2(n, W)$, con $V_1 := V(G_1)$ y $V_2 := V(G_2)$. Supongamos $V_1 = V_2 = [n]$. Para cada k entero positivo, definimos la variable aleatoria $X_k := |\mathcal{X}|$, donde \mathcal{X} se define igual que la ecuación (4.3). Dados $S \in V_1^{(k)}$ y $\phi : S \rightarrow V_2$, definimos $Y_{S,\phi}$ como la variable indicatriz de si $(S, \phi) \in \mathcal{X}$. De esta manera,

$$X_k = \sum_{S \in V_1^{(k)}} \sum_{\substack{\phi: S \rightarrow V_2 \\ \phi \text{ inyectiva}}} Y_{S,\phi}. \quad (5.1)$$

Por otro lado, para todo $S \in V_1^{(k)}$ y $\phi : S \rightarrow V_2$ inyectiva, se tiene que

$$\mathbb{E}[Y_{S,\phi}] = \mathbb{P}[Y_{S,\phi} = 1] = 2^{-\binom{k}{2}}.$$

El cálculo anterior resulta de que G_1 y G_2 son independientes y cada par de vértices conforma una arista de manera uniforme e independiente con probabilidad $\frac{1}{2}$. Luego, el evento donde un par $a, b \in V_1$ coincide en la existencia o no existencia de una arista en G_1 con su par asociado en G_2 , $\phi(a), \phi(b)$, ocurre con probabilidad $\frac{1}{2} \cdot W(x_{\phi(a)}, x_{\phi(b)}) + \frac{1}{2} \cdot (1 - W(x_{\phi(a)}, x_{\phi(b)})) = \frac{1}{2}$. Luego,

$$\begin{aligned} \mathbb{E}[X_k] &= \sum_{S \in V_1^{(k)}} \sum_{\substack{\phi: S \rightarrow V_2 \\ \text{inyectiva}}} \mathbb{E}[Y_{S,\phi}] \\ &= \sum_{S \in V_1^{(k)}} \sum_{\substack{\phi: S \rightarrow V_2 \\ \text{inyectiva}}} 2^{-\binom{k}{2}} \\ &= \binom{n}{k} \cdot \frac{n!}{(n-k)!} \cdot 2^{-\binom{k}{2}} \\ &= \binom{n}{k}^2 \cdot k! \cdot 2^{-\binom{k}{2}}. \end{aligned}$$

Como ha sido usual en este documento, la estructura de la demostración será definida por un argumento de primer y segundo momento.

COTA SUPERIOR: Notar que en el cálculo anterior el valor esperado de X_k coincide con el valor de la esperanza, visto en la ecuación (4.5), para el caso en que $G_1, G_2 \sim G(n, 1/2)$. Luego, sabemos que $\mathbb{E}[X_k] = o(1)$ para todo $\varepsilon > 0$ y $k = (4 + \varepsilon) \cdot \log_2 n$. Por desigualdad de Markov (Lema 2.2.3) se sigue que $\mathbb{P}[X_k > 0] = o(1)$. Así, con alta probabilidad, el tamaño del subgrafo inducido más grande en común entre G_1 y G_2 es a lo más $(4 + \varepsilon) \cdot \log_2 n$.

COTA INFERIOR: Veremos ahora el argumento del segundo momento. Sea $\varepsilon > 0$ y fijemos $k = (4 - \varepsilon) \cdot \log_2 n$. Notemos que ya conocemos una cota inferior para $\mathbb{E}[X_k]$, ya que la calculamos en la desigualdad (4.6). Entonces $\mathbb{E}[X_k] \xrightarrow[n]{n} \infty$. En lo que sigue, estudiamos la probabilidad de que no existan subgrafos inducidos en común, vía isomorfismos, de tamaño k entre G_1 y G_2 . Por Teorema 2.2.2 y cálculos análogos a los vistos en la desigualdad (4.9) se sigue que

$$\begin{aligned} \mathbb{P}[X_k = 0] &\leq \frac{1}{\mathbb{E}[X_k]} + \frac{1}{\mathbb{E}[X_k]^2} \sum_{\substack{0 \leq j_1, j_2 \leq k \\ \max\{j_1, j_2\} \geq 2}} \sum_{\substack{(S, \phi), (T, \varphi) \\ \in B_{j_1, j_2}}} \text{Cov}(Y_{S, \phi}, Y_{T, \varphi}) \\ &\leq \frac{1}{\mathbb{E}[X_k]} + \underbrace{\frac{1}{\mathbb{E}[X_k]^2} \sum_{\substack{0 \leq j_1, j_2 \leq k \\ \max\{j_1, j_2\} \geq 2}} \sum_{\substack{(S, \phi), (T, \varphi) \\ \in B_{j_1, j_2}}} \mathbb{E}[Y_{S, \phi} \cdot Y_{T, \varphi}]}_{(\clubsuit)}. \end{aligned}$$

donde B_{j_1, j_2} se define igual que en la igualdad (4.10). En lo que resta de la demostración, demostraremos que el término en (\clubsuit) tiende a 0 cuando n tiende a infinito.

Afirmación 4.

$$\mathbb{E}[Y_{S, \phi} \cdot Y_{T, \varphi}] \leq 2^{-2\binom{k}{2} + \binom{j_1}{2}} \cdot q^{\max\{0, \binom{j_1}{2} - \binom{j_2}{2}\}},$$

donde $q := \alpha^2 + (1 - \alpha)^2$.

