Método de Contêineres

Pedro Campos Araújo*

21 de fevereiro de 2017

*Instituto Nacional de Matemática Pura e Aplicada

1 Índice

2 Introdução

Estudar conjuntos independentes em hipergrafos pode parecer, a princípio, muito abstrato para alguém que acaba de começar a estudar Combinatória. O objetivo deste trabalho é mostrar como muitos problemas de contagem ou estruturais podem ser traduzidos para a linguagem de hipergrafos. O método de contêiner é versátil e pode ser aplicado tanto em problemas determinísticos quanto aleatórios.

O resultado principal deste trabalho mostra que se um hipergrafo uniforme apresenta uma certa regularidade, então existe uma maneira de aglomerar seus conjuntos independentes nos chamados contêineres. Em [1], Kleitman e Winston desenvolveram técnicas para calcular o número de grafos sem cópias de C_4, construindo contêineres para grafos.

Pela motivação histórica, o foco primeiramente serão grafos, um caso particular cuja facilidade vai além do fato de que é menos abstrato. Além da construção dos contêineres para grafos, será mostrada uma aplicação deles para provar teorema de Alon [2] sobre conjuntos independentes em um grafo d-regular com n vértices.

Após provar o resultado principal, focaremos em demonstrar três resultados no mundo esparso que se originaram de três teoremas clássicos de áreas distintas da combinatória.

2.1 Turán para $G_{n,p}$

Um dos resultados mais clássicos de Combinatória é o teorema de Mantel, que afirma que o máximo de arestas que um grafo com n vértices e sem triângulos pode ter é $\lceil \frac{n}{2} \rceil \lfloor \frac{n}{2} \rfloor$. O único grafo que atinge esse número é o bipartido completo com partes de tamanhos $\lceil \frac{n}{2} \rceil$ e $\lfloor \frac{n}{2} \rfloor$. Uma generalização natural deste teorema é determinar o número seguinte.

$$\mathrm{ex}(n,H) = \max\{e(G) : G \text{ é um grafo com n vértices e } H \nsubseteq G\}$$

Nesta linguagem, o Teorema de Mantel diz que $\mathrm{ex}(n,\triangle) = \lceil \frac{n}{2} \rceil \lfloor \frac{n}{2} \rfloor$. O próximo resultado em determinar esses números se deu a Turán, que provou que $ex(n, K_k)$ é atingido somente pelo grafo $(k-1)$-partido completo, distribuindo-se os vértices entre as partes de forma balanceada.

Para um grafo H qualquer, o teorema de Erdős, Stone [3] determina assintoticamente $ex(n,H)$:

Teorema 2.1 (Erdős-Stone)**.** *Seja H um grafo e $\chi(H)$ seu número cromático. Então*

$$ex(n,H) = \left(1 - \frac{1}{\chi(H)-1} - o(1)\right)\binom{n}{2}$$

A generalização mais natural do número extremal de um grafo é a que considera um grafo qualquer (não necessariamente um clique) como 'hospedeiro':

$$\mathrm{ex}(G,H) = \max\{e(G') : G' \subseteq G \text{ e } H \nsubseteq G\}$$

Assim, usando contêineres, pode-se provar provar um análogo ao Teorema de Turán para $G_{n,p}$, se p for grande o suficiente. Tal resultado foi provado por Conlon e Gowers [4] e Schacht [5]

Teorema 2.2 (Turán para $G_{n,p}$). *Fixe H um grafo com grau máximo pelo menos 2. Para todo $\delta \in (0,1)$, existe $C > 0$ tal que com alta probabilidade*

$$ex(G_{n,p},H) \leq \left(1 - \frac{1}{\chi(H)-1} + \delta\right)p\binom{n}{2},$$

para todo $p \geq Cn^{-1/m_2(H)}$, em que $m_2(H) = \max\left\{\frac{e(F)-1}{v(F)-2} : F \subset H \text{ e } v(F) \geq 2\right\}$

2.2 Teorema de Ramsey Aleatório

Se as arestas de K_n forem coloridas de duas cores, será que sempre é possível encontrar um triângulo monocromático? A resposta é afirmativa se $n \geq 6$, para qualquer coloração. De fato, substituirmos o triângulo por um K_k, então o Teorema de Ramsey diz que existe n_0 tal que se $n > n_0$ então para qualquer coloração das arestas de K_n existe uma cópia monocromática de K_k (e consequentemente o mesmo vale para qualquer grafo H com k vértices)

Assim como o número extremal, existe uma pergunta análoga se o grafo 'hospedeiro' não é K_n ou se o número de cores é maior que dois. O que nos leva à seguinte definição:

Definição 2.3. *Dados grafos G e H, escreve-se*

$$G \to (H)_r$$

se toda coloração das arestas de G com r cores contém uma cópia monocromática de H.

Com essa definição, podemos provar o teorema abaixo, provado originalmente por Rödl e Ruciński em [6], mas posteriormente também por Nenadov e Steger em [7], usando contêineres.

Teorema 2.4. *Seja H um grafo que contém um ciclo. Existem constantes $c = c(H, r) > 0$ e $C = C(H, r) > 0$ tal que*

$$\mathbb{P}[G_{n,p} \to (F)_r] = \begin{cases} 1 - o(1), & \text{se } p \geq Cn^{-1/m_2(H)} \\ o(1), & \text{se } p \leq cn^{-1/m_2(H)} \end{cases}$$

2.3 Anticadeias em $\mathcal{P}(n, p)$

Se $\mathcal{C}$ é uma coleção de subconjuntos de $[n]$ em que não existem $A, B \in \mathcal{C}$ tal que $A \subset B$, quão grande $\mathcal{C}$ pode ser? Sperner provou em que se esta condição é satisfeita ($\mathcal{C}$ é chamada de uma anticadeia), então

$$|C| \leq \binom{n}{\lfloor n/2 \rfloor},$$

em que a igualdade é atingida pela família de conjuntos de tamanho $\lfloor n/2 \rfloor$.

Erdős então generalizou esse resultado em [8] para famílias de conjuntos que não possuem uma cadeia de tamanho k e provou que o máximo de conjuntos em uma família dessas é atingido pelas $k - 1$ camadas do meio de $\mathcal{P}(n)$.

Definimos $\mathcal{P}(n, p)$ uma família aleatória de conjuntos em que cada elemento de $\mathcal{P}(n)$ está presente com probabilidade p, independentemente. Assim, temos o seguinte resultado, provado por Collares e Morris [9] e independentemente por Balogh, Mycroft e Treglown [10].

Teorema 2.5. *Seja $2 \leq k \in \mathbb{N}$ e $p = p(n)$ tal que $\lim_{n\to\infty} pn = +\infty$. Então com alta probabilidade a maior família $\mathcal{A} \subset \mathcal{P}(n, p)$ que não contém uma k-cadeia satisfaz*

$$|\mathcal{A}| = (k - 1 + o(1))p\binom{n}{n/2}.$$

3 Notações e Resultados básicos

Primeiramente, convencionamos que $[n] = \{1, ..., n\}$. Além disso, dado um grafo G, denotamos por $V(G)$ o conjunto de vértices e $E(G)$ o conjunto de arestas. Para $v \in V(G)$, o grau de v, $d(v)$, é a quantidade de arestas que contém esse vértice. Para um hipergrafo $\mathcal{H}$, e $A \subset V(\mathcal{H})$, $d(A)$ é a quantidade de arestas que contém esse conjunto. Além disso,

$$\Delta_l(\mathcal{H}) = \max\{d(A) : A \subset V(\mathcal{H}), |A| = l\}.$$

Para duas funções f e g com domínio $\mathbb{N}$, dizemos que $f(n) = o(g(n))$ ou $f(n) \ll g(n)$ se $\lim_{n\to\infty} f(n)/g(n) = 0$. Temos também que $f(n) = \Theta(g(n))$ se existem $c, C > 0$ tal que $cg(n) < f(n) < Cg(n)$.

Ao longo deste trabalho, usaremos as seguintes desigualdades:

$$\binom{n}{k} \leq \left(\frac{en}{k}\right)^k \tag{1}$$

$$\binom{n}{k} = \Theta(n^k) \tag{2}$$

Por último se X é uma variável aleatória com distribuição $Bin(n, p)$, temos pela desigualdade de Chernoff (mais detalhes em [11]) que

$$\mathbb{P}(X > (1+\varepsilon)pn) \leq \exp\left(-\frac{\varepsilon^2 pn}{4}\right). \tag{3}$$

4 Contêineres para Grafos

Quando se trabalha com conjuntos independentes de um grafo G, tem-se um problema um pouco mais fácil de se lidar do que em hipergrafos k-uniformes com $k > 2$. Isto se deve ao fato de que, ao se verificar que um dado vértice v pertence a um conjunto independente $I \subset G$, automaticamente pode-se concluir que nenhum vértice da vizinhança de v pertence a I. Tal conclusão não pode ser feita para hipergrafos e, por isso, a construção dos contêineres se torna mais complicada.

A ideia do Lema de contêineres para grafos é a seguinte: Dado um grafo G e um conjunto independente I, em cada passo pergunta-se para um vértice de G, em uma ordem a ser especificada, se ele pertence a I. Se for o caso, então joga-se fora do contêiner toda sua vizinhança. Caso contrário, joga-se o próprio vértice para fora do contêiner. Assim, espera-se que no final tenha-se um conjunto que contém I e de tamanho significativamente menor do que o grafo G .

Para formalizar essa ideia, segue a definição da ordem em que os vértices de G serão dispostos e em seguida o resultado principal desta seção.

Definição 4.1 (Ordem do grau máximo)**.** *Dado um grafo (ou hipergrafo) G, a ordem do grau máximo em $V(G)$ é definida da seguinte maneira:*

1. *Fixe uma ordenação total de $V(G)$.*

2. *Para cada $j \in \{1, \cdots, v(G)\}$, seja u_j o vértice de maior grau em $G[V(G) \setminus \{u_1, \cdots, u_{j-1}\}]$; em caso de empate, escolha o vértice que vem antes na ordem fixada em (1).*

3. *A ordem do grau máximo em $V(\mathcal{G})$ é $(u_1, \cdots, u_{v(\mathcal{G})})$.*

Lema 4.2 (Contêineres para Grafos)**.** *Para todo $c > 0$, existe $\delta > 0$ tal que vale a afirmação seguinte. Seja H um grafo com grau médio ρ, $\Delta(H) \leq c \cdot \rho$ e fixe $\tau := \frac{2\delta}{\rho}$. Além disso, assuma que $\rho \leq v(H)/8c$. Então existe uma coleção $\mathcal{C}$ de subconjuntos de $V(H)$, com*

$$|\mathcal{C}| \leq \binom{v(H)}{\lceil \tau \cdot v(H) \rceil} \text{ tal que}$$

(i) Para todo $I \in \mathcal{I}(H)$, existe $C \in \mathcal{C}$ tal que $I \subseteq C$

(ii) $|C| \leq (1 - \delta)v(H)$

Demonstração. Fixado um conjunto $I \in \mathcal{I}(G)$, um algoritmo será executado para a construção de um contêiner e em cada passo teremos $V(G) = A_i \cup S_i \cup X_i$. O leitor deve pensar em A_i como o conjunto de vértices a serem verificados, em S_i como os que já foram analisados e que pertencem a I e em X_i como o conjunto dos vértices que com certeza não estão em I e, por isso, não farão parte do contêiner.

Estabeleça $A_0 := V(G)$ e $S_0 = X_0 = \emptyset$ e execute o algoritmo para $i \geq 1$ até comando PARE:

1) Se $X_{i-1} \geq \delta v(G)$, defina $A := A_{i-1}$ e $S := S_{i-1}$ e PARE

2) Seja v_i o primeiro vértice de I na ordem do grau máximo em $G[A_{i-1}]$ e *guarde-o*, $S_i := S_{i-1} \cup \{v_i\}$

3) Defina $W_i := \{u : u < v_i \text{ na ordem do grau máximo em } G[A_{i-1}]\}$

4) Jogue fora os vértices que precedem v_i e também sua vizinhança, i.e., $X_i := X_{i-1} \cup W_i \cup N(v_i)$

5) Remova os vértices analisados do conjunto de vértices ativos; $A_i := A_{i-1} \setminus (X_i \cup \{v_i\})$

Ao final deste algoritmo, temos trivialmente que $I \subseteq S(I) \cup A(I) =: C(I)$. Mas para mostrar que a coleção de contêineres $\mathcal{C} := \{C(I) : I \in \mathcal{I}(H)\}$ não é muito grande, vamos mostrar que se I, I' são conjuntos independentes de H com $S(I) = S(I')$, então $A(I) = A(I')$ e que $|S(I)| \leq \lceil \tau \cdot v(H) \rceil$ para todo $I \in \mathcal{I}(H)$.

Com efeito, se $I, I' \in \mathcal{I}(H)$ passarem pelo algoritmo e em algum ponto um vértice de $I \triangle I'$ for escolhido no passo 2), então ele seria colocado somente em $S(I)$ ou em $S(I')$ o que nos faz concluir que $S(I) \neq S(I')$. Como o algoritmo não depende dos vértices do conjunto independente que não apareceram, então podemos concluir que

$$S(I) = S(I') \implies A(I) = A(I').$$

Para terminar a prova, temos que mostrar que para cada vértice colocado em S, coloca-se pelo menos $\rho/2$ vértices em X, o que implica que em no máximo $\tau \cdot v(H)$ passos, X terá tamanho pelo menos $\delta v(H)$.

