Uniwersytet im. Adama Mickiewicza
w Poznaniu

Katarzyna Rybarczyk

O PEWNYCH ZASTOSOWANIACH HIPERGRAFÓW LOSOWYCH

Praca magisterska
napisana pod kierunkiem
prof.dr hab. J. Jaworskiego
w Zakładzie Matematyki
Dyskretnej
na Wydziale Matematyki
i Informatyki

Poznań 2005

Spis treści

Wstęp

W pracy zajmiemy się klasyfikacją obiektów ze względu na posiadane własności. Rozważmy pewien zbiór obiektów $\mathcal{V}$ i zbiór własności $\mathcal{W}$. Zależności między obiektami i ich własnościami można przedstawić za pomocą grafu dwudzielnego, w którym każdy obiekt jest połączony krawędziami z własnościami, które posiada. Naturalnym założeniem jest przyjęcie, że obiekty są podobne, gdy zbiory ich cech się przecinają, a własności są podobne, gdy zbiory obiektów, które je posiadają mają niepusty przekrój. Wykorzystując graf dwudzielny o dwupodziale $\mathcal{V} \cup \mathcal{W}$ można w ten sposób zdefiniować dwa grafy przecięć $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ o zbiorach wierzchołków odpowiednio $\mathcal{V}$ i $\mathcal{W}$, w których elementy podobne są połączone krawędziami.

W pracy interesowały nas przede wszystkim wyniki asymptotyczne dotyczące spójności tych grafów przecięć i rozmiarów ich składowych spójności. Jako narzędzie posłuży nam odpowiednio zdefiniowany indukowany hipergraf losowy o zbiorze wierzchołków $\mathcal{W}$. Jest to hipergraf o krawędziach wyznaczonych przez zbiory sąsiadów (zbiory własności) poszczególnych obiektów.

Naszym celem była zatem analiza istniejących wyników dotyczących spójności i składowych spójności hipergrafów losowych oraz odpowiednie uogólnienie niektórych z nich. Następnym zadaniem było pokazanie w jaki sposób własności tradycyjnych hipergrafów losowych możemy wykorzystać dla hipergrafu indukowanego przez obiekty i własności. Ostatecznie pozwoliło to na udowodnienie silnych związków między strukturą hipergrafu indukowanego a strukturą grafów przecięć i sformułowanie wniosków dotyczących spójności grafów $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$. Określonym powyżej zadaniom odpowiada układ pracy.

W pierwszym rozdziale przedstawione zostały najważniejsze pojęcia i definicje dotyczące losowych grafów przecięć i hipergrafów.

W rozdziale drugim omówione są własności hipergrafów i asymptotyczna równoważność tradycyjnych modeli hipergrafów losowych. Podane zostały znane twierdzenia o asymptotycznej równoważności wyników dotyczących monotonicznych i wypukłych własności hipergrafów. Rozdział ten opiera się przede wszystkim na [9]. Wyniki te uogólniłam dla hipergrafów losowych, w których krawędzie o różnych mocach pojawiają się z różnym prawdopodobieństwem.

Rozdział trzeci poświęcony został spójności hipergrafów. Przedstawiam w nim znane twierdzenia dotyczące spójności hipergrafów pochodzące z prac [11] i [12]. Udowodniłam tam również twierdzenie o spójności hipergrafów losowych o zróżnicowanych mocach krawędzi.

W rozdziale czwartym podane zostały wyniki opisujące strukturę składowych spójności w hipergrafie losowym w zależności od średniego stopnia wierzchołka w hipergrafie losowym. Przedstawiam rezultaty pochodzące z pracy [16].

W rozdziale piątym zdefiniowany został indukowany grafem dwudzielnym hipergraf losowy. Omawiane są jego własności i związki z klasycznymi modelami hipergrafów losowych oraz udowodniona asymptotyczna równoważność hipergrafu indukowanego i klasycznego hipergrafu losowego.

Rozdział szósty zawiera wnioski i twierdzenia dotyczące spójności i struktury losowych grafów przecięć. Dzięki zebranym wynikom i uzyskanym uogólnieniom dotyczącym hipergrafów, w których krawędzie powstają ze zróżnicowanym prawdopodobieństwem, wykorzystując hipergraf indukowany udało się zanalizować strukturę losowych grafów $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$. Przy danej liczbie obiektów, własności i danym rozkładzie stopni w grafie dwudzielnym określone zostało prawdopodobieństwo spójności grafu $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$. Podałam również warunki, przy których graf $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ jest spójny asymptotycznie prawie na pewno. Udało mi się także udowodnić, że przy danej liczbie n obiektów, $m(n)$ własności i danym rozkładzie stopni $\mathcal{P}_{(n)}$ i pewnych założeniach o rozkładzie wielkość największej składowej spójności w losowych grafach przecięć zmienia się szybko z wielkości rzędu $\ln n$ do wielkości rzędu n.

1. Preliminaria

Definicja 1.1. Nieskierowanym grafem prostym (grafem) nazywamy parę $G(V, E) = \langle V, E \rangle$, składającą się z niepustego zbioru wierzchołków $V = V(G)$ oraz zbioru krawędzi $E = E(G)$ będących nieuporządkowanymi parami wierzchołków z V.

Jeśli wierzchołek $v \in V$ w grafie $G(V, E)$ należy do pewnej krawędzi $e \in E$, to mówimy, że jest z nią incydentny.

Definicja 1.2. Stopniem $d(v)$ wierzchołka $v \in V$ w grafie $G(V, E)$ nazywamy liczbę krawędzi grafu G incydentnych z v.

Szczególnym przypadkiem grafu jest graf dwudzielny, w którym krawędzie występują tylko między dwoma określonymi rozłącznymi podzbiorami wierzchołków.

Definicja 1.3. Grafem dwudzielnym nazywamy graf $G(V, E)$, którego zbiór wierzchołków można podzielić na dwa niepuste rozłączne podzbiory X,Y ($V = X \cup Y$ i $X \cap Y = \emptyset$) takie, że każda krawędź z $E(G)$ ma dokładnie jeden wierzchołek w X i dokładnie jeden wierzchołek w Y. Parę zbiorów (X, Y) nazywamy dwupodziałem grafu $G(V, E)$.

Grafy, które powstają w wyniku eksperymentu losowego nazywamy grafami losowymi. Przegląd wyników z tej rozwiniętej w ciągu ostatnich kilkudziesięciu lat dziedziny można znaleźć w książkach Bollobása [1] i Jansona, Łuczaka i Rucińskiego [9]. Grafy losowe znajdują szereg zastosowań. Jedną z dziedzin, w których można wykorzystać wyniki dotyczące grafów losowych jest klasyfikacja i analiza skupień. Rozważmy zbiór obiektów $\mathcal{V} = \{v_1, v_2, \ldots, v_m\}$ oraz zbiór własności $\mathcal{W} = \{w_1, w_2, \ \ldots, w_n\}$, które te obiekty mogą posiadać. Zainteresowani jesteśmy sklasyfikowaniem tych obiektów ze względu na podobieństwa. Na przykład ze względu na posiadanie co najmniej s cech wspólnych. Naturalnym sposobem na zobrazowanie relacji podobieństwa między obiektami jest graf "obiektów" o zbiorze wierzchołków $\mathcal{V}$, w którym dwa wierzchołki są połączone krawędzią, gdy odpowiadające im obiekty są podobne (mają co najmniej s cech wspólnych). Aby zanalizować powyższy graf można rozpatrzyć pomocniczy graf dwudzielny, w którym krawędzie będą łączyć obiekty i ich własności. Mianowicie dla zbiorów $\mathcal{V}$ i $\mathcal{W}$ możemy skonstruować graf dwudzielny $\mathcal{BG}(\mathcal{V} \cup \mathcal{W}, \mathcal{E})$ o zbiorze wierzchołków $V = \mathcal{V} \cup \mathcal{W}$, dwupodziale $(\mathcal{W}, \mathcal{V})$ i zbiorze krawędzi $E = \mathcal{E}$. W tym grafie dla obiektu v i własności w występuje krawędź (v, w) wtedy i tylko wtedy gdy v ma własność w. Zauważmy, że w grafie "obiektów" występuje krawędź między wierzchołkami (obiektami), gdy w grafie dwudzielnym zbiory sąsiadów wierzchołków odpowiadających obiektom mają co najmniej s elementów wspólnych. Graf "obiektów" jest zatem grafem przecięć pewnych zbiorów. Z powyższych rozważań wynika, że można zdefiniować pomocniczy losowy graf dwudzielny, który będzie punktem wyjścia do klasyfikacji obiektów losowo wybierających własności. [7].

Definicja 1.4. Niech $\mathcal{V} = \{v_1, v_2, \ldots, v_m\}$, $\mathcal{W} = \{w_1, w_2, \ldots, w_n\}$ będą zbiorami wierzchołków a $\mathcal{P}_{(n)} = (P_0, P_1, \ldots, P_n)$ dyskretnym rozkładem prawdopodobieństwa. Oznaczmy przez $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ graf dwudzielny o zbiorze wierzchołków $\mathcal{V} \cup \mathcal{W}$ taki,że:

1. każdy wierzchołek $v \in \mathcal{V}$, niezależnie od innych wierzchołków z $\mathcal{V}$, wybiera swój stopień $X^*(v)$, a potem, także niezależnie, zbiór sąsiadów $\Gamma(v)$;
2. wierzchołek $v \in \mathcal{V}$ wybiera stopień zgodnie z rozkładem $\mathcal{P}_{(n)})$, tzn.
$$\Pr\{X^*(v) = k\} = P_k, \quad k = 0, 1, \ldots, n;$$
3. wierzchołek $v \in \mathcal{V}$ o stopniu k wybiera zbiór sąsiadów jednostajnie spośród wszystkich k-elementowych podzbiorów zbioru $\mathcal{W}$.

Przedmiotem naszego zainteresowania będą dwa grafy losowe o zbiorach wierzchołków odpowiednio $\mathcal{V}$ i $\mathcal{W}$: $\mathcal{IG}_m^{aktywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{IG}_m^{pasywny}(m, \mathcal{P}_{(n)})$, [7]. Pierwszy z nich jest losowym odpowiednikiem przedstawionego wcześniej grafu "obiektów". Oba grafy związane są z grafem $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$.

Definicja 1.5. Aktywnym losowym grafem przecięć $\mathcal{IG}_m^{aktywny}(m, \mathcal{P}_{(n)})$ nazywamy graf o zbiorze wierzchołków $\mathcal{V} = \{v_1, v_2, \ldots, v_m\}$, w którym (v_i, v_j) należy do zbioru krawędzi wtedy i tylko wtedy, gdy v_i i v_j mają wspólnego sąsiada w $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$.

Definicja 1.6. Pasywnym losowym grafem przecięć $\mathcal{IG}_m^{pasywny}(m, \mathcal{P}_{(n)})$ nazywamy graf o zbiorze wierzchołków $\mathcal{W} = \{w_1, w_2, \ldots, w_n\}$, w którym (w_i, w_j) należy do zbioru krawędzi wtedy i tylko wtedy, gdy w_i i w_j mają wspólnego sąsiada w $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$.

W naturalny sposób grafy $\mathcal{IG}_m^{aktywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{IG}_n^{pasywny}(m, \mathcal{P}_{(n)})$ można uogólnić, jeśli dla ustalonego l założymy, że krawędzie powstają, gdy odpowiednio v_i i v_j lub w_i i w_j mają co najmniej l wspólnych sąsiadów. Z powyższych definicji wynika, że w grafie $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ dwa wierzchołki (obiekty) są połączone krawędzią, gdy mają co najmniej jedną wspólną własność (obiekty są podobne), tzn. zbiory przyporządkowanych im własności się przecinają. Natomiast w przypadku grafu $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ między dwoma wierzchołkami (własnościami) występuje krawędź, gdy istnieje pewien obiekt, który ma obie te własności. Natomiast dla uogólnionych grafów przecięć krawędź łączy obiekty w $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$, gdy mają co najmniej l wspólnych własności a w grafie $\mathcal{IG}^{pasywny}(m,\mathcal{P}_{(n)})$ łączy własności, gdy są wspólnymi własnościami co najmniej l obiektów.

Z każdym wierzchołkiem grafu $\mathcal{IG}^{aktywny}(m,\mathcal{P}_{(n)})$ (wierzchołkiem ze zbioru $\mathcal{V}$) możemy powiązać zbiór jego sąsiadów w grafie $\mathcal{BG}_{m,n}(m,\mathcal{P}_{(n)})$. Czyli wierzchołkom z grafu $\mathcal{IG}^{aktywny}(m,\mathcal{P}_{(n)})$ możemy przyporządkować pewne podzbiory zbioru $\mathcal{W}$, które tworzą rodzinę. Tak samo wierzchołkom z $\mathcal{W}$ przyporządkowujemy podzbiory zbioru $\mathcal{V}$ tworzące rodzinę zbiorów. Grafy $\mathcal{IG}^{aktywny}(m,\mathcal{P}_{(n)})$ i $\mathcal{IG}^{pasywny}(m,\mathcal{P}_{(n)})$ nazywamy grafami przecięć, gdyż krawędzie łączą w nich te wierzchołki, dla których przyporządkowane im zbiory (zbiory ich sąsiadów w $\mathcal{BG}_{m,n}(m,\mathcal{P}_{(n)})$) przecinają się.

Na rodzinę podzbiorów pewnego zbioru możemy spojrzeć jak na zbiór krawędzi pewnego hipergrafu. Definicja hipergrafu jest naturalnym uogólnieniem definicji grafu, z tą różnicą, że w hipergrafie krawędziami są nie tylko pary wierzchołków, lecz dowolne podzbiory zbioru wierzchołków.

Definicja 1.7. Hipergrafem $\mathcal{H}=\mathcal{H}(V,E)$ nazywamy parę $\langle V,E\rangle$ składającą się z niepustego zbioru wierzchołków $V=V(\mathcal{H})$ oraz zbioru krawędzi $E=E(\mathcal{H})$ będącego rodziną niepustych podzbiorów zbioru V.

Warto zauważyć, że w powyższej definicji nie występują zbiory puste. Będziemy również zakładać, że nie są uwzględnione krawędzie wielokrotne.

Dla ułatwienia zapisu, dla hipergrafów wprowadzimy oznaczenia: $e(\mathcal{H}) = |E(\mathcal{H})|$. W przypadku hipergrafów z krawędziami o co najwyżej s wierzchołkach przez $\overline{\mathcal{H}}=(m_1,m_2,\ldots,m_s)$ oznaczmy taki ciąg, w którym m_i oznacza liczbę krawędzi w grafie $\mathcal{H}$ o i wierzchołkach dla $i=1,2,\ldots,s$. Oczywiście, dla każdego i mamy: $0\leqslant m_i\leqslant\binom{n}{i}=N_i$. Dla tych ciągów, można wprowadzić naturalną relację częściowego porządku:

$$\underline{m}=(m_1,\ldots,m_s)\prec(m'_1,\ldots,m'_s)=\underline{m'}\Leftrightarrow\ \forall_{1\leqslant i\leqslant s}\, m_i\leqslant m'_i.$$

Jeśli w ciągu tylko dla jednego i, m_i jest różne od zera, wtedy wszystkie krawędzie hipergrafu są tej samej mocy a odpowiedni hipergraf nazywamy jednorodnym.

Definicja 1.8. Hipergraf $\mathcal{H}$ nazywamy s-jednorodnym, gdy wszystkie jego krawędzie są mocy s.

Z definicji wynika, że graf jest hipergrafem o dwuelementowych krawędziach, a dokładniej graf $G(V,E)$ jest 2-jednorodnym hipergrafem $\mathcal{H}(V,E)$.

Tak samo jak dla grafów, w przypadku hipergrafów możemy zdefiniować hipergrafy losowe, w których krawędzie powstają z pewnym prawdopodobieństwem. Dla hipergrafów losowych przyjmiemy oznaczenia jak w [10].

Definicja 1.9. $\mathcal{H}(s,n,M)$ jest to hipergraf losowy wybrany w sposób jednostajny spośród wszystkich hipergrafów s-jednorodnych na n oznaczonych wierzchołkach, o M krawędziach. ($\mathcal{H}(2,n,M)=G(n,M)$ jest grafem losowym)

Definicja 1.10. $\mathcal{H}(s,n,p)$ jest to s-jednorodny hipergraf losowy na n wierzchołkach, w którym dowolna krawędź pojawia się niezależnie od innych krawędzi z prawdopodobieństwem p. ($\mathcal{H}(2,n,p)=G(n,p)$ jest grafem losowym)

Niech $\mathcal{H}$ będzie dowolnym s-jednorodnym hipergrafem na n wierzchołkach o M krawędziach. Wtedy $\Pr(\mathcal{H}(s,n,p)=\mathcal{H}) = p^M(1-p)^{\binom{n}{s}-M}$. Czyli wszystkie hipergrafy o M krawędziach są równie prawdopodobne. Zatem dla dowolnego hipergrafu $\mathcal{H}$ na n wierzchołkach $\Pr\Big\{\mathcal{H}(s,n,p)=\mathcal{H}\ \Big|\mathcal{H}(s,n,p)=M\Big\} = \Pr\Big\{\mathcal{H}(s,n,M)=\mathcal{H}\Big\}$.

Powyższe definicje dotyczą hipergrafów jednorodnych. Uogólnieniem hipergrafów losowych s-jednorodnych są hipergrafy losowe o krawędziach o różnych mocach: $\mathcal{H}(\leqslant s,n,\underline{M})$ i $\mathcal{H}(\leqslant s,n,\underline{p})$.

Definicja 1.11. $\mathcal{H}(\leqslant s,n,\underline{M})$ dla $\underline{M}=(M_1,M_2,\ldots,M_s)$ jest hipergrafem losowym na n oznaczonych wierzchołkach, wybranym w sposób jednostajny spośród wszystkich hipergrafów $\mathcal{H}$ takich, że $\overline{\mathcal{H}}=(M_1,M_2,\ldots,M_s)$.

Definicja 1.12. $\mathcal{H}(\leqslant s,n,\underline{p})$, dla $\underline{p}=(p_1,p_2,\ldots,p_s)$, jest to hipergraf losowy na n wierzchołkach, w którym ustalona krawędź mocy i jest pojawia się niezależnie od innych krawędzi z prawdopodobieństwem p_i dla $i=1,2,\ldots,s$.

Podobnie jak w przypadku losowych hipergrafów s-jednorodnych dla ustalonego hipergrafu $\mathcal{H}$ takiego, że $\overline{\mathcal{H}}=(M_1,M_2,\ldots,M_s)$ prawdopodobieństwo $\Pr(\mathcal{H}(\leqslant s,n,\underline{M})=\mathcal{H})=\prod_{i=1}^{s} p_i^{M_i}(1-p_i)^{N_i-M_i}$. Czyli w modelu $\mathcal{H}(\leqslant s,n,\underline{p})$ wszystkie hipergrafy o dokładnie M_i krawędziach mocy i, $i=1,2,\ldots,s$, są równoprawdopodobne. Czyli dla dowolnego hipergrafu $\mathcal{H}$ mamy:

$$\Pr\Big\{\mathcal{H}(\leqslant s,n,\underline{p})=\mathcal{H}\ \Big|\ \overline{\mathcal{H}(\leqslant s,n,\underline{p})}=\underline{M}\Big\}=\Pr\Big\{\mathcal{H}(\leqslant s,n,\underline{M})=\mathcal{H}\Big\}.$$

Z definicji wynika, że hipergrafy $\mathcal{H}(\leqslant s,n,\underline{M})$ i $\mathcal{H}(\leqslant s,n,\underline{p})$ są sumami niezależnych jednorodnych hipergrafów losowych, gdzie przez sumę hipergrafów o tym samym zbiorze wierzchołków V rozumiemy tutaj taki hipergraf o wierzchołkach z V, który zawiera wszystkie krawędzie z sumowanych hipergrafów.

$$\begin{aligned} E(\mathcal{H}(\leqslant s,n,\underline{p})) &= \bigcup_{i=1}^{s} E(\mathcal{H}(i,n,p_i)) \\ E(\mathcal{H}(\leqslant s,n,\underline{M})) &= \bigcup_{i=1}^{s} E(\mathcal{H}(i,n,M_i)). \end{aligned} \tag{1.1}$$

W pracy skupimy się na własnościach związanych ze spójnością grafów przecięć $\mathcal{IG}^{aktywny}(m,\mathcal{P}_{(n)})$ i $\mathcal{IG}^{pasywny}(m,\mathcal{P}_{(n)})$.

Definicja 1.13. Niech $v,w\in V$ będą wierzchołkami grafu $G(V,E)$. Niezerowy ciąg $(v=v_0,e_0,v_1,e_1,\ldots,e_k,v_{k+1}=w)$ taki, że $v_0,v_1,\ldots,v_{k+1}\in V(G)$ oraz dla każdego $1\leqslant i\leqslant k$, $\{v_i,v_{i+1}\}=e_i\in E(G)$ będziemy nazywać (v,w)-spacerem.

Definicja 1.14. Niech $v, w \in V$ będą wierzchołkami grafu $G(V, E)$. (v, w)-ścieżką będziemy nazywać (v, w)-spacer, w którym wierzchołki i krawędzie są parami różne.

Fakt 1.15. Dla dowolnych różnych wierzchołków $v, w \in V$ w grafie $G(V, E)$ istnieje (v, w)-ścieżka wtedy i tylko wtedy, gdy istnieje (v, w)-spacer.

Definicja 1.16. Graf $G(V, E)$ nazywamy spójnym, gdy dla dowolnych dwóch wierzchołków $v, w \in V$ istnieje (v, w)-ścieżka.

Fakt 1.17. Graf $G(V, E)$ jest spójny wtedy i tylko wtedy, gdy dla dowolnych dwóch wierzchołków $v, w \in V$ istnieje (v, w)-spacer.

Definicja 1.18. Wierzchołek $v \in V$ stopnia zero ($d(v) = 0$) w grafie $G(V, E)$ nazywamy wierzchołkiem izolowanym.

Większość definicji dotyczących grafów ma naturalne uogólnienia dla hipergrafów.

Definicja 1.19. Niech $\mathcal{H}(V, E)$ będzie hipergrafem, $v, w \in V$. Hiper-spacerem między v i w w $\mathcal{H}(V, E)$ będziemy nazywać ciąg wierzchołków i krawędzi hipergrafu $(v = v_0, e_0, v_1, e_1, \ldots, e_k, v_{k+1} = w)$ taki, że $v_0, v_1, \ldots, v_{k+1} \in V(\mathcal{H})$ oraz dla każdego $1 \leqslant i \leqslant k$, $\{v_i, v_{i+1}\} \subseteq e_i \in E(\mathcal{H})$.

Definicja 1.20. Hiper-ścieżką między $v \in V$ i $w \in V$ w hipergrafie $\mathcal{H}(V, E)$ będziemy nazywać hiper-spacer, w którym wierzchołki i krawędzie są parami różne.

Fakt 1.21. Dla dowolnych różnych wierzchołków $v, w \in V$ w hipergrafie $\mathcal{H}(V, E)$ między v i w istnieje hiper-ścieżka wtedy i tylko wtedy, gdy istnieje hiper-spacer między tymi wierzchołkami.

Definicja 1.22. Hipergraf $\mathcal{H}(V, E)$ nazywamy spójnym, gdy dla dowolnych dwóch wierzchołków $v, w \in V$ istnieje hiper-ścieżka między v i w.

Definicja 1.23. Hipergraf $\mathcal{H}(V, E)$ jest spójny, gdy dla dowolnych dwóch wierzchołków $v, w \in V$ istnieje hiper-spacer między v i w.

Fakt 1.24. Hipergraf $\mathcal{H}(V, E)$ jest spójny, gdy dla dowolnego podziału zbioru wierzchołków V ma niepuste podzbiory V_1, V_2, istnieje krawędź $e \in E$ taka, że $V_1 \cap e \neq \emptyset$ i $V_2 \cap e \neq \emptyset$.

Tak samo jak w przypadku grafów, tak i dla hipergrafów wprowadzamy definicję stopnia wierzchołka.

Definicja 1.25. Stopniem krawędziowym $d(v)$ wierzchołka $v \in V$ w hipergrafie $\mathcal{H}$ nazywamy liczbę krawędzi, w których zawarty jest v.

Definicja 1.26. W hipergrafie $\mathcal{H}(V, E)$ i-stopniem krawędziowym $d_i(v)$ wierzchołka $v \in V$ nazywamy liczbę krawędzi o mocy i zawierających v.

Definicja 1.27. W hipergrafie $\mathcal{H}(V,E)$ i-stopniem wierzchołkowym $d_i^*(v)$ wierzchołka $v \in V$ nazywamy liczbę $d_i^*(v) = (i-1)\, d_i(v)$.

Definicja 1.28. Stopniem wierzchołkowym wierzchołka $v \in V$ w hipergrafie $\mathcal{H}(V,E)$ nazywamy liczbę $d^*(v) = \sum_{i=1,2\ldots} d_i^*(v) = \sum_{i=1,2\ldots}(i-1)\, d_i(v)$.

Oczywiście $d(v) = \sum_{i=1,2,\ldots} d_i(v)$. Zauważmy też, że jeśli wierzchołek $v \in V$ nie należy do żadnej krawędzi lub należy tylko do krawędzi jednoelementowej, to jego stopień wierzchołkowy jest równy zero. Dla hipergrafu 2-jednorodnego (grafu) stopień wierzchołkowy dowolnego wierzchołka jest równy stopniowi krawędziowemu tego wierzchołka i stopniowi tego wierzchołka w grafie.

Definicja 1.29. Wierzchołek $v \in V$ nazywamy wierzchołkiem izolowanym w hipergrafie $\mathcal{H}(V,E)$, gdy jego stopień wierzchołkowy jest równy zero.

Definicja 1.30. Krawędź $e \in E$ hipergrafu $\mathcal{H}(V,E)$ nazywamy krawędzią izolowaną, gdy dla dowolnej krawędzi $e' \in E$, takiej, że $e' \neq e$, $e \cap e' = \emptyset$.

Definicja 1.31. Podzbiór $S \subseteq V(\mathcal{H})$ nazywamy spójnym, gdy istnieje hiper-ścieżka między dowolną parą wierzchołków z S.

Definicja 1.32. Podzbiór $S \subseteq V(\mathcal{H})$ nazywamy izolowanym, gdy nie istnieje krawędź $e \in E(\mathcal{H})$ taka, że $S \cap e \neq \emptyset$ i $(V \setminus S) \cap e \neq \emptyset$.

Definicja 1.33. Podzbiór $S \subseteq V(\mathcal{H})$ nazywamy składową spójności, jeśli jest spójny i izolowany.

Fakt 1.34. Dwie różne składowe spójności hipergrafu są wierzchołkowo rozłączne.

Definicja 1.35. Hiperdrzewem rozpinającym na $S \subseteq V(\mathcal{H})$ w $\mathcal{H}(V,E)$ nazywamy spójny hipergraf $\mathcal{H}'(S,E)$, taki, że $E_s \subseteq E$ oraz usunięcie dowolnej krawędzi z E_s go rozspójni.