Demostración. Recordemos que $Y_{S, \phi}$ corresponde a la variable aleatoria indicatriz de si $G_1[S] \cong_\phi G_2[\phi(S)]$. Por ende, nos interesa calcular la probabilidad conjunta de que $G_1[S] \cong_\phi G_2[\phi(S)]$ y $G_1[T] \cong_\varphi G_2[\varphi(T)]$. Para esto, es necesario recurrir nuevamente a nuestro grafo auxiliar Γ . Recordemos la definición de Γ

$$V(H) := A \cup B, \text{ y}$$

$$E(H) := \{ab, \phi(a)\phi(b) : ab \in S^{(2)}\} \cup \{xy, \varphi(x)\varphi(y) : xy \in T^{(2)}\},$$

donde $A = S^{(2)} \cup T^{(2)}$ y $B = \phi(S)^{(2)} \cup \varphi(T)^{(2)}$. Sea \mathcal{C} el conjunto de las componentes conexas de Γ . Luego, ocupando la fórmula de la probabilidad conjunta (4.2) se obtiene que la probabilidad de que $G_1[S] \cong_\phi G_2[\phi(S)]$ y $G_1[T] \cong_\varphi G_2[\varphi(T)]$ está dada por

$$\begin{aligned}
 & \prod_{\substack{C \in \mathcal{C} \\ C = C^A \dot{\cup} C^B}} \left(\frac{1}{2}\right)^{|C^A|} \cdot \prod_{ij \in C^B} W(x_a, x_b) + \left(\frac{1}{2}\right)^{|C^A|} \prod_{ij \in C^B} (1 - W(x_a, x_b)) \\
 &= \prod_{\substack{C \in \mathcal{C} \\ C = C^A \dot{\cup} C^B}} 2^{-|C^A|} \left[\prod_{ij \in C^B} W(x_a, x_b) + \prod_{ij \in C^B} (1 - W(x_a, x_b)) \right] \\
 &= 2^{-|A|} \prod_{\substack{C \in \mathcal{C} \\ C = C^A \dot{\cup} C^B}} \left[\prod_{ij \in C^B} W(x_a, x_b) + \prod_{ij \in C^B} (1 - W(x_a, x_b)) \right].
 \end{aligned}$$

Así, usando que para todo $x, y > 0$ y $a \geq 1$ se cumple que $(x + y)^a \geq x^a + y^a$, se sigue que

$$\begin{aligned}
 \mathbb{E}[Y_{S,\phi} \cdot Y_{T,\varphi}] &= 2^{-|A|} \prod_{\substack{C \in \mathcal{C} \\ C = C^A \dot{\cup} C^B}} \left[\prod_{ij \in C^B} W(x_a, x_b) + \prod_{ij \in C^B} (1 - W(x_a, x_b)) \right] \\
 &\leq 2^{-|A|} \prod_{\substack{C \in \mathcal{C} \\ C = C^A \dot{\cup} C^B}} \left[\alpha^{|C^B|} + (1 - \alpha)^{|C^B|} \right] \\
 &= 2^{-|A|} \prod_{\substack{C \in \mathcal{C}^* \\ C = C^A \dot{\cup} C^B}} \left[\alpha^{|C^B|} + (1 - \alpha)^{|C^B|} \right] \\
 &= 2^{-|A|} \prod_{\substack{C \in \mathcal{C}^* \\ C = C^A \dot{\cup} C^B}} \left[(\alpha^2)^{\frac{|C^B|}{2}} + ((1 - \alpha)^2)^{\frac{|C^B|}{2}} \right] \\
 &\leq 2^{-|A|} \prod_{\substack{C \in \mathcal{C}^* \\ C = C^A \dot{\cup} C^B}} \left[\alpha^2 + (1 - \alpha)^2 \right]^{\frac{|C^B|}{2}}, \tag{5.2}
 \end{aligned}$$

donde \mathcal{C}^* es el conjunto de componentes conexas en Γ que poseen al menos 2 vértices en B y definimos los conjuntos $C^A := C \cap A$ y $C^B := C \cap B$.

Ahora, definamos los conjuntos A_1, A_2, B_1 y B_2 tales que en A_1 (respect. B_1) están los vértices en A (respect. B) que pertenecen a una componente conexa de Γ con un solo vértice en B , y en A_2 (respect. B_2) están los vértices en A (respect. B) que pertenecen a una componente conexa de Γ que posee al menos dos vértices en B . Luego, $A = A_1 \dot{\cup} A_2$ y $B = B_1 \dot{\cup} B_2$.

Notemos que las únicas componentes conexas en Γ con vértices en B_1 son aristas o caminos

de largo 3 con vértice intermedio en B , entonces

$$|B_1| \leq |A_1| = |A| - |A_2|$$

Además, como los elementos en A_2 y B_2 pertenecen a las componentes conexas en Γ que poseen al menos dos vértices en B , se satisface que

$$|B_2| \leq 2|A_2| = 2(|A| - |A_1|)$$

y como $|B| = |B_1| + |B_2|$, se obtiene que

$$|B_2| \geq |B| - |A_1| = |B| - |A| + |A_2| \geq |B| - |A| + \frac{|B_2|}{2}$$

y por ende

$$|B_2| \geq 2(|B| - |A|).$$

De esta forma, aplicando esto a la cota (5.2) obtenemos

$$\begin{aligned} \mathbb{E}[Y_{S,\phi} \cdot Y_{T,\varphi}] &\leq 2^{-|A|} \prod_{\substack{C \in \mathcal{C}^* \\ C = C^A \dot{\cup} C^B}} \left[\alpha^2 + (1 - \alpha)^2 \right]^{\frac{|C^B|}{2}} \\ &= 2^{-|A|} q^{\frac{|B_2|}{2}} \\ \boxed{\text{como } q < 1} &\leq 2^{-|A|} q^{\max\{0, |B| - |A|\}} \\ &= 2^{-2\binom{k}{2} + \binom{j_1}{2}} q^{\max\{0, \binom{j_1}{2} - \binom{j_2}{2}\}}. \end{aligned}$$