Para isso, vamos supor que H não é o grafo vazio, então $1 \leq \Delta(H) \leq c\rho$ e lembre que $\rho \leq e(H)/8c$. Note então que se $|S_i| \leq \tau \cdot v(H) \leq \delta v(H)$, $|X| \leq \delta v(H)$ e $|W_i| \leq \rho/2$, então podemos estimar quantas arestas foram jogadas foras no máximo, i.e:

$$\begin{aligned}
e(H[A \setminus W_i]) &\geq e(H) - (|S_i| + |X| + |W_i|) \cdot \Delta(H) \\
&\geq e(H) - (2\delta v(H) + \rho/2) \cdot c\rho \\
&\geq e(H) - (2\delta v(H) + \rho/2) \cdot 2c\frac{e(H)}{v(H)} \\
&\geq e(H) - 4\delta e(H) + e(H)/8 \geq e(H)/4
\end{aligned}$$

se $\delta \leq 1/32$. Em particular, o grau médio de $H[A \setminus W_i]$ é pelo menos $\rho/2$, o que implica que o próximo vértice na ordem, pertencente a I (pela definição de W_i) tem pelo menos $\rho/2$ vizinhos em $A \setminus W_i$. Assim, no próximo passo, $\rho/2$ vértices serão realocados para X.

Como cada contêiner foi indexado por um conjunto de tamanho menor que $\lceil \tau v(H) \rceil$, então é trivial a cota superior de $|\mathcal{C}|$. As outras propriedades são garantidas por construção.

$\square$

O que se consegue com este lema é que cada conjunto independente de H estará contido em um contêiner com um pouco menos de vértices que o próprio grafo H. Desta forma, será possível aplicar o lemma repetidas vezes, dado que se consiga escolher uma constante c de maneira esperta e garantir que em cada passo as hipóteses do lema sejam satisfeitas. Para ilustrar isto, segue a demonstração de um Teorema provado originalmente por Alon.

Teorema 4.3. *Seja G um grafo d regular com n vértices, com $d \gg \log n$. Então G tem no máximo $2^{n/2+o(n)}$ conjuntos independentes.*

Para provar este teorema, vamos demonstrar antes o seguinte lema de supersaturação, que fornecerá condições suficientes para se continuar aplicando o lema de contêineres.

Lema 4.4. *Seja G um grafo d regular com n vértices e $A \subset V(G)$ que satisfaz*

$$|A| \geq \frac{n}{2} + \varepsilon n, \text{ para algum } \varepsilon > 0.$$

Então $e(A) \geq \varepsilon dn$.

Demonstração. Ao somarmos os graus dos vértices de A, as arestas dentro de A serão contadas duas vezes, enquanto as arestas entre A e A^c serão contadas apenas uma vez. Então segue que:

$$\begin{gathered} 2e(A) + e(A, A^c) = \sum_{u \in A} d(u) = d|A| \\ 2e(A) \geq d|A| - (n - |A|)d \\ e(A) \geq d|A| - nd/2 \geq \varepsilon dn \end{gathered}$$

$\square$

Prova do Teorema 2.2: Vamos provar que, fixado $0 < \varepsilon < 1$, existe n_0 tal que para todo $n > n_0$, a quantidade de conjuntos independentes de G, $i(G)$, é no máximo $2^{n/2+2\varepsilon n}$. Fixado um $I \in \mathcal{I}(G)$, construiremos um contêiner com o seguinte algoritmo:

Fixe $c = \dfrac{2}{\varepsilon}$, $C_0 = V(G)$ e repita para $i \geq 1$ até atingir o comando PARE:

Se $|C_{i-1}| \geq \dfrac{n}{2} + \varepsilon n$, aplique o lema de contêineres em $G[C_{i-1}]$ e conjunto I e obtenha C_i, caso contrário PARE.

Se for provado que este algoritmo está bem definido, isto é, que as hipóteses do lema 2.1 são satisfeitas em cada passo, então se terá construído uma família $\mathcal{C}$ de subconjuntos de $V(G)$, chamados contêineres que satisfaz as seguintes propriedades:

(i) $|\mathcal{C}| \leq \binom{n}{\lceil (2\delta/\varepsilon) \cdot n \rceil}^{O(1)}$

(ii) Para todo $I \in \mathcal{I}(H)$, existe $C \in \mathcal{C}$ tal que $I \subseteq C$

(iii) $|C| \leq \dfrac{n}{2} + \varepsilon n$ para todo $C \in \mathcal{C}$

Com efeito, no primeiro passo, temos que o $\Delta(G) = d = \rho(G)$, e pode-se ver que:

$$\Delta(G) = d \leq \frac{2}{\varepsilon} \cdot d = c \cdot \rho(G)$$

Para $j > 1$, se a condição de continuidade do algoritmo for satisfeita, então pelo Lema 2.3, teremos que

$$\rho(G[C_{j-1}]) = \frac{e(C_{j-1})}{|C_{i-1}|} \geq \frac{\varepsilon d n}{2n} \geq \frac{\varepsilon d}{2}$$

Logo, podemos concluir que

$$\Delta(G[C_{j-1}]) \leq d = \frac{2}{\varepsilon} \cdot \frac{d\varepsilon}{2} \leq c \cdot \rho(G[C_{j-1}])$$

Finalmente, para terminar a demonstração do teorema, temos que $i(G)$ pode ser cotado pela soma do número de subconjuntos dos contêineres.

$$i(G) \leq \sum_{C \in \mathcal{C}} 2^{n/2+\varepsilon n}$$

Então só nós falta provar que

$$|\mathcal{C}| \leq 2^{\varepsilon n}$$

De fato, pela propriedade (i) da família $\mathcal{C}$ e usando a desigualdade (2), conclui-se que

$$|\mathcal{C}| \leq n^{O(\tau n)} = \exp(O(\tau n \log n)) \leq 2^{\varepsilon n},$$

pois $\tau = 4\delta/(\epsilon d)$ e $d \gg \log n$.

$\square$

5 Resultado Principal

Nesta seção, vamos apresentar a demonstração do resultado principal do trabalho. Entretanto, se o leitor não estiver familiarizado com o assunto, é recomendado que primeiro leia e entenda o enunciado e vá para a próxima seção para ver as aplicações.

Lema 5.1 (Contêineres)**.** *Para todo inteiro k e $c > 0$, existe $\delta > 0$ tal que vale a seguinte afirmação. Seja $q \in (0, 1)$ e suponha que $\mathcal{H}$ é um hipergrafo k-uniforme tal que para todo $l \in [k]$,*

$$\Delta_l(\mathcal{H}) \leq c \cdot q^{l-1} \tfrac{e(\mathcal{H})}{v(\mathcal{H})}.$$

Então existe uma família $\mathcal{S} \subseteq \binom{V(\mathcal{H})}{\leq (k-1)q \cdot v(\mathcal{H})}$ e funções $f_0 : \mathcal{S} \to \mathcal{P}(V(\mathcal{H}))$ e $g_0 : \mathcal{I}(\mathcal{H}) \to \mathcal{S}$ tal que para todo $I \in \mathcal{I}(\mathcal{H})$

$$g_0(I) \subseteq I \subseteq f_0(g_0(I)) \cup g_0(I) \qquad e \qquad |f_0(g_0(I))| \leq (1 - \delta) v(\mathcal{H}).$$

Além disso, se I, $I' \in \mathcal{I}(\mathcal{H})$, $g_0(I) \subseteq I'$ e $g_0(I') \subseteq I$, então $g_0(I) = g_0(I')$.

A princípio, o leitor pode não entender o que motiva a existência da família $\mathcal{S}$, de f_0 e g_0. Isto ficará evidenciadonas aplicações, mas uma pequena explicação é importante neste ponto. Primeiramente, é fácil ver a motivação do nome 'contêiner', pois $I \subset f_0(g_0(I)) \cup g_0(I)$. Mas então para que se precisa dessas funções g_0 e f_0 separadamente?

Note que $f_0(g_0(I))$ possivelmente tem tamanho com mesma ordem de grandeza que $v(\mathcal{H})$, enquanto os conjuntos da família $\mathcal{S}$ podem ser bem menores, dependendo de q. Como cada contêiner está indexado por um elemento de $\mathcal{S}$, isso possivelmente dá boas cotas sobre o número de contêineres.

A partir deste ponto, fixemos k, c e q como na proposição acima e podemos supor que $c \geq 1$. Dado $I \in \mathcal{H}$, usaremos um algoritmo para construir uma sequência $(B_{k-1}, \cdots, B_m)$ de subconjuntos de I para algum $m \in [k-1]$, tal que a cardinalidade de cada conjunto não supere $qv(\mathcal{H})$, e também uma sequência de hipergrafos $(\mathcal{H}_{k-1}, \cdots, \mathcal{H}_r)$, em que $r \in \{m, m+1\}$. Para descrever as propriedades que queremos nesses hipergrafos, precisamos definir os números Δ_l^i, em que $1 \leq l \leq i \leq k$ que serão importantes para se garantir o passo de indução.

Definição 5.2. *Para $l \in [k]$, seja $\Delta_l^k = \Delta_l(\mathcal{H})$ e para $i \in [k-1]$ e $l \in [i]$,*

$$\Delta_l^i = \max\{2 \cdot \Delta_{l+1}^{i+1}, q \cdot \Delta_l^{i+1}\} \tag{4}$$

Seja $b = qv(\mathcal{H})$ e para $i \in [k]$ seja $c_i = (ck2^{k+1})^{i-k}$. Queremos que a sequência de hipergrafos anteriormente citada tenha as seguintes propriedades:

(P1) $\mathcal{H}_i$ é i-uniforme e $V(\mathcal{H}_i) = V(\mathcal{H})$
(P2) I é um conjunto independente em $\mathcal{H}_i$
(P3) $\Delta_l(\mathcal{H}_i) \leq \Delta_l^i$ para todo $l \in [i]$
(P4) $e(\mathcal{H}_i) \geq c_i q^{k-i} e(\mathcal{H})$

Ao construir cada $\mathcal{H}_i$, será necessário identificar se uma aresta está saturada, ou seja, se algum subconjunto dela está contida em muitas outras arestas em $\mathcal{H}_i$. Assim, definiremos o seguintes conjuntos:

Definição 5.3. *Dado $i \in [k]$, um hipergrafo i-uniforme $\mathcal{G}$ e $l \in [i]$, seja*

$$M_l^i(\mathcal{G}) = \left\{ T \in \binom{V(\mathcal{G})}{l} : deg_{\mathcal{G}}(T) \geq \frac{\Delta_l^i}{2} \right\}.$$

5.1 O Algoritmo *Sychthe*

Como já ressaltado, na construção dos contêineres de grafos, sempre que o algoritmo verifica que um vértice está em um conjunto independente I, ele transfere todos os vértices da vizinhança para fora do contêiner, pois eles com certeza não pertencem a I. Tal afirmação não pode ser feita para o caso de um hipergrafo, mas o que se pode fazer é guardar a informação de quem está na vizinhança de dado vértice. A estratégia então é construir um hipergrafo novo em que as arestas são os conjuntos que, unidos ao vértice analisado, formam uma aresta no hipergrafo original.

Dado um hipergrafo $(i+1)$-uniforme $\mathcal{H}_{i+1}$ e um conjunto independente $I \in \mathcal{I}(\mathcal{H}_{i+1})$, seja $\mathcal{A}_{i+1}^{(0)} = \mathcal{H}_{i+1}$ e seja $\mathcal{H}_i^{(0)}$ o hipergrafo sem arestas e com conjunto de vértices $V(\mathcal{H}_{i+1})$. O leitor deve interpretar cada $\mathcal{A}_{i+1}^{(j)}$ como sendo o hipergrafo de vértices e arestas a serem analisados em cada passo e $\mathcal{H}_i^{(j+1)}$ como sendo um hipergrafo i-uniforme em construção, que guarda informações de acordo com quem está *ativo* em $\mathcal{A}_{i+1}^{(j)}$.

Um dos obstáculos para o bom funcionamento desse algoritmo é que, ao definir $\mathcal{H}_i$, poderia-se perder a regularidade que $\mathcal{H}_{i+1}$. Para lidar com tal problema, no passo (4) do algoritmo, retiraremos do conjunto de arestas ativas todas aquelas que possuem algum subconjunto desbalanceado, usando a definição (5.3). Para $j \in [b-1]$, o algoritmo é dado pelos seguintes passos:

1. Se $I \cap V(\mathcal{A}_{i+1}^{(j)}) = \emptyset$, então $\mathcal{H}_i = \mathcal{H}_i^{(0)}$, $A_i = \emptyset$ e $B_i = \{u_0, \cdots, u_{j-1}\}$ e pare o algoritmo.

2. Seja u_j o primeiro vértice de I na ordem do grau máximo em $V(\mathcal{A}_{i+1}^{(j)})$.

3. Seja $\mathcal{H}_i^{(j+1)}$ o hipergrafo com conjunto de vértices $V(\mathcal{H}_{i+1})$ e definido por:

$$\mathcal{H}_i^{(j+1)} = \mathcal{H}_i^{(j)} \cup \left\{ D \in \binom{V(\mathcal{H}_{i+1})}{i} : D \cup \{u_j\} \in \mathcal{A}_{i+1}^{(j)} \right\}.$$

4. Seja $\mathcal{A}_{i+1}^{(j+1)}$ o hipergrafo com conjunto de vértices $V(\mathcal{A}_{i+1}^{(j)}) \setminus W(u_j)$ (em que $W(u_j)$ é relativo à ordem do grau máximo de $V(\mathcal{A}_{i+1}^{(j)})$) definido por:

$$\mathcal{A}_{i+1}^{(j+1)} = \left\{ D \in \mathcal{A}_{i+1}^{(j)} : D \cap W(u_j) = \emptyset \text{ e } T \nsubseteq D \ \forall T \in \bigcup_{l=1}^{i} M_l^i(\mathcal{H}_i^{(j+1)}) \right\}.$$

Defina $\mathcal{H}_i = \mathcal{H}_i^{(b)}$, $A_i = V(\mathcal{A}_{i+1}^{(b)})$ e $B_i = \{u_0, \cdots, u_{b-1}\}$.

As próximas quatro proposições estabelecem propriedades sobre os hipergrafos construídos através desse algoritmo, a fim de provar o Lema 5.1.