Definicja 1.36. Cyklem w hipergrafie $\mathcal{H}(V,E)$ nazywamy ciąg wierzchołków i parami różnych krawędzi $(v_0, e_0, v_1, e_1, \ldots, v_k, e_k, v_0)$ taki, że $(v_0, e_0, v_1, e_1, \ldots, v_k)$ jest hiper-ścieżką oraz $\{v_k, v_0\} \subseteq e_k \in E(\mathcal{H})$.

Chociaż hiperdrzewo jest naturalnym uogólnieniem drzewa, grafu acyklicznego, zdecydowanie się od niego różni. Rozważmy bowiem pewne hiperdrzewo rozpinające na S. Niech hiperdrzewo to zawiera pewne krawędzie e i e', takie że $v, v', v'' \in e$ i $v', v'', v''' \in e'$ Dodatkowo niech e będzie jedyną krawędzią w hiperdrzewie do której należy v oraz e' jedyną krawędzią zawierającą v'''. Wtedy rzeczywiście usunięcie krawędzi e lub e' rozspójniłoby hiperdrzewo a hiperdrzewo rozpinające zawiera cykl (v', e, v'', e', v'), Widzimy więc, że hiperdrzewo to zawiera cykl, z którego nie można usunąć krawędzi.

Fakt 1.37. Hiperdrzewo rozpinające może zawierać wiele cykli.

Zauważmy, że w hiperdrzewie bez cykli dowolne dwie krawędzie przecinają się na co najwyżej jednym wierzchołku, gdyż gdyby istniały krawędzie $e', e'' \in E$ takie, że dla pewnych wierzchołków $v', v'' \in e' \cap e''$, to ciąg (v', e', v'', e'', v') byłby cyklem. Jeśli więc zaczniemy konstrukcję takiego drzewa bez cykli od pewnej krawędzi i kolejno będziemy dodawać do drzewa po jednej krawędzi, tak by zawsze otrzymać hipergraf spójny, to dodana krawędź będzie miała zawsze dokładnie jeden wierzchołek wspólny z wcześniej dołączonymi krawędziami (dwa wierzchołki wspólne zamykałyby pewien cykl). Zatem w każdym kroku, wraz z krawędzią o i wierzchołkach dodajemy dokładnie $i - 1$ wierzchołków do hiperdrzewa. Stąd wynika następujący fakt.

Fakt 1.38. Jeśli przez T_i oznaczymy liczbę krawędzi o i wierzchołkach w hiperdrzewie bez cykli, rozpinającym S, wtedy $\sum_i (i-1)T_i = |S| - 1$.

Z drugiej strony dodana i-wierzchołkowa krawędź, która do hiperdrzewa dołączyłaby tylko $i - l (2 \leqslant l \leqslant i)$ wierzchołków zamyka co najmniej $l - 1$ cykli.

Definicja 1.39. Niech $\mathcal{H}(V, E)$ będzie hipergrafem, $S \subseteq V$ jego składową spójności hipergrafu o dokładnie T_i krawędziach o mocy i, $i = 1, 2, \ldots$. Indeksem cyklicznym składowej spójności S nazywamy liczbę $(\sum_i (i-1)T_i - (|S| - 1))$.

Fakt 1.40. Składowa spójności o indeksie cyklicznym c ma co najmniej c cykli.

2. Własności monotoniczne

Jeśli pewien hipergraf $\mathcal{H}(V, E_1)$ jest spójny, to dowolny hipergraf $\mathcal{H}(V, E_2)$ taki, że $E_1 \subseteq E_2$ też jest spójny. Własności o takich cechach nazywamy własnościami monotonicznymi. Twierdzenia dotyczące spójności hipergrafów określają prawdopodobieństwo spójności dla hipergrafów $\mathcal{H}(\leqslant s, n, \underline{p})$. W praktyce, dla $\underline{m}$ zależnego od $\underline{p}$, przy n dążącym do nieskończoności, prawdopodobieństwo, że $\mathcal{H}(\leqslant s, n, \underline{p})$ ma pewną własność monotoniczną jest równe prawdopodobieństwu, że $\mathcal{H}(\leqslant s, n, \underline{m})$ ma tę własność.

Własności monotoniczne można zdefiniować ogólnie, nie tylko jako własności hipergrafów i grafów, lecz jako własności dowolnych rodzin podzbiorów. Tak na przykład przedstawiono to w książce [9]. W pracy będziemy używać oznaczeń tam wprowadzonych. Niech Γ oznacza dowolny skończony zbiór. Na przykład zbiór składający się z n pierwszych liczb naturalnych. Będziemy go oznaczać przez $[n] = \{1, 2, \ldots, n\}$. Niech $N = |\Gamma|$, $0 \leqslant p \leqslant 1$, $0 \leqslant M \leqslant N$. Zdefiniujemy dwa rodzaje podzbiorów losowych zbioru Γ: Γ_p będzie losowym podzbiorem Γ, do którego dowolny element z Γ jest dołączony z prawdopodobieństwem p niezależnie od innych elementów, natomiast zbiór Γ_M będzie losowym podzbiorem zbioru Γ wybranym w sposób jednostajny spośród wszystkich M-elementowych podzbiorów zbioru Γ.

Fakt 2.1. Niech Γ^s będzie zbiorem wszystkich s-elementowych podzbiorów pewnego zbioru n-elementowego ($\Gamma^s = \{A \subseteq [n] : |A| = k\}$). Jeżeli $V = [n]$ będzie zbiorem wierzchołków hipergrafu losowego, wtedy możemy utożsamić Γ^s_p z $\mathcal{H}(s, n, p)$ oraz Γ^s_M z $\mathcal{H}(s, n, M)$.

Wykorzystując powyższy fakt i zależność (1.1) można napisać:

Fakt 2.2. Hipergrafy $\mathcal{H}(\leqslant s, n, \underline{p})$ i $\mathcal{H}(\leqslant s, n, \underline{M})$ możemy utożsamić odpowiednio z sumami niezależnych zbiorów $\bigcup_{i=1}^{s} \Gamma^i_{p_i}$ i $\bigcup_{i=1}^{s} \Gamma^i_{M_i}$, gdzie Γ^i oznacza zbiór wszystkich i-elementowych podzbiorów zbioru n-elementowego $[n]$.

Definicja 2.3. Rodzina $\mathcal{Q}$ podzbiorów zbioru Γ ($\mathcal{Q} \subseteq 2^{\Gamma}$) nazywa się rosnąca, gdy dla $A \subseteq B$ jeżeli $A \in \mathcal{Q}$, to $B \in \mathcal{Q}$. Rodzina $\mathcal{Q}$ jest malejąca, gdy jest dopełnieniem własności rosnącej lub równoważnie, gdy dla dowolnych $A \subseteq B$ i $B \in \mathcal{Q}$ zachodzi $A \in \mathcal{Q}$.

Definicja 2.4. Własność $\mathcal{Q}$ nazywamy wypukłą, gdy z tego, że $A \subseteq B \subseteq C$ i $A, C \in \mathcal{Q}$ wynika, że $B \in \mathcal{Q}$.

Przyjmijmy za Γ wszystkie co najmniej 2-elementowe podzbiory zbioru n-elementowego ($V = [n]$). Wtedy dowolny skończony podzbiór zbioru Γ jest zbiorem krawędzi pewnego hipergrafu o zbiorze wierzchołków V. Załóżmy, że hipergraf ma własność $\mathcal{Q}$, gdy odpowiadający mu podzbiór losowy należy do $\mathcal{Q}$. W tym ujęciu rodzina $\mathcal{Q}$- podzbiorów zbioru Γ jest po prostu własnością hipergrafów o n wierzchołkach, o ile jest zamknięta ze względu na izomorfizmy hipergrafów. Przykładami własności rosnącej hipergrafów jest spójność hipergrafu oraz posiadanie największej składowej spójności o co najmniej k wierzchołkach (dla $k \in \mathbb{N}$). Natomiast przykładem własności wypukłej jest posiadanie dokładnie k wierzchołków izolowanych. Wykorzystując pewne cechy własności monotonicznych, udowodnimy równoważność hipergrafów $\mathcal{H}(s, n, p)$ z $\mathcal{H}(s, n, M)$ oraz ogólniej $\mathcal{H}(\leqslant s, n, \underline{p})$ z $\mathcal{H}(\leqslant s, n, \underline{m})$.

Lemat 2.5. Niech $\mathcal{Q}$ będzie rosnącą własnością podzbiorów zbioru Γ oraz $0 \leqslant M_1 \leqslant M_2 \leqslant N$ i $0 \leqslant p_1 \leqslant p_2 \leqslant 1$, wtedy

$$\Pr\{\Gamma_{p_1} \in \mathcal{Q}\} \leqslant \Pr\{\Gamma_{p_2} \in \mathcal{Q}\} \quad \text{i} \quad \Pr\{\Gamma_{M_1} \in \mathcal{Q}\} \leqslant \Pr\{\Gamma_{M_2} \in \mathcal{Q}\}$$

Dowód. Oznaczmy przez $p_0 = \frac{p_2 - p_1}{1 - p_1}$, wtedy Γ_{p_2} można rozpatrywać jako sumę dwóch niezależnie powstających losowych podzbiorów Γ_{p_0} i Γ_{p_1}. Prawdopodobieństwo, że ustalony element zbioru Γ należy do $\Gamma_{p_0} \cup \Gamma_{p_1}$ jest równe: $p_0 + p_1 - p_0 p_1 = \frac{p_2 - p_1}{1 - p_1} + p_1 - \frac{p_2 - p_1}{1 - p_1} p_1 = p_2$ oraz elementy pojawiają się niezależnie. Czyli $\Pr\{\Gamma_{p_2} \in \mathcal{Q}\} = \Pr\{\Gamma_{p_0} \cup \Gamma_{p_1} \in \mathcal{Q}\}$. Natomiast jeżeli $\Gamma_{p_1} \in \mathcal{Q}$, to $\Gamma_{p_0} \cup \Gamma_{p_1} \in \mathcal{Q}$, czyli

$$\Pr\{\Gamma_{p_1} \in \mathcal{Q}\} \leqslant \Pr\{\Gamma_{p_0} \cup \Gamma_{p_1} \in \mathcal{Q}\} = \Pr\{\Gamma_{p_2} \in \mathcal{Q}\}.$$

Γ_M może być rozpatrywany jako w M-ty podzbiór powstały w wyniku procesu polegającego na kolejnym dodawaniu elementów z Γ. Oznaczmy przez $\mathcal{Q}_M$ rodzinę wszystkich podzbiorów M-elementowych zbioru Γ, które są zawarte w $\mathcal{Q}$. Przez $\mathcal{Q}_M^*$ oznaczmy rodzinę $M+1$-elementowych nadzbiorów rodziny $\mathcal{Q}_M$. Ponieważ własność $\mathcal{Q}$ jest rosnąca, więc $\mathcal{Q}_M^* \subseteq \mathcal{Q}$. Z własności nadzbiorów mamy: $\Pr\{\Gamma_M \in \mathcal{Q}\} = |\mathcal{Q}_M| / \binom{N}{M} \leqslant |\mathcal{Q}_M^*| / \binom{N}{M+1} \leqslant \Pr\{\Gamma_{M+1} \in \mathcal{Q}\}$, stąd dla $M_1 \leqslant M_2$

$$\Pr\{\Gamma_{M_1} \in \mathcal{Q}\} \leqslant \Pr\{\Gamma_{M_2} \in \mathcal{Q}\}.$$

$\square$

Ponieważ $\mathcal{H}(\leqslant s, n, \underline{p})$ i $\mathcal{H}(\leqslant s, n, \underline{M})$ mogą być postrzegane jako sumy niezależnych podzbiorów $\Gamma_{p_i}^i$ i $\Gamma_{m_i}^i$, zatem identyczne argumenty jak w dowodzie lematu 2.5 prowadzą nas do wniosku, że:

Lemat 2.6. Niech $\mathcal{Q}$ będzie rosnącą własnością hipergrafów oraz $\underline{M} \prec \underline{M'}$ i $\underline{p} \prec \underline{p'}$, wtedy

$$\Pr\{\mathcal{H}(\leqslant s, n, \underline{p}) \in \mathcal{Q}\} \leqslant \Pr\{\mathcal{H}(\leqslant s, n, \underline{p'}) \in \mathcal{Q}\}$$
$$\text{i}$$
$$\Pr\{\mathcal{H}(\leqslant s, n, \underline{M}) \in \mathcal{Q}\} \leqslant \Pr\{\mathcal{H}(\leqslant s, n, \underline{M'}) \in \mathcal{Q}\}$$

Analogiczne nierówności dla własności malejących otrzymujemy powtarzając rozumowanie krok po kroku lub też korzystając z własności, że dopełnienie własności malejącej jest własnością rosnącą.

Twierdzenie 2.7. *Niech $\mathcal{Q}$ będzie własnością monotoniczną podzbiorów zbioru Γ, $0 \leqslant a \leqslant 1$ oraz $M = M(n)$ i $N = N(n)$ będą ciągami liczb całkowitych takimi, że $0 \leqslant M \leqslant N$. Wtedy jeżeli dla każdego ciągu $p = p(n) \in (0,1)$ takiego, że $p = M/N + O(\sqrt{M(N-M)/N^3})$ mamy $\Pr\{\Gamma_p \in \mathcal{Q}\} \to a$, to $\Pr\{\Gamma_M \in \mathcal{Q}\} \to a$, dla $n \to \infty$.*

Dowód. Rozważmy najpierw przypadek, gdy własność $\mathcal{Q}$ jest rosnąca. Niech C będzie dowolną dużą stałą. Wprowadźmy oznaczenia:

$$\begin{aligned}
p_0 &= \frac{M}{N}, \\
q_0 &= 1 - p_0, \\
p_+ &= \min\left\{p_0 + C\sqrt{\frac{p_0 q_0}{N}}\ ,\ 1\right\} \\
p_- &= \max\left\{p_0 - C\sqrt{\frac{p_0 q_0}{N}},\ 0\right\}
\end{aligned}$$

Ze wzoru na prawdopodobieństwo całkowite i monotoniczności $\mathcal{Q}$ mamy

$$\begin{aligned}
\Pr\{\Gamma_{p_+} \in \mathcal{Q}\} &= \sum_{M'=1}^{N} \Pr\{\Gamma_{M'} \in \mathcal{Q}\}\ \Pr\{|\Gamma_{p_+}| = M'\} \geqslant \\
&\geqslant \sum_{M' \geqslant M} \Pr\{\Gamma_{M'} \in \mathcal{Q}\}\ \Pr\{|\Gamma_{p_+}| = M'\} \geqslant \\
&\geqslant \Pr\{\Gamma_M \in \mathcal{Q}\}\ \Pr\{|\Gamma_{p_+}| \geqslant M\} = \\
&= \Pr\{\Gamma_M \in \mathcal{Q}\}\ (1 - \Pr\{|\Gamma_{p_+}| < M\}) = \\
&= \Pr\{\Gamma_M \in \mathcal{Q}\} - \Pr\{\Gamma_M \in \mathcal{Q}\}\ \Pr\{|\Gamma_{p_+}| < M)\} \geqslant \\
&\geqslant \Pr\{\Gamma_M \in \mathcal{Q}\} - \Pr\{|\Gamma_{p_+}| < M)\}
\end{aligned} \tag{2.1}$$

Podobnie

$$
\begin{aligned}
\Pr\{\Gamma_{p_-} \in \mathcal{Q}\} &= \sum_{M'=1}^{N} \mathbb{P}\{\Gamma_{M'} \in \mathcal{Q}\}\ \Pr\{|\Gamma_{p_-}| = M'\} \leqslant \\
&\leqslant \sum_{M' \leqslant M} \Pr\{\Gamma_{M'} \in \mathcal{Q}\}\ \Pr\{|\Gamma_{p_-}| = M'\} + \Pr\{|\Gamma_{p_-}| > M\} \leqslant \\
&\leqslant \Pr\{\Gamma_M \in \mathcal{Q}\}\ \Pr\{|\Gamma_{p_-}| \leqslant M\} + \Pr\{|\Gamma_{p_-}| > M\} \leqslant \\
&\leqslant \Pr\{\Gamma_M \in \mathcal{Q}\} + \Pr\{|\Gamma_{p_-}| > M\}
\end{aligned} \tag{2.2}
$$

Dla $M = 0$ i $p = 0$, Γ_M oraz Γ_p są zawsze podzbiorami pustymi (w Γ_M wybieramy z Γ podzbiór o mocy zero, natomiast w Γ_p prawdopodobieństwo wybrania dowolnego elementu z Γ jest równe zero). Mamy zatem, że dla $M = 0$ i $p = M/N + O(\sqrt{M(N-M)/N^3}) = 0$ zachodzi $\Gamma_M = \Gamma_p$, czyli oczywiście jeżeli $\Pr\{\Gamma_p \in \mathcal{Q}\} \to a$, to $\Pr\{\Gamma_M \in \mathcal{Q}\} \to a$, przy $n \to \infty$.
Podobnie, dla $M = N$ i $p = M/N + O(\sqrt{M(N-M)/N^3}) = 1$, $\Gamma_M = \Gamma_p = \Gamma$, stąd jeśli $\Pr\{\Gamma_p \in \mathcal{Q}\} \to a$, to $\Pr\{\Gamma_M \in \mathcal{Q}\} \to a$, przy $n \to \infty$.
Załóżmy zatem, że $1 \leqslant M \leqslant N-1$ i $N \geqslant 2$, wtedy

$$
Np_0q_0 = \frac{M(N-M)}{N} \geqslant \frac{1(N-1)}{N} \geqslant \frac{1}{2}.
$$

Wykorzystamy nierówności

$$
\begin{aligned}
p_+ &= \min\left\{p_0 + C\sqrt{\frac{p_0q_0}{N}}, 1\right\} \leqslant p_0 + C\sqrt{\frac{p_0q_0}{N}} \\
p_- &= \max\left\{p_0 - C\sqrt{\frac{p_0q_0}{N}}, 0\right\} \leqslant p_0
\end{aligned}
$$

oraz

$$
\begin{aligned}
1 - p_+ &\leqslant 1 - p_0 = q_0 \\
1 - p_- &\leqslant 1 - p_0 + C\sqrt{\frac{p_0q_0}{N}}
\end{aligned}
$$

oraz $|\Gamma_{p_+}|$ i $|\Gamma_{p_-}|$ mają rozkład dwumianowy, więc

$$
\begin{aligned}
\mathbb{E}|\Gamma_{p_+}| &= Np_+ \\
Var|\Gamma_{p_+}| &= Np_+(1-p_+) \leqslant N\left(p_0 + C\sqrt{\frac{p_0q_0}{N}}\right)q_0 \\
\mathbb{E}|\Gamma_{p_-}| &= Np_- \\
Var|\Gamma_{p_-}| &= Np_-(1-p_-) \leqslant Np_0\left(q_0 + C\sqrt{\frac{p_0q_0}{N}}\right).
\end{aligned}
$$

Z nierówności Czebyszewa:

$$\begin{aligned}
\Pr\{|\Gamma_{p_+}| < M\} &= \Pr\{Np_+ - |\Gamma_{p_+}| > Np_+ - M\} \leqslant \\
&\leqslant \Pr\{|\mathbb{E}|\Gamma_{p_+}| - |\Gamma_{p_+}|| > Np_+ - Np_0\} \leqslant \\
&\leqslant \frac{N\left(p_0 + C\sqrt{\frac{p_0 q_0}{N}}\right) q_0}{N^2(p_+ - p_0)^2} \leqslant \\
&\leqslant \frac{Np_0q_0 + C\sqrt{Np_0q_0}}{N^2C^2\frac{p_0q_0}{N}} = \\
&= \frac{1}{C^2} + \frac{1}{C\sqrt{Np_0q_0}} \leqslant \frac{1}{C^2} + \frac{\sqrt{2}}{C}
\end{aligned} \tag{2.3}$$

$$\begin{aligned}
\Pr\{|\Gamma_{p_-}| > M\} &= \Pr\{|\Gamma_{p_-}| - Np_- > M - Np_-\} \leqslant \\
&\leqslant \Pr\{||\Gamma_{p_+}| - Np_-| > Np_0 - Np_-\} \leqslant \\
&\leqslant \frac{Np_0\left(q_0 + C\sqrt{\frac{p_0 q_0}{N}}\right)}{N^2(p_0 - p_-)^2} \leqslant \\
&\leqslant \frac{Np_0q_0 + C\sqrt{Np_0q_0}}{N^2C^2\frac{p_0q_0}{N}} = \\
&= \frac{1}{C^2} + \frac{1}{C\sqrt{Np_0q_0}} \leqslant \frac{1}{C^2} + \frac{\sqrt{2}}{C}\ .
\end{aligned} \tag{2.4}$$

Z (2.1) i (2.2)

$\Pr\{\Gamma_{p_-} \in \mathcal{Q}\} - \Pr\{|\Gamma_{p_-}| < M\} \leqslant \Pr\{\Gamma_M \in \mathcal{Q}\} \leqslant \Pr\{\Gamma_{p_+} \in \mathcal{Q}\} + \Pr\{|\Gamma_{p_+}| < M\}$,

czyli

$$\Pr\{\Gamma_{p_-} \in \mathcal{Q}\} - \frac{1}{C^2} - \frac{\sqrt{2}}{C} \leqslant \Pr\{\Gamma_M \in \mathcal{Q}\} \leqslant \Pr\{\Gamma_{p_+} \in \mathcal{Q}\} + \frac{1}{C^2} + \frac{\sqrt{2}}{C}. \tag{2.5}$$

Ponieważ z założenia mamy, że

$$\lim_{n\to\infty} \Pr\{\Gamma_{p_-} \in \mathcal{Q}\} = \lim_{n\to\infty} \Pr\{\Gamma_{p_+} \in \mathcal{Q}\} = a,$$

więc ustalając, że $C \to \infty$ przy $n \to \infty$ z (2.5) wynika, że,

$$\lim_{n\to\infty} \Pr\{\Gamma_M \in \mathcal{Q}\} = a.$$

W przypadku własności malejącej dowód jest analogiczny. Zmieniają się jedynie oszacowania (2.1) i (2.2):

$$\begin{aligned}
\Pr\{\Gamma_{p_-} \in \mathcal{Q}\} &\geqslant \sum_{M' \leqslant M} \Pr\{\Gamma_{M'} \in \mathcal{Q}\} \Pr\{|\Gamma_{p_-}| = M'\} \geqslant \\
&\geqslant \Pr\{\Gamma_M \in \mathcal{Q}\}\ \Pr\{|\Gamma_{p_-}| \leqslant M\} = \\
&= \Pr\{\Gamma_M \in \mathcal{Q}\}\ (1 - \Pr\{|\Gamma_{p_-}| > M\}) \geqslant \\
&\geqslant \Pr\{\Gamma_M \in \mathcal{Q}\} - \Pr\{|\Gamma_{p_-}| > M\}
\end{aligned}$$

$$\begin{aligned}\Pr\{\Gamma_{p_+} \in \mathcal{Q}\} &\leqslant \sum_{M' \geqslant M} \Pr\{\Gamma_{M'} \in \mathcal{Q}\}\ \Pr\{|\Gamma_{p_+}| = M'\} + \Pr\{|\Gamma_{p_+}| < M\} \leqslant \\ &\leqslant \Pr\{\Gamma_M \in \mathcal{Q}\}\ \Pr\{|\Gamma_{p_+}| \geqslant M\} + \Pr\{|\Gamma_{p_+}| < M\} \leqslant \\ &\leqslant \Pr\{\Gamma_M \in \mathcal{Q}\} + \Pr\{|\Gamma_{p_+}| < M)\}\end{aligned}$$

□

Powyższe twierdzenie dotyczy podzbiorów Γ_p, do których każdy element z Γ jest dołączony z tym samym prawdopodobieństwem. Pewnym rodzinom podzbiorów typu Γ_p odpowiadają hipergrafy s-jednorodne $\mathcal{H}(s,n,p)$ oraz hipergrafy $\mathcal{H}(\leqslant s,n,\underline{p})$, w których wszystkie krawędzie powstają z równym prawdopodobieństwem (tzn. $\underline{p} = (p,\ldots,p)$).

Zauważmy jednak, że jeżeli w hipergrafie losowym $\mathcal{H}(\leqslant s,n,\underline{p})$ przyjmiemy, że wszystkie krawędzie powstają z tym samym prawdopodobieństwem p, to w przypadku zagadnienia spójności, przy wynikach asymptotycznych, znaczenie mają tylko krawędzie z największą liczbą wierzchołków. Dlatego warto uogólnić twierdzenie 2.7.

Twierdzenie 2.8. *Niech $\mathcal{Q}$ będzie własnością monotoniczną hipergrafów, $0 \leqslant a \leqslant 1$ oraz $M_i = M_i(n)$ i $N_i = \binom{n}{i}$ będą ciągami liczb całkowitych takimi, że $0 \leqslant M_i < N_i$ dla $i = 2\ldots s$. Wtedy jeżeli dla każdego ciągu $\underline{(p_2(n), p_3(n),\ldots,p_s(n))}$ takiego, że $p_i = p_i(n) \in (0,1)$ oraz $p_i = M_i/N_i + O(\sqrt{M_i(N_i-M_i)/N_i^3})$, dla $i = 2,\ldots,s$, mamy $\Pr\{\mathcal{H}(\leqslant s,n,\underline{p}) \in \mathcal{Q}\} \to a$, to $\Pr\{\mathcal{H}(\leqslant s,n,\underline{M}) \in \mathcal{Q}\} \to a$, dla $n \to \infty$.*

Dowód. Jak poprzednio rozważymy przypadek, gdy $\mathcal{Q}$ jest własnością rosnącą, gdyż dowód dla własności malejącej jest analogiczny. Niech C będzie dowolną dużą stałą. Oznaczmy przez:

$$p_{i0} = \frac{M_i}{N_i}, \qquad p_{i+} = \min\left\{p_{i0} + C\sqrt{\frac{p_{i0}q_{i0}}{N_i}}, 1\right\}$$
$$q_{i0} = 1 - p_{i0}, \qquad p_{i-} = \max\left\{p_{i0} - C\sqrt{\frac{p_{i0}q_{i0}}{N_i}}, 0\right\}.$$

Gdyby dla pewnego i $M_i = 0$, wtedy pomijamy ten indeks w dalszych rozważaniach, gdyż znaczenie mają tylko krawędzie o liczbie wierzchołków różnej od i.

Jak w 2.1 i 2.2 wykorzystamy wzór na prawdopodobieństwo całkowite i monotoniczność własności $\mathcal{Q}$.