□

Volviendo a (\clubsuit), por Afirmación 4 se obtiene

$$(\clubsuit) = \sum_{\substack{0 \leq j_1, j_2 \leq k \\ \max\{j_1, j_2\} \geq 2}} \sum_{(S,\phi), (T,\varphi) \in B_{j_1, j_2}} \frac{\mathbb{E}[Y_{S,\phi} \cdot Y_{T,\varphi}]}{\mathbb{E}[X_k]^2} \leq \sum_{\substack{0 \leq j_1, j_2 \leq k \\ \max\{j_1, j_2\} \geq 2}} g_{j_1, j_2}, \quad (5.3)$$

donde en la ecuación anterior definimos el término

$$g_{j_1, j_2} := |B_{j_1, j_2}| \frac{2^{-2\binom{k}{2} + \binom{j_1}{2}} q^{\max\{0, \binom{j_1}{2} - \binom{j_2}{2}\}}}{\binom{n}{k}^4 \cdot k!^2 \cdot 2^{-2\binom{k}{2}}}.$$

Notemos que cuando $2 \leq j_1 \leq j_2$, tenemos que $\max\{0, \binom{j_1}{2} - \binom{j_2}{2}\} = 0$, y con esto obtenemos en la cota (5.3) el mismo término (\spadesuit) de la desigualdad (4.13) para el caso en el que $G_1, G_2 \sim G(n, 1/2)$. Por ende, solo restaría acotar el término (\clubsuit) cuando

- (i) $j_1 \in \{0, 1\}$ y $j_2 \geq 2$,
- (ii) $j_2 = 0$ y $j_1 \geq 2$, y
- (iii) $1 \leq j_2 \leq j_1$ y $\max\{2, j_2\} \leq j_1$.

En el caso (i), fijamos $j_1 \in \{0, 1\}$, así

$$\begin{aligned} \sum_{\substack{0 \leq j_1, j_2 \leq k \\ \max\{j_1, j_2\} \geq 2}} g_{j_1, j_2} &= \sum_{2 \leq j_2 \leq k} \frac{|B_{j_1, j_2}| \cdot 1}{\binom{n}{k}^4 \cdot k!^2} \\ &= \sum_{2 \leq j_2 \leq k} \frac{\binom{n}{j_1} \binom{n-j_1}{k-j_1} \binom{n-k}{k-j_1} \binom{n}{j_2} \binom{n-j_2}{k-j_2} \binom{n-k}{k-j_2}}{\binom{n}{k}^4} \\ &\leq \sum_{2 \leq j_2 \leq k} \frac{n \cdot \binom{n}{k-j_1}^2 \binom{n}{j_2} \binom{n}{k-j_2}^2}{\binom{n}{k}^4} \\ &\stackrel{\text{Lema 2.1.2}}{\leq} \sum_{2 \leq j_2 \leq k} \frac{n \cdot \left(\frac{en}{k-j_1}\right)^{2(k-j_1)} \left(\frac{en}{j_2}\right)^{j_2} \left(\frac{en}{k-j_2}\right)^{2(k-j_2)}}{\binom{n}{k}^4} \\ &\stackrel{\text{caso (I) de Lema 2.1.3}}{\leq} \sum_{2 \leq j_2 \leq k} \frac{n \cdot \left(\frac{en}{k-j_1}\right)^{2(k-j_1)} \left(\frac{en}{j_2}\right)^{j_2} \left(\frac{en}{k-j_2}\right)^{2(k-j_2)}}{\left[\frac{1+o(1)}{\sqrt{2\pi k}} \left(\frac{en}{k}\right)^k\right]^4} \\ &\leq 4\pi^2 k^2 \sum_{2 \leq j_2 \leq k} \frac{n \cdot \left(\frac{en}{k-1}\right)^{2k} \left(\frac{en}{j_2}\right)^{j_2} \left(\frac{en}{k-j_2}\right)^{2(k-j_2)}}{\left(\frac{en}{k}\right)^{4k}} \end{aligned}$$

$$\begin{aligned}
 &= 4\pi^2 k^2 \sum_{2 \leq j_2 \leq k} n \left(\frac{en}{k-1} \right)^{2k} \binom{k}{j_2}^{j_2} \left(\frac{k}{k-j_2} \right)^{2(k-j_2)} \left(\frac{k}{en} \right)^{2k+j_2} \\
 &\stackrel{(4.15)}{\leq} 4\pi^2 k^2 \sum_{2 \leq j_2 \leq k} n \left(\frac{en}{k-1} \right)^{2k} \binom{k}{j_2}^{j_2} e^{2j_2} \left(\frac{k}{en} \right)^{2k+j_2} \\
 &= 4\pi^2 k^2 \sum_{2 \leq j_2 \leq k} n \left(\frac{k}{k-1} \right)^{2k} \left(\frac{k^2 e^2}{j_2 en} \right)^{j_2} \\
 &\leq 4\pi^2 k^2 n e^2 \sum_{2 \leq j_2 \leq k} \left(\frac{k^2 e^2}{j_2 en} \right)^{j_2} \\
 &\stackrel{\text{Lema 2.1.1}}{\leq} 4\pi^2 k^2 n e^2 \left(\frac{k^2 e}{2n} \right)^2 \\
 &= O(k^6 n^{-1}) \\
 &= o(1).
 \end{aligned}$$