Proposição 5.4. *Para todo $i \in [k-1]$ valem as seguintes afirmações:*
(a) $\mathcal{H}_i$ é i-uniforme e $V(\mathcal{H}_i) = V(\mathcal{H})$.
(b) Se $I \in \mathcal{I}(\mathcal{H}_{i+1})$, então $I \in \mathcal{I}(\mathcal{H}_i)$.
(c) $B_i \subseteq I \subseteq A_i \cup B_i$.
(d) O hipergrafo $\mathcal{H}_i$ e o conjunto A_i dependem somente de $\mathcal{H}_{i+1}$ e de B_i.

Demonstração. (a) Segue direto do algoritmo, pois $\mathcal{H}_k = \mathcal{H}$ e porque o conjunto de vértices é sempre o mesmo e as arestas de $\mathcal{H}_i$ são tomadas sempre de tamanho i, no passo (3) do algoritmo.

(b) Se I contém uma aresta de $\mathcal{H}_i$, então $D \cup u_j \in \mathcal{A}_{i+1}^{(j)}$, para algum $j \in [b-1]$. Da maneira como foi definido $\mathcal{A}_{i+1}^{(j)}$, vemos que de fato $D \cup u_j$ é uma aresta de $\mathcal{H}_{i+1}$. Como D é formado por vértices de I e $u_j \in I$, temos uma contradição.

(c) Obviamente $B_i \subseteq I$, pois no passo (2), u_j é sempre escolhido dentre os vértices de I. Podemos ver que $I \subseteq A_i \cup B_i$, pois podemos afirmar de modo mais geral que $I \subseteq V(\mathcal{A}_{i+1}^{(j)}) \cup \{u_0, \cdots, u_{j-1}\}$ para todo $j \in \{0\} \cup [b]$. Com efeito, para $j = 0$ é trivial e por indução vemos que é verdade, pelo passo (4).

(d) Note que no segundo passo, $W(u_j) \cap B_i = \{u_j\}$. Assim, no passo (3) podemos obter $\mathcal{H}_i^{(j+1)}$ apenas com B_i, $\mathcal{H}_i^{(j)}$ e $\mathcal{A}_{i+1}^{(j)}$, enquanto no passo (4) $\mathcal{A}_{i+1}^{(j+1)}$ pode ser recuperado com $\mathcal{A}_{i+1}^{(j)}$ e $\mathcal{H}_i^{(j+1)}$. E isso completa a prova por indução, exceto se $|B_i| < b$, que não gera problemas, pois nesse caso $\mathcal{H}_i$ e A_i são vazios.

□

Proposição 5.5. *Suponha que ao executar o Algoritmo Scythe com as entradas $(\mathcal{H}_{i+1}, I)$ e $(\mathcal{H}_{i+1}, I')$, as saídas sejam $(A_i, B_i, \mathcal{H}_i)$ e $(A_i', B_i', \mathcal{H}_i')$, respectivamente. Se $B_i \subseteq I'$ e $B_i' \subseteq I$, então $(A_i, B_i, \mathcal{H}_i) = (A_i', B_i', \mathcal{H}_i')$.*

Demonstração. Vamos supor primeiramente, que $|B_i| < b$. Nesse caso, o algoritmo vai parar no passo (1) em algum momento e $A_i = \emptyset$. Pelo lema anterior, $B_i = I$. Como $B_i' \subseteq I = B_i$, temos que $|B_i'| < b$ e pelo mesmo argumento, então $B_i' = I'$ e $B_i = B_i'$. Por outro lado, se $|B_i| = |B_i'| = b$ e $B_i \neq B_i'$, então ao colocar os vértices dos conjuntos na ordem do grau máximo em $\mathcal{H}_{i+1}$, podemos escolher j como o menor número tal que $u_j \neq u_j'$ e podemos supor, sem perda de generalidade que u_j vem antes na ordem. Assim, quando o algoritmo fosse executado com a entrada $(\mathcal{H}_{i+1}, I')$ e ele já tivesse encontrado o conjunto $\{u_0, \cdots, u_{j-1}\}$ ele iria procurar o primeiro vértice em I na ordem do grau máximo de $\mathcal{A}_{i+1}^{(j)}$ e esse vértice seria u_j, ao invés de u_j', o que é uma contradição. Sendo assim, como podemos concluir em ambos os casos que $B_i = B_i'$, então pelo lema anterior podemos concluir que $(A_i, B_i, \mathcal{H}_i) = (A_i', B_i', \mathcal{H}_i')$.

□

O próximo resultado é essencial para mostrar que o hipergrafo construído pelo algoritmo preserva alguma regularidade de $\mathcal{H}_{i+1}$, estabelecida pelos valores Δ_l^i.

Proposição 5.6. *Se $\Delta_{l+1}(\mathcal{H}_{i+1}) \leq \Delta_{l+1}^{i+1}$ para algum $l \in [i]$, então $\Delta_l(\mathcal{H}_i) \leq \Delta_l^i$.*

Demonstração. Seja $T \in \binom{V(\mathcal{H}_{i+1})}{l}$. Se para algum $j \in [m]$, tivermos que

$$\deg_{\mathcal{H}_i^{(j)}}(T) \geq \frac{\Delta_l^i}{2},$$

então, pelo passo (4) do algoritmo, nenhuma aresta de $\mathcal{A}_{i+1}^{(j)}$ vai conter T, o que implica que $\deg_{\mathcal{H}_i}(T) = \deg_{\mathcal{H}_i^{(j)}}(T)$. Além disso, as arestas adicionadas em $\mathcal{H}_i^{(j-1)}$ para se construir $\mathcal{H}_i^{(j)}$, são os conjuntos D tal que $D \cup \{u_j\} \in \mathcal{A}_{i+1}^{(j-1)}$ no passo (3). Logo, temos que

$$\deg_{\mathcal{H}_i^{(j)}}(T) - \deg_{\mathcal{H}_i^{(j-1)}}(T) \leq \deg_{\mathcal{H}_{i+1}}(T \cup \{u_j\}) \leq \Delta_{l+1}(\mathcal{H}_{i+1})$$
$$\deg_{\mathcal{H}_i^{(j)}}(T) \leq \deg_{\mathcal{H}_i^{(j-1)}}(T) + \Delta_{l+1}(\mathcal{H}_{i+1})$$

Logo, podemos concluir que

$$\Delta_l(\mathcal{H}_i) \leq \frac{\Delta_l^i}{2} + \Delta_{l+1}(\mathcal{H}_{i+1}) \leq \frac{\Delta_l^i}{2} + \Delta_{l+1}^{i+1} \leq \Delta_l^i,$$

por causa de (4). □

Agora vamos provar que ao se aplicar o algoritmo, ou ele vai fornecer condições para que ele seja aplicado de novo ou ele terá terminado de maneira satisfatória.

Proposição 5.7. $\Delta_1^i \leq c2^k q^{k-i}\dfrac{e(\mathcal{H})}{v(\mathcal{H})}$ *para todo* $i \in \{1, \ldots, k\}$.

Demonstração. A partir da Definição 2.4, é fácil provar por indução que para todo $i \in [k]$ e $l \in [i]$ existe $d \in [k-i]$ tal que

$$\Delta_i^l = 2^d q^{k-i-d} \Delta_{d+l}(\mathcal{H}).$$

Usando a hipótese de que $\Delta_l(\mathcal{H}) \leq c \cdot q^{l-1}\frac{e(\mathcal{H})}{v(\mathcal{H})}$, segue que

$$\Delta_1^l \leq \max_{0 \leq d \leq k-1} \{2^d q^{k-i-d} \Delta_{d+1}(\mathcal{H})\}$$

$$\leq \max_{0 \leq d \leq k-1} \left\{2^d q^{k-i-d} c \cdot q^d \frac{e(\mathcal{H})}{v(\mathcal{H})}\right\} \leq c2^k q^{k-i}\frac{e(\mathcal{H})}{v(\mathcal{H})}.$$

□

Proposição 5.8. *Seja* $i \in [k-1]$ *e suponha que* $e(\mathcal{H}_{i+1}) \geq c_{i+1}q^{k-(i+1)}e(\mathcal{H})$ *e que* $\Delta_l(\mathcal{H}_{i+1}) \leq \Delta_l^{i+1}$ *para todo* $l \in [i+1]$. *Então uma das seguintes afirmações é verdadeira*

i) $e(\mathcal{H}_i) \geq \dfrac{q}{c \cdot 2^{k+1}k} e(\mathcal{H}_{i+1}) \geq c_i q^{k-i} e(\mathcal{H})$

ii) $|A_i| \leq (1 - c_i)v(\mathcal{H})$

Demonstração. Note que podemos supor que o algoritmo não para no passo (1), pois neste caso, $|A_i| = 0$. Agora será necessário usar a propriedade mais essencial da ordem do grau máximo, que é que se os vértices de um hipergrafo $\mathcal{G}$ estão na ordem $v_1, v_2, ..., v_{v(\mathcal{G})}$, então para todo $s \in [v(\mathcal{G})]$, no sub-hipergrafo induzido por $\{v_s, ..., v_{v(\mathcal{G})}\}$ o vértice v_l tem o maior grau e isso implica, em particular, que seu grau é maior que o grau médio desse hipergrafo.

Assim, para $j \in \{0, ..., b-1\}$, quando u_j é encontrado no passo (2) do algoritmo, então seu grau no hipergrafo $\tilde{\mathcal{A}}_{i+1}^{(j)}$ (como no parágrafo acima) é maior que o grau médio. Assim, temos que

$$e(\mathcal{H}_i^{(j+1)}) - e(\mathcal{H}_i^{(j)}) = \deg_{\mathcal{A}_{i+1}^{(j)}}(u_j) \geq \deg_{\tilde{\mathcal{A}}_{i+1}^{(j)}}(u_j) \geq$$

$$\frac{(i+1)e(\tilde{\mathcal{A}}_{i+1}^{(j)})}{v(\tilde{\mathcal{A}}_{i+1}^{(j)})} \geq \frac{(i+1)e(\mathcal{A}_{i+1}^{(j+1)})}{v(\mathcal{H})},$$

em que a última desigualdade é verdade, pois na definição de $\mathcal{A}_{i+1}^{(j+1)}$, nenhuma aresta contém elementos de $W(u_j)$.

Agora, se $(i+1)e(\mathcal{A}_{i+1}^{(j+1)}) \geq e(\mathcal{H}_{i+1})$ para todo $j \in \{0, ..., b-1\}$, então podemos somar o resultado acima sobre todos os j e, lembrando que $\mathcal{H}_i^{(0)}$ é vazio e $\mathcal{H}_i = \mathcal{H}_i^{(b)}$, então

$$e(\mathcal{H}_i) \geq \sum_{j=0}^{b-1} \frac{(i+1)e(\mathcal{A}_{i+1}^{(j+1)})}{v(\mathcal{H})} \geq b \cdot \frac{e(\mathcal{H}_{i+1})}{v(\mathcal{H})} = q \cdot e(\mathcal{H}_{i+1}),$$

o que nos dá o resultado desejado. Caso contrário, como $e(\mathcal{A}_{i+1}^{(b)}) \leq e(\mathcal{A}_{i+1}^{(j+1)})$ para todo $j \in \{0, ..., b-1\}$, então podemos supor que

$$e(\mathcal{A}_{i+1}^{(b)}) < \frac{e(\mathcal{H}_{i+1})}{i+1}.$$

Intuitivamente, o que isto quer dizer é que muitas arestas foram jogadas fora, o que pode significar duas coisas: ou $|M_l^i(\mathcal{H}_i)|$ é grande demais para algum $l \in [i]$, o que exclui todas as arestas que contém esses elementos, ou então os conjuntos $W(u_j)$ são muito grandes. Para formalizar isso, vamos analisar quantas arestas são jogadas fora em cada iteração, i.e, $e(\mathcal{A}_{i+1}^{(j)}) - e(\mathcal{A}_{i+1}^{(j+1)})$. Este valor é no máximo

$$|W(u_j)| \cdot \Delta_1(\mathcal{H}_{i+1}) + \sum_{l=1}^{i} |M_l^i(\mathcal{H}_i^{(j+1)}) \setminus M_l^i(\mathcal{H}_i^{(j)})| \cdot \Delta_l(\mathcal{H}_{i+1}). \qquad (5)$$

Esta desigualdade se deve ao fato de que no passo (4) do algoritmo uma aresta é jogada fora ou porque ela continha algum elemento de $W(u_j)$ ou porque ela continha um elemento de $M_l^i(\mathcal{H}_i^{(j+1)})$, lembrando de se descontar os que já haviam sido retirados por este motivo no passo anterior.

Note que, como $\mathcal{A}_{i+1}^{(0)} = \mathcal{H}_{i+1}$, $\mathcal{H}_i^{(b)} = \mathcal{H}_i$ e $\Delta_l(\mathcal{H}_{i+1}) \leq \Delta_l^{i+1}$, então somando (5) sobre j, temos que

$$e(\mathcal{H}_{i+1}) - e(\mathcal{A}_{i+1}^{(b)}) \leq \sum_{j=0}^{b-1} |W(u_j)| \cdot \Delta_1^{i+1} + \sum_{l=1}^{i} |M_l^i(\mathcal{H}_i)| \cdot \Delta_l^{i+1} \qquad (6)$$

Se $|M_l^i(\mathcal{H}_i)| \cdot \Delta_l^{i+1} < \frac{1}{2(i+1)} e(\mathcal{H}_{i+1})$ para todo $l \in [i]$, usando o fato de que $e(\mathcal{A}_{i+1}^{(b)}) \leq e(\mathcal{H}_{i+1})/(i+1)$ então (6) se torna

$$\sum_{j=0}^{b-1} |W(u_j)| \geq \frac{i}{2(i+1)\Delta_1^{i+1}} \cdot e(\mathcal{H}_{i+1}) \geq \frac{1}{4\Delta_1^{i+1}} \cdot e(\mathcal{H}_{i+1}).$$

Se $M_l^i(\mathcal{H}_i)$ é grande para algum $l \in [i]$, então devemos ser capazes de dar uma cota inferior sobre o número de arestas de $\mathcal{H}_i$. Por outro lado, se os conjuntos $W(u_j)$ são grandes, então quer dizer que 'jogamos fora' muitos vértices, então podemos dizer alguma coisa sobre A_i.