(2.6)

$$\begin{aligned}
&\Pr\{\mathcal{H}(\leqslant s,n,\underline{p_+})\in\mathcal{Q}\} = \\
&\quad= \sum_{\underline{m}} \Pr\{\mathcal{H}(\leqslant s,n,\underline{m})\in\mathcal{Q}\}\ \Pr\{\overline{\mathcal{H}(\leqslant s,n,\underline{p_+})}=\underline{m}\} \geqslant \\
&\quad\geqslant \sum_{\underline{M}\prec\underline{m}} \Pr\{\mathcal{H}(\leqslant s,n,\underline{m})\in\mathcal{Q}\}\ \Pr\{\overline{\mathcal{H}(\leqslant s,n,\underline{p_+})}=\underline{m}\} \geqslant \\
&\quad\geqslant \Pr\{\mathcal{H}(\leqslant s,n,\underline{M})\in\mathcal{Q}\}\ \Pr\{\underline{M}\prec\overline{\mathcal{H}(\leqslant s,n,\underline{p_+})}\} = \\
&\quad= \Pr\{\mathcal{H}(\leqslant s,n,\underline{M})\in\mathcal{Q}\}\ (1-\Pr\{\underline{M}\not\prec\overline{\mathcal{H}(\leqslant s,n,\underline{p_+})}\}) \geqslant \\
&\quad\geqslant \Pr\{\mathcal{H}(\leqslant s,n,\underline{M})\in\mathcal{Q}\} - \Pr\{\underline{M}\not\prec\overline{\mathcal{H}(\leqslant s,n,\underline{p_+})}\}.
\end{aligned}$$

Oraz

(2.7)

$$\begin{aligned}
&\Pr\{\mathcal{H}(\leqslant s,n,\underline{p_-})\in\mathcal{Q}\} = \\
&\quad= \sum_{\underline{m}} \Pr\{\mathcal{H}(\leqslant s,n,\underline{m})\in\mathcal{Q}\}\Pr\{\overline{\mathcal{H}(\leqslant s,n,\underline{p_-})}=\underline{m}\} \leqslant \\
&\quad\leqslant \sum_{\underline{m}\prec\underline{M}} \Pr\{\mathcal{H}(\leqslant s,n,\underline{m})\in\mathcal{Q}\}\ \Pr\{\overline{\mathcal{H}(\leqslant s,n,\underline{p_-})}=\underline{M}\} + \\
&\qquad\qquad + \Pr\{\overline{\mathcal{H}(\leqslant s,n,\underline{p_-})}\not\prec\underline{M}\} \leqslant \\
&\quad\leqslant \Pr\{\mathcal{H}(\leqslant s,n,\underline{M})\in\mathcal{Q}\} + \Pr\{\overline{\mathcal{H}(\leqslant s,n,\underline{p_-})}\not\prec\underline{M}\}.
\end{aligned}$$

Ponieważ $E(\mathcal{H}(\leqslant s,n,\underline{p_+})) = \bigcup_{0\leqslant i\leqslant s}\Gamma^i_{p_{i+}}$ oraz $E(\mathcal{H}(\leqslant s,n,\underline{p_-})) = \bigcup_{0\leqslant i\leqslant s}\Gamma^i_{p_{i-}}$, więc korzystając z (2.3) i (2.4) mamy:

$$\Pr\{\underline{M}\not\prec\overline{\mathcal{H}(\leqslant s,n,\underline{p_+})}\} \leqslant \Pr\{\exists_i|\Gamma^i_{p_{i+}}|<M_i\} \leqslant s\left(\frac{1}{C^2}+\frac{\sqrt{2}}{C}\right)$$

i

$$\Pr\{\overline{\mathcal{H}(\leqslant s,n,\underline{p_-})}\not\prec\underline{M}\} \leqslant \Pr\{\exists_i|\Gamma^i_{p_{i-}}|>M_i\} \leqslant s\left(\frac{1}{C^2}+\frac{\sqrt{2}}{C}\right)$$

Czyli

$$\begin{aligned}
&\Pr\{\mathcal{H}(\leqslant s,n,\underline{p_-})\in\mathcal{Q}\} - \frac{s}{C^2} - \frac{s\sqrt{2}}{C} \leqslant \Pr\{\mathcal{H}(\leqslant s,n,\underline{M})\in\mathcal{Q}\} \leqslant \\
&\qquad\qquad \leqslant \Pr\{\mathcal{H}(\leqslant s,n,\underline{p_+})\in\mathcal{Q}\} + \frac{s}{C^2} + \frac{s\sqrt{2}}{C}
\end{aligned}$$

Ostatecznie przy $C\to\infty$ i $n\to\infty$ z założeń wynika, że:

$$\lim_{n\to\infty}\Pr\{\mathcal{H}(\leqslant s,n,\underline{M})\in\mathcal{Q}\} = a$$

$\square$

Lemat 2.9. Niech $\mathcal{Q} \subseteq 2^{\Gamma}$ będzie wypukłą własnością podzbiorów Γ oraz niech M_1, M, M_2 będą liczbami całkowitymi spełniającymi warunek: $M_1 \leqslant M \leqslant M_2$. Wtedy $\Pr\{\Gamma_M \in \mathcal{Q}\} \geqslant \Pr\{\Gamma_{M_1} \in \mathcal{Q}\} + \Pr\{\Gamma_{M_2} \in \mathcal{Q}\} - 1$.

Dowód. Każda własność wypukła $\mathcal{Q}$ jest częścią wspólną własności rosnącej

$$\mathcal{Q}' = \{B \in \Gamma : \exists_{A \in \mathcal{Q}}\ A \subseteq B\}$$

i malejącej

$$\mathcal{Q}'' = \{B \in \Gamma : \exists_{C \in \mathcal{Q}}\ B \subseteq C\},$$

gdyż: $B \in \mathcal{Q}' \cap \mathcal{Q}'' \Rightarrow \exists_{\substack{A \in \mathcal{Q} \\ C \in \mathcal{Q}}} A \subseteq B \subseteq C \Rightarrow B \in \mathcal{Q}$ oraz $\mathcal{Q} \subseteq \mathcal{Q}' \cap \mathcal{Q}''$. Czyli dla $M_1 \leqslant M \leqslant M_2$, korzystając z monotoniczności, $\mathcal{Q}'$ i $\mathcal{Q}''$ otrzymujemy:

$$\begin{aligned}\Pr\{\Gamma_M \in \mathcal{Q}\} &= \Pr\{\Gamma_M \in \mathcal{Q}' \cap \mathcal{Q}''\} = \\ &= \Pr\{\Gamma_M \in \mathcal{Q}'\} + \Pr\{\Gamma_M \in \mathcal{Q}''\} - \Pr\{\Gamma_M \in \mathcal{Q}' \cup \mathcal{Q}''\} \geqslant \\ &\geqslant \Pr\{\Gamma_{M_1} \in \mathcal{Q}'\} + \Pr\{\Gamma_{M_2} \in \mathcal{Q}''\} - 1 \geqslant \\ &\geqslant \Pr\{\Gamma_{M_1} \in \mathcal{Q}\} + \Pr\{\Gamma_{M_2} \in \mathcal{Q}\} - 1\end{aligned}$$

□

Twierdzenie 2.10. *Niech* $\mathcal{Q} \subseteq 2^{\Gamma}$ *będzie własnością wypukłą podzbiorów zbioru* Γ *oraz niech* $M = M(n)$ *i* $N = N(n)$ *będą ciągami liczb całkowitych takimi, że* $0 \leqslant M(n) \leqslant N(n)$. *Wtedy jeżeli* $\Pr\{\Gamma_{\frac{M}{N}} \in \mathcal{Q}\} \to 1$, *dla* $n \to \infty$, *to* $\Pr\{\Gamma_M \in \mathcal{Q}\} \to 1$.

Dowód. Ciągom prawdopodobieństw $\Pr\{\Gamma_{\frac{M}{N}} \in \mathcal{Q}\}$ i $\Pr\{\Gamma_M \in \mathcal{Q}\}$ odpowiadają ciągi M, $N - M$ i $Var|\Gamma_{\frac{M}{N}}|$. $|\Gamma_{\frac{M}{N}}|$ ma rozkład dwumianowy, czyli $Var|\Gamma_{\frac{M}{N}}| = N\frac{M}{N}\left(1 - \frac{M}{N}\right) = \frac{M(N-M)}{N}$. Zauważmy, że dla $M \leqslant N/2$ mamy $M \geqslant \frac{M(N-M)}{N} \geqslant (N-M)/2$ oraz dla $M \geqslant N/2$ zachodzi nierówność $N - M \geqslant \frac{M(N-M)}{N} \geqslant M/2$. Czyli dla dowolnego ograniczonego podciągu ciągu $\frac{M(N-M)}{N}$ odpowiadający mu podciąg ciągu M lub ciągu $N - M$ jest ograniczony. Natomiast dla podciągu rozbieżnego do nieskończoności rozbieżne są także M i $N - M$. Wystarczy zatem rozpatrzyć trzy przypadki podciągów: gdy M jest ograniczony, gdy $N - M$ jest ograniczony oraz gdy $\frac{M(N-M)}{N}$ jest rozbieżny do nieskończoności. Będą one odpowiadać podciągom ciągu $\Pr\{\Gamma_M \in \mathcal{Q}\}$.

Rozważmy najpierw przypadek, podciągu takiego, że M lub $N - M$ jest ograniczone. Wtedy

$$\begin{aligned}\Pr\{|\Gamma_{\frac{M}{N}}| = M\} &= \binom{N}{M}\left(\frac{M}{N}\right)^M\left(\frac{N-M}{N}\right)^{N-M} \geqslant \\ &\geqslant \left(\frac{N}{M}\right)^M\left(\frac{M}{N}\right)^M\left(\left(1 + \frac{M}{N-M}\right)^{\frac{N-M}{M}}\right)^{-M} \geqslant e^{-M}\end{aligned}$$

oraz

$$\begin{aligned}\Pr\{|\Gamma_{\frac{M}{N}}| = M\} &= \tbinom{N}{N-M} \left(\tfrac{N-M}{N}\right)^{N-M} \left(\tfrac{M}{N}\right)^{M} \geqslant \\ &\geqslant \left(\left(1 + \tfrac{N-M}{M}\right)^{\frac{M}{N-M}}\right)^{-(N-M)} \geqslant e^{-(N-M)}.\end{aligned}$$

Czyli, ponieważ M lub $N-M$ jest ograniczony, więc dla pewnej stałej $0 < C \leqslant 1$ zachodzi nierówność: $\Pr\{|\Gamma_{\frac{M}{N}}| = M\} \geqslant C$. Stąd

$$\begin{aligned}1 \geqslant \Pr\{\Gamma_M \in \mathcal{Q}\} &= 1 - \frac{1}{C}\, C \Pr\{\Gamma_M \notin \mathcal{Q}\} \geqslant \\ &\geqslant 1 - \frac{1}{C}\, \Pr\{\Gamma_M \notin \mathcal{Q}\} \Pr\{|\Gamma_{\frac{M}{N}}| = M\} \geqslant 1 - \frac{1}{C} \Pr\{\Gamma_{\frac{M}{N}} \notin \mathcal{Q}\}.\end{aligned}$$

Jeśli $\lim_{n\to\infty} \Pr\{\Gamma_{\frac{M}{N}} \in \mathcal{Q}\} = 1$, to $\lim_{n\to\infty} \Pr\{\Gamma_{\frac{M}{N}} \notin \mathcal{Q}\} = 0$, czyli z powyższej nierówności wynika, że $\lim_{n\to\infty} \Pr\{\Gamma_M \in \mathcal{Q}\} = 1$

Załóżmy teraz, że $\frac{M(N-M)}{N} \xrightarrow[n\to\infty]{} \infty$. Niech M_1 maksymalizuje $\Pr\{\Gamma_{M'} \in \mathcal{Q}\}$ wśród $M' \leqslant M$ oraz M_2 maksymalizuje $\Pr\{\Gamma_{M'} \in \mathcal{Q}\}$ wśród $M' \geqslant M$.

$$\begin{aligned}\Pr\{\Gamma_{\frac{M}{N}} \in \mathcal{Q}\} &\leqslant \sum_{M' \leqslant M} \Pr\{\Gamma_{M'} \in \mathcal{Q}\}\ \Pr\{|\Gamma_{\frac{M}{N}}| = M'\} + \Pr\{|\Gamma_{\frac{M}{N}}| > M\} \leqslant \\ &\leqslant \Pr\{\Gamma_{M_1} \in \mathcal{Q}\}\ \Pr\{|\Gamma_{\frac{M}{N}}| \leqslant M\} + \Pr\{|\Gamma_{\frac{M}{N}}| > M\} \\ \Pr\{\Gamma_{\frac{M}{N}} \in \mathcal{Q}\} &\leqslant \sum_{M' \geqslant M} \Pr\{\Gamma_{M'} \in \mathcal{Q}\}\ \Pr\{|\Gamma_{\frac{M}{N}}| = M'\} + \Pr\{|\Gamma_{\frac{M}{N}} < M\} \leqslant \\ &\leqslant \Pr\{\Gamma_{M_2} \in \mathcal{Q}\}\ \Pr\{|\Gamma_{\frac{M}{N}}| \geqslant M\} + \Pr\{|\Gamma_{\frac{M}{N}}| < M\}\end{aligned} \tag{2.8}$$

Z centralnego twierdzenia granicznego

$$\lim_{n\to\infty} \Pr\left\{\frac{|\Gamma_{\frac{M}{N}}| - M}{\sqrt{N(N-M)}} \leqslant 0\right\} = \int_{-\infty}^{0} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt = \frac{1}{2}, \text{czyli}$$

$$\lim_{n\to\infty} \Pr\{|\Gamma_{\frac{M}{N}}| \leqslant M\} = \lim_{n\to\infty} \Pr\{|\Gamma_{\frac{M}{N}}| < M\} = \frac{1}{2} \text{ oraz}$$

$$\lim_{n\to\infty} \Pr\{|\Gamma_{\frac{M}{N}}| \geqslant M\} = \lim_{n\to\infty} \Pr\{|\Gamma_{\frac{M}{N}}| > M\} = \frac{1}{2}.$$

Zatem z założeń i (2.8)

$$1 = \lim_{n\to\infty} \Pr\{\Gamma_{\frac{M}{N}} \in \mathcal{Q}\} \leqslant \frac{1}{2} \liminf_{n\to\infty} \Pr\{\Gamma_{M_1} \in \mathcal{Q}\} + \frac{1}{2} \leqslant 1$$

$$1 = \lim_{n\to\infty} \Pr\{\Gamma_{\frac{M}{N}} \in \mathcal{Q}\} \leqslant \frac{1}{2} \liminf_{n\to\infty} \Pr\{\Gamma_{M_2} \in \mathcal{Q}\} + \frac{1}{2} \leqslant 1,$$

Czyli

$$\lim_{n\to\infty} \Pr\{\Gamma_{M_1} \in \mathcal{Q}\} = 1 \text{ i } \lim_{n\to\infty} \Pr\{\Gamma_{M_2} \in \mathcal{Q}\} = 1$$

$M_1 \leqslant M \leqslant M_2$, czyli z lematu 2.9 dla własności wypukłej $\mathcal{Q}$ wynika, że

$$1 = \lim_{n\to\infty} \Pr\{\Gamma_{M_1} \in \mathcal{Q}\} + \lim_{n\to\infty} \Pr\{\Gamma_{M_2} \in \mathcal{Q}\} - 1 \leqslant \lim_{n\to\infty} \Pr\{\Gamma_M \in \mathcal{Q}\} \leqslant 1.$$

Podsumowując, każdy podciąg ciągu $\Pr\{\Gamma_M \in \mathcal{Q}\}$ jest zbieżny do jedynki, stąd $\lim_{n\to\infty} \Pr\{\Gamma_M \in \mathcal{Q}\} = 1$. □

O własności hipergrafu będziemy mówili, że zachodzi asymptotycznie prawie na pewno, gdy przy n dążącym do nieskończoności prawdopodobieństwo, że hipergraf ma tą własność dąży do 1.

Wniosek 2.11. Niech $\mathcal{Q}$ będzie własnością monotoniczną podzbiorów zbioru Γ, $0 \leqslant M(n) \leqslant N(n)$, wtedy:

(i) jeśli $\Pr\{\Gamma_{\frac{M}{N}} \in \mathcal{Q}\} \to 1$, dla $n \to \infty$, to $\Pr\{\Gamma_M \in \mathcal{Q}\} \to 1$,
(ii) jeśli $\Pr\{\Gamma_{\frac{M}{N}} \in \mathcal{Q}\} \to 0$, dla $n \to \infty$, to $\Pr\{\Gamma_M \in \mathcal{Q}\} \to 0$.

Dowód. Ponieważ każda własność monotoniczna jest wypukła, więc pierwsza część wynika bezpośrednio z twierdzenia 2.10. Natomiast dopełnienie własności monotonicznej jest też własnością monotoniczną, czyli

$$\begin{aligned}\lim_{n\to\infty} \Pr\{\Gamma_{\frac{M}{N}} \in \mathcal{Q}\} = 0 \Rightarrow \lim_{n\to\infty} \Pr\{\Gamma_{\frac{M}{N}} \notin \mathcal{Q}\} = 1 \Rightarrow \\ \Rightarrow \lim_{n\to\infty} \Pr\{\Gamma_M \notin \mathcal{Q}\} = 1 \Rightarrow \lim_{n\to\infty} \Pr\{\Gamma_M \in \mathcal{Q}\} = 0\end{aligned}$$

□

3. Spójność hipergrafów

Graf jest spójny, gdy dowolne dwa wierzchołki są połączone ścieżką, czyli gdy istnieje ciąg wierzchołków połączonych krawędziami. Chcemy sklasyfikować obiekty ze zbioru $\mathcal{V}$ ze względu na podobieństwo używając do tego celu grafy przecięć $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$, w których elementy podobne są połączone krawędziami. Pytanie o spójność wydaje się zatem naturalnym pytaniem dotyczącym tych grafów. Zarówno graf $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ jak i $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ jest grafem przecięć podzbiorów pewnego zbioru. Dlatego prawdopodobieństwa spójności grafów $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ są ściśle związane z prawdopodobieństwem spójności pewnych hipergrafów.

Spójność jest własnością monotoniczną. Z własnościami monotonicznymi grafów i hipergrafów losowych związane jest pojęcie funkcji progowych. Jednym z największych odkryć teorii grafów losowych było spostrzeżenie, że dla własności monotonicznych istnieją funkcje progowe $\hat{p}(n)$. W przypadku własności rosnących dla p rzędowo mniejszych od funkcji progowej $\hat{p}(n)$ danej własności, graf $G(n, p)$ nie ma własności asymptotycznie prawie na pewno, a dla p rzędowo większych od $\hat{p}(n)$ asymptotycznie prawie na pewno ma tą własność. Analogicznie dla własności malejących. Funkcja progowa spójności grafu pokrywa się z funkcją progową istnienia wierzchołków izolowanych.

Oczywiście spójny hipergraf nie posiada wierzchołków izolowanych. Okazuje się, że w przypadku klasycznych hipergrafów losowych można powiedzieć więcej. Zniknięcie ostatniego wierzchołka izolowanego jest asymptotycznie równoważne spójności hipergrafu prawie na pewno. Dlatego istotne jest oszacowanie liczby wierzchołków izolowanych za pomocą metody momentów. Podamy klasyczny wynik (tw 4.4.5 w [3]):

Lemat 3.1. Niech Z będzie zmienną losową z dystrybuantą wyznaczoną jednoznaczne przez swoje momenty. Jeśli $X_1, X_2, \ldots$ są zmiennymi losowymi o skończonych momentach takich, że $\mathbb{E}X_n^k \to \mathbb{E}Z^k$ przy $n \to \infty$ dla dowolnego k naturalnego, wtedy X_n jest słabo zbieżna (zbieżna według dystrybuanty) do Z.

Momentami silniowymi zmiennej losowej X o skończonych momentach nazywamy liczby

$$\mathbb{E}(X)_k = \mathbb{E}[X(X-1)\ldots(X-k+1)] \text{ dla } k \geqslant 1 \text{ oraz } \mathbb{E}(X)_0 = 1.$$

Ponieważ $X(X-1)\ldots(X-k+1)$ jest wielomianem stopnia k zmiennej X o wszystkich współczynnikach niezerowych oraz wartość oczekiwana jest funkcją liniową, więc dla dowolnej zmiennej losowej X o skończonych momentach, momenty silniowe $\mathbb{E}(X)_k = \mathbb{E}[X(X-1)\dots(X-k+1)]$ wyznaczają jednoznacznie momenty zwykłe i na odwrót. Czyli $\mathbb{E}X_n^k \to \mathbb{E}Z^k$ dla dowolnego k wtedy i tylko wtedy gdy $\mathbb{E}(X_n)_k \to \mathbb{E}(Z)_k$ dla dowolnego k. Czyli powyższy lemat jest równoważny:

Lemat 3.2. Niech Z będzie zmienną losową z dystrybuantą wyznaczoną jednoznaczne przez swoje momenty. Jeśli $X_1, X_2, \dots$ są zmiennymi losowymi o skończonych momentach takich, że $\mathbb{E}(X_n)_k \to \mathbb{E}(Z)_k$ przy $n \to \infty$ dla dowolnego k naturalnego, wtedy X_n jest słabo zbieżna (zbieżna według rozkładu) do Z.

Dla zmiennych $X = \sum_{i=1}^n \xi_i$, będących sumą n zmiennych zero-jedynkowych ξ_i, $1 \leqslant i \leqslant n$, stosując indukcję można łatwo dowieść, że

$$(\sum_{i=1}^{n} \xi_i)_k = \sum_{i_1,\dots,i_k}^{*} \xi_{i_1}\xi_{i_2}\dots\xi_{i_k},$$

gdzie $\sum_{i_1,\dots,i_k}^{*}$ oznacza sumę po wszystkich ciągach indeksów $i_1, \dots, i_k$. Zatem dla takich zmiennych losowych mamy:

$$\mathbb{E}(\sum_{i=1}^{n} \xi_i)_k = \mathbb{E}\left(\sum_{i_1,\dots,i_k}^{*} (\xi_{i_1}\xi_{i_2}\dots\xi_{i_k})\right) = \sum_{i_1,\dots,i_k}^{*} \Pr\{\xi_{i_1} = 1, \dots, \xi_{i_k} = 1\}. \tag{3.1}$$

Najczęściej będziemy wykorzystywali następujący wniosek z lematu 3.2

Wniosek 3.3. Niech X_n będzie sumą n zero-jedynkowych zmiennych losowych ξ_i (tzn. $X_n = \sum_{i=1}^n \xi_i$). Jeśli dla pewnego $\lambda > 0$

$$\mathbb{E}(\sum_{i=1}^{n} \xi_i)_k = \sum_{i_1,\dots,i_k}^{*} \Pr\{\xi_{i_1} = 1, \dots, \xi_{i_k} = 1\} \to (\lambda)^k$$

przy $n \to \infty$, to X_n jest słabo zbieżna (zbieżna według dystrybuanty) do zmiennej losowej o rozkładzie Poissona z parametrem λ.

Zauważmy, że dowolna krawędź należąca do ścieżki musi posiadać co najmniej dwa różne wierzchołki. Krawędzie jednoelementowe nie mogą zatem należeć do żadnej ścieżki, czyli dodanie lub usunięcie krawędzi jednoelementowej z hipergrafu nie przyczyni się do powstania ani zniknięcia żadnej ścieżki. Wynika stąd, że krawędzie o jednym wierzchołku nie wpływają na spójność hipergrafu. W dalszych rozważaniach będziemy więc zakładać, że w hipergrafach losowych krawędzie mają co najmniej dwa wierzchołki.

W [11] Wojciech Kordecki określił, z jakim prawdopodobieństwem losowy hipergraf jednorodny $\mathcal{H}(s, n, p)$ jest spójny przy n dążącym do nieskończoności.

Twierdzenie 3.4 (spójność hipergrafu s-jednorodnego). *Niech μ oznacza liczbę izolowanych wierzchołków w s-jednorodnym hipergrafie losowym. Jeżeli*

$$p\,(n) = \frac{\ln n + c + o(1)}{\binom{n-1}{s-1}} = \frac{(s-1)!(\ln n + c + o(1))}{n^{s-1}}, \tag{3.2}$$

to s-jednorodny hipergraf ma $\mu + 1$ składowych spójności z prawdopodobieństwem dążącym do 1 dla $n \to \infty$, a zmienna losowa μ jest zbieżna według dystrybuanty do zmiennej losowej o rozkładzie Poissona z parametrem e^{-c}.

Dowód. Po pierwsze zauważmy , że dla $p = p(n)$ dążącego do zera mamy $1 - p = e^{(-p - O(p^2))}$. Z drugiej strony:

$$O(p^2) = O\left(\ln^2 n / \binom{n-1}{s-1}\right) / \binom{n-1}{s-1} = o(1)/\binom{n-1}{s-1}.$$

Dlatego w tym przypadku można przyjąć, że $1 - p = e^{-p}$.Udowodnimy, że rozkładem granicznym (zbieżność według dystrybuanty) zmiennej losowej μ jest rozkład Poissona z parametrem e^{-c}. W tym celu zastosujemy metodę momentów (wniosek 3.3). Niech $K = K(n)$ oznacza zbiór wierzchołków izolowanych. Wtedy $|K(n)| = \mu(n) = \mu$. Oznaczmy przez ξ_i zero-jedynkową zmienną losową:

$$\xi_i = \begin{cases} 1, & \text{gdy } v_i \in K \\ 0, & \text{gdy } v_i \notin K \end{cases} \tag{3.3}$$

$$\text{wtedy } \mu = \sum_{i=1}^{n} \xi_i.$$

Ponieważ

$$\begin{aligned} \binom{n}{s} - \binom{n-k}{s} &= \binom{n-1}{s-1} + \binom{n-2}{s-1} + \ldots + \binom{n-k}{s-1} = \\ &= \frac{(n-1)^{s-1} + (n-2)^{s-1} + \ldots + (n-k)^{s-1}}{(s-1)!} + o\left(\frac{n^{s-1}}{\ln n}\right) = \\ &= \frac{kn^{s-1}}{(s-1)!} + o\left(\frac{n^{s-1}}{\ln n}\right), \end{aligned}$$

więc

$$\begin{aligned} & p\left(\binom{n}{s} - \binom{n-k}{s}\right) = \\ &= \frac{(s-1)!(\ln n + c + o(1))}{n^{s-1}} \left(\frac{kn^{s-1}}{(s-1)!} + o\left(\frac{n^{s-1}}{\ln n}\right)\right) = (\ln n + c)k + o(1). \end{aligned} \tag{3.4}$$

Czyli

$$\mathbb{E}(\sum_{i=1}^{n}\xi_i)_k = \sum_{i_1,\ldots,i_k}^{*} \mathbb{E}(\xi_{i_1}\xi_{i_2}\ldots\xi_{i_k}) = \sum_{i_1,\ldots,i_k}^{*} \Pr(\xi_{i_1}=1,\ldots,\xi_{i_k}=1) =$$
$$= k!\binom{n}{k}(1-p)^{\binom{n}{s}-\binom{n-k}{s}} = (n)_k\, e^{-p\left(\binom{n}{s}-\binom{n-k}{s}\right)} = (n)_k\, e^{-(\ln n+c)k+o(1)} =$$
$$(n)_k\frac{e^{-kc}}{n^k}\,(1+o(1)) \to (e^{-c})^k$$

Czyli z wniosku 3.3 wynika, że

$$\Pr\{\mu=k\} - e^{-e^{-c}}\frac{e^{-ck}}{k!} \longrightarrow 0$$

dla każdego k przy $n\to\infty$.