En el caso (ii), fijamos $j_2 = 0$ y $2 \leq j_1 \leq k-1$. Luego, $\max\{0, \binom{j_1}{2} - \binom{j_2}{2}\} = \binom{j_1}{2}$. Además, recordemos que por hipótesis $q = \frac{1}{\sqrt{2}}$, por ende $2q = \sqrt{2}$. Así, obtenemos

$$\begin{aligned}
 \sum_{\substack{0 \leq j_1, j_2 \leq k \\ \max\{j_1, j_2\} \geq 2}} g_{j_1, j_2} &= \sum_{2 \leq j_1 \leq k} \frac{|B_{j_1, 0}| (2q)^{\binom{j_1}{2}}}{\binom{n}{k}^4 \cdot k!^2} \\
 &= \sum_{2 \leq j_1 \leq k} \frac{\binom{n}{j_1} \binom{n-j_1}{k-j_1} \binom{n-k}{k-j_1} \binom{n}{k} \binom{n-k}{k} 2^{\frac{1}{2} \binom{j_1}{2}}}{\binom{n}{k}^4} \\
 &\leq \sum_{2 \leq j_1 \leq k} \frac{\binom{n}{j_1} \binom{n-j_1}{k-j_1} \binom{n-k}{k-j_1} 2^{\frac{1}{2} \binom{j_1}{2}}}{\binom{n}{k}^2} \\
 &= \sum_{2 \leq j_1 \leq k} \frac{\binom{n}{k} \binom{k}{j_1} \binom{n-k}{k-j_1} 2^{\frac{1}{2} \binom{j_1}{2}}}{\binom{n}{k}^2} \\
 &= \sum_{i=2}^k c_i
 \end{aligned}$$

donde en la ecuación anterior definimos, para todo $i \in \{2, \dots, k\}$,

$$c_i := \frac{\binom{l}{i} \binom{n-k}{k-i} 2^{\frac{1}{2} \binom{i}{2}}}{\binom{n}{k}}.$$

Ahora, veremos que para n suficientemente grande y para todo $i \in \{2, \dots, k\}$, $c_2 \geq c_i$. En efecto,

$$\begin{aligned}
 \frac{c_i}{c_2} &= \frac{\binom{k}{i} \binom{n-k}{k-i} 2^{\frac{1}{2} \binom{i}{2}}}{\binom{k}{2} \binom{n-k}{k-2} \cdot \sqrt{2}} \\
 &= \frac{2}{i!} \cdot \frac{(n-2k+2)!}{(n-2k+i)!} \cdot \left[\frac{(k-2)!}{(k-i)!} \right]^2 \cdot 2^{\frac{i(i-1)}{4} - \frac{1}{2}} \\
 &\leq \frac{1}{(n-2k)^{i-2}} \cdot k^{2(i-2)} \cdot 2^{\frac{i(i-1)}{4} - \frac{1}{2}} \\
 &= \left[\frac{k^2 \cdot 2^{\frac{(i+1)}{4}}}{n-2k} \right]^{i-2} \\
 &\leq \left[\frac{k^2 \cdot 2^{\frac{(k+1)}{4}}}{n-2k} \right]^{i-2}.
 \end{aligned}$$

Como $k = (4 - \varepsilon) \cdot \log_2 n$, entonces $2^{\frac{k+1}{4}} = 2^{\frac{1}{4}} n^{1-\frac{\varepsilon}{4}}$. Además, para n suficientemente grande se tiene que $k \leq \frac{n}{4}$, entonces

$$\frac{a_i}{a_2} \leq \underbrace{\left[\frac{2^{\frac{1}{4}} \cdot n^{1-\frac{\varepsilon}{2}}}{n} \right]^{i-2}}_{= O(n^{-\frac{\varepsilon}{4}})} < 1. \tag{5.4}$$

Luego, se sigue que

$$\sum_{\substack{0 \leq j_1, j_2 \leq k \\ \max\{j_1, j_2\} \geq 2}} g_{j_1, j_2} \leq k \cdot c_2 = k \cdot \frac{\binom{k}{2} \binom{n-k}{k-2} \sqrt{2}}{\binom{n}{2}},$$

donde el último término sabemos que tiende a 0 cuando $n \rightarrow \infty$, ya que coincide con la expresión (3.10) (salvo un factor constante), para el estudio de $\omega(G(n, 1/2))$. Por último, veamos el caso (iii), donde fijamos $1 \leq j_2 \leq k$ y $\max\{2, j_2\} \leq j_2$.