Caso 1: $|M_l^i(\mathcal{H}_i)| \geq \frac{1}{2(i+1)\Delta_l^{i+1}} \cdot e(\mathcal{H}_{i+1})$ para algum $l \in [i]$

Neste caso, como $\deg_{\mathcal{H}_i}(T) \geq \Delta_l^i/2$ para todo $T \in M_l^i(\mathcal{H}_i)$, então

$$e(\mathcal{H}_i) = \binom{i}{l}^{-1} \sum_{T \in \binom{V(\mathcal{H})}{l}} \deg_{\mathcal{H}_i}(T) \geq \frac{|M_l^i(\mathcal{H}_i)| \cdot \Delta_l^i}{2\binom{i}{l}}.$$

Com isso, terminamos a análise deste caso $|M_l^i(\mathcal{H}_i)$ e o fato de que, pela definição, $\Delta_l^i \geq q\Delta_l^{i+1}$, pois com isto obtemos

$$e(\mathcal{H}_i) \geq \frac{e(\mathcal{H}_{i+1})}{4(i+1)\binom{i}{l}} \cdot \frac{\Delta_l^i}{\Delta_l^{i+1}} \geq \frac{q}{2^{i+2}(i+1)} \cdot e(\mathcal{H}_{i+1}) \geq \frac{q}{c2^{k+1}k} \cdot e(\mathcal{H}_{i+1}),$$

já que $\binom{i}{l} \leq 2^i$, $(i+1) \leq k$ e $c \geq 1$.

Caso 2: $\sum_{j=0}^{b-1} |W(u_j)| \geq \frac{1}{4\Delta_1^{i+1}} \cdot e(\mathcal{H}_{i+1})$ Note que na j-ésima iteração do algoritmo, os vértices jogados fora são exatamente $W(u_j)$. Como $v(\mathcal{A}_{i+1}^{(0)}) = v(\mathcal{H})$ e $A_i = v(\mathcal{A}_{i+1}^{(b)})$, então

$$v(\mathcal{H}) - |A_i| = v(\mathcal{A}_{i+1}^{(0)}) - v(\mathcal{A}_{i+1}^{(b)}) = \sum_{j=0}^{b-1} |W(u_j)| \geq \frac{e(\mathcal{H}_{i+1})}{4\Delta_1^{i+1}}.$$

Por outro lado, temos que $e(\mathcal{H}_{i+1}) \geq c_{i+1}q^{k-(i+1)}e(\mathcal{H})$ e, pela proposição 5.7, que $\Delta_1^{i+1} \leq c2^k q^{k-(i+1)} \cdot \frac{e(\mathcal{H})}{v(\mathcal{H})}$, o que resulta em

$$v(\mathcal{H}) - |A_i| \geq \frac{c_{i+1}v(\mathcal{H})}{c2^{k+2}} \geq c_i v(\mathcal{H}).$$

$\square$

5.2 Prova do Lema 5.1

Seja k um inteiro e $c > 0$. Além disso, seja $q \in (0,1)$ e $\mathcal{H}$ um hipergrafo k-uniforme que satisfaz as hipóteses do Lema 5.1. A prova do Teorema é só aplicar iterativamente o algoritmo de foice até que algum dos A_i seja pequeno o suficiente ou até chegar em um hipergrafo 1-uniforme. Mais precisamente, defina $\delta = (ck2^{k+1})^{-k}$, $\mathcal{H}_k = \mathcal{H}$ e $i = k-1$, fixe $I \in \mathcal{I}(\mathcal{H})$ e execute o seguinte algoritmo:

1) Aplique o algoritmo Scythe a I e $\mathcal{H}_{i+1}$ e obtenha $\mathcal{H}_i$, A_i e B_i.

2) Se $|A_i| \leq (1-\delta)v(\mathcal{H})$, defina $m = i$ e

$$g_0(I) = B_{k-1} \cup ... \cup B_m \qquad f_0^*(I) = A_m$$

e PARE.

3) Se $i > 1$, $i \leftarrow (i-1)$ e volte para o passo 1). Caso contrário, defina $m = 1$,

$$g_0(I) = B_{k-1} \cup ... \cup B_m \qquad f_0^*(I) = \big\{v \in V(\mathcal{H}_1) : \{v\} \notin e(\mathcal{H}_1)\big\}$$

e PARE.

Note que enquanto este algoritmo estiver sendo executado, os hipergrafos obtidos em cada passo satisfazem as propriedades (P1)-(P4), a serem listadas novamente abaixo:

(P1) $\mathcal{H}_i$ é i-uniforme e $V(\mathcal{H}_i) = V(\mathcal{H})$
(P2) I é um conjunto independente em $\mathcal{H}_i$
(P3) $\Delta_l(\mathcal{H}_i) \leq \Delta_l^i$ para todo $l \in [i]$
(P4) $e(\mathcal{H}_i) \geq c_i q^{k-i} e(\mathcal{H})$

Note que $\mathcal{H}_k$ satisfaz essas propriedades naturalmente pelas definições e enquanto o passo de indução é dado pelas proposições 5.4, que garante que I continua sendo um conjunto independente em todos os passos, pela proposição 5.6 que garante que (P3) se mantém e pela proposição 5.8 que diz que enquanto $|A_{i+1}| > (1-\delta)v(\mathcal{H}) > (1-c_{i+1})v(\mathcal{H})$, então $\mathcal{H}_i$ tem a propriedade (P4). Agora, se definimos

$$\mathcal{S} = \{g_0(I) : I \in \mathcal{I}(\mathcal{H})\},$$

precisamos provar que $f_0(S) := f_0^*(I)$, para I, tal que $g_0(I) = S$ está bem definida. Mas isso é garantido por indução pela propriedade (d) da Proposição 5.4. Para terminar a prova, note que

(i) $S \leq (k-1)pv(\mathcal{H})$, pois $|B_i| \leq pv(\mathcal{H})$ para todo $i \in \{m, ..., k-1\}$.

(ii) $g_0(S) \subset I \subset g_0(S) \cup f_0(g_0(S))$, pela Proposição 5.4 (c) e pelo fato de I ser um conjunto independente em $\mathcal{H}_1$, caso $m = 1$.

(iii) $|f_0(g_0(S))| \leq (1-\delta)v(\mathcal{H})$. Se para algum $i \in [k-1]$, $|A_i| \leq (1-\delta v(\mathcal{H})$, então é trivial. Caso contrário, temos que

$$|\big\{v \in V(\mathcal{H}_1) : \{v\} \in e(\mathcal{H}_1)\big\}| \geq \frac{e(\mathcal{H}_1)}{\Delta_1(\mathcal{H}_1)} \geq \delta v(\mathcal{H}),$$

em que a última desigualdade segue das proposições 5.7 e 5.8, além das definições de c_1 e δ.

(iv) $g_0(I') \subset I$ e $g_0(I) \subset I'$ implica que $g_0(I) = g_0(I')$, que segue por indução garantida pela Proposição 5.5.

6 Turán para grafos aleatórios

Teorema 6.1. *Seja H um grafo com $\Delta(H) \geq 2$. Para todo $\delta > 0$ existe $C > 0$ tal que $p \geq Cn^{-1/m_2(H)}$ então com alta probabilidade*

$$ex(G_{n,p}, H) \leq \left(1 - \frac{1}{\chi(H)-1} + \delta\right)p\binom{n}{2}.$$

Para provar tal teorema, vamos contruir uma classe de contêineres adequados. Para isso, precisamos primeiro da seguinte definição:

Definição 6.2 (Hipergrafo de cópias de H)**.** *Dado um grafo H, definimos o hipergrafo $e(H)$-uniforme com $V(\mathcal{H}) = E(K_n)$ e as arestas são*

$$E(\mathcal{H}) = \left\{H' \in \binom{E(K_n)}{e(H)} : H' \text{ é uma cópia de } H\right\}$$

A princípio, esta definição pode parecer confusa, mas note que se I é um conjunto independente de $\mathcal{H}$, então ele é, na verdade, um conjunto de arestas (ou pode ser enxergado como um grafo) que não contém uma cópia de H.

Assim, quando os contêineres forem construídos, teremos que cada grafo sem cópias de H estará contido em um contêiner e que o número de contêineres não será grande demais.

Para o resto desta seção, H sempre vai ser um grafo com $\Delta(H) \geq 2$ e $\mathcal{F}_\varepsilon \subset \mathcal{P}(V(\mathcal{H}))$ é família de subconjuntos de $V(\mathcal{H})$ que contém menos que $\varepsilon n^{v(H)}$ arestas em $\mathcal{H}$, i.e, grafos com n vértices com menos que $\varepsilon n^{v(H)}$ cópias de H.

Proposição 6.3. *Seja $\mathcal{H}$ o hipergrafo de cópias de H e defina $q := n^{-1/m_2(H)}$. Para todo $\delta > 0$, existem $\varepsilon, C' > 0$, uma família $\mathcal{S} \subseteq \mathcal{P}(V(\mathcal{H}))$ e funções $f : \mathcal{S} \to \mathcal{F}_\varepsilon$ e $g : \mathcal{I}(\mathcal{H}) \to \mathcal{S}$ que satisfazem as seguintes propriedades:*

(i) $g(I) \subseteq I \subseteq g(I) \cup f(g(I))$, $\forall I \in \mathcal{I}(\mathcal{H})$

(ii) $|S| \leq C'qv(\mathcal{H})$, $\forall S \in \mathcal{S}$

(iii) $|f(S)| \leq \left(1 - \dfrac{1}{\chi(H)-1} + \delta\right)v(\mathcal{H})$, $\forall S \in \mathcal{S}$

O que esta proposição nos dá é que não só cada grafo sem cópias de H está contido em um contêiner e que eles são indexados pela família $\mathcal{S}$, além de cotar o tamanho dos contêineres.

Para motivar a prova desta proposição, vamos demonstrar o Teorema 6.1 a partir dela. O leitor deve lembrar que se $\mathcal{H}$ é o hipergrafo de cópias de H, então $V(\mathcal{H}) = E(K_n)$, logo podemos enxergar qualquer subconjunto de $V(\mathcal{H})$ como um grafo. Tal abuso é feito para não sobrecarregar o leitor com notação.

Demonstração do Teorema 6.1. Seja $\delta' = \frac{\delta}{3}$ e C' obtida pela proposição 6.3 aplicada a δ'. Então temos $\mathcal{S}$, uma família de suconjuntos de $V(\mathcal{H})$ e funções $f : \mathcal{S} \to \mathcal{F}_\varepsilon$ e $g : \mathcal{I}(\mathcal{H}) \to \mathcal{S}$ que satisfazem as propriedades da proposição 6.3.

Para o resto da prova, considere $C = C'/\delta^3$, $p \geq C/n^{1/m_2(H)}$. Além disso, seja $m = \left(1 - \frac{1}{\chi(H)-1} + \delta\right)pv(\mathcal{H})$ e $\mathcal{I}_m(\mathcal{H}) = \{I \in \mathcal{I}(\mathcal{H}) : |I| = m\}$. Temos que

$$\begin{aligned}
\mathbb{P}(ex(G_{n,p}, H) \geq m) &= \mathbb{P}(\exists I \in \mathcal{I}_m(\mathcal{H}) : I \subseteq G_{n,p}) \\
&\leq \sum_{S \in \mathcal{S}} \mathbb{P}(\exists I \in \mathcal{I}_m(\mathcal{H}) : I \subseteq G_{n,p} \text{ e } g(I) = S) \\
&\leq \sum_{S \in \mathcal{S}} \mathbb{P}(S \subseteq G_{n,p}) \cdot \mathbb{P}(|G_{n,p} \cap f(S)| \geq m - |S|)
\end{aligned}$$

Agora, cada termo da multiplicação acima será analisado separadamente, começando pelo segundo termo:

$$|S| \leq C'qv(\mathcal{H}) = C\delta^3 n^{-1/m_2(H)}v(\mathcal{H}) \leq \delta^3 pv(\mathcal{H}) \leq \delta' pv(\mathcal{H})$$

A última desigualdade segue do fato de que podemos, sem perda de generalidade, tomar δ suficientemente pequeno de forma que $\delta^3 < \delta/3$. Seja $\alpha = 1 - \dfrac{1}{\chi(H) - 1}$. Então

$$m = (\alpha + 3\delta')pv(\mathcal{H})$$
$$m - |S| \geq (\alpha + 2\delta')pv(\mathcal{H})$$

Por outro lado, pela propriedade (iii), $|f(S)| \leq (\alpha + \delta')v(\mathcal{H})$ e, já que $|G_{n,p} \cap f(S)| =_d \mathrm{Bin}(|f(S)|, p)$, temos que

$$\mathbb{E}(|G_{n,p} \cap f(S)|) \leq (\alpha + \delta')pv(\mathcal{H})$$

Logo, a probabilidade de $G_{n,p} \cap f(S)$ ter pelo menos $\delta' pv(\mathcal{H})$ mais arestas que a esperança, pode ser cotada por Chernoff:

$$\mathbb{P}(|G_{n,p} \cap f(S)| \geq m - |S|) \leq \exp\left(- \frac{(\delta')^2 pv(\mathcal{H})}{4} \right) = \exp\left(- \frac{\delta^2 pv(\mathcal{H})}{36} \right)$$

No restante desta demonstração, fica clara a utilidade da família $\mathcal{S}$, pois ela dá boas cotas para a quantidade de contêineres. Como não tem-se muita informação sobre ela, além do tamanho de cada conjunto $S \in \mathcal{S}$, vamos substituir o somatório indexado sobre esta família pelo somatório sobre todos os conjuntos de tamanho menor que $\delta^3 pv(\mathcal{H})$:

$$\begin{aligned}
\sum_{S \in \mathcal{S}} \mathbb{P}(S \subseteq G_{n,p}) &\leq \sum_{s=0}^{\delta^3 pv(\mathcal{H})} \binom{v(\mathcal{H})}{s} p^s \\
&\leq \sum_{s=0}^{\delta^3 pv(\mathcal{H})} \left(\frac{epv(\mathcal{H})}{s} \right)^s \\
&\leq pv(\mathcal{H}) \left(\frac{e}{\delta^3} \right)^{\delta^3 pv(\mathcal{H})}
\end{aligned}$$

A última desigualdade segue do fato de que a função $(k/s)^s$ é crescente em $(0, k/e)$. Se as desigualdades obtidas até agora forem agregadas, falta pouco para concluir a prova:

$$\begin{aligned}\mathbb{P}(\mathrm{ex}(G_{n,p}, H) \geq m) &\leq \exp\left[pv(\mathcal{H})\left(\frac{\log(pv(\mathcal{H}))}{pv(\mathcal{H})} + \delta^3 - 3\delta^3\log\delta - \frac{\delta^2}{36}\right)\right] \\ &\leq \exp(-\delta^3 pv(\mathcal{H})).\end{aligned}$$

A última desigualdade é verdade se tomamos δ suficientemente pequeno e n suficientemente grande, contanto que provemos que $\lim_{n\to\infty} pv(\mathcal{H}) = +\infty$.