Teraz wystarczy udowodnić, że hipergraf losowy z prawdopodobieństwem dążącym do 1 ma jedną dużą składową spójności a poza nią tylko wierzchołki izolowane. Dla ułatwienia wprowadzimy dwie definicje. Podzbiór W zbioru wierzchołków hipergrafu V będziemy nazywać zamkniętym, jeżeli w hipergrafie nie będzie istnieć żadna krawędź pomiędzy $W\backslash V$ a W. Natomiast w przypadku, gdy $W\subset K$ lub $V\backslash W\subset K$, podzbiór W będziemy nazywać prostym. Zauważmy, że jeśli W jest zamknięty, to oczywiście jego dopełnienie jest także zbiorem zamkniętym i na odwrót. Ponieważ warunek na to, że zbiór jest prosty jest symetryczny, więc mamy także, że W jest prosty wtedy i tylko wtedy, gdy $V\setminus W$ jest prosty.

Niech A będzie zdarzeniem, że w hipergrafie istnieje zbiór zamknięty, który nie jest prosty. Przez A_k oznaczmy zdarzenie, że istnieje taki k-elementowy zbiór wierzchołków. Zbiór W o mocy $k\leqslant n/2$ jest zamknięty i nie prosty wtedy i tylko wtedy gdy jego dopełnienie o mocy $n-k\geqslant n/2$ jest zbiorem zamkniętym i nie prostym. Stąd mamy, że wystarczy rozpatrzyć zdarzenia A_k, dla których $k\leqslant n/2$, a istnienie zbioru zamkniętego, który nie jest prosty dla pewnego $k\leqslant n/2$ jest równoważne temu, że wierzchołki hipergrafu można podzielić na co najmniej dwie nietrywialne składowe spójności. Dla $W\subset V$,$|W|=k$ mamy:

$$\Pr\{W \text{ jest zamknięty}\} = (1-p)^{\left[\binom{n}{s}-\binom{k}{s}-\binom{n-k}{s}\right]}$$

$$\Pr\{W \text{ jest prosty}\} = P\{W\subset K \text{ lub } V\backslash W\subset K\} =$$
$$= (1-p)^{\binom{n}{s}-\binom{n-k}{s}} + (1-p)^{\binom{n}{s}-\binom{k}{s}} - (1-p)^{\binom{n}{s}}$$

Zauważmy, że każdy zbiór prosty jest zamknięty, czyli

$$\begin{aligned}\Pr\{A_k\} &= \binom{n}{k}\left(\Pr\{W \text{ jest zamknięty}\} - \Pr\{W \text{ jest prosty}\}\right) = \\ &= \binom{n}{k}(1-p)^{\binom{n}{s}}\left((1-p)^{-\binom{k}{s}}-1\right)\left((1-p)^{-\binom{n-k}{s}}-1\right) = \\ &= \binom{n}{k}e^{-p\binom{n}{s}}\left(e^{p\binom{k}{s}}-1\right)\left(e^{p\binom{n-k}{s}}-1\right)\end{aligned}$$

Wykorzystamy nierówności:

$$\frac{\binom{n}{s}-\binom{n-k}{s}-\binom{k}{s}}{\binom{n-1}{s-1}} \geqslant \frac{k}{s} \quad , \quad \frac{\binom{n}{s}-\binom{n-k}{s}-\binom{k}{s}}{\binom{n-1}{s-1}} = k - O\left(\frac{k^2}{n}\right),$$

$$\frac{\binom{n}{s}-\binom{n-k}{s}}{\binom{n-1}{s-1}} = k - O\left(\frac{k^2}{n}\right).$$

Ustalmy pewne

$$\epsilon \in (0, 1/s) \text{ i } \delta \in (1/s, 1).$$

Mamy wtedy:

$$\begin{aligned}&\Pr\{\text{Istnieje podzbiór zamknięty, który nie jest prosty}\} \leqslant \\ &\sum_{k=s}^{n/2}\Pr\{A_k\} = \sum_{k=s}^{n^{1-\delta}}\Pr\{A_k\} + \sum_{k=n^{1-\delta}}^{n^{1-\epsilon}}\Pr\{A_k\} + \sum_{k=n^{1-\epsilon}}^{n/2}\Pr\{A_k\} \leqslant \\ &\leqslant \sum_{k=s}^{n^{1-\delta}}\binom{n}{k}e^{-\frac{\binom{n}{s}-\binom{n-k}{s}}{\binom{n-1}{s-1}}(\ln n+c)}\left(e^{\frac{\binom{k}{s}}{\binom{n-1}{s-1}}(\ln n+c)}-1\right) + \\ &+ \sum_{k=n^{1-\delta}}^{n^{1-\epsilon}}\binom{n}{k}e^{-\frac{\binom{n}{s}-\binom{n-k}{s}-\binom{k}{s}}{\binom{n-1}{s-1}}(\ln n+c)} + \\ &+ \sum_{k=n^{1-\epsilon}}^{n/2}\binom{n}{k}e^{-\frac{\binom{n}{s}-\binom{n-k}{s}-\binom{k}{s}}{\binom{n-1}{s-1}}(\ln n+c)} \leqslant\end{aligned} \tag{3.5}$$

$$
\begin{aligned}
&\leqslant \sum_{k=s}^{n^{1-\delta}} \frac{e^k n^k}{k^k} \frac{e^{-ck+O(k^2/n)}}{n^{k-O(k^2/n)}} \left(e^{O\left(\frac{n^{s(1-\delta)}}{n^{s-1}}\right)(\ln n + c)} - 1 \right) + \\
&+ \sum_{k=n^{1-\delta}}^{n^{1-\epsilon}} \frac{e^k n^k}{k^k} \frac{e^{-ck+O(k^2/n)}}{n^{k-O(k^2/n)}} + \sum_{k=n^{1-\epsilon}}^{n/2} \frac{e^k n^k}{k^k} \frac{e^{-kc/s}}{n^{k/s}} \leqslant \\
&\leqslant \sum_{k=s}^{n^{1-\delta}} \left(\frac{e^{O(1)} n^{O(k/n)}}{k} \right)^k \left(e^{O(1/n^{s\delta-1}) \ln n} e^{o(1)} - 1 \right) + \\
&+ \sum_{k=n^{1-\delta}}^{n^{1-\epsilon}} \left(\frac{e^{O(1)} n^{O(k/n)}}{k} \right)^k + \sum_{k=n^{1-\epsilon}}^{n/2} \left(\frac{e^{O(1)} n^{1-1/s}}{n^{1-\epsilon}} \right)^k \leqslant \\
&\leqslant (e^{o(1)} - 1) \sum_{k=s}^{n^{1-\delta}} \left(\frac{e^{O(1)} n^{1/n^{\delta}}}{k} \right)^k + \sum_{k=n^{1-\delta}}^{n^{1-\epsilon}} \left(\frac{e^{O(1)} n^{1/n^{\epsilon}}}{n^{1-\delta}} \right)^{n^{1-\epsilon}} + \\
&\sum_{k=n^{1-\epsilon}}^{n/2} \left(\frac{e^{O(1)}}{n^{1/s-\epsilon}} \right)^{n/2} = o(1)
\end{aligned}
$$

Czyli z prawdopodobieństwem dążącym do 1 hipergraf nie ma zbioru zamkniętego, który nie jest prosty a tym samym ma $\mu + 1$ składowych. $\square$

Wniosek 3.5. Dla $p(n) = \frac{\ln n + c + o(1)}{\binom{n-1}{s-1}}$

$$\lim_{n\to\infty} \Pr\{\mathcal{H}(s,n,p) \text{jest spójny}\} = \lim_{n\to\infty} \Pr\{\mu = 0\} = e^{-e^{-c}} \tag{3.6}$$

Wniosek 3.6. Jeżeli

$$\lim_{n\to\infty} [n^{s-1}p - (s-1)! \ln n] = \infty,$$

to

$$\lim_{n\to\infty} \Pr\{\mathcal{H}(s,n,p) \text{ jest spójny}\} = 1. \tag{3.7}$$

Natomiast, jeśli

$$\lim_{n\to\infty} [n^{s-1}p - (s-1)! \ln n] = -\infty,$$

to

$$\lim_{n\to\infty} \Pr\{\mathcal{H}(s,n,p) \text{ jest spójny}\} = 0 \tag{3.8}$$

Ponieważ spójność jest własnością monotoniczną, więc korzystając z twierdzeń z rozdziału 2 możemy podać wniosek dotyczący hipergrafu $\mathcal{H}(s,n,M)$.

Wniosek 3.7. Dla

$$M(n) = \frac{n(\ln n + c + o(1))}{s}$$

prawdopodobieństwo spójności $\mathcal{H}(s,n,M)$ jest równe $e^{-e^{-c}}$ przy $n \to \infty$.

Dowód. Dla $M = M(n) = \frac{n(\ln n + c + o(1))}{s}$ oraz $N = \binom{n}{s}$ mamy:

$$\begin{aligned}
\frac{M(N-M)}{N^3} &= \frac{\binom{n}{s}\frac{(\ln n+c+o(1))}{\binom{n-1}{s-1}}\left(\binom{n}{s} - \binom{n}{s}\frac{(\ln n+c+o(1))}{\binom{n-1}{s-1}}\right)}{\binom{n}{s}^3} = \\
&= \frac{1}{\binom{n-1}{s-1}^2}\frac{(\ln n + c + o(1))\left(\binom{n-1}{s-1} - (\ln n + c + o(1))\right)}{\binom{n}{s}} \leqslant \\
&\leqslant \frac{1}{\binom{n-1}{s-1}^2}\left(\frac{s}{n}(\ln n + c + o(1))\right) = \frac{o(1)}{\binom{n-1}{s-1}^2}.
\end{aligned} \tag{3.9}$$

Ponieważ $p = \frac{M}{N} + O(\sqrt{\frac{M(N-M)}{N^3}}) = \frac{\ln n + c + o(1)}{\binom{n-1}{s-1}}$, czyli z wniosku 3.5 dla dowolnego takiego p mamy:

$$\lim_{n\to\infty} \Pr\{\mathcal{H}(s,n,p) \text{ jest spójny}\} = \Pr\{\mu = 0\} = e^{-e^{-c}}.$$

Z twierdzenia 2.7 prawdopodobieństwo spójności $\mathcal{H}(s,n,M)$ dąży do $e^{-e^{-c}}$ przy $n \to \infty$. $\square$

Fakt 3.8. Z M zdefiniowanym jak we wniosku 3.7, jeśli $M' = M + o(n)$, wtedy przy $n \to \infty$ prawdopodobieństwo spójności $\mathcal{H}(s,n,M)$ jest równe prawdopodobieństwu spójności $\mathcal{H}(s,n,M')$.

W [12] Kordecki uogólnił wynik dotyczący hipergrafów s-jednorodnych i podał twierdzenie dotyczące dowolnych hipergrafów. W tym przypadku dowód przebiega analogicznie. Ponieważ zakładamy, że prawdopodobieństwa powstania krawędzi mogą być zróżnicowane, więc dla ujednolicenia zapisu w [12] Kordecki wprowadził dodatkowe oznaczenia.

Dla określonego podzbioru $\{v_1, \ldots, v_k\} \subseteq V(\mathcal{H})$ zbioru wierzchołków hipergrafu losowego $\mathcal{H}$, przez $p_{v_1,\ldots,v_k}$ oznaczmy prawdopodobieństwo pojawienia się krawędzi $\{v_1, \ldots, v_k\}$, przez $q_{v_1,\ldots,v_k} = 1 - p_{v_1,\ldots,v_k}$-prawdopodobieństwo, że ta krawędź się nie pojawi w hipergrafie losowym. Jak w dowodzie dla hipergrafu s--jednostajnego K oznacza zbiór wierzchołków izolowanych hipergrafu, a zmienna losowa μ określa moc tego zbioru. Dla określonego zbioru $A \subseteq V(\mathcal{H})$ i danego wierzchołka $v \in A$

$$q_v(A) = \prod_{r=1}^{s-1}\prod_{k=1}^{r}\prod_{\substack{\{v,v_1,\ldots,v_k\}\subset A\\ \{v_{k+1},\ldots,v_r\}\subset V\setminus A}} q_{v,v_1,\ldots,v_r}$$

jest prawdopodobieństwem, że dowolna krawędź zawierająca v nie zawiera żadnych innych wierzchołków z A. Dodatkowo:

$$Q_{vk} = \min_{|A|=k} q_v(A)$$

Prawdopodobieństwo, że żadna krawędź hipergrafu nie zawiera wierzchołka $v \in V$, czyli prawdopodobieństwo, że wierzchołek v jest izolowany oznaczamy przez

$$Q_v = \Pr(v \in K) = \prod_{r=1}^{s-1} \prod_{\{v,v_1,\ldots,v_r\}\subset V} q_{v,v_1,\ldots,v_r}.$$

Niech

$$\lambda = \sum_{v\in V} Q_v = \sum_{v\in V} \Pr\{v \in K\},$$

czyli λ jest wartością oczekiwaną liczby wierzchołków izolowanych.

Twierdzenie 3.9 (Spójność hipergrafu w przypadku ogólnym). *Jeśli*

$$\max_{v\in V} Q_v = o(1), \tag{3.10}$$

$$\lambda = O(1), \tag{3.11}$$

$$\forall_i \sum_{1\leqslant k\leqslant n/2} \frac{1}{k!}\left(\sum_{v\in V} \frac{Q_v}{Q_{vi}}\right)^k - \left(e^{-\lambda} - 1\right) = o(1), \tag{3.12}$$

wtedy hipergraf ma $\mu + 1$ składowych spójności z prawdopodobieństwem dążącym do 1 dla $n \to \infty$ oraz

$$\Pr(\mu = k) - e^{-\lambda}\frac{\lambda^k}{k!} \longrightarrow 0 \tag{3.13}$$

dla każdego k przy $n \to \infty$.

Dowód. Z założeń wynikają następujące oszacowania. Ponieważ

$$\lambda^k = \left(\sum_{v\in V} Q_v\right)^k = \sum_{r_1+\ldots+r_n=k} \frac{k!}{r_1!\ldots r_n!} Q_{v_1}^{r_1}\ldots Q_{v_n}^{r_n}$$

stąd

$$\sum_{i_1,\ldots,i_k}^{*} Q_{v_{i_1}}\ldots Q_{v_{i_k}} - \lambda^k = o(1) \tag{3.14}$$

Natomiast z (3.12) mamy, że

$$\sum_{1\leqslant i_1<\ldots<i_k\leqslant n} \frac{Q_{v_{i_1}}}{Q_{v_{i_1}k}}\cdots\frac{Q_{v_{i_k}}}{Q_{v_{i_k}k}} - Q_{v_{i_1}}\ldots Q_{v_{i_k}} = \frac{o(1)}{k!}, \tag{3.15}$$

gdzie o(1) dąży jednostajnie do 0 względem $k \leqslant n/2$. Czyli z (3.14) i (3.15)

$$\sum_{i_1,\dots,i_k}^{*} \frac{Q_{v_{i_1}}}{Q_{v_{i_1}k}} \cdots \frac{Q_{v_{i_k}}}{Q_{v_{i_k}k}} - \lambda^k = o(1) \tag{3.16}$$

Niech $W = \{v_{i_1}, \dots, v_{i_k}\} \subseteq V$ będzie k-elementowym podzbiorem zbioru wierzchołków hipergrafu losowego.

$$\Pr\{W \subseteq K\} \geqslant Q_{v_{i_1}} \dots Q_{v_{i_k}}. \tag{3.17}$$

Z drugiej strony dla dowolnego wierzchołka $v \in W$

$$\Pr\{\text{w hipergrafie nie ma krawędzi } e \text{ takich, że } e \cap W = \{v\}\ \} = \frac{Q_v}{q_v(W)} \leqslant \frac{Q_v}{Q_{vk}}.$$

Czyli

$$\Pr\{W \subseteq K\} \leqslant \frac{Q_{v_{i_1}}}{Q_{v_{i_1}k}} \cdots \frac{Q_{v_{i_k}}}{Q_{v_{i_k}k}}. \tag{3.18}$$

Oznaczmy zmienną losową ξ_i jak w (3.3), wtedy $\mu = \sum_{i=1}^{n} \xi_i$ oraz $\mathbb{E}(\xi_{i_1}\xi_{i_2}\dots\xi_{i_k}) = \Pr\{W \subseteq K\}$. Z (3.17) i (3.18)

$$Q_{v_{i_1}} \dots Q_{v_{i_k}} \leqslant \mathbb{E}(\xi_{i_1}\xi_{i_2}\dots\xi_{i_k}) \leqslant \frac{Q_{v_{i_1}}}{Q_{v_{i_1}k}} \cdots \frac{Q_{v_{i_k}}}{Q_{v_{i_k}k}}, \text{ więc}$$

$$\sum_{i_1,\dots,i_k}^{*} Q_{v_{i_1}} \dots Q_{v_{i_k}} \leqslant \sum_{i_1,\dots,i_k}^{*} \mathbb{E}(\xi_{i_1}\xi_{i_2}\dots\xi_{i_k}) \leqslant \sum_{i_1,\dots,i_k}^{*} \frac{Q_{v_{i_1}}}{Q_{v_{i_1}k}} \cdots \frac{Q_{v_{i_k}}}{Q_{v_{i_k}k}}$$

Z (3.14) i (3.16)

$$\sum_{i_1,\dots,i_k}^{*} \mathbb{E}(\xi_{i_1}\xi_{i_2}\dots\xi_{i_k}) \to \lambda^k \text{ przy } n \to \infty,$$

więc z 3.3

$$\Pr\{\mu = k\} \to e^{-\lambda}\frac{\lambda^k}{k!} \text{ przy } n \to \infty.$$

Zdefiniujmy zbiór prosty i zbiór zamknięty jak w dowodzie twierdzenia 3.4. Ponieważ W jest zamknięty, gdy nie ma żadnej krawędzi o niepustym przekroju z W i $V \setminus W$ oraz krawędzie o dokładnie jednym wierzchołku w W pojawiają się niezależnie, więc

$$\begin{aligned}\Pr\{W \text{ jest zamknięty}\} &\leqslant \\ &\leqslant \prod_{v \in W} \Pr\{\text{w hipergrafie nie ma krawędzi } e \text{ takich, że } e \cap W = \{v\}\ \} \leqslant \\ &\leqslant \frac{Q_{v_{i_1}}}{Q_{v_{i_1}k}} \cdots \frac{Q_{v_{i_k}}}{Q_{v_{i_k}k}}.\end{aligned}$$

Zbiór W jest prosty, gdy $W \subseteq K$ lub $V \setminus W \subseteq K$, czyli

$$\Pr\{W \text{ jest prosty}\} \geqslant \Pr\{W \subseteq K\} \geqslant Q_{v_{i_1}} \dots Q_{v_{i_k}}.$$

Ostatecznie, ponieważ każdy zbiór prosty jest zamknięty, więc

$$\begin{aligned} &\Pr\{\text{W jest zamknięty i nie jest prosty}\} = \\ &= \Pr\{\text{W jest zamknięty}\} - \Pr\{\text{W jest prosty}\} \leqslant \\ &\leqslant \frac{Q_{v_{i_1}}}{Q_{v_{i_1}k}} \dots \frac{Q_{v_{i_k}}}{Q_{v_{i_k}k}} - Q_{v_{i_1}} \dots Q_{v_{i_k}} \end{aligned} \tag{3.19}$$

Czyli z (3.19) i (3.15)

$$\begin{aligned} &\Pr\{\text{W hipergrafie istnieje zbiór zamknięty, który nie jest prosty}\} \leqslant \\ &\leqslant \sum_{k=1}^{n/2} \sum_{1 \leqslant i_1 < \ldots < i_k \leqslant n} \left(\frac{Q_{v_{i_1}}}{Q_{v_{i_1}k}} \dots \frac{Q_{v_{i_k}}}{Q_{v_{i_k}k}} - Q_{v_{i_1}} \dots Q_{v_{i_k}} \right) = \frac{o(1)}{k!} = \sum_{k=1}^{n/2} \frac{o(1)}{k!} = o(1) \end{aligned}$$

□

Twierdzenie 3.9 nie podaje konkretnego prawdopodobieństwa wyboru krawędzi, przy którym hipergraf byłby spójny asymptotycznie prawie na pewno. Poza tym sprawdzenie warunku (3.12) może być skomplikowane. Okazuje się jednak, że dowód twierdzenia 3.4 w prosty, bezpośredni sposób można uogólnić i podać analogiczne twierdzenie dla hipergrafu $\mathcal{H}(\leqslant s, n, \underline{p})$, gdzie $\underline{p} = (p_2, p_3, \dots, p_s)$ oraz p_j oznaczają odpowiednio prawdopodobieństwo powstania określonej krawędzi o j wierzchołkach dla $j = 2, \dots, s$.

Twierdzenie 3.10 (Spójność hipergrafu losowego). *Niech μ oznacza liczbę izolowanych wierzchołków w hipergrafie losowym $\mathcal{H}(\leqslant s, n, \underline{p})$, w którym $\underline{p} = (p_2, p_3, \dots, p_s)$ oraz*

$$p_j(n) = a_j(n) \frac{\ln n + c_j(n)}{\binom{n-1}{j-1}} = a_j(n) \frac{(j-1)!(\ln n + c_j(n))}{n^{j-1}}, \tag{3.20}$$

dla $j = 1, 2, \dots, s$, $\sum_{j=2}^{s} a_j(n) = 1$, $\sum_{j=2}^{s} a_j(n)\, c_j(n) = c + o(1)$,$c_j = o(\ln n)$ i $0 \leqslant a_j(n) \leqslant 1$. Wtedy $\mathcal{H}(\leqslant s, n, \underline{p})$ ma $\mu + 1$ składowych spójności z prawdopodobieństwem dążącym do 1 dla $n \to \infty$, a zmienna losowa μ jest zbieżna według dystrybuanty do zmiennej losowej o rozkładzie Poissona z parametrem e^{-c}.

Dowód. Jak poprzednio oszacujemy liczbę wierzchołków izolowanych w hipergrafie $\mathcal{H}(\leqslant s, n, \underline{p})$. Określamy zmienną losową μ jak w (3.3). Jak w (3.4) mamy

$$\begin{aligned} &p_j \left(\binom{n}{j} - \binom{n-k}{j} \right) = \\ &\quad = a_k \frac{(j-1)!(\ln n + c_j))}{n^{j-1}} \left(\frac{kn^{j-1}}{(j-1)!} + o\left(\frac{n^{j-1}}{\ln n} \right) \right) = a_k(\ln n + c_j)k + o(1), \end{aligned}$$

zatem

$$\mathbb{E}(\sum_{i=1}^{n} \xi_i)_k = \sum\nolimits^*_{i_1,\ldots,i_k} \mathbb{E}(\xi_{i_1}\xi_{i_2}\ldots\xi_{i_k}) = \sum\nolimits^*_{i_1,\ldots,i_k} \Pr\{\xi_{i_1}=1,\ldots,\xi_{i_k}=1\} =$$

$$= k!\binom{n}{k}\prod_{j=2}^{s}(1-p_j)^{\binom{n}{j}-\binom{n-k}{j}} = (n)_k\, e^{-\sum_{j=2}^{s} p_j\left(\binom{n}{j}-\binom{n-k}{j}\right)} =$$

$$(n)_k\, e^{-\sum_{j=2}^{s}(a_j(\ln n+c_j)k+o(1))} = (n)_k\frac{e^{-kc}}{n^k}(1+o(1)) \to (e^{-c})^k.$$

Z wniosku 3.3 wynika zatem, że podobnie jak w przypadku hipergrafu s -jednorodnego rozkładem granicznym (zbieżność według dystrybuanty) zmiennej losowej μ jest rozkład Poissona z parametrem e^{-c}.
Wystarczy teraz udowodnić, że poza największą składową spójności $\mathcal{H}(\leqslant s, n, \underline{p})$ posiada tylko wierzchołki izolowane.
Ponieważ dla dowolnego podzbioru k-elementowego $W \subseteq V$

$$\Pr\{W \text{ jest zamknięty}\} = \prod_{j=2}^{s}(1-p_j)^{\left[\binom{n}{j}-\binom{k}{j}-\binom{n-k}{j}\right]}$$

$$\Pr\{W \text{ jest prosty}\} = \Pr\{W \subset K \text{ lub } V\backslash W \subset K\} =$$

$$= \prod_{j=2}^{s}(1-p_j)^{\binom{n}{j}-\binom{n-k}{j}} + \prod_{j=2}^{s}(1-p_j)^{\binom{n}{j}-\binom{k}{j}} - \prod_{j=2}^{s}(1-p_j)^{\binom{n}{j}}, \text{ więc}$$

$$\Pr\{A_k\} = \binom{n}{k}(\Pr\{W \text{ jest zamknięty}\} - \Pr\{W \text{ jest prosty}\}) =$$

$$= \binom{n}{k}\prod_{j=2}^{s}(1-p_j)^{\binom{n}{j}}\left(\prod_{j=2}^{s}(1-p_j)^{-\binom{k}{j}}-1\right)\left(\prod_{j=2}^{s}(1-p_j)^{-\binom{n-k}{j}}-1\right) =$$

$$= \binom{n}{k}e^{-\sum_{j=2}^{s}p_j\binom{n}{j}}\left(e^{\sum_{j=2}^{s}p_j\binom{k}{j}}-1\right)\left(e^{\sum_{j=2}^{s}p_j\binom{n-k}{j}}-1\right)$$

Czyli dla dużych n, podobnie jak w dowodzie dla hipergrafu s-jednorodnego mamy:

$$\sum_{j=2}^{s} a_j\frac{\binom{n}{j}-\binom{n-k}{j}-\binom{k}{j}}{\binom{n-1}{j-1}}(\ln n+c_j) \geqslant \sum_{j=2}^{s} a_j\frac{k}{j}(\ln n+c_j) \geqslant$$

$$\geqslant \frac{k}{s}\sum_{j=2}^{s} a_j(\ln n+c_j) = \frac{k}{s}(\ln n+c),$$

$$\sum_{j=2}^{s} a_j \frac{\binom{n}{j} - \binom{n-k}{j} - \binom{k}{j}}{\binom{n-1}{j-1}} (\ln n + c_j) = \sum_{j=2}^{s} a_j \left(k - O\left(\frac{k^2}{n}\right)\right) (\ln n + c_j) \geqslant$$
$$\geqslant \left(k - O\left(\frac{k^2}{n}\right)\right) (\ln n + c),$$
$$\sum_{j=2}^{s} a_j \frac{\binom{n}{j} - \binom{n-k}{j}}{\binom{n-1}{j-1}} (\ln n + c_j) \geqslant \left(k - O\left(\frac{k^2}{n}\right)\right) (\ln n + c)$$

oraz dla $k \leqslant n^{1-\delta}$

$$\sum_{j=2}^{s} a_j \frac{\binom{k}{j}}{\binom{n-1}{j-1}} (\ln n + c_j) \leqslant \sum_{j=2}^{s} a_j\, O\left(\frac{n^{j-j\delta}}{n^{j-1}}\right) (\ln n + c_j) \leqslant O\left(\frac{1}{n^{2\delta-1}}\right) (\ln n + c),$$

więc jeśli weźmiemy $\epsilon \in (0, 1/s)$ i $\delta \in (1/2, 1)$ oraz powtórzymy rozumowanie jak w (3.5) otrzymamy, że z prawdopodobieństwem dążącym do 1 $\mathcal{H}(\leqslant s, n, \underline{p})$ ma $\mu + 1$ składowych spójności. □