$$\sum_{\substack{1 \leq j_2 \leq j_1 \leq k \\ \max\{j_1, j_2\} \geq 2}} g_{j_1, j_2} = \sum_{\substack{1 \leq j_2 \leq j_1 \leq k \\ \max\{2, j_2\} \leq j_1}} \frac{|B_{j_1, j_2}| 2^{\binom{j_1}{2}} q^{\binom{j_1}{2} - \binom{j_2}{2}}}{\binom{n}{k}^4 \cdot k!^2}$$

$$\begin{aligned}
 \boxed{\text{por Afirmación 2}} &\leq 4\pi^2 k^2 \sum_{\substack{1 \leq j_2 \leq j_1 \leq k \\ \max\{2, j_2\} \leq j_1}} (2q)^{\binom{j_1}{2}} b^{\binom{j_2}{2}} \left(\frac{ek}{n}\right)^{j_1+j_2} \left(\frac{k}{j_1}\right)^{j_1} \left(\frac{k}{j_2}\right)^{j_2} \\
 &= 4\pi^2 k^2 \sum_{1 \leq j_2 \leq k} b^{\binom{j_2}{2}} \left(\frac{el^2}{nj_2}\right)^{j_2} \sum_{\max\{2, j_2\} \leq j_1 \leq k-1} \left(\frac{ek^2}{nj_1}\right)^{j_1} (2q)^{\binom{j_1}{2}} \\
 &\leq 4\pi^2 k^2 \sum_{1 \leq j_2 \leq k} \left(\frac{ek^2 b^{\frac{j_2}{2}}}{nj_2}\right)^{j_2} \sum_{\max\{2, j_2\} \leq j_1 \leq k-1} \left(\frac{ek^2 (2q)^{\frac{j_1}{2}}}{nj_1}\right)^{j_1}
 \end{aligned}$$

donde $b := q^{-1}$. Ahora, recordemos que por hipótesis $q = \frac{1}{\sqrt{2}}$, por ende $2q = \sqrt{2}$ y así $(2q)^{\frac{j_1}{2}} = 2^{\frac{j_1}{4}} \leq 2^{\frac{k}{4}} = n^{1-\frac{\varepsilon}{4}}$. Luego, para n suficientemente grande tenemos que

$$\begin{aligned}
 \sum_{\substack{1 \leq j_2 \leq j_1 \leq k \\ \max\{j_1, j_2\} \geq 2}} g_{j_1, j_2} &\leq 4\pi^2 k^2 \sum_{1 \leq j_2 \leq k} \left(\frac{ek^2 b^{\frac{j_2}{2}}}{nj_2}\right)^{j_2} \sum_{\max\{2, j_2\} \leq j_1 \leq k-1} \left(\frac{ek^2}{n^{\frac{\varepsilon}{4}} j_1}\right)^{j_1} \\
 \boxed{\text{Lema 2.1.1}} &\leq 8\pi^2 k^2 \sum_{1 \leq j_2 \leq k} \left(\frac{el^2 b^{\frac{j_2}{2}}}{nj_2}\right)^{j_2} \cdot \left(\frac{ek^2}{n^{\frac{\varepsilon}{4}} j_2}\right)^{j_2} \\
 &= 8\pi^2 k^2 \sum_{1 \leq j_2 \leq k} \left(\frac{e^2 k^4 \cdot b^{\frac{j_2}{2}}}{n^{1+\frac{\varepsilon}{4}} j_2^2}\right)^{j_2} \\
 \boxed{b^{\frac{j_2}{2}} \leq n^{1-\frac{\varepsilon}{4}}} &\leq 8\pi^2 k^2 \sum_{1 \leq j_2 \leq k} \left(\frac{e^2 k^4}{n^{\frac{\varepsilon}{2}}}\right)^{j_2} \\
 \boxed{\text{Lema 2.1.1}} &\leq 8\pi^2 k^2 \cdot \left(\frac{e^2 k^4}{n^{\frac{\varepsilon}{2}}}\right) \\
 &= o(1).
 \end{aligned}$$

Luego, con alta probabilidad existe un subgrafo inducido en común entre G_1 y G_2 cuyo tamaño es menor a $(4 - \varepsilon) \cdot \log_2 n$, lo que concluye la demostración. \square

Observemos lo siguiente. Ya vimos que para el estudio del típico valor de $L(G_1, G_2)$ cuando $G_1 \sim G(n, p_1)$ y $G_2 \sim G(n, p_2)$, uno de los estudios que mencionamos consideraban la siguiente restricción para las constantes $p_1, p_2 \in (0, 1)$

$$\max \left\{ p_1^2 p_2 + (1 - p_1)^2 (1 - p_2), p_1 p_2^2 + (1 - p_1)(1 - p_2)^2 \right\} \leq [p_1 p_2 + (1 - p_1)(1 - p_2)]^{3/2}. \quad (5.5)$$

Veamos para qué valores de p_2 se alcanza la igualdad en la última expresión, cuando fijamos

$p_1 = \frac{1}{2}$. Tenemos por un lado que

$$\left(\frac{1}{2}p_2 + \frac{1}{2}(1 - p_2)\right)^{3/2} = \left(\frac{1}{2}\right)^{3/2} = \frac{\sqrt{2}}{4}$$

y a su vez

$$\max\left\{\frac{1}{2}p_2^2 + \frac{1}{2}(1 - p_2)^2, \frac{1}{4}p_2 + \frac{1}{4}(1 - p_2)\right\} = \frac{1}{2} \max\left\{p_2^2 + (1 - p_2)^2, \frac{1}{2}\right\} = \frac{1}{2} [p_2^2 + (1 - p_2)^2].$$

De esta forma, al tomar la igualdad en la expresión (5.5) cuando $p_1 = \frac{1}{2}$ obtenemos

$$p_2^2 + (1 - p_2)^2 = \frac{\sqrt{2}}{2},$$

que corresponde a exactamente la misma ecuación con la cual definimos el valor de α en el Teorema 5.1.1. Esto nos dice que nuestra restricción coincide exactamente con la restricción que imponen Diamantidis, Konstantopoulus y Yuan para el cálculo de $L(G_1, G_2)$ con $G_1 \sim G(n, p_1)$ y $G_2 \sim G(n, p_2)$, cuando fijamos $p_1 = \frac{1}{2}$. Entonces, en este caso fue posible obtener restricciones para los valores que toma W similares a las que se imponen para las intensidades de aristas en el caso homogéneo. La razón por la cual impusimos esta restricción para W fue para poder realizar el estudio bajo el “buen funcionamiento” del método del segundo momento. Esto calza también con la motivación de imponer la restricción (5.5) en el estudio de Diamantidis, Konstantopoulus y Yuan.