Com efeito, como estamos considerando que $\Delta(H) \geq 2$, então, apenas pela definição, $m_2(H) \geq 1$, logo

$$p \geq Cn^{-1} \implies \lim_{n\to\infty} pv(\mathcal{H}) = +\infty$$

$\square$

Como já foi evidenciado na seção 4, construiremos os contêineres iterando o Lema de Contêineres um número finito de vezes. Para que isso seja possível, será preciso provar que em cada iteração, exceto possivelmente no final, o Contêiner obtido mantém pelo menos uma proporção constante do grau médio do hipergrafo de cópias de H.

Isso nos leva então ao seguinte lema de supersaturação, provado originalmente por Erdős e Simonovits [12]:

Lema 6.4 (Supersaturação de Erdős-Simonovits)**.** *Para todo $\delta > 0$, existe $\varepsilon > 0$ tal que a afirmação seguinte é verdadeira, para n grande o suficiente. Se G é um grafo que satisfaz*

$$e(G) \geq \left(1 - \frac{1}{\chi(H) - 1} + \delta\right)\binom{n}{2},$$

então G contém pelo menos $\varepsilon n^{v(H)}$ cópias de H.

Demonstração. Pelo Teorema 2.1, podemos escolher um k inteiro tal que

$$ex(k, H) \geq \left(1 - \frac{1}{\chi(H) - 1} - \frac{\delta}{2}\right)\binom{k}{2}.$$

Fixado este k, a prova segue para qualquer $n > k$. Seja S um subconjunto aleatório de $V(G)$ escolhido uniformemente entre $\binom{V(G)}{k}$ e $X = e(S)$. Por um lado, temos que

$$\mathbb{E}[X] = \sum_{S' \in \binom{V(G)}{k}} e(S')\mathbb{P}(S = S') = \frac{1}{\binom{n}{k}} \sum_{e \in E(G)} \left|\left\{S' \in \binom{V(G)}{k} : e \subset S'\right\}\right|$$

$$\mathbb{E}[X] \geq \left(1 - \frac{1}{\chi(H)-1} + \delta\right)\frac{\binom{n}{2}\binom{n-2}{k-2}}{\binom{n}{k}} \tag{7}$$

Por outro lado, separando os eventos $Y = [X \geq \mathrm{ex}(k,H)]$ e o seu complementar, podemos estimar a esperança por

$$\mathbb{E}[X] \leq \mathbb{P}(Y)\binom{k}{2} + \mathbb{P}(Y^{\mathsf{c}})\ \mathrm{ex}(k,H)$$

$$\mathbb{E}[X] \leq \mathbb{P}(Y)\binom{k}{2} + \mathrm{ex}(k,H). \tag{8}$$

Juntando (7) e (8), temos

$$\mathbb{P}(Y) \geq \left(1 - \frac{1}{\chi(H)-1} + \delta\right)\frac{\binom{n}{2}\binom{n-2}{k-2}}{\binom{k}{2}\binom{n}{k}} - \left(1 - \frac{1}{\chi(H)-1} - \frac{\delta}{2}\right)$$

Note que, expandindo os termos binomiais, tem-se que

$$\frac{\binom{n}{2}\binom{n-2}{k-2}}{\binom{k}{2}\binom{n}{k}} = 1.$$

Pode-se concluir facilmente então que

$$|\{S' \in \binom{V(G)}{k} : e(S') \geq \mathrm{ex}(k,H)\}| \geq \frac{\delta}{2}\binom{n}{k}.$$

Além disso, fixada uma cópia de H em G, H',

$$|\{S' \in \binom{V(G)}{k} : H' \subseteq S'\}| = \binom{n-k}{k-v(H)}$$

Portanto, por (2) consegue-se constantes c e c' tal que o número de cópias de H em G pode ser estimado por baixo por

$$\frac{\frac{\delta}{2}\binom{n}{k}}{\binom{n-k}{k-v(H)}} \geq \frac{\delta}{2}\frac{cn^k}{c'n^{k-v(H)}} = \left(\frac{\delta c}{2c'}\right)n^{v(H)}$$

□

Agora, precisamos provar também o hipergrafo de cópias de H satisfaz as hipóteses do Lema de Contêineres.

Proposição 6.5. *Seja $\mathcal{H}$ o hipergrafo de cópias de H. Existe $c > 0$ tal que se $q = n^{-1/m_2(H)}$, então*

$$\Delta_l(\mathcal{H}) \leq cq^{l-1}\frac{e(\mathcal{H})}{v(\mathcal{H})}.$$

Demonstração. Note que

$$v(\mathcal{H}) = \binom{n}{2} = \Theta(n^2)$$
$$e(\mathcal{H}) = \frac{(v(H)!)}{|Aut(H)|}\binom{n}{v(H)} = \Theta(n^{v(H)})$$
$$\implies \frac{e(\mathcal{H})}{v(\mathcal{H})} \geq c' n^{v(H)-2}. \tag{9}$$

Além disso, segue da definição de $m_2(H)$ que para qualquer $H' \subseteq H$ é verdade que $d(H') \leq m_2(H)$. Então para n suficientemente grande, temos que

$$(n^{1/d(H')-1/m_2(H)})^{e(H')-1} \geq 1.$$

Assim, é fácil concluir que

$$q^{e(H')-1} n^{v(H')-2} \geq 1 \quad \forall H' \subseteq H. \tag{10}$$

Pela definição de $\Delta_l(\mathcal{H})$ e usando (2), temos que existe $c'' > 0$ tal que

$$\Delta_l(\mathcal{H}) = \max\left\{\binom{n}{v(H)-v(H')} : H' \subseteq H \text{ com } e(H') = l\right\}$$
$$\Delta_l(\mathcal{H}) \leq c'' \max\{n^{v(H)-v(H')} : H' \subseteq H \text{ com } e(H') = l\}$$

Logo, juntando a equação acima com (9) e (10), terminamos a prova, pois

$$\begin{aligned}
\Delta_l(\mathcal{H}) &\leq c'' p^{l-1} \max_{H' \subseteq H : e(H') = l} \left(\frac{n^{v(H)}}{p^{l-1} n^{v(H')}}\right) \\
&\leq \frac{c''}{c'} p^{l-1} \max_{H' \subseteq H : e(H') = l} \left(\frac{1}{p^{e(H')-1} n^{v(H')-2}}\right) \cdot \frac{e(\mathcal{H})}{v(\mathcal{H})} \\
&\leq \frac{c''}{c'} p^{l-1} \frac{e(\mathcal{H})}{v(\mathcal{H})}.
\end{aligned}$$

□

Para concluir esta seção, só falta usar os últimos dois resultados para provar a Proposição 6.3.

Demonstração da Proposição 6.3. Dado que já temos o Lema de Contêineres, a prova agora se dá basicamente igual no caso de Grafos, na seção 2. Fixado $\delta > 0$, seja $\varepsilon > 0$ dado pelo Lema 6.4. Seja $\mathcal{H}$ o hipergrafo de cópias de H.

Fixado $I \in \mathcal{I}(\mathcal{H})$, defina $S_0 := \emptyset$ e $C_0 := V(\mathcal{H})$ e execute o seguinte algoritmo para $i \geq 1$:

1) Se $|C_i| \leq \left(1 - \dfrac{1}{\chi(H)-1} + \delta_{5.1}\right) v(\mathcal{H})$, então defina $g(I) := S_1 \cup \cdots S_i$ e $f(g(I)) := C_i$ e PARE.

2) Obtenha S_i e C_i usando o Lema 5.1 no conjunto $I \cap C_i$, em $\mathcal{H}[C_i]$ e com constantes $c_{5.1} = c/\varepsilon$ e $5_{3.1} = n^{-1/m_2(H)}$, em que c é obtida pela Proposição 6.5.

Primeiramente, vale ressaltar que a função f está bem definida pelo mesmo motivo que a função f_0 está bem definida no Lema 5.1, que é porque, em cada passo, a construção de S_i e C_i depende somente dos vértices de I que aparecem até acabar a construção dos contêineres. Assim, se existem I, $I' \in \mathcal{I}(\mathcal{H})$ e $g(I) = g(I')$, então eles terão o mesmo conjunto C_i.

É preciso verificar então que em cada iteração as hipóteses do Lema de Contêineres são satisfeitas. Com efeito, se no passo i, a condição de parada em 1) não é satisfeita, então o Lema 4.5 diz que

$$\frac{e(\mathcal{H}[C_i])}{v(\mathcal{H}[C_i])} \geq \varepsilon \cdot \frac{e(\mathcal{H})}{v(\mathcal{H})}.$$

Assim, podemos concluir que

$$\Delta(\mathcal{H}[C_i]) \leq \Delta(\mathcal{H}) \leq \frac{c}{\varepsilon} \cdot \varepsilon \cdot \frac{e(\mathcal{H})}{v(\mathcal{H})} \leq \frac{c}{\varepsilon} \cdot \frac{e(\mathcal{H}[C_i])}{v(\mathcal{H}[C_i])}.$$

Para terminar a prova, note que a propriedade (i) é satisfeita por construção e a (iii) simplesmente porque podemos supor, sem perda de generalidade, que $\delta_{5.1} \leq \delta$. Assim, continua-se aplicando contêineres até que a propriedade (iii) seja satisfeita.

Quanto à propriedade (ii), o Lema de Contêineres nos dá que $|S_i| = O(qv(\mathcal{H}))$ para todo i, com a mesma constante. Então basta provar que precisaremos de uma quantidade finita de iterações. Com efeito, temos que

$$|C_i| \leq (1-\delta)^i v(\mathcal{H}).$$

Assim, em no máximo $\lceil \log_{(1-\delta)} (1 - \frac{1}{\chi(H)-1} + \delta) \rceil = O(1)$ iterações do lema, o processo terminará e assim pode-se afirmar que $|S| = O(qv(\mathcal{H}))$.

□

Quando p é suficientemente pequeno, o próximo resultado mostrará que a proporção entre cópias de H e arestas é tão pequena, que pode-se deletar uma aresta de cada cópia de H. Para esta demonstração, será fixado $H' \subseteq H$ que maximiza $m(H)$ e $H'' \subseteq H$ que maximiza $m_2(H)$.

Teorema 6.6. *Seja H um grafo com $\Delta(H) \geq 2$. Para todo $\delta \in (0, \frac{1}{2(\chi(H)-1)})$ existe c tal que se $p \leq cn^{-1/m_2(H)}$, então com alta probabilidade*

$$ex(G_{n,p}, H) \geq \left(1 - \frac{1}{\chi(H) - 1} + \delta\right) p \binom{n}{2}.$$

Demonstração. Como a demonstração deste resultado não usa contêineres, faremos apenas um esboço da demonstração, dividindo-a em três casos.

1) $p \ll n^{-1/d(H')}$

   Nesta situação, o número esperado de cópias de H' tende para 0. Usando a desigualdade de Markov, com alta probabilidade $G_{n,p}$ não tem cópias de H' e consequentemente não tem cópias de H.

2) $n^{-2} \ll p \ll n^{-1/m_2(H)}$

   Primeiramente, note que como $\Delta(H) \geq 2$, então $d(H') \geq 2/3$, o que implica que $n^{-1/d(H')} \geq n^{-3/2} >> n^{-2}$. Assim, os casos 1 e 2 não são disjuntos.

   Além disso, neste caso $\mathbb{E}(e(G_{n,p})) >> 1$, logo Chernoff pode ser usado para dizer que com alta probabilidade $e(G_{n,p}) \geq (1 - \delta)p\binom{n}{2}$.

   Para terminar, podemos afirmar que o número esperado de cópias de H'' é $o(pn^2)$ e assim, usar desigualdade de Markov para afirmar que, com alta probabilidade, $G_{n,p}$ tem no máximo $\frac{\delta}{2}p\binom{n}{2}$ cópias de H''. Assim, pode-se remover cada cópia apenas retirando-se uma aresta.

3) $n^{-1/d(H')} \ll p \leq cn^{-1/m_2(H)}$

   Note que para mostrar que este caso também tem interseção com o anterior, basta mostrar que $d(H') \geq d_2(H')$, o que é equivalente a mostrar que $v(H') \leq 2e(H')$, o que é verdade, pois H tem um vértice de grau pelo menos 2.