Wniosek 3.11. Dla dowolnego hipergrafu losowego $\mathcal{H}(\leqslant s, n, \underline{p})$, gdzie $\underline{p} = (p_2, p_3, \ldots, p_s)$ oraz

$$p_j(n) = a_j(n) \frac{\ln n + c_j(n)}{\binom{n-1}{j-1}} = a_j(n) \frac{(j-1)!(\ln n + c_j(n) + o(1))}{n^{j-1}}, ; \tag{3.21}$$

gdzie $\sum_{j=2}^{s} a_j(n) = 1$, $\sum_{j=2}^{s} a_j(n) c_j(n) = c + o(1)$ i $0 \leqslant a_j(n) \leqslant 1$ dla $j = 1, 2, \ldots, s$, mamy

$$\lim_{n\to\infty} \Pr\{\mathcal{H}(\leqslant s, n, \underline{p}) \text{jest spójny}\} = \lim_{n\to\infty} \Pr\{\mu = 0\} = e^{-e^{-c}}. \tag{3.22}$$

Wniosek 3.12. Dla dowolnego hipergrafu losowego $\mathcal{H}(\leqslant s, n, \underline{p})$, gdzie $\underline{p(n)} = (p_2(n), p_3(n), \ldots, p_s(n))$, jeśli

$$\exists_{n_0} \forall_{n>n_0}\; b \leqslant \sum_{j=2}^{s} \frac{n^{j-1} p_j(n)}{(j-1)!} - \ln n \leqslant c,$$

to

$$e^{-e^{-b}} \leqslant \underline{\lim}_{n\to\infty} \Pr\{\mathcal{H}(\leqslant s, n, \underline{p}) \text{jest spójny}\} \leqslant$$
$$\leqslant \overline{\lim}_{n\to\infty} \Pr\{\mathcal{H}(\leqslant s, n, \underline{p}) \text{ jest spójny}\} \leqslant e^{-e^{-c}}.$$

Dowód. Udowodnimy prawą nierówność. Jeśli podstawimy za:

$$a_j(n) = \frac{\frac{n^{j-1} p_j(n)}{(j-1)!}}{\sum_{i=2}^{s} \frac{n^{i-1} p_i(n)}{(i-1)!}} \quad \text{i} \quad c_j(n) = c,$$

Wtedy

$$\frac{n^{j-1}p_j(n)}{(j-1)!} = a_j(n)\sum_{i=2}^{s}\frac{n^{i-1}p_i(n)}{(i-1)!} \leqslant a_j(n)(\,c+\ln n)$$

$$p_j(n) \leqslant a_j(n)\frac{(j-1)!(c+\ln n)}{n^{j-1}} = p_j'(n)$$

Hipergraf $\mathcal{H}(\leqslant s, n, \underline{p'})$, gdzie $\underline{p'(n)} = (p_2'(n), p_3'(n), \ldots, p_s'(n))$ spełnia założenia wniosku 3.11. Ponieważ spójność jest własnością monotoniczną oraz dla dużych n i każdego j, $p_j(n) \leqslant p_j'(n)$, więc z wniosku 3.11 mamy

$$\overline{\lim_{n\to\infty}} \Pr\{\mathcal{H}(\leqslant s, n, \underline{p}) \text{ jest spójny}\} \leqslant \overline{\lim_{n\to\infty}} \Pr\{\mathcal{H}(\leqslant s, n, \underline{p'} \text{ jest spójny}\} = e^{-e^{-c}}.$$

Dowód lewej nierówności jest analogiczny. □

Wniosek 3.13. Dla dowolnego hipergrafu losowego $\mathcal{H}(\leqslant s, n, \underline{p})$, gdzie $\underline{p(n)} = (p_2(n), p_3(n), \ldots, p_s(n))$, jeśli

$$\lim_{n\to\infty}\left(\sum_{j=2}^{s}\frac{n^{j-1}p_j(n)}{(j-1)!} - \ln n\right) = \infty,$$

to

$$\lim_{n\to\infty} \Pr\{\mathcal{H}(\leqslant s, n, \underline{p}) \text{ jest spójny}\} = 1$$

oraz jeśli

$$\lim_{n\to\infty}\left(\sum_{j=2}^{s}\frac{n^{j-1}p_j(n)}{(j-1)!} - \ln n\right) = -\infty,$$

to

$$\lim_{n\to\infty} \Pr\{\mathcal{H}(\leqslant s, n, \underline{p}) \text{ jest spójny}\} = 0.$$

Dowód. W pierwszym przypadku dobieramy ciąg $b(n)$, taki,że $b(n) \to \infty$ przy $n \to \infty$ oraz dla dowolnego n, $b(n) \leqslant \sum_{j=2}^{s}\frac{n^{j-1}p_j(n)}{(j-1)!} - \ln n$. Wtedy zbieżność granicy do jedynki jest bezpośrednią konsekwencją poprzedniego wniosku. W drugim przypadku rozumujemy analogicznie. □

Uwaga 3.14. W powyższym wniosku można zastąpić sumę $\sum_{j=2}^{s}\frac{n^{j-1}p_j(n)}{(j-1)!}$ sumą $\sum_{j=2}^{s}\binom{n-1}{j-1}p_j(n)$

Wniosek 3.15. Dla dowolnego hipergrafu losowego $\mathcal{H}(\leqslant s, n, \underline{m})$, gdzie $\underline{m} = (m_2, m_3, \ldots, m_s)$ oraz

$$m_j(n) = \frac{a_j(n)\, n\, (\ln n + c_j(n))}{j};$$

gdzie $\sum_{j=2}^{s} a_j(n) = 1$, $\sum_{j=2}^{s} a_j(n)c_j(n) = c + o(1)$ i $0 \leqslant a_j(n) \leqslant 1$ dla $j = 1, 2, \ldots, s$, mamy

$$\lim_{n\to\infty} \Pr\{\mathcal{H}(\leqslant s, n, \underline{m})\text{jest spójny}\} = e^{-e^{-c}}.$$

Dowód. Po przeprowadzeniu obliczeń jak w (3.9) dla każdego p_i z osobna, wniosek wynika bezpośrednio z wniosku 3.11 i twierdzenia 2.8. $\square$

Argumentując identyczne jak w przypadku wniosków 3.12 i 3.13 możemy udowodnić,że

Wniosek 3.16.

$$\lim_{n\to\infty} \Pr\{\mathcal{H}(\leqslant s, n, \underline{m})\text{jest spójny}\} = \begin{cases} 0 & \text{, gdy } \frac{\sum_{j=2}^{s} j\, m_j(n)}{n} - \ln n \to -\infty \\ e^{-e^{-c}} & \text{, gdy } \frac{\sum_{j=2}^{s} j\, m_j(n)}{n} - \ln n \to c \\ 1 & \text{, gdy } \frac{\sum_{j=2}^{s} j\, m_j(n)}{n} - \ln n \to \infty \end{cases}$$

Fakt 3.17. Jeśli $m'_j = m_j + o(n)$ dla dowolnego $0 \leqslant j \leqslant s$ i m_j określonego jak wyżej, wtedy przy $n \to \infty$ prawdopodobieństwo spójności $\mathcal{H}(\leqslant s, n, \underline{m})$ jest równe prawdopodobieństwu spójności $\mathcal{H}(\leqslant s, n, \underline{m'})$.

4. Przejście fazowe

W poprzednim rozdziale pokazaliśmy, że prawdopodobieństwo spójności hipergrafu jest ściśle powiązane z prawdopodobieństwem, że każdy wierzchołek ma sąsiada (nie jest izolowany). Okazuje się, że własność posiadania sąsiada jest bardzo istotna w przypadku własności związanych ze spójnością. Jeśli analizujemy wielkość składowym spójności hipergrafów losowych, to ciekawe są te hipergrafy, w których średnio każdy wierzchołek ma dokładnie jednego sąsiada. A mianowicie, gdy średni stopień wierzchołkowy wierzchołka w hipergrafie przekracza jeden, moc największej składowej spójności zamienia się szybko z wielkości rzędu $\ln n$ do wielkości rzędu n. Czyli mamy do czynienia z tak zwanym przejściem fazowym. Przejście fazowe daje szersze spojrzenie na zagadnienie spójności hipergrafów. Daje też odpowiedź na pytania dotyczące największej składowej spójności grafów $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$.

Średnim stopniem i-wierzchołkowym hipergrafów $\mathcal{H}(\leqslant s, n, \underline{p})$ i $\mathcal{H}(\leqslant s, n, \underline{m})$ nazywamy odpowiednio:

$$\alpha_i = (i-1)\binom{n-1}{i-1}p_i \qquad \text{oraz} \qquad \alpha_i = (i-1)\frac{i\, m_i}{n}. \tag{4.1}$$

Natomiast średni stopień wierzchołkowy jest równy: $\alpha = \sum_{i=2}^{s} \alpha_i$. J.Schmidt--Pruzan i E.Shamir udowodnili w [16]:

Twierdzenie 4.1. *Niech C i ϵ będą dwoma stałymi dodatnimi takimi, że $C \geqslant 1$ i $0 < \epsilon < 1$ i niech $\mathcal{H}(\leqslant s, n, \underline{p})$ ma średni stopień wierzchołkowy $\alpha < 1$. Wtedy istnieją stałe $\overline{K}_{\alpha,C}$ i $\underline{K}_{\alpha,\epsilon}$ takie, że z prawdopodobieństwem $1 - o(n^{-C})$, największa składowa spójności w $\mathcal{H}(\leqslant s, n, \underline{p})$ ma mniej niż $\overline{K}_{\alpha,C} \ln n$ wierzchołków i z prawdopodobieństwem $1 - o(n^{-\epsilon})$ więcej niż $\underline{K}_{\alpha,\epsilon} \ln n$ wierzchołków. Co więcej indeks cykli składowych spójności jest co najwyżej k z prawdopodobieństwem $1 - o(n^{-k+o(1)})$.*

Twierdzenie 4.2. *Jeśli hipergraf $\mathcal{H}(\leqslant s, n, \underline{p})$ ma średni stopień wierzchołkowy $\alpha = 1$, wtedy z prawdopodobieństwem $1 - o(1)$ największa składowa spójności ma mniej niż $\omega_1(n)n^{2/3}$ i więcej niż $n^{2/3}/\omega(n)$ wierzchołków, gdzie $\omega(n)$ jest dowolną funkcją dążącą do nieskończoności z n i $\ln n = O(\omega_1(n))$.*

Twierdzenie 4.3. *Z prawdopodobieństwem $1 - o(1)$ największa składowa spójności hipergrafu $\mathcal{H}(\leqslant s, n, \underline{p})$ o średnim stopniu $\alpha > 1$ zawiera więcej niż $\beta_\alpha n/s$ wierzchołków, gdzie $0 < \beta_\alpha < 1$ jest pewną stałą niezależną od s i n.*

Aby udowodnić powyższe twierdzenia J.Schmidt-Pruzan i E.Shamir posłużyli się szeregiem lematów. Przez $C(K)$ oznaczymy zmienną losową określającą liczbę składowych spójności o K wierzchołkach w $\mathcal{H}(\leqslant s, n, \underline{p})$. Natomiast $T(K)$ będzie określać liczbę izolowanych hiperdrzew bez cykli na K wierzchołkach w $\mathcal{H}(\leqslant s, n, \underline{p})$.

Lemat 4.4. W hipergrafie $\mathcal{H}(V,E) \in \mathcal{H}(\leqslant s, n, \underline{p})$ ze średnim stopniem wierzchołkowym $\alpha \leqslant 1$ dla $1 \leqslant K_1 < K_2 \leqslant n$ mamy:

$$\Pr\{C([K_1, K_2]) > 0\} \geqslant 1 - \frac{1}{\mathbb{E}(C([K_1, K_2])}\ , \text{gdzie } \ C([K_1, K_2]) = \sum_{K=K_1}^{K_2} C(K).$$

Dowód. Dowód zostanie przeprowadzony przy wykorzystaniu metody drugiego momentu ([6],[14]). W przypadku, gdy $K_1 = K_2 = K$, mamy

$$\begin{aligned}\Pr\{C(K) \leqslant 0\} &\leqslant \Pr\{|C(K) - \mathbb{E}(C(K))| \geqslant \mathbb{E}(C(K))\} \leqslant \\ &\leqslant \frac{\mathrm{Var}(C(K))}{\mathbb{E}^2(C(K))} = \frac{\mathbb{E}(C(K)^2)}{\mathbb{E}^2(C(K))} - 1.\end{aligned}$$

Zatem wystarczy udowodnić,że

$$\frac{\mathbb{E}(C(K)^2)}{\mathbb{E}^2(C(K))} \leqslant 1 + \frac{1}{\mathbb{E}(C(K))}.$$

Ponumerujmy wszystkie K- elementowe podzbiory zbioru V, indeksami $1 \leqslant i \leqslant \binom{n}{K}$. Wtedy:

$$C(K) = \sum_{i=1}^{\binom{n}{K}} X_i, \text{ gdzie } X_i = X_i^K = \begin{cases} 1 & \text{gdy i-ty podzbiór K-elementowy} \\ & \text{jest składową spójności,} \\ 0 & \text{w przeciwnym przypadku.}\end{cases}$$

Ponieważ dowolne dwie różne składowe spójności są wierzchołkowo rozłączne, więc $\mathbb{E}(X_iX_j)$ jest większe od zera tylko wtedy, gdy zbiory o indeksach i oraz j są rozłączne. Istnieje dokładnie $\binom{n}{K}\binom{n-K}{K}$ takich par zbiorów. W przypadku rozłącznych par zdarzenia $X_i = 1$ oraz $X_j = 1$ są zależne, ponieważ istnieją podzbiory zbioru wierzchołków (potencjalne krawędzie) o niepustym przekroju zarówno z i-tym jak i j-tym zbiorem. Jeśli w hipergrafie podzbiór taki będzie krawędzią, wtedy ani i-ty ani j-ty zbiór nie jest składową spójności. Ponieważ takich k-elementowych potencjalnych krawędzi jest:

$$\begin{aligned}\binom{n}{k} - 2\binom{n-K}{k} + \binom{n-2K}{k} &= \binom{n}{k} - \binom{n-K}{k} - \left(\binom{n-K}{k} - \binom{n-2K}{k}\right) = \\ &= \sum_{l=1}^{K}\binom{n-l}{k-1} - \sum_{l=1}^{K}\binom{n-K-l}{k-1} = \sum_{l=1}^{K}\sum_{r=1}^{K}\binom{n-l-r}{k-2} \leqslant \\ &\leqslant K^2\binom{n-2}{k-2},\end{aligned}$$

mamy zatem:

$$\begin{aligned}\mathbb{E}(X_iX_j) &= \mathbb{E}(X_i)\mathbb{E}(X_j)\prod_{k=2}^{s}(1-p_k)^{-\left(\binom{n}{k}-2\binom{n-K}{k}+\binom{n-2K}{k}\right)} = \\ &\leqslant \mathbb{E}(X_i)\mathbb{E}(X_j)\prod_{k=2}^{s}(1-p_k)^{-K^2\binom{n-2}{k-2}}.\end{aligned}$$

Natomiast:

$$\binom{n-K}{K}/\binom{n}{K} = \prod_{k=0}^{K-1}\frac{n-K-k}{n-k} \leqslant \left(\frac{n-K}{K}\right)^K$$

oraz

$$\prod_{k=2}^{s}(1-p_k)^{K\binom{n-2}{k-2}} \geqslant 1-\sum_{k=2}^{s}\frac{K}{n}p_k\left((k-1)\binom{n-1}{k-1}\right) = 1-\frac{\alpha K}{n}.$$

Ostatecznie zakładając, że $1 \leqslant i,j \leqslant \binom{n}{K}$

$$\begin{aligned}\frac{\mathbb{E}(C(K)^2)}{\mathbb{E}^2(C(K))} &= \frac{\mathbb{E}(\sum_{i,j}X_iX_j)}{\left(\sum_i\mathbb{E}(X_i)\right)^2} = \frac{\sum_{i\neq j}\mathbb{E}(X_iX_j)+\sum_i\mathbb{E}(X_i^2)}{\left(\sum_i\mathbb{E}(X_i)\right)^2} \leqslant \\ &\leqslant \frac{\binom{n}{K}\binom{n-K}{K}\mathbb{E}(X_i)\mathbb{E}(X_j))}{\binom{n}{K}\binom{n}{K}\mathbb{E}^2(X_i)}\prod_{k=2}^{s}(1-p_k)^{-K^2\binom{n-2}{k-2}} + \frac{\binom{n}{K}\mathbb{E}(X_i)}{\binom{n}{K}\binom{n}{K}\mathbb{E}^2(X_i)} \leqslant \\ &\leqslant \frac{\binom{n-K}{n}}{\binom{n}{K}}\left(1-\frac{\alpha K}{n}\right)^K + \frac{1}{\binom{n}{K}\mathbb{E}(X_i)} \leqslant 1+\frac{1}{\mathbb{E}(C(K))}\end{aligned}$$

W przypadku, gdy $K_1 < K_2$ oznaczmy $K_1 \leqslant K', K'' \leqslant K_2$, $1 \leqslant i \leqslant \binom{n}{K'}$ i $1 \leqslant j \leqslant \binom{n}{K''}$, $X_i^{s'} = X_i$, $X_j^{K''} = X_j$.

$$(C([K_1,K_2]))^2 = \left(\sum_{K_1\leqslant K'\leqslant K_2}\sum_i X_i^{K'}\right)^2 = \sum_{K'}\sum_{K''}\sum_{i,j}X_i^{K'}X_j^{K''}.$$

Dla ustalonych K' i K'' takich par indeksów (i,j), że $\mathbb{E}(X_iX_j)$ jest różne od zera jest tyle, ile par rozłącznych podzbiorów zbioru V o mocach K' i K'', czyli dokładnie $\binom{n}{K'}\binom{n-K'}{K''}$. Natomiast k-elementowych krawędzi między tymi podzbiorami jest:

$$\begin{aligned}\binom{n}{k}-\binom{n-K'}{k}-\binom{n-K''}{k}+\binom{n-K'-K''}{k} &= \\ &= \binom{n}{k}-\binom{n-K'}{k}-\left(\binom{n-K'}{k}-\binom{n-K'-K''}{k}\right) = \\ &= \sum_{l=1}^{K'}\binom{n-l}{k-1}-\sum_{l=1}^{K''}\binom{n-K'-l}{k-1} = \sum_{l=1}^{K'}\sum_{r=1}^{K''}\binom{n-l-r}{k-2} \leqslant \\ &\leqslant K'K''\binom{n-2}{k-2}.\end{aligned}$$

Czyli dokładniej możemy oszacować:

$$\frac{\mathbb{E}(C([K_1,K_2])^2)}{\mathbb{E}^2(C([K_1,K_2]))} = \frac{\mathbb{E}(\sum_{K'}\sum_{K''}\sum_{i,j}(X_iX_j))}{\left(\sum_{K'}\sum_i \mathbb{E}(X_i)\right)^2} =$$

$$= \frac{\mathbb{E}(\sum_{K'}\sum_{K''}\sum_{i\neq j}(X_iX_j)) + \mathbb{E}(\sum_{K'}\sum_{K''}\sum_{i=j}(X_iX_j))}{\left(\sum_{K'}\sum_i \mathbb{E}(X_i)\right)^2} =$$

$$= \frac{\sum_{K'}\sum_{K''}\binom{n}{K'}\binom{n-K'}{K''}\mathbb{E}(X_iX_j)}{\left(\sum_{K'}\binom{n}{K'}\mathbb{E}(X_i)\right)^2} + \frac{\sum_{K'}\sum_i \mathbb{E}(X_i)}{\left(\sum_{K'}\sum_i \mathbb{E}(X_i)\right)^2} \leqslant$$

$$= \frac{\sum_{K'}\sum_{K''}\binom{n}{K'}\mathbb{E}(X_i)\mathbb{E}(X_j)\binom{n-K'}{K''}\prod_{k=2}^{s}(1-p_k)^{-K'K''\binom{n-2}{k-2}}}{\sum_{K'}\sum_{K''}\binom{n}{K'}\binom{n}{K''}\mathbb{E}(X_i)\mathbb{E}(X_j)} +$$

$$+ \frac{1}{\mathbb{E}(C[K_1,K_2])} \leqslant$$

$$\leqslant \frac{\sum_{K'}\sum_{K''}\binom{n}{K'}\mathbb{E}(X_i)\mathbb{E}(X_j)\binom{n}{K''}}{\sum_{K'}\sum_{K''}\binom{n}{K'}\binom{n}{K''}\mathbb{E}(X_i)\mathbb{E}(X_j)} + \frac{1}{\mathbb{E}(C[K_1,K_2])} =$$

$$= 1 + \frac{1}{\mathbb{E}(C[K_1,K_2]))}$$

Ostatecznie

$$\Pr\{C([K_1,K_2]) > 0\} =$$

$$= 1 - \Pr\{C([K_1,K_2]) \leqslant 0\} =$$

$$= 1 - \Pr\{\mathbb{E}(C([K_1,K_2])) - C([K_1,K_2]) \geqslant \mathbb{E}(C([K_1,K_2]))\} \geqslant$$

$$\geqslant 1 - \Pr\{|\mathbb{E}(C([K_1,K_2]) - C([K_1,K_2]| \geqslant \mathbb{E}(C([K_1,K_2])\} \geqslant$$

$$\geqslant 1 - \frac{Var(C([K_1,K_2]))}{(\mathbb{E}(C([K_1,K_2]))^2} =$$

$$= 1 - \frac{\mathbb{E}(C([K_1,K_2])^2)}{\mathbb{E}(C([K_1,K_2]))^2} + \frac{\mathbb{E}(C([K_1,K_2]))^2}{\mathbb{E}(C([K_1,K_2]))^2} \geqslant$$

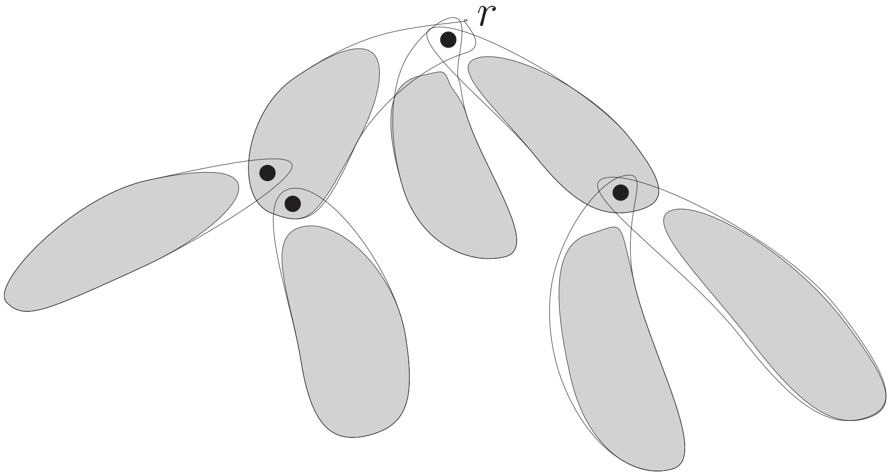
$$\geqslant 2 - 1 - \frac{1}{\mathbb{E}(C([K_1,K_2]))} = 1 - \frac{1}{\mathbb{E}(C([K_1,K_2]))}$$

□

Zauważmy, że dokładnie ten sam dowód jaki przedstawiliśmy dla dowolnych składowych spójności można przeprowadzić dla spójnych hiperdrzew na K wierzchołkach.

Lemat 4.5. Liczba hiperdrzew bez cykli o krawędziach o mocy co najwyżej s na $[K+1]$ oznaczonych wierzchołkach jest równa:

$$\sum_{\{T_i|\sum_i(i-1)T_i=K\}} \frac{(K+1)!(K+1)^{(\sum T_i-1)}}{(K+1)\prod_i(T_i)!(i-1)!^{T_i}},$$



Rysunek 4.1. Przykład podziału zbioru wierzchołków hiperdrzewa na podzbiory i korzeń.

gdzie $i = 2, 3, \ldots, s$ oraz T_i oznacza liczbę krawędzi o mocy i w hiperdrzewie.

Dowód. Niech $T(V, E)$ będzie dowolnym spójnym hiperdrzewem bez cykli na [K+1] wierzchołkach ($|V(T)| = K + 1$). W hiperdrzewie tym przez T_i oznaczmy liczbę krawędzi o mocy i, czyli $\sum_i T_i = |E(T)|$. Przypomnijmy, że w hiperdrzewie bez cykli dowolne dwie krawędzie przecinają się na co najwyżej jednym wierzchołku. Ponieważ dowolne dwie krawędzie przecinają się na co najwyżej jednym wierzchołku, więc jeśli ustalimy pewien wierzchołek r ze zbioru wierzchołków, wtedy krawędzie spójnego hiperdrzewa bez cykli w jednoznaczny sposób wyznaczają podział zbioru wierzchołków na $\sum_i T_i$ podzbiorów (krawędzi bez jednego wierzchołka) i korzeń r (rys.4.1). Z podziału tego wynika, że $K+1 = |V| = \sum_i (i-1)T_i + 1$, czyli

$$\sum_i (i-1)T_i = K.$$

Natomiast z definicji stopnia krawędziowego w dowolnym hipergrafie mamy równość $\sum_{v \in V} d(v) = \sum_i iT_i$. Ostatecznie dla hiperdrzew bez cykli z dwóch ostatnich równości wynika równość:

$$\sum_{v \in V} (d(v) - 1) = \sum_i i\, T_i - K - 1 = \sum_i T_i - 1.$$

Każde hiperdrzewo z ustalonym korzeniem wyznacza zatem podział wierzchołków na pewne podzbiory o określonych własnościach. Z drugiej strony, każdy podział zbioru wierzchołków na korzeń i podzbiory wyznacza pewne hiperdrzewo bez cykli. Danemu podziałowi nie można przyporządkować jednoznacznie hiperdrzewa, gdyż różne hiperdrzewa mogą wyznaczyć ten sam podział. Określimy dokładną liczbę hiperdrzew o danym podziale. W tym celu posłużymy się procedurą konstruującą hiperdrzewa.