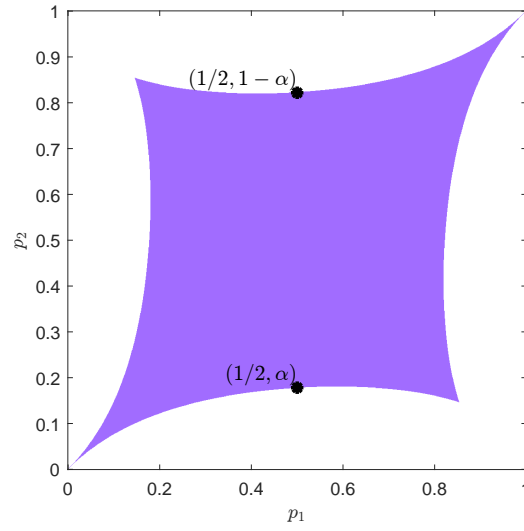


Figura 5.1: Región de los valores $p_1, p_2 \in (0, 1)$ junto a las coordenadas $(\frac{1}{2}, \alpha)$ y $(\frac{1}{2}, 1 - \alpha)$.

Conclusiones y trabajo futuro

En este trabajo pudimos estudiar el tamaño típico del clique más grande en un grafo aleatorio y el tamaño típico del subgrafo inducido en común más grande entre dos grafos aleatorios G_1 y G_2 , en distintos casos dependiendo de si los grafos son o no generados por un modelo de grafos aleatorios homogéneo o no homogéneo. En particular, obtuvimos un resultado que no se había visto antes en la literatura, donde vimos que si $G_1 \sim G(n, 1/2)$ y $G_2 \sim G(n, W)$ son grafos aleatorios independientes, para cierto grafón W , entonces con alta probabilidad $L(G_1, G_2) = (4 + o(1)) \cdot \log_2 n$. Para obtener este resultado, notamos que una parte importante de los cálculos fueron similares a los realizados para obtener el valor de $L(G_1, G_2)$, con $G_1, G_2 \sim G(n, 1/2)$, debido a que el cálculo de la esperanza de nuestra variable aleatoria X_k coincidía en ambos casos. Sin embargo, si quisiéramos extender el Teorema 5.1.1 al caso en el que $G_1 \sim G(n, p)$, para cualquier constante $p \in (0, 1)$, los cálculos se verían drásticamente complejizados, de manera similar a los cálculos vistos por Doležal, Hladký y Máthé para determinar el tamaño del clique más grande en un grafo aleatorio no homogéneo.

A partir de este último resultado y de haber estudiado el trabajo de Doležal, Hladký y

Máthé [14], damos la siguiente conjetura.

Conjetura 1. *Dados W un grafón adecuado con $0 < \text{ess inf } W \leq \text{ess sup } W < 1$, $p \in (0, 1)$ cierta constante fija, $G_1 \sim G(n, p)$ y $G_2 \sim G(n, W)$ independientes, se tiene que con alta probabilidad*

$$L(G_1, G_2) \approx 2\omega(G(n, \tilde{W})), \quad (6.1)$$

donde $\tilde{W} = \tilde{W}(p, W)$ es el grafón definido por $\tilde{W} \equiv pW + (1 - p)(1 - W)$.

Cuando hablamos de un grafón adecuado y cierta constante p , nos referimos a las condiciones suficientes para que funcione el método del segundo momento. De la fórmula vista por Surya, Warnke y Zhu [30] (4.18), podemos intuir que la fórmula del tamaño típico del subgrafo inducido en común más grande entre dos grafos aleatorios no homogéneos $G(n, p)$, $G(n, W)$ será diferente si p y W son tales que no funciona el método del segundo momento. Por otro lado, con la aproximación (6.1) nos referimos a que las variables aleatorias $L(G_1, G_2)$ y $2\omega(G(n, \tilde{W}))$ toman el mismo valor con alta probabilidad.

Notemos que la Conjetura 1 se satisface en el Teorema 5.1.1, ya que por un lado, del Teorema 3.3.1 tenemos que

$$\frac{\omega(G(n, \tilde{W}))}{\log n} = (1 + o(1)) \cdot \kappa(\tilde{W}),$$

donde $\tilde{W} \equiv \frac{1}{2}W + \frac{1}{2}(1 - W) = \frac{1}{2}$, y ya vimos en la igualdad (3.16) que en este caso $\kappa(\tilde{W}) = \frac{2}{\log 2}$.