   Se $X_{H''}$ representa o número de cópias de H'' em $G_{n,p}$, então é fácil ver que $\mathbb{E}(X_{H''}) \leq n^{v(H'')}p^{e(H'')}$. Agora será necessário calcular a variância de $X_{H''}$ para usar a desigualdade de Chebyshev. Para isso, é mais fácil olhar $X_{H''} = X_1 + ... + X_m$, em que cada uma dessas variáveis é a indicadora de uma cópia específica de H'' estar em $G_{n,p}$. Assim, temos que

$$Var(X_{H''}) = \sum_{i,j \leq m} (\mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j)).$$

Na expressão acima, se duas cópias de H'' não tem interseção de aresta, então os dois termos se anulam, então, podemos escrever a variância da seguinte forma:

$$Var(X_{H''}) = \sum_{F \subseteq H'' : e(F) \geq 1} c(F, H'') n^{2v(H'')-v(F)} p^{2e(H'')-e(F)}$$

em que $c(F, H'')$ não dependem de p e n, pois representa só a quantidade de possíveis extensões de F, dado que se escolheu $v(H'') - v(F)$ outros vértices. Assim, podemos concluir que

$$Var(X_{H''}) = O\left(\frac{n^{2v(H'')} p^{2e(H'')}}{\min_{F \subseteq H'' : e(F) \geq 1} n^{v(F)} p^{e(F)}} \right) = o\big(n^{2v(H'')} p^{2e(H'')}\big)$$

em que a última igualdade se deve ao fato de que $p \gg n^{-1/d(H')}$ e que H' maximiza a densidade d. Juntando o cálculo da váriância com a esperança, podemos usar Chebyshev:

$$\mathbb{P}(X_{H''} \geq 2n^{v(H'')} p^{e(H'')}) \leq \frac{Var(X_{H''})}{n^{2v(H'')} p^{2e(H'')}} = o(1).$$

Por outro lado, podemos escolher c apropriadamente tal que se $p \leq cn^{-1/m_2(H)}$, então

$$2n^{v(H'')} p^{e(H'')} \leq \frac{\delta}{2} p \binom{n}{2},$$

o que conclui a demonstração, ao se remover uma aresta de cada cópia de H''.

□

7 Teorema de Ramsey Aleatório

Dado que já construímos contêineres para o hipergrafo de cópias de um grafo H fixo, a demonstração do resultado desta seção segue sem dificuldades, a começar por uma supersaturação simples.

Lema 7.1. *Para todo grafo H e toda constante $r \geq 2$ existe $\alpha > 0$ e n_0 tal que para todo $n \geq n_0$, toda coloração das arestas de K_n com r cores contém pelo menos $\alpha n^{v(H)}$ cópias monocromáticas de H.*

Demonstração. Por Teoria de Ramsey básica, sabe-se que existe $n_0 := n_0(H, r)$ tal que para todo $n \geq n_0$ toda coloração das arestas de K_n com r cores contém uma cópia monocromática de H. Entretanto, cada cópia de H está contido em $\binom{n-v(H)}{n_0-v(H)}$ conjuntos de tamanho n_0. Assim, a quantidade de cópias monocromáticas de H é pelo menos

$$\binom{n}{n_0}\binom{n-v(H)}{n_0-v(H)}^{-1} = \Omega\left(\frac{n^{n_0}}{n^{n_0-v(H)}}\right) = \Omega(n^{v(H)})$$

□

Note que na demonstração do Teorema de Turán para $G_{n,p}$, não foi usado o fado de que $f(g(I)) \in \mathcal{F}_\varepsilon$ para todo $I \in \mathcal{I}(\mathcal{H})$. Entretanto, neste caso esse detalhe terá relevância, contida no seguinte resultado.

Teorema 7.2. *Seja $r \geq 2$ e H um grafo com $\Delta(H) \geq 2$. Existe $C = C(H, r)$ tal que se $p \geq Cn^{-1/m_2(H)}$, então*

$$\mathbb{P}[G_{n,p} \to (F)_r^e] = 1 - o(1).$$

Demonstração. Seja $c : E(G_{n,p}) \to [r]$ uma coloração que não contém cópias monocromáticas de H. Então para todo $j \in [r]$, $I_j := c^{-1}(j)$ é um grafo sem cópias de H. Assim, existem as funções f e g dadas pela proposição 6.3 que satisfazem

$$g(I_j) \subset I_j \subset g(I_j) \cup f(g(I_j)).$$

Além disso, se definimos $C_j := g(I_j) \cup f(g(I_j))$, então o número de cópias de H em C_j é no máximo $\varepsilon n^{v(H)}$, para algum $\varepsilon > 0$, dado pela supersaturação que, sem perda de generalidade, pode ser tomado pequeno suficiente.

Seja $\alpha > 0$ dado pelo Lema 7.1 aplicado ao grafo H e número de cores $r + 1$. Vamos definir uma coloração $c' : E(K_n) \to [r + 1]$ dada por $c(e) = \{\min j : e \in C_j\}$, em que $C_{r+1} := E(K_n) \setminus (C_1 \cup ... \cup C_r)$.

Podemos supor então que $\varepsilon \leq \alpha/2r$. Assim, como $C_1 \cup ... \cup C_r$ tem no máximo $\alpha/2 \cdot n^{v(H)}$ cópias de H, então C_{r+1} tem pelo menos a mesma quantidade de cópias de H.

Como cada aresta está contida em no máximo $\binom{n-2}{v(H)-2}$ cópias de H, então segue que existe $\delta > 0$ tal que C_{r+1} tem pelo menos δn^2 arestas. O importante desta conclusão é que $G_{n,p}$ não intersecta nenhuma dessas arestas.

Podemos cotar a probabilidade de $G_{n,p} \not\to (F)_r^e$ aglomerando todas as colorações tal que $(g(I_j))_{j\in[r]} = (S_j)_{j\in[r]} \in \mathcal{S}^r$, e cotar a probabilidade de $[S_1 \cup ... \cup S_r \subset G_{n,p}] \cap [C_{r+1} \cap G_{n,p} = \emptyset]$. Assim, o evento de existir tal coloração tem probabilidade menor ou igual a

$$\sum_{S_1,...,S_r \in \mathcal{S}} \mathbb{P}([S_1 \cup ... \cup S_r \subset G_{n,p}] \cap [G_{n,p} \cap C_{r+1} = \emptyset]) \leq$$

$$\sum_{|S| \leq r \cdot cqn^2} \mathbb{P}(S \subset G_{n,p}) \cdot (1-p)^{\delta n^2} \leq$$

$$\sum_{s=1}^{r \cdot cqn^2} \binom{n^2}{s} p^s \cdot (1-p)^{\delta n^2} \leq$$

$$\sum_{s=1}^{r \cdot cqn^2} \left(\frac{epn^2}{s}\right)^s \cdot (1-p)^{\delta n^2}$$

Para concluir, podemos supor que $p = C \cdot q$, pois a probabilidade de $G_{n,p} \not\to (F)^e_r$ é não crescente em p, então temos que, para C grande suficiente,

$$\mathbb{P}(G_{n,p} \not\to (F)^e_r) \leq r \cdot cqn^2 \left(\frac{eC}{rc}\right)^{r \cdot cqn^2} \cdot e^{-\delta Cqn^2} = o(1).$$

□

8 Sperner em $\mathcal{P}(n,p)$

Definição 8.1 (Hipergrafo de k-cadeias)**.** *Dado $\mathcal{F} \subset \mathcal{P}(n)$, $\mathcal{G}_k[\mathcal{F}]$ é o hipergrafo cujo conjunto de vértices é $\mathcal{F}$ e arestas definidas por*

$$\Big\{\{F_1, ..., F_k\} \subset \mathcal{F} : F_1 \supsetneq ... \supsetneq F_k\Big\}$$

Ao contrário das aplicações anteriores do Teorema de Contêineres, nesse caso $\mathcal{F}$ pode ter tamanho 2^n, logo poderia ter um grau médio tão pequeno que seria difícil garantir que as hipóteses do Teorema de Contêineres fossem satisfeitas. Para lidar com tal problema, precisaremos provar que se a família $\mathcal{F}$ é grande o suficiente, não só ela contém muitas k-cadeias, mas contém uma quantidade significativa delas que estão bem distribuídas.

Lema 8.2 (Supersaturação balanceada)**.** *Para todo $k \geq 2$ e $\alpha > 0$, existe $\delta = \delta(\alpha, k) > 0$ tal que a afirmação seguinte é verdadeira. Seja $n \in \mathbb{N}$ e $\mathcal{F} \subset \mathcal{P}(n)$ com $|\mathcal{F}| \geq (k-1+\alpha)\binom{n}{n/2}$ e suponha que existe m tal que $2\delta^{-1} \leq m \leq \binom{|F|}{|G|}$, para todos F,$G \in \mathcal{F}$ com $F \supsetneq G$. Então existe $\mathcal{H} \subset \mathcal{G}_k[\mathcal{F}]$ que satisfaz:*

(i) $e(\mathcal{H}) \geq \delta^k m^{k-1} \binom{n}{n/2}$

(ii) $\Delta_l(\mathcal{H}) \leq (\delta m)^{k-l}$ *para todo* $1 \leq l \leq k$.

Dizemos que uma família $\mathcal{A} \subset \mathcal{P}(n)$ é saturada se $d(\mathcal{A}) = \lfloor (\delta m)^{k-|\mathcal{A}|} \rfloor$, ou seja, se não podemos extendê-la sem violar a propriedade *(ii)* da supersaturação balanceada. Entretanto, uma família pode ser não saturada mas conter uma subfamília saturada. Isso nos levar a definir que $\mathcal{A} \subset \mathcal{P}(n)$ é *boa* se não contém subfamília saturada e *ruim* caso contrário.

Lema 8.3. *Para qualquer $\mathcal{A} \subset \mathcal{P}(n)$ boa, existem no máximo $2^{|\mathcal{A}|} \cdot 2\delta km$ conjuntos $F \in \mathcal{P}(n)$ para os quais $\{F\} \cup \mathcal{A}$ é ruim e F não é saturado.*

Demonstração. Note que, como $\mathcal{A}$ é bom, se F é tal que $\{F\} \cup \mathcal{A}$ é ruim, então existe $\mathcal{B} \subset A$ tal que $\{F\} \cup \mathcal{B}$ é saturado, i.e, F pertence ao seguinte conjunto

$$S(\mathcal{B}) := \{F \in \mathcal{P}(n) : d(\{F\} \cup \mathcal{B}) = \lfloor (\delta m)^{k-|\mathcal{B}|-1} \rfloor\}.$$

Agora só é necessário usar uma dupla contagem para cotar tamanho deste conjunto.

$$|S(\mathcal{B})| \cdot \lfloor (\delta m)^{k-|\mathcal{B}|-1} \rfloor = \sum_{F \in S(\mathcal{B})} d(\{F\} \cup \mathcal{B})$$

Por outro lado, como cada aresta que contem $\mathcal{B}$ contém no máximo k conjuntos de $S(\mathcal{B})$, então o lado direito da inequação acima satisfaz

$$\sum_{F \in S(\mathcal{B})} d(\{F\} \cup \mathcal{B}) \leq kd(\mathcal{B}) \leq k(\delta m)^{k-|\mathcal{B}|}$$

Juntando as duas inequações e notando que $\lfloor x \rfloor \geq x/2$ para todo $x \geq 2$, temos que

$$|S(\mathcal{B})| \leq 2\delta km.$$

Assim, para se chegar à afirmação do lema, basta somar o resultado acima para todos os subconjuntos de $\mathcal{A}$.

□

A estratégia para provar a supersaturação será achar uma k-cadeia por vez. Para isso, será útil achar um conjunto F_1 que tem uma certa densidade alta de cadeias em que ele é o primeiro elemento da cadeia e, dado que se construiu uma cadeia de tamanho $l < k$, tentar excluir os casos em que se tenta estender essa cadeia e dá errado, através do Lema 8.3. Tendo isso em mente, vamos à seguinte definição.

Definição 8.4. *Seja $F_1 \in \mathcal{F}$. A densidade de l-cadeias de F_1 é dada por*

$$c_l(F_1) := \sum_{\substack{F_2,\cdots,F_l \in \mathcal{F} \\ F_1 \supsetneq F_2 \supsetneq \ldots \supsetneq F_l}} \binom{|F_1|}{|F_2|}^{-1} \cdots \binom{|F_{l-1}|}{|F_l|}^{-1}$$

Os dois seguintes resultados comparam $c_i(F)$ e $c_j(F)$, quando $i \neq j$ e dá cotas que independem de n, que vão ser úteis para a demonstração da supersaturação.