Ustawmy wierzchołki z [K+1]-elementowego zbioru w ciąg $\{r, v_1, v_2, \ldots, v_K\}$, ustalmy liczby T_i takie, że $\sum_i (i-1)T_i = K$ oraz dowolny ciąg nieujemnych liczb całkowitych $(d_r, d_{v_1}, \ldots, d_{v_K})$ takich, że $d_r + \sum_{j=1}^{K} d_{v_j} = \sum_i T_i - 1$. Dla tak ustalonych ciągów i danego podziału zbioru wierzchołków na korzeń i $\sum_i T_i$ podzbiorów możemy przeprowadzić procedurę konstruującą spójne hiperdrzewo bez cykli na $[K+1]$ wierzchołkach:

1. Wszystkie $\sum_i T_i$ zbiory są "wolne" i "nie zamrożone". Wszystkie wierzchołki są "wolne".
2. "Zablokuj" korzeń i wszystkie wierzchołki v takie, że $d_v = 0$.(wierzchołek "zablokowany"$\neq$"wolny")
3. (*While*) Dopóki istnieje "wolny" wierzchołek, powtarzaj (Konstrukcja poddrzewa zakorzenionego na "wolnym" wierzchołku):
   (*do begin*)
   a) Weź pierwszy na liście "wolny" wierzchołek v należący do "wolnego" zbioru. "Zamroź" zbiór zawierający wierzchołek v . Dla wierzchołka v wybierz d_v sąsiadujących podzbiorów spośród "wolnych" i "nie zamrożonych" zbiorów. "Zablokuj" wszystkie wybrane zbiory.
   b) (*While*) Dopóki istnieje "wolny" wierzchołek w "zablokowanym" zbiorze powtarzaj:
      (*do begin*)
      Dla v_f "wolnego" wierzchołka w "zablokowanym" zbiorze wybierz d_{v_f} sąsiadujących podzbiorów spośród "wolnych" i "nie zamrożonych". "Zablokuj" wybrane podzbiory i wierzchołek v_f.
      (*end*)
   "Odmroź" zbiór zawierający v.
   (*end*)
4. Wszystkie "wolne" podzbiory (jest ich dokładnie $\sum_i T_i - \sum_{j=1}^{K} d_{v_j} = d_r + 1$) zostają sąsiadami korzenia r.

Trzeba zwrócić uwagę na to, że procedura nie może zaczynać się od korzenia, gdyż w pierwszym kroku korzeń mógłby "wybrać" same podzbiory z wszystkimi wierzchołkami stopnia jeden. W takim przypadku procedura nie mogłaby być kontynuowana.

W punkcie 3a) pierwszy z listy wierzchołek v wybiera sąsiadujące podzbiory na $\binom{\sum T_i - 1}{d_v}$ sposobów. Drugi wierzchołek w wybiera na $\binom{\sum T_i - 1 - d_v}{d_w}$ etc. Ostatecznie procedura dla określonego ciągu stopni generuje: $\binom{\sum_i T_i - 1}{d_r, d_{v_1}, \ldots, d_{v_K}}$ różnych hiperdrzew. Czyli określony podziału zbioru wierzchołków na korzeń i T_i podzbiorów o mocy $i-1$, wyznacza dokładnie

$$\sum_{\substack{(d_r,d_{v_1},\ldots,d_{v_K})\\ d_r\geqslant 0,\ d_{v_i}\geqslant 0}} \binom{\sum_i T_i - 1}{d_r, d_{v_1}, \ldots, d_{v_K}} = (K+1)^{\sum_i T_i - 1} \tag{4.2}$$

spójnych hipergrafów bez cykli.

Dla ustalonych liczb T_i oznaczających liczbę krawędzi o mocy i, takich, że $\sum_i (i-1)T_i = K$ i dla danego korzenia r możemy określić dokładnie

$$\frac{(K+1)!}{\prod_i (T_i)!(i-1)!^{T_i}}, \tag{4.3}$$

różnych podziałów zbioru wierzchołków. Czyli, ponieważ korzeń można wybrać na $K+1$ sposobów z (4.2) i (4.3) wynika, że wszystkich hiperdrzew na oznaczonych wierzchołkach jest:

$$\sum_{\{T_i|\sum_i (i-1)T_i = K\}} \frac{(K+1)!(K+1)^{(\sum T_i - 1)}}{(K+1)\prod_i (T_i)!(i-1)!^{T_i}}.$$

□

Lemat 4.6. Niech $T(K)$ określa liczbę izolowanych drzew acyklicznych na K wierzchołkach w hipergrafie $\mathcal{H}(\leqslant s, n, \underline{p})$ takiego, że $p_i = \alpha_i / \left((i-1)\binom{n}{i-1}\right)$ dla $i = 2, 3, \ldots, s$ oraz $\sum_{i=2}^{s} \alpha_i = \alpha$, wtedy

(4.4)

$$\text{dla } \alpha < 1 \begin{cases} \mathbb{E}(T(K)) = o(n^{-C}) & \text{dla } K \geqslant \overline{K}_{\alpha,C} s \ln n = C\overline{K}_\alpha s \ln n \\ \mathbb{E}(T(K)) \geqslant \theta(n^\delta) & \text{dla } K < \underline{K}_\delta \ln n \end{cases}$$

(4.5)

$$\text{dla } \alpha = 1 \begin{cases} \mathbb{E}(T(K)) = o\left(n^{-2/3} e^{-\omega^3(n)}\right) & \text{dla } K \geqslant \omega(n) n^{2/3} \\ \mathbb{E}(\sum_{\mathcal{J}} T(K)) \geqslant \theta(\omega(n)^{1,5}) & \text{dla } \mathcal{J} = \left\{K | \frac{1}{2} n^{2/3} < \omega(n) K < n^{2/3}\right\} \end{cases}$$

(4.6)

$$\text{dla } \alpha > 1 \quad \mathbb{E}(T(K)) = o(e^{-\beta(K/s)}) \text{ dla } \underline{K}'_\alpha s \ln n < K < \overline{K}'_\alpha n/s.$$

Dowód. Niech $K' = K + 1$.
Zauważmy, że w hipergrafie $\mathcal{H}(\leqslant s, n, \underline{p})$ określone hiperdrzewo na K' wierzchołkach zawierające dokładnie T_i krawędzi o i wierzchołkach jest izolowane z prawdopodobieństwem:

$$\prod_i p_i^{T_i} \prod_i (1-p_i)^{\sum_{j=1}^{i} \binom{K'}{j}\binom{n-K'}{i-j} - T_i}$$

Pierwszy iloczyn określa prawdopodobieństwo powstania hiperdrzewa w hipergrafie $\mathcal{H}(\leqslant s, n, \underline{p})$ a drugi prawdopodobieństwo, że to hiperdrzewo jest izolowane. Dodatkowo ponieważ $p_i = \frac{\alpha_i(i-1)!}{(i-1)(n)_{i-1}} = O(1/n^{i-1})$ oraz $T_i \leqslant K' \leqslant n$, więc $1 \leqslant (1-p_i)^{-T_i} \leqslant (1-p_i)^{-n} = e^{O(1)}$. Czyli $(1-p_i)^{-T_i} = e^{O(1)}$.
Możemy także oszacować:

$$
\begin{aligned}
\sum_{j=1}^{i} \binom{K'}{j}\binom{n-K'}{i-j} &= \binom{n}{i} - \binom{n-K-1}{i} = \sum_{j=1}^{K+1} \binom{n-j}{i-1} = \\
&= (K+1)\binom{n}{i-1} - \sum_{j=1}^{K+1}\left(\binom{n}{i-1} - \binom{n-j}{i-1}\right) = \\
&= K\binom{n}{i-1} + O(n^{i-1}) - \sum_{j=1}^{K+1}\sum_{l=1}^{j}\binom{n-j-l}{i-2} = \\
&= K\binom{n}{i-1} - \frac{(K+1)(K+2)}{2}\binom{n}{i-2} + \sum_{j=1}^{K+1}\sum_{l=1}^{j}\left(\binom{n}{i-2} - \binom{n-j-l}{i-2}\right) + O(n^{i-1}) = \\
&= K\binom{n}{i-1} - \frac{K^2}{2}\binom{n}{i-2} + O(K^3 n^{i-3}) + O(n^{i-1}).
\end{aligned}
$$

Mamy też: $p_i^2 \sum_{j=1}^{i} \binom{K'}{j}\binom{n-K'}{i-j} = O(n^{-2(i-1)}\ n^i) = O(1)$ oraz $O(K^3 n^{i-3})p_i = O(K^3/n^2)$
Będziemy wykorzystywać to, że $e^{O(1)} = \theta(1)$ oraz to, że w powyższych oszacowaniach $O(1)$ nie zależy od wartości T_i. Zatem z lematu 4.5 stosując podstawienie $p_i = \frac{\alpha_i(i-1)!}{(i-1)(n)_{i-1}}$:

(4.7)

$$
\begin{aligned}
\mathbb{E}(T(K')) = \sum_{\{T_i | \sum_i (i-1)T_i = K'-1\}} \frac{(n)_{K'}}{K'!} \frac{(K')!(K')^{(\sum T_i - 1)}}{(K')\prod_i (T_i)!(i-1)!^{T_i}} \\
\prod_i p_i^{T_i} \prod_i (1-p_i)^{\sum_{j=1}^{i}\binom{K'}{j}\binom{n-K'}{i-j} - T_i} = \\
= \sum_{\{T_i | \sum_i (i-1)T_i = K'-1\}} \frac{(n)_{K'} \prod_i K'^{T_i}}{K'^2} \prod_i \frac{p_i^{T_i}}{(T_i)!(i-1)^{T_i}!} \\
\prod_i e^{-p_i K\binom{n}{i-1} + p_i \frac{K^2}{2}\binom{n}{i-2}} e^{O(1)} e^{-O(K^3/n^2)} = \\
= \sum_{\{T_i | \sum_i (i-1)T_i = K'-1\}} \frac{1}{K'^2} \frac{(n)_{K'}}{\prod_i (n)_{i-1}^{T_i}} \prod_i \frac{K'^{T_i}\alpha_i^{T_i}}{(i-1)^{T_i} T_i!} \\
e^{O(1)} e^{-O(K^3/n^2)} \prod_i e^{-p_i K\binom{n}{i-1} + p_i \frac{K^2}{2}\binom{n}{i-2}} =
\end{aligned}
$$

$$= \theta(1) \sum_{\{T_i | \sum_i (i-1)T_i = K\}} \frac{n}{K^2} \frac{(n)_K}{\prod_i (n)_{i-1}^{T_i}}$$
$$e^{-O(K^3/n^2)} \prod_i e^{-Kp_i\binom{n}{i-1} + \frac{K^2}{2} p_i \binom{n}{i-2}} \prod_i \frac{(K\alpha_i)^{T_i}}{(i-1)^{T_i}(T_i)!} =$$
$$= \theta(1) \sum_{\{T_i | \sum_i (i-1)T_i = K\}} \frac{n}{K^2} \frac{n^K e^{-K^2/(2n) - O(K^3/n^2)}}{n^{\sum_i (i-1)T_i} e^{-O(\sum_i (i-1)^2/2n)T_i}}$$
$$e^{-O(K^3/n^2)} e^{-K \sum_i \alpha_i/(i-1)} e^{\frac{K^2}{2} \sum_i \alpha_i/(n-i+2)} \prod_i \frac{(K\alpha_i)^{T_i}}{(i-1)^{T_i}(T_i)!} =$$
$$= \theta(1) \sum_{\{T_i | \sum_i (i-1)T_i = K\}} \frac{n}{K^2} e^{-\frac{K^2}{2n}(1-\alpha)} e^{-O(\frac{K^3}{n^2})} e^{(s-1)\frac{K}{2n}} e^{-K \sum_i \alpha_i/(i-1)} \prod_i \frac{(K\alpha_i)^{T_i}}{(i-1)^{T_i}(T_i)!}.$$

Ostatecznie

$$\mathbb{E}(T(K')) = \sum_{\{T_i | \sum_i (i-1)T_i = K\}} \rho(K, \alpha_i) f\left(\vec{T}\right), \text{ gdzie} \tag{4.8}$$

$$\rho(K, \alpha_i) = \theta(1) \frac{n}{K^2} e^{\frac{K^2}{2n}(1-\alpha)} e^{O(\frac{K^3}{n^2})} e^{(s-1)\frac{K}{2n}} e^{-K \sum_i \alpha_i/(i-1)} \text{ oraz} \tag{4.9}$$

$$f\left(\vec{T}\right) = \prod_i \frac{(K\alpha_i)^{T_i}}{(i-1)^{T_i}(T_i)!}, \text{ dla } \vec{T} = (T_2, \ldots, T_s). \tag{4.10}$$

Oszacujemy najpierw wyrażenie $f\left(\vec{T}\right)$. Przyjmijmy za

$$T_i^0 = \frac{(\alpha_i + \epsilon_i)K}{i-1}, \text{ gdzie} \tag{4.11}$$

$$\epsilon_i = \alpha_i(e^{(i-1)\beta} - 1) = \alpha_i e^{(i-1)\beta} - \alpha_i \tag{4.12}$$

oraz β jest jedynym rzeczywistym rozwiązaniem równania:

$$\sum_i \alpha_i e^{(i-1)\beta} = 1. \tag{4.13}$$

Rozwiązanie takie istnieje, gdyż dla dodatnich α_i funkcja $g(\beta) = \sum_i \alpha_i e^{(i-1)\beta}$ zmiennej β jest ściśle rosnąca oraz dąży do zera przy β dążącym do minus nieskończoności oraz do nieskończoności dla β dążącego do nieskończoności. Mamy zatem dla określonego β

$$\frac{\alpha_i}{\alpha_i + \epsilon_i} = \frac{\alpha_i}{\alpha_i e^{(i-1)\beta}} = e^{-(i-1)\beta}.$$

Udowodnimy, że $f(\vec{T}) \leqslant f(\vec{T^0})$ dla dowolnego $\vec{T}$ takiego, że $\sum_i (i-1)T_i = K$. Oczywiście $\sum_i (i-1)T_i^0 = K$. Podzielmy zbiór indeksów $I = \{2, \ldots s\}$ na dwa podzbiory tak, żeby do L należały takie indeksy l dla, których $T_l < T_l^0$ oraz do J należały takie indeksy j, że $T_j \geqslant T_j^0$. Wtedy dla pewnych dodatnich c_l i c_j

$$\forall_{l\in L}\ T_l + c_l = T_l^0 \qquad \text{oraz} \qquad \forall_{j\in J}\ T_j - c_j = T_j^0\ . \tag{4.14}$$

Wtedy

$$\sum_l (l-1)T_l + \sum_j (j-1)T_j = K = \sum_l (l-1)(T_l + c_l) + \sum_j (j-1)(T_j - c_j),\ \text{czyli}$$

$$\sum_l (l-1)c_l - \sum_j (j-1)c_j = 0. \tag{4.15}$$

Czyli

$$\begin{aligned}
f\left(\vec{T}\right) &= \prod_{l\in L} \frac{(K\alpha_l)^{T_l^0 - c_l}}{(l-1)^{T_l^0 - c_l}(T_l^0 - c_l)!} \prod_{j\in J} \frac{(K\alpha_j)^{T_j^0 + c_j}}{(j-1)^{T_j^0 + c_j}(T_j^0 + c_j)!} = \\
&= \prod_{i\in I} \frac{(K\alpha_i)^{T_i^0}}{(i-1)^{T_i^0}(T_i^0)!} \frac{\prod_{j\in J}\left(\frac{K\alpha_j}{j-1}\right)^{c_j}}{\prod_{l\in L}\left(\frac{K\alpha_l}{l-1}\right)^{c_l}} \frac{\prod_{l\in L}(T_l^0)_{c_l}}{\prod_{j\in J}(T_j^0 + c_j)_{c_j}} = \\
&= f\left(\vec{T^0}\right) \frac{\prod_{j\in J}\left(\frac{K\alpha_j}{j-1}\right)^{c_j}}{\prod_{l\in L}\left(\frac{K\alpha_l}{l-1}\right)^{c_l}} \frac{\prod_{l\in L}(T_l^0)^{c_l}}{\prod_{j\in J}(T_j^0)^{c_j}} e^{-O\left(\sum_l c_l^2/T_l^0\right)} = \\
&= f\left(\vec{T^0}\right) \frac{\prod_{j\in J}\left(\frac{K\alpha_j}{j-1}\right)^{c_j}}{\prod_{l\in L}\left(\frac{K\alpha_l}{l-1}\right)^{c_l}} \frac{\prod_{l\in L}\left(\frac{(\alpha_l+\epsilon_l)K}{l-1}\right)^{c_l}}{\prod_{j\in J}\left(\frac{(\alpha_j+\epsilon_j)K}{j-1}\right)^{c_j}} e^{-O\left(\sum_l c_l^2/T_l^0\right)} = \\
&= f\left(\vec{T^0}\right) \frac{\prod_{j\in J}\left(\frac{\alpha_j}{\alpha_j+\epsilon_j}\right)^{c_j}}{\prod_{l\in L}\left(\frac{\alpha_l}{\alpha_l+\epsilon_l}\right)^{c_l}} e^{-O\left(\sum_l c_l^2/T_l^0\right)} = \\
&= f\left(\vec{T^0}\right) \frac{\prod_{j\in J} e^{-(i-1)\beta c_j}}{\prod_{l\in L} e^{-(l-1)\beta c_l}} e^{-O\left(\sum_l c_l^2/T_l^0\right)} = \\
&= f\left(\vec{T^0}\right) e^{\beta\left(\sum_{l\in L}(l-1)c_l - \sum_{j\in J}(j-1)c_j\right)} e^{-O\left(\sum_l c_l^2/T_l^0\right)} = \\
&= f\left(\vec{T^0}\right) e^{-O\left(\sum_l c_l^2/T_l^0\right)} \leqslant f\left(\vec{T^0}\right).
\end{aligned}$$

Zatem z powyższej równości i (4.8), wynika, że wystarczy oszacować wyrażenie $\rho(K,\alpha_i)f\left(\overrightarrow{T^0}\right)$.

$$\begin{aligned}
\rho(K,\alpha_i)f\left(\overrightarrow{T^0}\right) &= \rho(K,\alpha_i)\prod_{i\in I}\frac{(K\alpha_i)^{T_i^0}}{(i-1)^{T_i^0}(T_i^0)!} \leqslant \\
&\leqslant \rho(K,\alpha_i)\prod_{i\in I}\left(\frac{K\alpha_i}{i-1}\right)^{T_i^0}(T_i^0)^{-T_i^0}e^{T_i^0} = \\
&= \rho(K,\alpha_i)\prod_{i\in I}\left(\frac{\alpha_i}{\alpha_i+\epsilon_i}\right)^{\frac{K(\alpha_i+\epsilon_i)}{i-1}}e^{\frac{K(\alpha_i+\epsilon_i)}{i-1}} = \\
&= \rho(K,\alpha_i)\prod_{i\in I}\left(1-\frac{\epsilon_i}{\alpha_i+\epsilon_i}\right)^{\frac{K(\alpha_i+\epsilon_i)}{i-1}}e^{\frac{K(\alpha_i+\epsilon_i)}{i-1}} \leqslant \\
&\leqslant \rho(K,\alpha_i)\prod_{i\in I}e^{-\left(\frac{K\epsilon_i}{i-1}+\frac{K\epsilon_i^2}{2(i-1)(\alpha_i+\epsilon_i)}\right)+\frac{K(\alpha_i+\epsilon_i)}{i-1}} = \\
&= \rho(K,\alpha_i)\,e^{\sum_{i\in I}\frac{K}{i-1}\left(\alpha_i-\frac{\epsilon_i^2}{2(\alpha_i+\epsilon_i)}\right)}
\end{aligned} \tag{4.16}$$

Zauważmy dodatkowo, że z (4.12), (4.13) i nierówności Cauchy-Schwarza

$$\begin{aligned}
\sum_{i\in I}\frac{\epsilon_i^2}{\alpha_i+\epsilon_i} = \sum_{i\in I}\alpha_i\frac{\left(e^{(i-1)\beta}-1\right)^2}{e^{(i-1)\beta}} &= \sum_{i\in I}\alpha_i\left(e^{(i-1)\beta}-2+e^{-(i-1)\beta}\right) = \\
&= 1-2\sum_{i\in I}\alpha_i+\left(\sum_{i\in I}\alpha_ie^{-(i-1)\beta}\right)\left(\sum_{i\in I}\alpha_ie^{(i-1)\beta}\right) \geqslant \\
&\geqslant 1-2\sum_{i\in I}\alpha_i+\left(\sum_{i\in I}\alpha_i\right)^2 = \left(1-\sum_{i\in I}\alpha_i\right)^2 = (1-\alpha)^2
\end{aligned}$$

Czyli ostatecznie mamy

$$\begin{aligned}
\rho(K,\alpha_i)f\left(\overrightarrow{T^0}\right) &\leqslant \rho(K,\alpha_i)e^{\sum_{i\in I}\frac{K\alpha_i}{i-1}-\frac{K}{s-1}(1-\alpha)^2} \leqslant \\
&\leqslant \theta(1)\frac{n}{K^2}e^{-\frac{K^2}{2n}(1-\alpha)}e^{-O(\frac{K^3}{n^2})}e^{(s-1)\frac{K}{2n}}e^{-K\sum_i\frac{\alpha_i}{(i-1)}} \\
&\qquad e^{\sum_{i\in I}\frac{K\alpha_i}{i-1}-\frac{K}{s-1}(1-\alpha)^2} \leqslant \\
&\leqslant \theta(1)\frac{n}{K^2}e^{-\frac{K^2}{2n}(1-\alpha)}e^{(s-1)\frac{K}{2n}-\frac{K}{s-1}(1-\alpha)^2} \leqslant \\
&\leqslant \theta(1)\frac{n}{K^2}e^{-\frac{K^2}{2n}(1-\alpha)}e^{-\frac{K}{s-1}(1-\alpha)^2\left(1-\frac{(s-1)^2}{2n(1-\alpha)^2}\right)} = \\
&= \theta(1)\frac{n}{K^2}e^{-\frac{K^2}{2n}(1-\alpha)}e^{-\frac{K}{s-1}(1-\alpha)^2(1-o(1))}
\end{aligned} \tag{4.17}$$

Rozważmy teraz trzy przypadki:

1. $\alpha < 1$
Zbiór $\{T_i | \sum_{i=2}^{s}(i-1)T_i = K\}$ ma co najwyżej $\binom{K}{s-1}$ elementów. Mianowicie ciągów $(\sum_{i=2}^{s} T_i, \sum_{i=3}^{s} T_i, \ldots, \sum_{i=s}^{s} T_i)$, takich, że $\sum_{i=2}^{s} T_i + \sum_{i=3}^{s} T_i + \ldots + \sum_{i=s}^{s} T_i = K$ jest co najwyżej tyle ile sum s niezerowych składników równych K, dokładnie co najwyżej $\binom{K-1}{s-2} \leqslant \binom{K}{s-1}$. Czyli szacowana suma (4.7) ma co najwyżej $\binom{K}{s-1}$ składników. Stąd oraz z (4.17) mamy dla dużych n i dla $K = \overline{K}_{\alpha,C} s \ln n$, gdzie $\overline{K}_{\alpha,C} = CK_\alpha = C\frac{4}{(1-\alpha)^2}$, $C \geqslant 1$:

$$\begin{aligned}
\mathbb{E}(T(K')) &= \sum_{\{T_i | \sum_i (i-1)T_i = K'-1\}} O(1)\frac{n}{K^2} e^{-\frac{K}{2(s-1)}(1-\alpha)^2} = \\
&= O(1)\frac{n}{K^2}\binom{K}{s-1} e^{-\frac{K}{2(s-1)}(1-\alpha)^2} = \\
&= O(1)n\frac{(\overline{K}_{\alpha,C} s \ln n)^{s-3} e^{s-1}}{(s-1)^{s-1}} n^{-2C\frac{s}{s-1}} = \\
&= O(1)(\ln n)^{s-3} n^{-2C+1-\frac{2}{s-1}} = \\
&= O(1)\frac{(\ln n)^{s-3}}{n^{\frac{2}{s-1}}} n^{-C} = o(n^{-C})
\end{aligned} \tag{4.18}$$

Natomiast dla $K < \underline{K}_{\delta,\alpha} \ln n$, gdzie $0 < \delta < 1$, $\underline{K}_{\delta,\alpha} = \frac{1-\delta}{1+\beta}$ (warto zauważyć, że dla $\alpha < 1$ z (4.13) $-\ln \alpha/(s-1) \leqslant \beta \leqslant -\ln \alpha$), z (4.16) składnik sumy (4.7) przyporządkowany $\overrightarrow{T^0}$ można oszacować.

(4.19)

$$\begin{aligned}
\rho(K,\alpha_i) f\left(\overrightarrow{T^0}\right) &= \rho(K,\alpha_i)\prod_{i\in I}\frac{(K\alpha_i)^{T_i^0}}{(i-1)^{T_i^0}(T_i^0)!} \geqslant \\
&\geqslant \theta(1)\frac{n}{K^2} e^{\frac{K^2}{2n}(1-\alpha)} e^{O(\frac{K^3}{n^2})} e^{(s-1)\frac{K}{2n}} e^{-K\sum_i \alpha_i/(i-1)} \prod_{i\in I}\left(\frac{K\alpha_i}{(i-1)T_i^0}\right)^{T_i^0} = \\
&= \theta(1)\frac{n}{K^2} e^{o(1)} e^{-K\sum_i \alpha_i/(i-1)} \prod_{i\in I}\left(\frac{\alpha_i}{\alpha_i + \epsilon_i}\right)^{T_i^0} \geqslant \\
&\geqslant \theta(1)\frac{n}{K^2} e^{-\alpha K} \prod_{i\in I} e^{-(i-1)\beta T_i^0} = \theta(1)\frac{n}{K^2} e^{-\alpha K} e^{-K\beta} \geqslant \\
&\geqslant \theta(1)\frac{e^{K(1-\alpha)}}{K^2} n e^{-K(1+\beta)} = \theta(1)\frac{e^{K(1-\alpha)}}{K^2} n e^{-(1-\delta)\ln n} > \theta(n^\delta)
\end{aligned}$$

Skoro jeden składnik sumy jest większy od $\theta(n^\delta)$, więc tym bardziej cała suma $\mathbb{E}(T(K)) > \theta(n^\delta)$.