Luego,

$$\frac{\omega(G(n, \tilde{W}))}{\log n} = (1 + o(1)) \cdot \frac{2}{\log 2} \stackrel{\text{Teorema 5.1.1}}{=} \frac{L(G_1, G_2)}{\log n}.$$

La Conjetura 1 nos entrega una relación directa entre el estudio del tamaño del subgrafo inducido en común más grande entre $G(n, p)$ y $G(n, W)$ y el estudio del tamaño del clique más grande en un un grafo $G(n, \tilde{W})$, con \tilde{W} un grafón que se obtiene como un reescalamiento de p y W . Más aún, gracias al trabajo realizado por Doležal, Hladký y Máthé, sabemos que en la ecuación (6.1) podemos ocupar el Teorema 3.3.1, que nos provee de una fórmula explícita para el valor de $\omega(G(n, \tilde{W}))$.

Para resolver el problema de determinar el tamaño típico del subgrafo inducido en común

más grande (bajo isomorfismos) entre dos grafos aleatorios no homogéneos, un camino natural a seguir sería estudiar el valor de $L(G_1, G_2)$ cuando

- (I) $G_1 \sim G(n, p)$ y $G_2 \sim G(n, W)$, con $p \in (0, 1)$ y W un grafón,
- (II) G_1 y G_2 son generados por un modelo bloque estocástico, y luego
- (III) $G_1 \sim G(n, W_1)$ y $G_2 \sim G(n, W_2)$, con W_1 y W_2 grafones.

Este orden de trabajo sigue una línea similar a la definida por esta tesis, ya que va avanzando a medida que el resultado es cada vez más general.

Para todo par de grafones $H_1 : [0, 1]^2 \rightarrow [0, 1]$, $H_2 : [0, 1]^2 \rightarrow [0, 1]$, definimos el grafón auxiliar $\hat{W}_{H_1, H_2} \equiv H_1 H_2 + (1 - H_1)(1 - H_2)$. Aquí, podríamos pensar que para determinar $L(G_1, G_2)$ con $G_1 \sim G(n, W_1)$, $G_2 \sim G(n, W_2)$ independientes, donde W_1 y W_2 son grafones que admiten el buen funcionamiento del argumento de primer y segundo momento, una conjetura tentativa sería decir que con alta probabilidad

$$L(G_1, G_2) \approx 2\omega(G(n, \hat{W}_{W_1, W_2})). \quad (6.2)$$

Sin embargo, daremos un ejemplo sencillo que ilustra por qué esta conjetura no es precisa para este caso. Tomemos $W_1 : [0, 1]^2 \rightarrow [0, 1]$ y $W_2 : [0, 1]^2 \rightarrow [0, 1]$ tales que para todo $(x, y) \in [0, 1]^2$,

$$W_1(x, y) = \begin{cases} \frac{1}{3}, & \text{si } (x, y) \in \left[0, \frac{1}{2}\right] \times \left(\frac{1}{2}, 1\right] \\ \frac{2}{3}, & \text{si no,} \end{cases} \quad \text{y} \quad W_2(x, y) = \begin{cases} \frac{1}{3}, & \text{si } (x, y) \in \left[\frac{1}{2}, 1\right] \times \left[0, \frac{1}{2}\right) \\ \frac{2}{3}, & \text{si no.} \end{cases}$$

Luego, para $(x, y) \in [0, 1]^2$, salvo un conjunto de medida nula,

$$\hat{W}_{W_1, W_2}(x, y) = \begin{cases} \frac{4}{9}, & \text{si } (x, y) \in \left[0, \frac{1}{2}\right] \times \left(\frac{1}{2}, 1\right] \text{ o } (x, y) \in \left[\frac{1}{2}, 1\right] \times \left[0, \frac{1}{2}\right), \\ \frac{5}{9}, & \text{si no.} \end{cases}$$

Por otro lado, no es difícil ver que W_1 y W_2 son *isomorfos c.t.p.* [26, p. 121], es decir, existe una biyección que preserva medida $\varphi : [0, 1] \rightarrow [0, 1]$ tal que para todo $(x, y) \in [0, 1]^2$, salvo un

conjunto de medida nula, $W_1(x, y) = W_2(\varphi(x), \varphi(y))$. Denotaremos por W_2^φ al grafón dado por la composición entre W_2 y φ .

Además, notemos que $\hat{W}_{W_1, W_2^\varphi} = \hat{W}_{W_1, W_1} = \frac{5}{9}$ c.t.p., y por lo tanto $\hat{W}_{W_1, W_2} \leq \hat{W}_{W_1, W_1}$ c.t.p.. Más, aún, en un conjunto de medida no nula $X \subset [0, 1]^2$ se tiene que $\hat{W}_{W_1, W_2} < \hat{W}_{W_1, W_1}$. Luego, podemos realizar el mismo razonamiento visto en Sección 3.3 para obtener desigualdades análogas a las vistas en la ecuación (3.19). Así, tendríamos que con alta probabilidad el tamaño del clique máximo en un grafo $G(n, \hat{W}_{W_1, W_2})$ sería menor al tamaño del clique máximo en un grafo $G(n, \hat{W}_{W_1, W_1})$. Esto último nos permite intuir que quizás la fórmula en (6.2) no nos entregaría un valor óptimo para determinar el tamaño típico del subgrafo inducido en común más grande entre dos grafos aleatorios no homogéneos y, siguiendo la idea de este ejemplo, sería necesario incorporar el rol de las posibles biyecciones entre W_1 y W_2 o en su defecto, las transformaciones tales que W_1 y W_2 sean *débilmente isomorfos* [26, Cáp. 7, Sec. 3].