Proposição 8.5. *Para quaisquer $1 \leq i < j \leq k$, tem-se que*

$$\sum_{F \in \mathcal{F}} \frac{1}{\binom{n}{|F|}} (c_i(F) - c_j(F)) \leq \max_{s \in \mathbb{N}} \binom{s}{i} - \binom{s}{j}$$

Demonstração. Uma cadeia maximal é uma cadeia de subconjuntos de $[n]$ $C_1 \supsetneq C_2 \supsetneq \ldots \supsetneq C_n$ tal que $|C_i| = i$. Fixada uma l-cadeia $F_1 \supsetneq F_2 \supsetneq \ldots \supsetneq F_l$, o número de cadeias maximais que a contém é dado por

$$(n-|F_1|)!(|F_1 \backslash F_2|)! \cdots (|F_{l-1} \backslash F_l|)!|F_l|! = n! \binom{n}{|F_1|}^{-1} \binom{|F_1|}{|F_2|}^{-1} \cdots \binom{F_{l-1}}{F_l}^{-1}.$$

Assim, se denotamos X_l o número l-cadeias contidas em uma cadeia maximal escolhida uniforme e aleatoriamente entre as $n!$ possíveis, temos que

$$\mathbb{E}(X_l) = \sum_{\substack{F_1, F_2,\cdots,F_l \in \mathcal{F} \\ F_1 \supsetneq F_2 \supsetneq \ldots \supsetneq F_l}} \binom{n}{|F_1|}^{-1} \binom{|F_1|}{|F_2|}^{-1} \cdots \binom{F_{l-1}}{F_l}^{-1} = \sum_{F_1 \in \mathcal{F}} \frac{c_l(F_1)}{\binom{n}{|F_1|}}$$

Por outro lado, se uma dada cadeia maximal contém s elementos de $\mathcal{F}$, então essa cadeia maximal contém exatamente $\binom{s}{l}$ l-cadeias. Assim, podemos concluir que

$$X_i - X_j \leq \max_{s \in \mathbb{N}} \binom{s}{i} - \binom{s}{j}$$
$$\mathbb{E}(X_i) - \mathbb{E}(X_j) \leq \max_{s \in \mathbb{N}} \binom{s}{i} - \binom{s}{j}$$

□

Lema 8.6. *Para todo $F_1 \in \mathcal{F}$ e $1 \leq l < k$, $c_l(F_1) \leq c_k(F_1) + 4^k$*

Demonstração. Só pela definição de $c_l(F_1)$, vale a seguinte igualdade:

$$c_l(F_1) = \sum_{F_2 \in \mathcal{F}: F_2 \subsetneq F_1} \frac{c_{l-1}(F_2)}{\binom{|F_1|}{|F_2|}}$$

Note também que a mesma prova da Proposição 8.5 vale para $\mathcal{F} \cap \mathcal{P}(F_1)$, logo

$$\begin{aligned} c_l(F_1) - c_k(F_1) &= \sum_{F_2 \in \mathcal{F} : F_2 \subsetneq F_1} \frac{1}{\binom{|F_1|}{|F_2|}} (c_{l-1}(F_2) - c_{k-1}(F_2)) \\ &\leq \max_{s \in \mathbb{N}} \binom{s}{l-1} - \binom{s}{k-1}. \end{aligned}$$

Pela propriedade de $\binom{a}{b}$ aumentar quando b se aproxima de $a/2$, o lado direito da inequação acima é negativo quando $s \geq 2k$. Assim, temos que

$$\max_{s \in \mathbb{N}} \binom{s}{l-1} - \binom{s}{k-1} \leq \max_{s \leq 2k} \binom{s}{l-1} = \binom{2k}{l-1} \leq 4^k.$$

□

O seguinte Lema é um resultado do tipo Princípio das Casas de Pombo e afirma que é possível achar F com $c_k(F)$ "grande". O importante é que essa cota inferior seja estritamente maior que zero, mas mais do que isso, que não dependa da família $\mathcal{F}$, como ficará evidenciado posteriormente.

Lema 8.7. *Se $0 < \alpha < 1$ e $|\mathcal{F}| \geq (k - 1 + \alpha)\binom{n}{n/2}$, então existe $F \in \mathcal{F}$ com $c_k(F) \geq \frac{\alpha}{k}$.*

Demonstração. Podemos usar a Proposição 8.5 no caso em que $l = 1$ e, como $c_1(F) = 1$, notamos que

$$\sum_{F \in \mathcal{F}} \frac{1}{\binom{n}{|F|}} (1 - c_k(F)) \leq \max_{s \in \mathbb{N}} \binom{s}{1} - \binom{s}{k} = k - 1.$$

Por outro lado, se $c_k(F) < \alpha/k$ para todo $F \in \mathcal{F}$, teremos que

$$\sum_{F \in \mathcal{F}} \frac{1}{\binom{n}{|F|}} (1 - c_k(F)) > \sum_{F \in \mathcal{F}} \frac{1}{\binom{n}{n/2}} (1 - \frac{\alpha}{k}) \geq (k - 1 + \alpha) \cdot \frac{k - \alpha}{k} > k - 1,$$

em que a última desigualdade é obtida através da análise de uma simples inequação do segundo grau.

□

Demonstração da supersaturação. O Teorema de supersaturação será provado encontrando-se uma aresta em $\mathcal{G}_k[\mathcal{F}]$ de cada vez. Mais precisamente, se $\mathcal{H} \subset \mathcal{G}_k[\mathcal{F}]$ e valem as seguintes propriedades

(i) $e(\mathcal{H}) < \delta^k m^{k-1}\binom{n}{n/2}$

(ii) $\Delta_l(\mathcal{H}) \leq (\delta m)^{k-l}$ para todo $1 \leq l \leq k$;

então existe $f \in \mathcal{G}_k[\mathcal{F}] \setminus \mathcal{H}$ tal que f é boa.

Primeiramente, pelo mesmo argumento de dupla contagem do Lema 8.3, se S é a coleção de conjuntos saturados, então S satisfaz

$$\begin{aligned} |S|\lfloor(\delta m)^{k-1}\rfloor &\leq k \cdot e(\mathcal{H}) \\ \implies |S| &\leq 2\delta k\binom{n}{n/2}. \end{aligned}$$

Assim, tomando δ suficientemente pequeno, podemos excluir esse conjunto de $\mathcal{F}$ e ter, por exemplo, pelo menos $(k-1+\alpha/2)\binom{n}{n/2}$ elementos na família, mas por simplicidade, o mesmo α será mantido.

Lembrando que uma aresta em $\mathcal{G}_k[\mathcal{F}]$ é uma k-cadeia de elementos de $\mathcal{P}(n)$, vamos começar com F_1 de mínima cardinalidade tal que $c_k(F_1) \geq \alpha/k$. Assim, vamos tentar construir uma k-cadeia procurando conjunto por conjunto e contar as vezes que o processo der errado, i.e, olhar para as cadeias críticas de tamanho menor que k.

Note que se uma k-cadeia é ruim, então existe o menor l tal que os l primeiros conjuntos dessa cadeia formam uma cadeia crítica. Assim, a expressão

$$\sum_{\substack{F_2,\cdots,F_k\in\mathcal{F} \\ F_1\supsetneq F_2\supsetneq\ldots\supsetneq F_k \text{ ruim}}} \binom{|F_1|}{|F_2|}^{-1}\cdots\binom{|F_{k-1}|}{|F_k|}^{-1}$$

pode ser agrupada por tais l's, de forma que ela é cotada por cima por

$$\sum_{l=2}^{k}\left(\sum_{\substack{F_2,\cdots,F_l\in\mathcal{F} \\ F_1\supsetneq F_2\supsetneq\ldots\supsetneq F_l \text{ crítica}}} \binom{|F_1|}{|F_2|}^{-1}\cdots\binom{|F_{l-1}|}{|F_l|}^{-1}\cdot c_{k-l-1}(F_l)\right).$$

Note também que, pela minimalidade de F_1 e pelo Lema 8.6, na expressão acima temos que $c_{k-l-1}(F_{l+1}) \leq c_k(F_{l+1}) + 4^k \leq \alpha/k + 4^k = 2^{O(k)}$. Usando o Lema 8.3, para o somatório de dentro, temos que

$$\sum_{\substack{F_2,\cdots,F_l\in\mathcal{F} \\ F_1\supsetneq F_2\supsetneq\ldots\supsetneq F_l \text{ crítica}}} \binom{|F_1|}{|F_2|}^{-1}\cdots\binom{|F_{l-1}|}{|F_l|}^{-1}\cdot c_{k-l-1}(F_l) \leq 2^l\cdot 2\delta k\cdot c_l(F_1)\cdot 2^{O(k)}.$$

Juntando tudo e usando que $c_l(F_1) \leq c_k(F_1) + 4^k$, podemos concluir que

$$\sum_{\substack{F_2,\cdots,F_k\in\mathcal{F}\\ F_1\supsetneq F_2\supsetneq\ldots\supsetneq F_k \text{ ruim}}} \binom{|F_1|}{|F_2|}^{-1}\cdots\binom{|F_{k-1}|}{|F_k|}^{-1} = \delta\cdot 2^{O(k)}\cdot(c_k(F_1)+4^k) \leq \frac{c_k(F_1)}{2}.$$

A última desigualdade vale pois $c_k(F_1) \geq \alpha/k$, i.e, está a uma distância positiva do 0 que não depende de nada além de α e k e assim δ pode ser escolhido suficientemente pequeno dependendo apenas dos mesmos parâmetros.

Note que a demonstração agora acabou, pois se a densidade de cadeias ruins que começam em F_1 é menor do que a densidade total, então existe uma cadeia boa.

□

Agora que provamos a supersaturação balanceada, vamos ver que este problema apresenta outra particularidade não vista em Turán e Ramsey em $G(n,p)$, que é a existência de dois regimes na hora de iterar o Teorema de Contêineres.

Lema 8.8. *Para todo $k \geq 2$ e $\alpha > 0$ existe $c = c(\alpha, k) > 0$ tal que a seguinte afirmação é verdadeira. Seja $n \in \mathbb{N}$ suficientemente grande e $\mathcal{F} \subset \mathcal{P}(n)$ que satisfaz $(k-1+\alpha)\binom{n}{n/2} \leq |\mathcal{F}| \leq 3k\binom{n}{n/2}$. Então existe $\mathcal{H} \subset \mathcal{G}_k[\mathcal{F}]$ que satisfaz*

$$\Delta_l(\mathcal{H}) \leq \frac{c}{n^{l-1}}\cdot\frac{e(\mathcal{H})}{|\mathcal{F}|}$$

para todo $1 \leq l \leq k$.

Demonstração. Primeiramente, note que podemos retirar todos os conjuntos de tamanhoa menor que $n/3$ de $\mathcal{F}$ e garantir, por exemplo, que $|\mathcal{F}| \geq (k-1+\alpha/2)\binom{n}{n/2}$, mas por simplicidade, o mesmo α será mantido. Desta maneira, temos que para todo $F, G \in \mathcal{F}$ tal que $F \supsetneq G$, vale

$$\binom{|F|}{|G|} \geq |F| \geq n/3.$$

Aplicando a supersaturação com $m = n/3$, existe $\mathcal{H} \subset \mathcal{G}_k[\mathcal{F}]$ que satisfaz

(i) $e(\mathcal{H}) \geq \delta^k (n/3)^{k-1}\binom{n}{n/2}$

(ii) $\Delta_l(\mathcal{H}) \leq (\delta\cdot n/3)^{k-l}$ para todo $1 \leq l \leq k$.

Para terminar a prova do lema, baste mostrar que existe $c := c(\alpha, k) > 0$ tal que as seguintes desigualdades são verdadeiras para todo $1 \leq l \leq k$:

$$\frac{c}{n^{l-1}}\cdot\frac{e(\mathcal{H})}{|\mathcal{F}|} \geq \frac{c}{n^{l-1}}\cdot\frac{\delta^k(n/3)^{k-1}\binom{n}{n/2}}{3k\binom{n}{n/2}} \geq (\delta m)^{k-l} \geq \Delta_l(\mathcal{H}),$$

o que fácil de ser verificado se $c = k \cdot (3/\delta)^k$. $\square$

Lema 8.9. *Para todo $k \geq 2$ existe $c = c(k) > 0$ tal que a seguinte afirmação é verdadeira. Seja $n \in \mathbb{N}$ suficientemente grande e $\mathcal{F} \subset \mathcal{P}(n)$ que satisfaz $|\mathcal{F}| \geq 3k\binom{n}{n/2}$. Então existe $\mathcal{H} \subset \mathcal{G}_k[\mathcal{F}]$ que satisfaz*

$$\Delta_l(\mathcal{H}) \leq \frac{c}{n^{3(l-1)}} \cdot \frac{e(\mathcal{H})}{|\mathcal{F}|}$$

para todo $1 \leq l \leq k$.

Demonstração. Considere uma partição arbitrária $\mathcal{F} = \mathcal{F}_0 \cup \mathcal{F}_1 ... \cup \mathcal{F}_t$, em que para todo $i \in [t]$, $\mathcal{F}_i = 3k\binom{n}{n/2}$ e $\mathcal{F}_0 < 3k\binom{n}{n/2}$. Fixe $i \in [t]$ e note que pelo princípio das casas de pombo, pelo menos $k\binom{n}{n/2}$ desses conjuntos tem tamanhos com mesmo resto módulo 3.

Alé disso, remova de $\mathcal{F}_i$ todos os conjuntos de tamanho menor que $n/3$ e obtenha $\mathcal{F}'_i$, de tamanho $(k - o(1))\binom{n}{n/2}$. Veja que se para todos $F, G \in \mathcal{F}'_i$ tal que $F \supsetneq G$, vale que

$$\binom{|F|}{|G|} \geq \binom{|G| + 3}{|G|} = \binom{|G| + 3}{3} \geq \binom{n/3}{3}.$$

Note também que o termo $o(1)$ no tamanho de $\mathcal{F}'_i$ é o mesmo para todo i, então podemos supor, por exemplo que $\mathcal{F}'_i \geq (k - 1 + 1/2)\binom{n}{n/2}$ para todo $i \in [t]$ e obter, com $m = \binom{n/3}{3}$, $\delta = \delta(k)$ e $\mathcal{H}_i \subset \mathcal{G}_k[\mathcal{F}'_i]$ tal que

$$e(\mathcal{H}_i) = \lceil \delta^k m^{k-1}\tbinom{n}{n/2} \rceil$$

$$\Delta_l(\mathcal{H}_i) \leq (\delta m)^{k-l} = \frac{k}{\delta^l m^{l-1}} \cdot \frac{\delta^k m^{k-1}\binom{n}{n/2}}{k\binom{n}{n/2}} \leq \frac{c'}{(n^3)^{l-1}} \cdot \frac{e(\mathcal{H}_i)}{|\mathcal{F}'_i|}.$$

Se definimos $\mathcal{H} = \mathcal{H}_1 \cup ... \cup \mathcal{H}_t$, podemos concluir a demonstração ao notar que $e(\mathcal{H}) = t \cdot e(\mathcal{H}_i)$ e $|\mathcal{F}| \leq 5t \cdot |\mathcal{F}'_i|$ para todo $i \in [t]$, pois como as arestas de $\mathcal{H}$ são apenas as contidas em cada $\mathcal{H}_i$, então

$$\Delta_l(\mathcal{H}) \leq \max_{i \in [t]}\{\Delta_l(\mathcal{H}_i)\} \leq \frac{c'}{(n^3)^{l-1}} \cdot \frac{\max_i(e(\mathcal{H}_i))}{\min_i |\mathcal{F}'_i|} \leq \frac{c}{(n^3)^{l-1}} \cdot \frac{e(\mathcal{H})}{|\mathcal{F}|}.$$

$\square$

Agora, tudo está preparado para enunciar e provar o Lema de Contêineres apropriado para este problema.