2. $\alpha = 1$

W tym przypadku $T_i^0 = K/(i-1)$. W pracy [16] zostało pokazane:

$$
\begin{aligned}
\mathbb{E}(T(S)) &= \sum_{\left\{c_i, c_j \mid \sum (l-1)c_l = \sum (j-1)c_j\right\}} \rho(S\alpha_i) f(\overrightarrow{T^0}) g(\vec{c}) = \\
&= \sum_{\left\{c_i, c_j \mid \sum (l-1)c_l = \sum (j-1)c_j\right\}} \theta(1)\frac{n}{K^2} \prod_{i=2}^{t} \frac{1}{\sqrt{T_i^0}} e^{\sum c_l^2/T_l^0} e^{-O(K^3/n^2)} = \\
&= \theta(1)\frac{n}{K^2} e^{-o(K^3/n^2)} \sum_{\left\{c_i, c_j \mid \sum (l-1)c_l = \sum (j-1)c_j\right\}} \prod_{i=2}^{t} \frac{1}{\sqrt{T_i^0}} e^{\sum c_l^2/T_l^0} = \\
&= \theta(1)\frac{n}{K^2} \frac{1}{\sqrt{K}} e^{-O(S^3/n^2)}
\end{aligned}
$$

Dla $K > \omega(n) n^{2/3}$ po podstawieniu do powyższej równości mamy, że

$$\mathbb{E}(T(K)) = o(n^{-2/3} e^{-\omega^3(n)}).$$

Natomiast dla $\mathcal{J} = [n^{2/3}/2\omega(n), n^{2/3}/\omega(n)]$

$$\sum_{S \in \mathcal{J}} \mathbb{E}(T(K) = \theta(\omega(n)^{1.5})$$

3. $\alpha > 1$

Kontynuując (4.17) dla dużych n mamy

$$\frac{n}{K^2} e^{-\left\{\frac{(\alpha-1)^2}{8(s-1)} - \frac{(\alpha-1)K}{2n}\right\}K}$$

Dla $K \leqslant \frac{\beta'(\alpha-1)n}{4(s-1)} = \frac{\overline{K}'_\alpha n}{t}$, gdzie $\beta' < 1$, wyrażenie w nawiasach jest większe niż

$$\frac{(1-\beta')(\alpha-1)^2}{8(s-1)} = \frac{\beta''}{s}.$$

Czyli

$$\mathbb{E}(T(K)) = O(1)\binom{K}{s-2}\frac{n}{K^2} o(e^{-\beta'' S/s}).$$

Zatem ostatecznie dla $\frac{3}{\beta''} s \ln n = \underline{K}'_\alpha s \ln n < K \leqslant \overline{K}'_\alpha \frac{n}{s}$ i $\beta = \beta''/3$

$$\mathbb{E}(T(K)) = O(1) K^{s-4} e^{-\beta K/s} n \frac{1}{n} o(e^{-\beta K/s}) = o(e^{-\beta K/s}).$$

$\square$

Lemat 4.7. Niech $C_k(K)$ oznacza liczbę składowych spójności o indeksie cyklicznym k w hipergrafie $\mathcal{H} \in \mathcal{H}(\leqslant s, n, \underline{p})$ o średnim stopniu wierzchołkowym α, wtedy

$$\mathbb{E}(C_k(K)) \leqslant \frac{O(\frac{K^2}{n})^k}{k!}\mathbb{E}(T(K+k)).$$

Dowód. Jeśli w trakcie procedury konstruowania hiperdrzewa rozpinającego na K wierzchołkach i korzeniu r dodana w kolejnym kroku krawędź miałaby więcej niż jeden wierzchołek wspólny z dołączonymi wcześniej krawędziami, to powstałby co najmniej jeden cykl. Można podobną procedurą spróbować skonstruować z dowolnej składowej S ($|S| = K$) hiperdrzewo bez cykli. Po prostu w momencie gdy dodana krawędź ma kilka wierzchołków wspólnych z krawędziami wcześniej dołączonymi sklonować te wierzchołki tak, aby nowa krawędź miała dokładnie jeden wierzchołek wspólny. Jeśli składowa miała indeks cykliczny k, to w trakcie tak zmodyfikowanej procedury dodamy dokładnie k sklonowanych wierzchołków i ostatecznie otrzymamy hiperdrzewo bez cykli na $K + k$ wierzchołkach.

Z każdego hiperdrzewa na oznaczonych wierzchołkach możemy skonstruować nie więcej niż $\binom{K+k}{k}$ składowych na K oznaczonych wierzchołkach tak, aby dla każdej składowej istniało hiperdrzewo, z którego ona powstaje. Mianowicie ustawione w ciąg wierzchołki hiperdrzewa dzielimy na K rozłącznych niepustych przedziałów na $\binom{K+k-1}{k-1} \leqslant \binom{K+k}{k}$ sposobów i sklejamy wierzchołki z tych samych przedziałów. Powstaje co najwyżej $\binom{K+k}{k}$ składowych o indeksie cyklicznym k. W drugą stronę podczas zmodyfikowanej procedury dla składowej S, jeśli klonowanemu wierzchołkowi i jego klonom damy kolejne indeksy, to powstanie hiperdrzewo generujące S. Czyli

$$C_k(K)\binom{n}{K}^{-1} \leqslant \binom{K+k}{k}T(K+k)\binom{n}{K+k}^{-1},$$

gdyż $C_k(K)\binom{n}{K}^{-1}$ określa liczbę spójnych składowych na ustalonych K wierzchołkach a $T(K+k)\binom{n}{K+k}^{-1}$ jest równe liczbie hiperdrzew bez cykli na ustalonych $K + k$ wierzchołkach.

Prawdopodobieństwo, że konkretna składowa spójności na K wierzchołkach o T_i krawędziach o mocy i wystąpi w $\mathcal{H}(\leqslant s, n, \underline{p})$ jest równe:

$$\prod_i p_i^{T_i}(1-p_i)^{\sum_{j=1}^{i}\binom{K}{j}\binom{n-K}{i-j}-T_i}.$$

Prawdopodobieństwo, że wystąpi izolowane hiperdrzewo bez cykli na ustalonych $K + k$ wierzchołkach wynosi

$$\prod_i p_i^{T_i}(1-p_i)^{\sum_{j=1}^{i}\binom{K+k}{j}\binom{n-K-k}{i-j}-T_i}.$$

Czyli

$$
\begin{aligned}
&\mathbb{E}(C_k(K)) \leqslant \\
&\leqslant \frac{\binom{n}{K}}{\binom{n}{K+k}} \binom{K+k}{k} \mathbb{E}(T(K+k)) \prod_i (1-p_i)^{\sum_{j=1}^{i} \binom{K}{j}\binom{n-K}{i-j} - \binom{K+k}{j}\binom{n-K-k}{i-j}} \leqslant \\
&\leqslant \frac{(n)_K (K+k)!(K+k)_k}{(n)_{K+k} K!k!} \mathbb{E}(T(K+k)) \prod_i (1-p_i)^{\binom{n}{i} - \binom{n-K}{i} - \binom{n}{i} + \binom{n-K-k}{i}} \leqslant \\
&\leqslant \frac{(K+k)_k^2}{(n-K)_k\ k!} \mathbb{E}(T(K+k)) \prod_i (1-p_i)^{-k\binom{n-1}{i-1}} \leqslant \\
&\leqslant \frac{(K+k)_k^2}{(n-K)_k\ k!} e^{\left(k \sum_i \frac{\alpha_i}{i-1}\right)} \mathbb{E}(T(K+k) \leqslant \frac{\left(O\left(K^2/n\right)\right)^k}{k!} \mathbb{E}(T(K+k)).
\end{aligned}
$$

□

Lemat 4.8. Niech Z będzie zmienną losową określającą liczbę cykli w $\mathcal{H}(\leqslant s, n, \underline{p})$ ze średnim stopniem wierzchołkowym $\sum \alpha_i = \alpha \leqslant 1$, wtedy

$$
\mathbb{E}(Z) = \begin{cases} O(\ln n) & \text{dla } \alpha = 1 \\ O(1) & \text{dla } \alpha < 1 \end{cases}
$$

oraz dla dowolnej stałej C istnieje stała K_C taka, że prawdopodobieństwo, że $\mathcal{H}(\leqslant s, n, \underline{p})$ zawiera więcej niż $(K_C \ln n\ )$ cykli jest mniejsze niż $\mathbb{E}(Z) n^{-C}$.

Dowód. W grafie $\mathcal{H}(\leqslant s, n, \underline{p})$ wartość oczekiwana cykli o c krawędziach wynosi :

$$
\binom{n}{c}(c-1)! \left(1 - \prod_i (1-p_i)^{\binom{n-2}{i-2}}\right)^c \leqslant
$$

$$
\leqslant \frac{n}{n-1} \frac{(n-1)^c}{c!}(c-1)! \left(1 - 1 + \sum_i p_i \binom{n-2}{i-2}\right)^c \leqslant \frac{2}{c}(\sum_i \alpha_i) = \frac{2\alpha^c}{c}, \text{czyli}
$$

$$
\mathbb{E}(Z) \leqslant 2 \sum_{c=2}^{n} \frac{\alpha^c}{c},
$$

więc rzeczywiście $\mathbb{E}(Z) = O(\ln n)$ dla $\alpha = 1$ oraz $\mathbb{E}(Z) = O(1)$ dla $\alpha < 1$.

Możemy podzielić wszystkie cykle na $2\mathbb{E}(Z)$ grup takich, że wartość oczekiwana liczby cykli w każdej grupie jest co najwyżej $1/2$. Rozłączne krawędziowo cykle powstają niezależnie. Czyli mamy, że prawdopodobieństwo powstania więcej niż $C \ln n$ cykli rozłącznych w dowolnej grupie jest równe $o(n^{-C})$. Zatem prawdopodobieństwo, że w sumie istnieje $K_c \ln n$ rozłącznych cykli jest równe $o(\mathbb{E}(Z) n^{-C})$. Ponieważ nie rozłącznych cykli jest więc tym bardziej prawdopodobieństwo, że w składowej jest więcej niż $K_c \ln n$ dowolnych cykli jest równe $o(\mathbb{E}(Z) n^{-C})$. □

Dowód twierdzenia 4.1. Dla $\alpha < 1$ z lematu 4.8 z prawdopodobieństwem $1 - o(n^{-C})$ $\mathcal{H}(\leqslant s, n, \underline{p})$ ma co najwyżej $O(\ln n)$ cykli, czyli z lematu 4.7

$$\mathbb{E}(C(K)) \leqslant \sum_{k \leqslant O(\ln n)} \frac{\{O(K^2/n)\}^k}{k!} \mathbb{E}(T(K+k)) + o(n^{-C})$$

Z lematu 4.6 dla $K = C\overline{K}_\alpha t \ln n$ mamy, że $\mathbb{E}(T(K)) = o(n^{-C})$, więc:

$$\mathbb{E}(C(K)) \leqslant \sum_{k \leqslant O(\ln n)} \frac{\{O(K^2/n)\}^k}{k!} o(n^{-(K+k)/(\overline{K}_\alpha t \ln n)}) + o(n^{-C}) = o(n^{-C(1-o(1))})$$

Czyli

$$\Pr\{C(K) > 0\} \leqslant \mathbb{E}(C(K)) = o(n^{-C(1-o(1))}).$$

Ostatecznie dla $K = C\overline{K}_\alpha t \ln n$

$$\Pr\{C(K) = 0\} = 1 - o(n^{-C(1-o(1))}).$$

Natomiast dla $K < \underline{K}_{\delta,\alpha} \ln n$, z lematu 4.6 mamy $\mathbb{E}(C(K)) \geqslant \mathbb{E}(T(K)) \geqslant \theta(n^\delta)$, czyli z lematu 4.4

$$\Pr(C(K) > 0) = 1 - O(n^{-\delta})$$

Dodatkowo z lematu 4.7 dla K takich, że $\mathbb{E}(C(K)) \neq o(1)$ (czyli dla pewnych $K = O(\ln n)$)

$$\mathbb{E}(C(K)_{k'}) \leqslant \frac{\{O(K^2/n)\}^{k'}}{k'!} \mathbb{E}(T(K+k')) \leqslant \frac{\{O(K^2/n)\}^{k'}}{k'!} \frac{n}{K}.$$

Stąd $\sum_{k' \geqslant k+1} \mathbb{E}(C(K)_{k'}) = o(n^{-k+o(1)})$ .Czyli dla dowolnej liczby naturalnej k indeks cykliczny spójnej składowej jest co najwyżej k z prawdopodobieństwem $1 - o(n^{-k+o(1)})$. $\square$

Dowód twierdzenia 4.2. Dla $\alpha = 1$ z lematu 4.8 z prawdopodobieństwem $1 - o(1)$ $\mathcal{H}(\leqslant s, n, \underline{p})$ zawiera $O(\ln n)$ cykli. Czyli z lematu 4.6 dla $n^{1/3} \geqslant \omega(n) > \ln n$

$$\begin{aligned}
&\sum_{\{K | K = \omega(n) n^{2/3}\}} \mathbb{E}(C(K)) \leqslant \\
&\leqslant \sum_{\{K | K = \omega(n) n^{2/3}\}} \sum_{k \leqslant O(\ln n)} \frac{\{O(K^2/(n-K))\}^k}{k!} \mathbb{E}(T(K+k)) + o(1) \leqslant \\
&\leqslant \sum_{\{K | K = \omega(n) n^{2/3}\}} \sum_{k \leqslant O(\ln n)} \frac{\{O(K^2/(n-K))\}^k}{k!} o\left(n^{-2/3} e^{-\omega^3(n)}\right) + o(1) \leqslant \\
&\leqslant \sum_{\{K | K = \omega(n) n^{2/3}\}} \ln n (\ln n!)^{-1} \; o(1) \left(O(1) \frac{\omega^2(n) n^{2/3}}{(n^{1/3} - \omega(n)) e^{\omega^2(n)}} \right)^{\ln n} = o(1)
\end{aligned}$$

Ostatecznie prawdopodobieństwo, że istnieje składowa spójności o więcej niż $\omega(n)n^{2/3}$ wierzchołkach dąży do zera, czyli

$$\Pr\{\forall_{K>\omega(n)n^{2/3}} C(K)=0\} = 1-o(1)$$

Natomiast dla $\mathcal{J} = \{K | n^{2/3}/2 < K\omega(n) < n^{2/3}\}$ dla dowolnego ciągu $\omega(n)$ dążącego do nieskończoności z lematu 4.6

$$\mathbb{E}\left(\sum_{\mathcal{J}} C(K)\right) \geqslant \mathbb{E}\left(\sum_{\mathcal{J}} T(K)\right) = \infty.$$

Zatem z lematu 4.4

$$\Pr\left\{\sum_{\mathcal{J}} C(K) > 0\right\} = 1-o(1)$$

□

Dowód twierdzenia 4.3. Dla $\alpha > 1$ z lematu 4.6

$$\begin{aligned}
\sum_{K=\underline{K}'_\alpha t\ln n}^{\min\{\overline{K}'_\alpha,\beta\}n/t} \mathbb{E}(C(K)) &\leqslant \sum_{K=\underline{K}'_\alpha t\ln n}^{\min\{\overline{K}'_\alpha,\beta\}n/t} \sum_k \frac{\{O(K^2/n)\}^k}{k!}\mathbb{E}(T(K+k)) \leqslant \\
&\leqslant \sum_{K=\underline{K}'_\alpha t\ln n}^{\min\{\overline{K}'_\alpha,\beta\}n/t} e^{K^2/n} o(e^{-\beta K/t}) = \\
&= \sum_{K=\underline{K}'_\alpha t\ln n}^{\min\{\overline{K}'_\alpha,\beta\}n/t} o\left(e^{-K(\beta/t-K/n)}\right) = o(1)
\end{aligned}$$

Czyli z prawdopodobieństwem dążącym do zera $\mathcal{H}(\leqslant s,n,\underline{p})$ nie ma składowej spójności o K wierzchołkach dla $\underline{K}'_\alpha t\ln n \leqslant K \leqslant \min\{\overline{K}'_\alpha,\beta\}n/t$, więc z twierdzenia 4.2 ma prawie na pewno pewną składową o $K \geqslant \min\{\overline{K}'_\alpha,\beta\}n/t$ wierzchołkach. Ostatecznie w hipergrafie $\mathcal{H}(\leqslant s,n,\underline{p})$ największa składowa spójności ma $\theta(n)$ wierzchołków z prawdopodobieństwem dążącym do jeden.

□

5. Losowe Grafy Przecięć i Hipergrafy

Jak już zauważyliśmy graf $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ jest grafem przecięć pewnej rodziny zbiorów. Z każdym wierzchołkiem tego grafu możemy powiązać pewien zbiór podzbiór zbioru $\mathcal{W}$ - zbiór sąsiadów tego wierzchołka w grafie $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$. Zbiór ten jest podzbiorem zbioru $\mathcal{W}$. Zanalizujemy teraz pewne własności hipergrafu, którego krawędzie są wyznaczone przez zbiory sąsiadów wierzchołków z $\mathcal{V}$. Porównamy także strukturę tego hipergrafu losowego do struktury klasycznego hipergrafu losowego $\mathcal{H}(\leqslant s, n, \underline{m})$.

W tym celu przyjrzymy się dokładniej rodzinie zbiorów wyznaczonych przez wierzchołki z $\mathcal{V}$. Z definicji grafu $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ wynika, że każdemu wierzchołkowi $v \in \mathcal{V}$ przypisujemy dokładnie jeden podzbiór zbioru $\mathcal{W}$. Jest to zbiór sąsiadów wierzchołka v w $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$. W szczególności, gdy wierzchołek ma stopień równy zero, podzbiór ten jest zbiorem pustym, natomiast kiedy stopień jest równy jeden podzbiór jest jednoelementowy. Ponieważ wybory zbioru sąsiadów są niezależne, możliwe jest, że jeden podzbiór zbioru $\mathcal{W}$ jest przypisany do więcej niż jednego wierzchołka z $\mathcal{V}$. Ostatecznie graf $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ wyznacza pewną rodzinę indeksowanych wierzchołkami z $\mathcal{V}$ podzbiorów zbioru $\mathcal{W}$, w której niektóre zbiory mogą się powtarzać. W grafie $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ dwa wierzchołki są połączone krawędzią, jeśli odpowiadające im podzbiory z tej rodziny mają niepusty przekrój (przecinają się).

Podsumowując, graf $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$wyznacza pewną indeksowaną rodzinę podzbiorów zbioru $\mathcal{W}$, a dokładniej ciąg podzbiorów, w którym dopuszczamy powtórzenia oraz zbiory puste. Rodzina ta zawiera informacje o strukturze grafu $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$.

Z drugiej strony każda rodzina niepustych podzbiorów pewnego zbioru, w której zbiory się nie powtarzają jest zbiorem krawędzi pewnego hipergrafu. Jak pokażemy, z wielu własności hipergrafów można wyprowadzić wnioski dotyczące grafów $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$. W tym celu zdefiniujemy hipergraf indukowany przez graf $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$.

Definicja 5.1. Niech $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ będzie grafem zdefiniowanym w 1.4. Oznaczmy przez $\mathcal{H}(\mathcal{BG})$ hipergraf $\mathcal{H}(\mathcal{BG}) = \langle \mathcal{W}, E \rangle$ taki, że $\emptyset \neq \mathcal{S} \in E$ wtedy i tylko wtedy, gdy $\mathcal{S} = \Gamma(v)$ dla pewnego wierzchołka $v \in \mathcal{V}$.

Wierzchołki w $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ są połączone krawędzią, gdy odpowiadające im krawędzie w $\mathcal{H}(\mathcal{BG})$ przecinają się. Z hipergrafu $\mathcal{H}(\mathcal{BG})$ możemy uzyskać pewien graf utożsamiając z każdą krawędzią $\mathcal{H}(\mathcal{BG})$ jeden wierzchołek tego grafu i łącząc wierzchołki, gdy odpowiadające im krawędzie się przecinają w $\mathcal{H}(\mathcal{BG})$. Z tego grafu możemy dostać graf $\mathcal{IG}^{aktywny}(m,\mathcal{P}_{(n)})$ przez dodanie wierzchołków izolowanych i sklonowanie niektórych wierzchołków. Oczywiście na graf $\mathcal{IG}^{pasywny}(m,\mathcal{P}_{(n)})$ możemy też spojrzeć "przez" hipergraf na zbiorze wierzchołków $\mathcal{V}$ indukowany przez $\mathcal{BG}_{m,n}(m,\mathcal{P}_{(n)})$. Łatwiej jednak będzie nam analizować własności $\mathcal{IG}^{pasywny}(m,\mathcal{P}_{(n)})$ wykorzystując również hipergraf $\mathcal{H}(\mathcal{BG})$. Zauważmy, że $\mathcal{IG}^{pasywny}(m,\mathcal{P}_{(n)})$ powstaje z $\mathcal{H}(\mathcal{BG})$ poprzez połączenie wszystkich tych par wierzchołków z $\mathcal{W}$, które należą do tej samej krawędzi w $\mathcal{H}(\mathcal{BG})$.

Punktem wyjścia konstrukcji hipergrafu $\mathcal{H}(\mathcal{BG})$ jest graf $\mathcal{BG}_{m,n}(m,\mathcal{P}_{(n)})$. Ten graf dwudzielny powstaje w wyniku eksperymentu losowego, więc hipergraf $\mathcal{H}(\mathcal{BG})$ jest hipergrafem losowym. Aby zanalizować własności hipergrafu $\mathcal{H}(\mathcal{BG})$ warto więc porównać go do pewnego klasycznego modelu hipergrafu losowego. Zauważmy, że z definicji $\mathcal{BG}_{m,n}(m,\mathcal{P}_{(n)})$ i $\mathcal{H}(\mathcal{BG})$ wynika, że wybory krawędzi w $\mathcal{H}(\mathcal{BG})$ są równie prawdopodobne. Wydaje się zatem, że w przypadku grafu $\mathcal{BG}_{m,n}(m,\mathcal{P}_{(n)})$ z rozkładem zdegenerowanym $\mathcal{P}_{(n)}=(P_s)$ hipergraf $\mathcal{H}(\mathcal{BG})$ powinien mieć podobne własności do $\mathcal{H}(s,n,m)$ a w ogólnym przypadku $\mathcal{H}(\mathcal{BG})$ podobne cechy do $\mathcal{H}(\leqslant s,n,\underline{m})$ dla pewnego ciągu $\underline{m}$. Trzeba jednak wspomnieć o tym, że w grafie dwudzielnym $\mathcal{BG}_{m,n}(m,\mathcal{P}_{(n)})$ niektóre wierzchołki $\mathcal{V}$ mogą mieć stopień zero oraz dwa wierzchołki z $\mathcal{V}$ mogą mieć ten sam zbiór sąsiadów. Może się więc zdarzyć, że $\mathcal{H}(\mathcal{BG})$ będzie miał mniej niż m krawędzi. Ważne jest także to, że liczba wierzchołków z $\mathcal{V}$ wybierających dokładnie i sąsiadów jest pewną zmienną losową. Zauważmy, że w $\mathcal{BG}_{m,n}(m,\mathcal{P}_{(n)})$ mogą też występować wierzchołki stopnia jeden i odpowiadające im krawędzie jedoelementowe w $\mathcal{H}(\mathcal{BG})$.

W przypadku grafu $\mathcal{BG}_{m,n}(m,\mathcal{P}_{(n)})$ z rozkładem zdegenerowanym $\mathcal{P}_{(n)}=(P_s)$ (gdy dla pewnego $s\in\{1,2\dots,n\}$, $P_s=1$) każdy wierzchołek z $\mathcal{V}$ ma dokładnie s sąsiadów. Czyli hipergraf $\mathcal{H}(\mathcal{BG})$ jest s-jednorodny. Prawdopodobieństwo, że tak określony hipergraf $\mathcal{H}(\mathcal{BG})$ ma dokładnie m krawędzi jest równe:

$$\Pr\{e(\mathcal{H}(\mathcal{BG}))=m\}=\frac{(N_s)_m}{N_s^m}\leqslant\frac{N_s^m e^{-(m^2/2N_s)}}{N_s^m}=e^{-(m^2/2N_s)}$$

Z powyższej nierówności wynika, że przy $m^2/2N_s\to\infty$ hipergraf $\mathcal{H}(\mathcal{BG})$ asymptotycznie prawie na pewno ma mniej niż m krawędzi. Na przykład dla m rzędu $n\ln n$ oraz $N_2=\binom{n}{2}$, prawdopodobieństwo, że $\mathcal{H}(\mathcal{BG})$ ma dokładnie m krawędzi dąży do zera przy $n\to\infty$. Wiadomo przecież, że im więcej razy wybieramy z powtórzeniami, tym większe prawdopodobieństwo, że pewna krawędź zostanie wybrana kilkakrotnie. Z drugiej strony wiemy także, że dla m małych w stosunku do N_s prawdopodobieństwo powtórzenia wyboru dąży do zera wraz z n.

Natomiast w hipergrafie $\mathcal{H}(\mathcal{BG})$ indukowanym przez $\mathcal{BG}_{m,n}(m,\mathcal{P}_{(n)})$ z rozkładem $\mathcal{P}_{(n)}$ liczba krawędzi o i wierzchołkach (i=2,3,...,s) jest zmienną losową. Czyli, nawet w przypadku, gdy pominiemy krawędzie puste i jednowierzchołkowe, nie można bezpośrednio podać takiego $\underline{m}$, dla którego $\mathcal{H}(\mathcal{BG})$ i $\mathcal{H}(\leqslant s,n,\underline{m})$ są asymptotycznie sobie równe. Pokażemy więc, przy jakich założeniach i z jakim prawdopodobieństwem $\mathcal{H}(\mathcal{BG})$ i $\mathcal{H}(\leqslant s, n, \underline{m})$ mają podobne własności. Zmienna losowa określająca liczbę wierzchołków o stopniu i, dla określonego $i = 2, \ldots, s$, w grafie $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ ma rozkład dwumianowy z prawdopodobieństwem sukcesu P_i, więc spodziewamy się, że $\mathcal{H}(\mathcal{BG})$ i $\mathcal{H}(\leqslant s, n, \underline{m})$ będą asymptotycznie takie same dla ciągu $\underline{m} = (P_2m, \ldots, P_sm)$.

Oszacujemy najpierw liczbę krawędzi w hipergrafie $\mathcal{H}(\mathcal{BG})$. Z [7] wiemy, że:

Fakt 5.2. Dla ustalonego wierzchołka $v \in \mathcal{V}$ oraz dowolnego zbioru $\mathcal{S} \subseteq \mathcal{W}$, takiego, że $|\mathcal{S}| = k$

$$p_k = \Pr\{\Gamma(v) = \mathcal{S}\} = \frac{\Pr\{\Gamma(v) = |\mathcal{S}|\}}{\binom{n}{|\mathcal{S}|}} = \frac{P_k}{\binom{n}{k}}.$$

Fakt 5.3. Dla ustalonego zbioru $\mathcal{S} \subseteq \mathcal{W}$ prawdopodobieństwo, że jest on zbiorem cech pewnego wierzchołka jest równe

$$\Pr\{\exists_{v\in\mathcal{V}}\ \Gamma(v) = \mathcal{S}\} = 1 - \left(1 - \frac{P_k}{\binom{n}{k}}\right)^m.$$

Fakt 5.4. W przypadku gdy $\mathcal{P}_{(n)} = (P_k)$ jest rozkładem jednopunktowym, dla zbioru k-elementowego $\mathcal{S} \subseteq \mathcal{W}$

$$p_k = \Pr\{\Gamma(v) = \mathcal{S}\} = \frac{1}{\binom{n}{k}} \quad \text{i} \quad \Pr\{\exists_{v\in\mathcal{V}}\ \Gamma(v) = \mathcal{S}\} = 1 - \left(1 - \frac{1}{\binom{n}{k}}\right)^m.$$

Te proste fakty wykorzystamy przy oszacowaniu liczby krawędzi w $\mathcal{H}(\mathcal{BG})$. W tym celu wprowadzimy zmienną losową określającą liczbę krawędzi o i wierzchołkach w hipergrafie $\mathcal{H}(\mathcal{BG})$.