Bibliografía

- [1] Noga Alon and Joel H. Spencer, *The probabilistic method*, third ed., Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., Hoboken, NJ, 2008, With an appendix on the life and work of Paul Erdős.
- [2] Krishna B. Athreya and Soumendra N. Lahiri, *Measure theory and probability theory*, Springer Texts in Statistics, Springer, New York, 2006.
- [3] B. Bollobás and A. Thomason, *Threshold functions*, *Combinatorica* **7** (1987), no. 1, 35–38.
- [4] B. Bollobás and P. Erdős, *Cliques in random graphs*, *Mathematical Proceedings of the Cambridge Philosophical Society* **80** (1976), no. 3, 419–427.
- [5] Vincenzo Bonnici, Rosalba Giugno, Alfredo Pulvirenti, Dennis Shasha, and Alfredo Ferro, *A subgraph isomorphism algorithm and its application to biochemical data*, *BMC Bioinformatics* **14** (2013), 1–13.
- [6] Peter J. Cameron, *The random graph*, arXiv (2013).
- [7] Sourav Chatterjee and Persi Diaconis, *Isomorphisms between random graphs*, *J. Combin. Theory Ser. B* **160** (2023), 144–162.

-
- [8] Václav Chvátal and David Sankoff, *Longest common subsequences of two random sequences*, J. Appl. Probability **12** (1975), 306–315.
- [9] Donatello Conte, Pasquale Foggia, Carlo Sansone, and Mario Vento, *Thirty years of graph matching in pattern recognition*, International Journal of Pattern Recognition and Artificial Intelligence **18** (2004), no. 03, 265–298.
- [10] Diane J Cook and Lawrence B Holder, *Substructure discovery using minimum description length and background knowledge*, Journal of Artificial Intelligence Research **1** (1993), 231–255.
- [11] Shagnik Das, *A brief note on estimates of binomial coefficients*, 2015.
- [12] Luc Devroye and Nicolas Fraiman, *Connectivity of inhomogeneous random graphs*, Random Structures Algorithms **45** (2014), no. 3, 408–420.
- [13] Dimitris Diamantidis, Takis Konstantopoulos, and Linglong Yuan, *A combinatorial approach to phase transitions in random graph isomorphism problems*, arXiv (2024).
- [14] Martin Doležal, Jan Hladký, and András Máthé, *Cliques in dense inhomogeneous random graphs*, Random Structures Algorithms **51** (2017), no. 2, 275–314.
- [15] Hans-Christian Ehrlich and Matthias Rarey, *Maximum common subgraph isomorphism algorithms and their applications in molecular science: a review*, Wiley Interdisciplinary Reviews: Computational Molecular Science **1** (2011), no. 1, 68–79.
- [16] P. Erdős and A. Rényi, *On random graphs. I*, Publ. Math. Debrecen **6** (1959), 290–297.
- [17] P. Erdős and A. Rényi, *Asymmetric graphs*, Acta Math. Acad. Sci. Hungar. **14** (1963), 295–315.
- [18] Nicolas Fraiman and Dieter Mitsche, *The diameter of inhomogeneous random graphs*, Random Structures Algorithms **53** (2018), no. 2, 308–326.
- [19] E. N. Gilbert, *Random graphs*, Ann. Math. Statist. **30** (1959), 1141–1144.

-
- [20] G. R. Grimmett and C. J. H. McDiarmid, *On colouring random graphs*, Mathematical Proceedings of the Cambridge Philosophical Society **77** (1975), no. 2, 313–324.
- [21] Paul W. Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt, *Stochastic blockmodels: First steps*, Social Networks **5** (1983), no. 2, 109–137.
- [22] Svante Janson, *Graphons, cut norm and distance, couplings and rearrangements*, New York Journal of Mathematics. NYJM Monographs, vol. 4, State University of New York, University at Albany, Albany, NY, 2013.
- [23] Svante Janson, Tomasz Łuczak, and Ilkka Norros, *Large cliques in a power-law random graph*, Journal of Applied Probability **47** (2010), no. 4, 1124–1135.
- [24] Marcos Kiwi, Martin Loeb, and Jiří Matoušek, *Expected length of the longest common subsequence for large alphabets*, Adv. Math. **197** (2005), no. 2, 480–498.
- [25] Théo Lenoir, *Isomorphisms between random d -hypergraphs*, arXiv (2024).
- [26] László Lovász, *Large networks and graph limits*, American Mathematical Society Colloquium Publications, vol. 60, American Mathematical Society, Providence, RI, 2012.
- [27] László Lovász and Balázs Szegedy, *Limits of dense graph sequences*, J. Combin. Theory Ser. B **96** (2006), no. 6, 933–957.
- [28] David W. Matula, *On the complete subgraphs of a random graph*, Proc. Second Chapel Hill Conf. on Combinatorial Mathematics and its Applications (Univ. North Carolina, Chapel Hill, N.C., 1970), University of North Carolina, Chapel Hill, NC, 1970, pp. 356–369.
- [29] Dan Romik, *The surprising mathematics of longest increasing subsequences*, Institute of Mathematical Statistics Textbooks, vol. 4, Cambridge University Press, New York, 2015.
- [30] Erlang Surya, Lutz Warnke, and Emily Zhu, *Isomorphisms between dense random graphs*, arXiv (2023).

- [31] Martin Vingron and Michael S. Waterman, *Sequence alignment and penalty choice: Review of concepts, case studies and implications*, Journal of Molecular Biology **235** (1994), no. 1, 1–12.