Proposição 8.10. *Para todo $k \geq 2$ e $\varepsilon > 0$, existem $K = K(k, \varepsilon) > 0$, uma família $\mathcal{S} \subset \mathcal{P}(V(\mathcal{G}_k))$ e funções $f : \mathcal{S} \to \mathcal{P}(V(\mathcal{G}_k))$ e $g : \mathcal{I}(\mathcal{G}_k) \to \mathcal{S}$ tais que*

(i) $|S| \leq \frac{K}{n}\binom{n}{n/2}$, *para todo* $S \in \mathcal{S}$

(ii) A quantidade de elementos de $\mathcal{S}$ *de tamanho* s *é no máximo*

$$\left(\frac{K\binom{n}{n/2}}{s}\right)^s \exp\left(\frac{K}{n}\binom{n}{n/2}\right)$$

(iii) $f(g(I)) \leq (k - 1 + \varepsilon)\binom{n}{n/2}$ *para todo* $I \in \mathcal{I}(\mathcal{G}_k)$

(iv) $g(I) \subset I \subset g(I) \cup f(g(I))$ *para todo* $I \in \mathcal{I}(\mathcal{G}_k)$

Antes de provar esta proposição, vamos provar o Teorema 6.1, em que ficará evidenciado a necessidade da propriedade (ii), que não foi necessária nas outras aplicações mostradas neste trabalho.

Demonstração do Teorema 6.1. Primeiramente, fixe $\varepsilon > 0$ e K dado pela proposição 6.9 e note que podemos tomar n suficientemente grande tal que $pn \geq K/\varepsilon$, pela hipótese do teorema. Além disso, defina $m = (k - 1 + 3\varepsilon)p\binom{n}{n/2}$.

Se $I \subset \mathcal{P}(n,p)$ e $|I| = m$, então $g(I) \subset \mathcal{P}(n,p)$ e, além disso, como $|g(I)| \leq \varepsilon p\binom{n}{n/2}$ (pela cota inferior de pn), então $|\mathcal{P}(n,p) \cap f(g(I))| \geq (k - 1 + 2\varepsilon)p\binom{n}{n/2}$. Então a probabilidade de existir $I \in \mathcal{I}(\mathcal{G}_k)$ de tamanho m em $\mathcal{P}(n,p)$ é menor que:

$$\sum_{S \in \mathcal{S}} \mathbb{P}(S \in \mathcal{P}(n,p)) \cdot \mathbb{P}(|f(S) \cap \mathcal{P}(n,p)| \geq (k - 1 + 2\varepsilon)p\tbinom{n}{n/2}).$$

Como $|f(S) \cap \mathcal{P}(n,p)|$ é uma variável aleatória de distribuição binomial cuja esperança é menor que $(k - 1 + \varepsilon)p\binom{n}{n/2}$, então segue pela desigualdade de Chernoff que

$$\mathbb{P}(|f(S) \cap \mathcal{P}(n,p)| \geq (k - 1 + 2\varepsilon)p\tbinom{n}{n/2}) \leq \exp\left(-\frac{\varepsilon^2 p}{3}\binom{n}{n/2}\right)$$

Para cotar o outro termo, vamos usar a propriedade (ii) da proposição 6.9:

$$\sum_{S \in \mathcal{S}} \mathbb{P}(S \in \mathcal{P}(n,p)) \leq \sum_{s=0}^{\frac{K}{n}\binom{n}{n/2}} \left(\frac{K\binom{n}{n/2}}{s}\right)^s \exp\left(\frac{K}{n}\binom{n}{n/2}\right) \cdot p^s$$

Como a função $(k/x)^x$ é crescente em $(0, k/e)$ e $K/n \leq Kp/e$ para n suficientemente grande, então o lado direito da inequação é menor que

$$\frac{K}{n}\binom{n}{n/2}\exp\left(pK \cdot \frac{\log(pn)}{pn}\binom{n}{n/2} + pK \cdot \frac{1}{pn}\binom{n}{n/2}\right)$$

E agora só falta ver que, como $pn >> log(pn) >> 1$ então pode-se escolher n suficientemente grande para que a seguinte desigualdade seja verdade:

$$\sum_{S \in \mathcal{S}} \mathbb{P}(S \in \mathcal{P}(n,p)) \leq \exp\left(\frac{\varepsilon^2 p}{6}\binom{n}{n/2}\right)$$

Assim, juntando as desigualdades () e (), temos que

$$\mathbb{P}\big(\alpha_k(\mathcal{P}(n,p)) \geq (k-1+3\varepsilon)p\tbinom{n}{n/2}\big) \leq \exp\left(-\frac{\varepsilon^2 p}{6}\binom{n}{n/2}\right) \xrightarrow{n\to\infty} 0.$$

□

Os lemas 8.7 e 8.8 separam as iterações do lema de contêineres em dois regimes, de acordo com o tamanho da família em que o lema está sendo aplicado. Por isso, sempre que formos aplicar o lema, consideraremos c a maior constante dada pelos dois lemas e com isso, obteremos em cada caso um δ diferente. Sendo assim, consideraremos um único δ, como sendo o mínimo obtido nas duas situações.

Por fim, em cada um dos regimes, temos um q diferente, o que nos leva a definir a função $q : \mathcal{P}(\mathcal{P}(n)) \to \mathbb{R}$ definida por

$$q(C) = \begin{cases} n^{-1}, \text{ se } |C| \leq 3k\binom{n}{n/2} \\ n^{-3}, \text{ caso contrário} \end{cases}$$

Demonstração da Proposição 8.9. Definimos $C_0 = \mathcal{P}(n)$ e, dado $I \in \mathcal{I}(\mathcal{G}_k)$, execute o algoritmo para $i \geq 0$ até chegar ao comando PARE:

1. Se $|C_i| < (k-1+\varepsilon)\binom{n}{n/2}$, PARE
2. Obtenha $\mathcal{H}_i$ pela supersaturação aplicada a C_i
3. Aplique lema de contêineres com q e c, definidos anteriormente, ao hipergrafo $\mathcal{H}_i$ e obtenha funções g_i e f_i
4. Defina $I_i := I \cap C_i$ e obtenha $S_{i+1} := g_i(I_i)$
5. Defina $C_{i+1} := f_i(I_i) \setminus S_i$.

O lema de contêineres nos dá que em cada passo da iteração, obtemos as seguintes propriedades:

a) $I \subset C_i \cup S_1 \cup ... S_i$

b) Os conjuntos $C_i, S_1, ..., S_i$ são dois a dois disjuntos

c) $|C_{i+1}| \leq (1-\delta)|C_i|$

d) $|S_{i+1}| \leq (k-1) \cdot q(C_i)|C_i|$.

Se definimos m o número de iterações até o algoritmo parar, teremos que

$$(1-\delta)^m \cdot 2^n < (k-1+\varepsilon)\binom{n}{n/2}$$

Temos que $m = O(\log n)$, por uma simples aplicação da fórmula de Stirling, ao ver que

$$\binom{n}{n/2} \sim \frac{1}{\sqrt{2\pi n}} \cdot 2^n.$$

Para provar a propriedade (i), vamos analisar separadamente os dois regimes. Para isso, definimos $2 \leq m_0 \leq m$, como sendo o primeiro índice tal que $|C_{m_0}| \leq 3k\binom{n}{n/2}$. Usando a propriedade c) acima e fórmula de Stirling novamente, temos que

$$\sum_{i=1}^{m_0-1} (k-1) \cdot q(C_i)|C_i| \leq (k-1)\frac{m \cdot 2^n}{n^3} \ll \frac{1}{n^2}\binom{n}{n/2}. \tag{11}$$

Para o segundo regime, note para se chegar de um conjunto de tamanho menor ou igual a $3k\binom{n}{n/2}$ a um de tamanho $(k-1+\varepsilon)\binom{n}{n/2}$ retirando-se uma proporção δ de cada vez, precisa-se apenas de uma quantidade finita de iterações. Assim, é fácil ver que

$$\sum_{i=m_0}^{m} q(C_i)|C_i| = \frac{O(1)}{n}\binom{n}{n/2}, \tag{12}$$

o que conclui a prova da propriedade (i). Como a propriedade (iii) é garantida por construção, basta agora provar a propriedade (ii). Para isso, vemos que todo $S \in \mathcal{S}$ satisfaz que $S = S_1 \cup ... \cup S_{m_0(S)} \cup ... \cup S_{m(S)}$.

Para estimar quantos conjuntos de tamanho s tem em $\mathcal{S}$, vamos fixar $\hat{m}$, $\hat{m}_0$ e $(s_i)_{i=1}^{\hat{m}}$ com $\sum_i s_i = s$ e os conjuntos $S \in \mathcal{S}$ que satisfazem que $m(S) = \hat{m}$, $m_0(S) = \hat{m}_0$ e $S_i = s_i$ para todo $i \in [\hat{m}]$. Lembrando agora que C_{i+1} só depende dos conjuntos C_i e S_i, podemos cotar superiormente a quantidade de tais conjuntos por

$$\prod_{i=1}^{\hat{m}} \binom{|C_i|}{s_i} \leq \prod_{i=1}^{\hat{m}_0-1} \binom{2^n}{s_i} \prod_{i=\hat{m}_0}^{\hat{m}} \binom{3k\binom{n}{n/2}}{s_i}.$$

Vamos lidar com cada termo do lado direito da inequação separadamente. Usando (11) o primeiro termo pode ser fácilmente cotado

$$\prod_{i=1}^{\hat{m}_0-1} 2^{s_i n} \ll \exp\left(\frac{1}{n}\binom{n}{n/2}\right).$$

Para o segundo termo, vamos definir $s' = \sum_{i=\hat{m}_0}^{\hat{m}} s_i$ e ver que

$$\prod_{i=\hat{m}_0}^{\hat{m}} \binom{3k\binom{n}{n/2}}{s_i} \leq \prod_{i=\hat{m}_0}^{\hat{m}} \left(\frac{3ek\binom{n}{n/2}}{s_i}\right)^{s_i} = \left(3ek\binom{n}{n/2}\right)^{s'} \prod_{i=\hat{m}_0}^{\hat{m}} \left(\frac{1}{s_i}\right)^{s_i}.$$

Se provarmos que

$$\prod_{i=\hat{m}_0}^{\hat{m}} \left(\frac{1}{s_i}\right)^{s_i} \leq \left(\frac{1}{s'}\right)^{s'} \cdot \exp\left(\frac{O(1)}{n}\binom{n}{n/2}\right),$$

então o problema acabou, pois podemos substituir s' por s e obter uma cota superior. Mas isso é um resultado direto do fato de que a função $x \log x$ é convexa, da desigualdade de Jensen e de (12):

$$\sum_{i=\hat{m}_0}^{\hat{m}} s_i \log s_i \geq s' \log s' - s' \log(\hat{m} - \hat{m}_0 + 1)$$

$$\begin{aligned}
\sum_{i=\hat{m}_0}^{\hat{m}} s_i \log\left(\frac{1}{s_i}\right) &\leq s' \log\left(\frac{1}{s'}\right) + s' \log(\hat{m} - \hat{m}_0 + 1) \\
&\leq s' \log\left(\frac{1}{s'}\right) + \frac{O(1)}{n}\binom{n}{n/2}.
\end{aligned}$$

Todas as possibilidades de escolhas de $\hat{m}$ e $\hat{m}_0$ são no máximo $O(\log n)$, enquanto a quantidade de sequências de s_i é no máximo

$$\binom{s+\hat{m}-1}{s} \leq \binom{2s}{s} \leq \exp\left(\frac{O(1)}{n}\binom{n}{n/2}\right),$$

se $s \geq \hat{m}$. Caso contrário, esse número é trivialmente menor que $\hat{m}^{\hat{m}}$ que é desprezível perto do termo exponencial presente na propriedade (ii) e logo pode ser englobado na constante.

$\square$

Referências

[1] D.J. Kleitman and K.J. Winston. On the number of graphs without 4 -cycles. *Discrete Math*, 41(1982):167–172.

[2] N. Alon. Independent sets in regular graphs and sum-free subsets of finite groups. *Israel J. Math*, 73(1991):247–256.

[3] P. Erdős and A.H. Stone. On the structure of linear graphs. *Bulletin of the American Mathematical Society*, 52(1946):1087–1091.

[4] D. Conlon and W.T. Gowers. Combinatorial theorems in sparse random sets. *Annals of Mathematics*, 184(2016):367–454.

[5] M. Schacht. Extremal results for random discrete structures. *Annals of Mathematics*, 184(2)(2016):333–365.

[6] V. Rödl and A. Ruciński. Threshold functions for ramsey properties. *J. Amer. Math. Soc*, 8(1995):917–942.

[7] R. Nenadov and A. Steger. A short proof of the random ramsey theorem. *Combinatorics, Probability, and Computing*, 25(2016):130–144.

[8] P. Erdős. On a lemma of littlewood and offord. *Bull. Amer. Math. Soc*, 51(1945):898–902.

[9] M. Collares and R. Morris. Maximum-size antichains in random sets. *Random Structures & Algorithms, to appear.*

[10] R. Mycroft J. Balogh and A. Treglown. A random version of sperner's theorem. *J. Combin. Theory, Ser. A*, 128(2014):104–110.

[11] J.H. Spencer N. Alon. *The Probabilistic Method.* The probabilistic method, third ed., Wiley-Interscience Series in Discrete Math. and Optimization, John Wiley & Sons Inc., Hoboken, NJ, 2008, With an appendix on the life and work of Paul Erdős.

[12] P. Erdős and M. Simonovits. Supersaturated graphs and hypergraphs. *JCombinatorica*, 3(1983):181–192.