Fakt 5.5. Niech $i \in \{2, 3 \ldots, s\}$ Dla ustalonej krawędzi $e_i \subseteq V$ o i wierzchołkach mamy:

$$\Pr\{e_i \in \mathcal{H}(\mathcal{BG})\} = \Pr\{\exists_{v\in\mathcal{V}}\ \Gamma(v) = e_i\} = 1 - \left(1 - \frac{P_i}{\binom{n}{i}}\right)^m.$$

Fakt 5.6. Niech $\mathcal{H}(\mathcal{BG})$ będzie zdefiniowany jak w 5.1 a X_i oznacza zmienną losową określającą liczbę krawędzi o i wierzchołkach w $\mathcal{H}(\mathcal{BG})$, wtedy

$$\begin{aligned}\mathbb{E}\,X_i &= \sum_{e_i\subseteq V} \Pr\{e_i \in E(\mathcal{H}(\mathcal{BG}))\} = N_i\left(1 - \left(1 - \frac{P_i}{N_i}\right)^m\right) = \\ &= N_i - N_i + N_i\binom{m}{1}\frac{P_i}{N_i} - \binom{m}{2}\frac{P_i^2}{N_i} + O\left(\frac{m^3}{N_i^2}\right) = \\ &= P_i m - O\left(\frac{m^2}{N_i}\right).\end{aligned}$$

Pokażemy, że liczba krawędzi o i wierzchołkach w $\mathcal{H}(\mathcal{BG})$ niewiele różni się od jej wartości oczekiwanej. Określimy najpierw, ile krawędzi wyznaczają wierzchołki o stopniu i w $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$, jeśli jest ich wiele w stosunku do liczby krawędzi $N_i = \binom{n}{i}$.

Lemat 5.7. Niech $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$będzie zdefiniowany jak w 1.4, niech także $\mathcal{P}_{(n)} = (P_0, \ldots, P_s)$ oraz X_i określa liczbę krawędzi o i wierzchołkach w $\mathcal{H}(\mathcal{BG})$ dla $i = 2, 3, \ldots, s$. Załóżmy, że dokładnie $m_i = m_i(n)$ wierzchołków z $\mathcal{V}$ ma stopień i, wtedy dla $\omega(n)$, takiego, że $m_i^2/N_i = o(\omega(n))$ mamy:

$$\Pr\{X_i \leqslant m_i - \omega(n)\} = o(1).$$

Dowód. Najpierw oszacujemy prawdopodobieństwo, że wierzchołki o stopniu i wyznaczają dokładnie k krawędzi w $\mathcal{H}(\mathcal{BG})$. Spośród wszystkich wierzchołków o stopniu i wybieramy k, które będą wyznaczać różne krawędzie w $\mathcal{H}(\mathcal{BG})$. Następnie pozostałe wierzchołki będą wybierać jedną z tych k wcześniej wyznaczonych krawędzi. Czyli:

$$\Pr\{X_i = k\} \leqslant \frac{\binom{m_i}{k} k^{m_i - k} (N_i)_k}{N_i^{m_i}}$$

Zatem ponieważ $m_i^2/N_i = o(\omega(n))$, więc dla dużych n:

$$\Pr\{X_i \leqslant m_i(n) - \omega(n)\} = \sum_{t=\omega(n)}^{m_i} \Pr\{X_i = m_i(n) - t\} \leqslant$$

$$\leqslant \sum_{t(n)=\omega(n)}^{m_i} \frac{e^t m_i^t}{t^t} \frac{(m_i - t)^t}{N_i^t} \frac{(N_i)_{m_i - t}}{N_i^{m_i - t}} \leqslant \sum_{t(n)=\omega(n)}^{m_i} \left(\frac{e\, m_i^2}{t\, N_i} \right)^t \leqslant$$

$$\leqslant \sum_{t(n)=\omega(n)}^{m_i} \left(\frac{e\, m_i^2}{\omega(n)\, N_i} \right)^{\omega(n)} \left(\frac{e\, m_i^2}{\omega(n)\, N_i} \right)^{t - \omega(n)} =$$

$$= \left(\frac{e\, m_i^2}{\omega(n)\, N_i} \right)^{\omega(n)} \sum_{t=0}^{m_i - \omega(n)} \left(\frac{e\, m_i^2}{\omega(n)\, N_i} \right)^t = o(1).$$

□

Wniosek 5.8. Jeśli w grafie $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$dokładnie m_i wierzchołków z $\mathcal{V}$ wybrało stopień i, wtedy dla dowolnego $\omega(n)$ takiego, że $m_i^2/N_i = o(\omega(n))$ mamy:

$$\Pr\{m_i - \omega(n) < X_i \leqslant m_i\} = 1 - o(1).$$

Jeśli dodatkowo $m_i^2/N_i = o(1)$, to

$$\Pr\{X_i = m_i\} = \Pr\{m_i - 1 < X_i \leqslant m_i\} = 1 - o(1).$$

Wniosek 5.9. W przypadku grafu $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ z rozkładem zdegenerowanym $\mathcal{P} = (P_s)$, $P_s = 1$ jeśli $m^2/N_s = o(1)$, wtedy

$$\Pr\{e(\mathcal{H}(\mathcal{BG})) = m\} = 1 - o(1).$$

Ostatecznie z powyższego wniosku wynika, że dla rozkładu zdegenerowanego i odpowiedniego m, $\mathcal{H}(\mathcal{BG})$ i $\mathcal{H}(s, n, m)$ asymptotycznie prawie na pewno są takie same. Powyższy wniosek można też otrzymać bezpośrednio szacując prawdopodobieństwo (5).

Określmy teraz dokładniej liczbę wierzchołków wyznaczających krawędzie o i wierzchołkach. Dla grafu $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$, przez Y_i oznaczmy zmienną losową określającą liczbę wierzchołków z $\mathcal{V}$ wybierających dokładnie i wierzchołków (i=0,1,...,s). Oznaczmy także:

$$m_i = \mathbb{E}\, Y_i = mP_i.$$

Lemat 5.10. Niech $\mathcal{H}(\mathcal{BG})$ będzie zdefiniowany jak w 5.1, wtedy dla $\omega_i(n)$ takich, że $\sqrt{m(n)} = o(\omega_i(n))$ mamy:

$$\Pr\{\forall_{0 \leqslant i \leqslant s} |Y_i - m_i| \leqslant \omega_i(n)\} = 1 - o(1)$$

Dowód. Z nierówności Czebyszewa dla dowolnego $0 \leqslant i \leqslant s$:

$$\Pr\{|Y_i - m_i| > \omega_i(n)\} \leqslant \frac{m\, P_i(1 - P_i)}{\omega_i^2(n)} \leqslant \frac{m}{\omega_i^2(n)} = o(1)$$

Czyli

$$\Pr\{\exists_{0 \leqslant i \leqslant s} |Y_i - m_i| > \omega_i(n)\} \leqslant \sum_{0 \leqslant i \leqslant s} \Pr\{|Y_i - m_i| > \omega_i(n)\} = o(1)$$

□

Dla własności hipergrafu $\mathcal{Q}$, która nie zależy od krawędzi pustych i jednoelementowych możemy sformułować lemat.

Lemat 5.11. Niech $\mathcal{H}(\mathcal{BG})$ będzie zdefiniowany jak w 5.1 dla rozkładu $\mathcal{P}_{(n)} = (P_0, \ldots, P_s)$, $m_i = mP_i$, dla każdego $i = 2, 3, \ldots, s$ oraz $\omega_i(n)$ są takie, że $(m_i^2/N_i + \sqrt{m}) = o(\omega_i(n))$.
Jeśli dla dowolnych $|\omega_i'(n)| \leqslant \omega_i(n)$ hipergraf $\mathcal{H}(\leqslant s, n, \underline{m'})$ taki, że $\underline{m'} = (m_2 + \omega_2'(n), \ldots, m_s + \omega_s'(n))$ ma własność $\mathcal{Q}$ z prawdopodobieństwem dążącym do a ($a \in [0, 1]$) przy $n \to \infty$, to $\mathcal{H}(\mathcal{BG})$ ma tę własność z prawdopodobieństwem dążącym do a przy $n \to \infty$.

Dowód. Oznaczmy przez:

$$\underline{m_d} = (m_2 - \omega_2, \ldots, m_s - \omega_s) \text{ oraz } \underline{m_g} = (m_2 + \omega_2, \ldots, m_s + \omega_s)$$

Konsekwencją lematów 5.7 i 5.10 jest:

$$\Pr\{\underline{m_d} \prec \overline{\mathcal{H}(\mathcal{BG})} \prec \underline{m_g}\} = 1 - o(1).$$

Zatem bezpośrednio ze wzoru na prawdopodobieństwo całkowite oraz ponieważ dla każdego $\underline{m_d} \prec \underline{m} \prec \underline{m_g}$ prawdopodobieństwo zdarzenia, że $\mathcal{H}(\leqslant s, n, \underline{m})$ ma własność $\mathcal{Q}$ dąży do a mamy:

$$\begin{aligned}\Pr\{\mathcal{H}(\mathcal{BG}) \in \mathcal{Q}\} &= \sum_{\underline{m}} \Pr\{\mathcal{H}(\leqslant s, n, \underline{m}) \in \mathcal{Q}\} \Pr\{\overline{\mathcal{H}(\mathcal{BG})} = \underline{m}\} = \\ &= \sum_{\underline{m_d} \prec \underline{m} \prec \underline{m_g}} \Pr\{\mathcal{H}(\leqslant s, n, \underline{m}) \in \mathcal{Q}\} \Pr\{\overline{\mathcal{H}(\mathcal{BG})} = \underline{m}\} + o(1) = \\ &= (a + o(1))(1 - o(1)) + o(1) = a + o(1).\end{aligned}$$

$\square$

Jeśli dla określonych $\underline{m}$ znamy prawdopodobieństwo spójności losowych hipergrafów $\mathcal{H}(\leqslant s, n, \underline{m})$, lemat 5.11 może posłużyć do wyznaczenia prawdopodobieństwa spójności hipergrafu $\mathcal{H}(\mathcal{BG})$ przy n dążącym do nieskończoności.

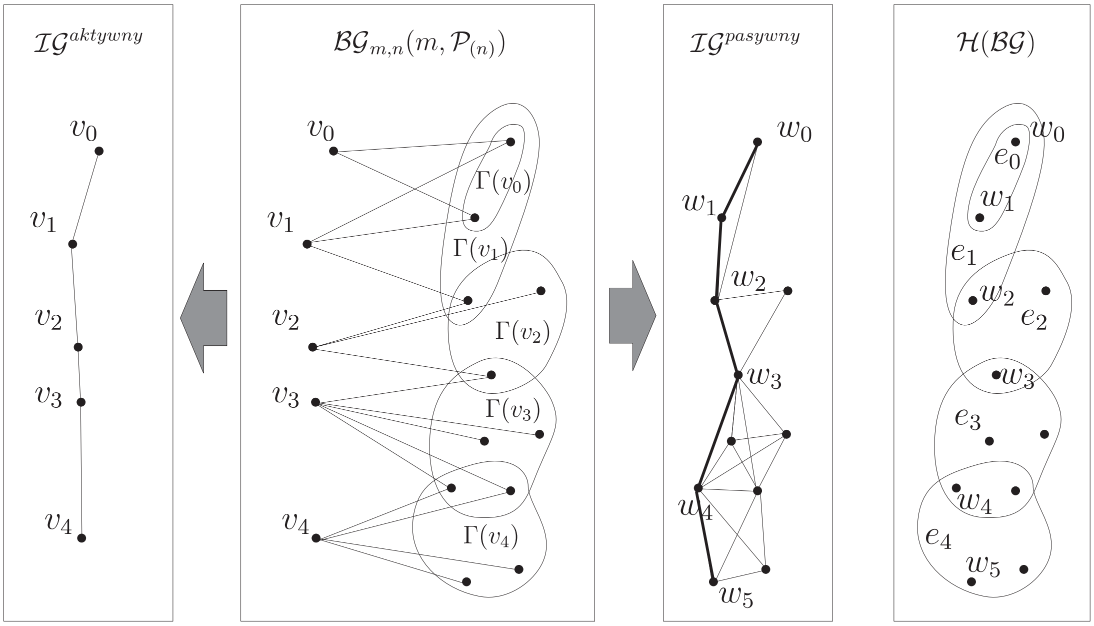
6. Struktura Losowych Grafów Przecięć

Definiując graf $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ łączyliśmy krawędzią obiekty ze zbioru $\mathcal{V}$, jeżeli były do siebie podobne (posiadały co najmniej jedną wspólną własność). Dlatego istotna dla klasyfikacji obiektów z $\mathcal{V}$ jest struktura składowych spójności grafu $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$. Ta z kolei związana jest z istnieniem odpowiednich spacerów między wierzchołkami określonych podzbiorów zbioru wierzchołków. Podobnie w przypadku elementów z $\mathcal{W}$ interesujące są własności związane ze spójnością grafu $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$. Podamy asymptotyczne wyniki dotyczące prawdopodobieństwa spójności grafów przecięć w zależności od struktury grafu $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ oraz pokażemy jak zmienia się struktura składowych spójności tych grafów.

Za narzędzie, dzięki któremu z twierdzeń dotyczących hipergrafów wyciągniemy wnioski dla grafów $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$, posłuży nam hipergraf $\mathcal{H}(\mathcal{BG})$. Udowodniliśmy już w rozdziale piątym, że niektóre własności hipergrafu $\mathcal{H}(\mathcal{BG})$ wynikają z własności hipergrafu $\mathcal{H}(\leqslant s, n, \underline{m})$. Teraz skupimy się na związkach między hipergrafem $\mathcal{H}(\mathcal{BG})$ a grafami przecięć $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$. Ponieważ o spójności grafów i hipergrafów decydują spacery, więc dokładniej zanalizujemy wzajemne zależności między istnieniem hiper-spacerów w $\mathcal{H}(\mathcal{BG})$ i spacerami w grafach przecięć.

Dowolny spacer $(v_0, e_0, v_1, e_1, \ldots, e_k, v_{k+1})$ w grafie $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ wyznacza ciąg podzbiorów zbioru $\mathcal{W}$: $(\Gamma(v_0), \Gamma(v_1), \ldots, \Gamma(v_k))$. W ciągu tym, z definicji grafu $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$, $\Gamma(v_{i-1}) \cap \Gamma(v_i) \neq \emptyset$ dla dowolnego $1 \leqslant i \leqslant k$. Ciąg ten odpowiada ciągowi krawędzi w hipergrafie $\mathcal{H}(\mathcal{BG})$. Ponieważ zbiory sąsiadów kolejnych wierzchołków (kolejne krawędzie) mają niepuste przekroje, więc można wybrać wierzchołki z $\mathcal{W}$ takie, że $w_0 \in \Gamma(v_0)$, $w_{k+1} \in \Gamma(v_k)$ oraz $w_i \in \Gamma(v_{i-1}) \cap \Gamma(v_i)$ dla $1 \leqslant i \leqslant k+1$. Zatem w $\mathcal{H}(\mathcal{BG})$ istnieje hiper-spacer $(w_0, \Gamma(v_0), w_1, \ldots, \Gamma(v_k), w_{k+1})$ o krawędziach wyznaczonych przez ścieżkę w $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$, o pierwszej krawędzi $\Gamma(v_0)$ i ostatniej krawędzi $\Gamma(v_k)$.

W drugą stronę, jeśli dla pary wierzchołków v i v' grafu $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ w hipergrafie $\mathcal{H}(\mathcal{BG})$ istnieje hiper-spacer $(w_0, e_0, w_1 \ldots, e_k, w_{k+1})$ taki, że $e_0 = \Gamma(v)$ i $e_k = \Gamma(v')$, to wyznacza ona ciąg krawędzi $(e_0, e_1, \ldots, e_k)$ hipergrafu $\mathcal{H}(\mathcal{BG})$. W ciągu tym dla każdego $1 \leqslant i \leqslant k$ krawędzie e_{i-1} i e_i mają niepusty przekrój ($w_i \in e_{i-1} \cap e_i \neq \emptyset$). Ponieważ każdej krawędzi hipergrafu $\mathcal{H}(\mathcal{BG})$ odpowiada co najmniej jeden wierzchołek z $\mathcal{V}$, więc spacerowi $(w_0, e_0, \ldots, e_k, w_{k+1})$ możemy przyporządkować ciąg $(v = v_0, v_1 \ldots, v_k = v')$

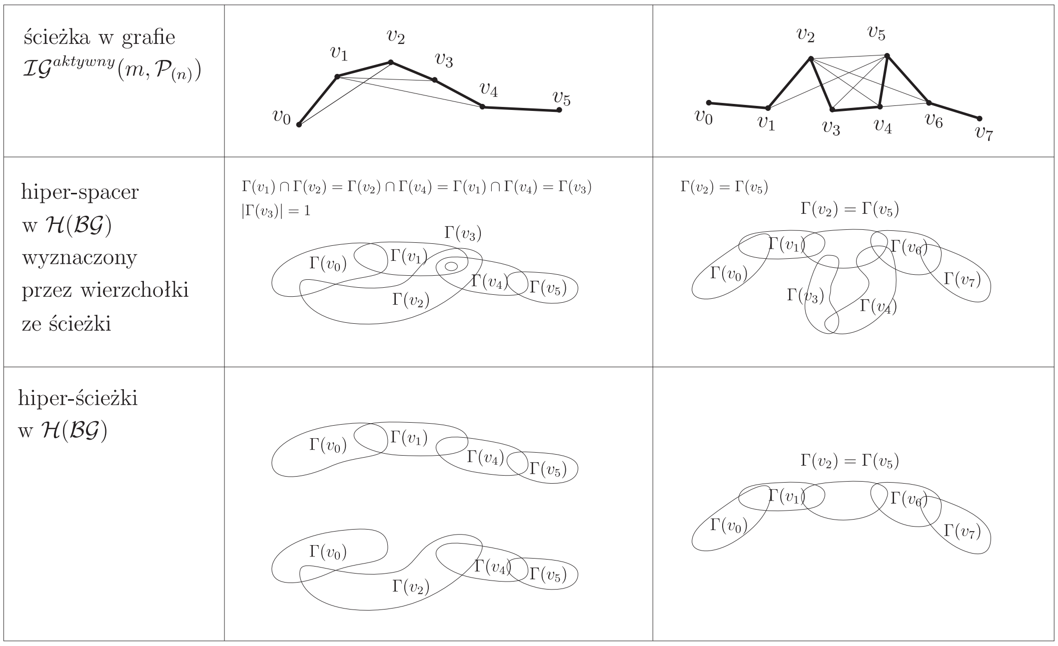


Rysunek 6.1.

wierzchołków w $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ taki, że $\Gamma(v_i) = e_i$ dla $0 \leqslant i \leqslant k$ oraz $\Gamma(v_{i-1}) \cap \Gamma(v_i) = e_{i-1} \cap e_i \neq \emptyset$ dla $1 \leqslant i \leqslant k$. Z definicji grafu $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ mamy zatem, że w ciągu $(v_0, \ldots, v_k)$ wierzchołki v_{i-1} i v_i są połączone krawędzią dla $1 \leqslant i \leqslant k$. Wynika stąd, że dowolnemu hiper-spacerowi w hipergrafie $\mathcal{H}(\mathcal{BG})$ odpowiada pewien spacer w grafie $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$. Co więcej dla dowolnych dwóch wierzchołków z grafu $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$, jeśli w $\mathcal{H}(\mathcal{BG})$ istnieje hiper-spacer między odpowiadającymi im krawędziami, to w $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ istnieje spacer między tymi wierzchołkami.

Podobnie dla grafu $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ - dla dowolnego spaceru w pasywnym losowym grafie przecięć łatwo skonstruować hiper-spacer w hipergrafie $\mathcal{H}(\mathcal{BG})$. Niech $(w_0, e_0, w_1 \ldots, e_k, w_{k+1})$ będzie spacerem w grafie $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$. W spacerze tym, dla $0 \leqslant i \leqslant k$, każda krawędź $\{w_i, w_{i+1}\}$ z grafu $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ odpowiada co najmniej jednemu wspólnemu sąsiadowi wierzchołków w_i i w_{i+1}. Jeśli v_i jest wspólnym sąsiadem wierzchołków w_i i w_{i+1}, to $w_i, w_{i+1} \in \Gamma(v_i)$ dla $0 \leqslant i \leqslant k$. Czyli $(w_0, \Gamma(v_0), w_1, \ldots, \Gamma(v_k), w_{k+1})$ jest hiper-spacerem łączącym wierzchołki w_0 i w_{k+1} w hipergrafie $\mathcal{H}(\mathcal{BG})$.

Rozważmy dowolny hiper-spacer $(w_0, e_0, \ldots, e_k, w_{k+1})$ w $\mathcal{H}(\mathcal{BG})$. Wyznacza on ciąg wierzchołków $(w_0, \ldots, w_{k+1})$ z $\mathcal{W}$ taki, że dla $0 \leqslant i \leqslant k$, $\{w_i, w_{i+1}\} \subseteq e_i$. Z definicji hipergrafu $\mathcal{H}(\mathcal{BG})$ wynika, że w ciągu $(w_0, \ldots, w_{k+1})$ następujące po sobie wierzchołki mają wspólnego sąsiada w $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$odpowiadającego krawędzi, do której te wierzchołki należą. Czyli dla każdego $0 \leqslant i \leqslant k$ para $\{w_i, w_{i+1}\}$



Rysunek 6.2.

jest krawędzią w grafie $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$. Wyznaczony więc przez hiper-spacer ciąg wierzchołków grafu $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ jest zatem spacerem tym w grafie.

Możemy zatem podsumować powyższe rozważania.

Fakt 6.1. Niech $v, v' \in \mathcal{V}$. Wtedy w grafie $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ istnieje (v, v')-spacer wtedy i tylko wtedy gdy istnieje hiper-spacer w $\mathcal{H}(\mathcal{BG})$ o pierwszej krawędzi $\Gamma(v)$ i ostatniej krawędzi $\Gamma(v')$.

Fakt 6.2. Niech $w, w' \in \mathcal{W}$. Wtedy w grafie $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ istnieje (w, w')--spacer wtedy i tylko wtedy gdy w $\mathcal{H}(\mathcal{BG})$ istnieje hiper-spacer łączący wierzchołki w i w'.

Z definicji spójności i faktu 6.2 mamy:

Fakt 6.3. Graf $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ jest spójny wtedy i tylko wtedy, gdy $\mathcal{H}(\mathcal{BG})$ jest spójny.

Fakt 6.4. Dowolny podzbiór $S \subseteq \mathcal{W}$ jest składową spójności hipergrafu $\mathcal{H}(\mathcal{BG})$ wtedy i tylko wtedy, gdy jest składową spójności grafu $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$.

Zauważmy, że $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{H}(\mathcal{BG})$ mają te same zbiory wierzchołków izolowanych. W przypadku grafu $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ wierzchołkami izolowanymi są te wierzchołki z $\mathcal{V}$, które mają stopień zero w $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ oraz te, które odpowiadają krawędziom izolowanym w $\mathcal{H}(\mathcal{BG})$ (o ile nie są to krawędzie odpowiadające więcej niż dwóm wierzchołkom z $\mathcal{V}$).

Graf $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ można uzyskać z hipergrafu $\mathcal{H}(\mathcal{BG})$ rozpatrując tylko krawędzie i ich przecięcia, dlatego na jego spójność nie wpływają wierzchołki izolowane w $\mathcal{H}(\mathcal{BG})$, które nie są krawędziami. Istotne znaczenie ma jednak struktura składowych spójności hipergrafu $\mathcal{H}(\mathcal{BG})$. Zatem z faktu 6.1:

Fakt 6.5. $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ jest spójny wtedy i tylko wtedy gdy w grafie $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ w zbiorze $\mathcal{V}$ nie ma wierzchołków o stopniu zero oraz hipergraf $\mathcal{H}(\mathcal{BG})$ poza największą składową spójności posiada tylko wierzchołki izolowane nie będące krawędziami jednoelementowymi.

Fakt 6.6. Dowolnej składowej spójności S hipergrafu $\mathcal{H}(\mathcal{BG})$ takiej, że $|S| > 1$, odpowiada dokładnie jedna składowa spójności grafu $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$, w której zawarte są wierzchołki odpowiadające krawędziom ze składowej S.

Twierdzenie 6.7. *Niech $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})i$ będzie zdefiniowany jak w 1.4 oraz $\mathcal{P}_{(n)} = (P_0, P_1, \ldots, P_s)$, wtedy:*

$$\lim_{n\to\infty} \Pr\{\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)}) \textit{ jest spójny}\} =$$
$$= \begin{cases} 0, & \textit{gdy } \lim_{n\to\infty} \frac{\left(\sum_{i=2}^{s} i\, P_i\right) m}{n} - \ln n = -\infty \\ e^{-e^{-c}}, & \textit{gdy } \lim_{n\to\infty} \frac{\left(\sum_{i=2}^{s} i\, P_i\right) m}{n} - \ln n = c \\ 1, & \textit{gdy } \lim_{n\to\infty} \frac{\left(\sum_{i=2}^{s} i\, P_i\right) m}{n} - \ln n = \infty \end{cases}$$

Dowód. Oznaczmy przez $m_i = m\, P_i$ oraz załóżmy, że $m = O(n \ln n)$, wtedy jeśli podstawimy za $\omega(n) = \sqrt{n} \ln n$, mamy: $\sqrt{m} + m^2/N_2 = O(\sqrt{n \ln n}) + O(\ln^2 n) = o(\omega(n))$ oraz $\omega(n) = o(n)$. Zatem twierdzenie wynika bezpośrednio z 5.11, 3.16, 3.17 i 6.3. Zauważmy, że dla $C > \sum_{i=2}^{s} i\, P_i$, jeśli $m = Cn \ln n$, to $\mathcal{H}(\mathcal{BG})$ jest spójny z prawdopodobieństwem dążącym do 1, czyli dla większych m, z monotoniczności własności spójności hipergraf $\mathcal{H}(\mathcal{BG})$ jest spójny asymptotycznie prawie na pewno. □

Wniosek 6.8. Niech $\mathcal{P}_{(n)} = (P_s)$ będzie rozkładem jednopunktowym, dla pewnego $s \geqslant 2$, $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ będzie grafem zdefiniowanym w 1.4, wtedy:

(i) jeśli $m = \frac{n(\ln n + c + o(1))}{s}$, to $\mathcal{IG}_n^{pasywny}(m, \mathcal{P}_{(n)})$ jest spójny z prawdopodobieństwem dążącym do $e^{-e^{-c}}$;

(ii) jeśli $m = \frac{n(\ln n + \omega(n))}{s}$, gdzie $\omega(n) \xrightarrow[n\to\infty]{} \infty$, to $\mathcal{IG}_n^{pasywny}(m, \mathcal{P}_{(n)})$ jest spójny z prawdopodobieństwem dążącym do 1;

(iii) jeśli $m = \frac{n(\ln n + \omega(n))}{s}$, gdzie $\omega(n) \xrightarrow[n\to\infty]{} -\infty$, to $\mathcal{IG}_n^{pasywny}(m, \mathcal{P}_{(n)})$ jest spójny z prawdopodobieństwem dążącym do 0.

Jeśli w grafie $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ żaden z wierzchołków z $\mathcal{V}$ nie ma stopnia 0, to spójność hipergrafu $\mathcal{H}(\mathcal{BG})$ implikuje spójność grafu $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$. Jednakże nie zawsze, gdy $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ jest spójny, to $\mathcal{H}(\mathcal{BG})$ jest spójny. Zauważmy także, że własność posiadania dokładnie jednej składowej spójności nie jest własnością monotoniczną. Nie podamy zatem bezpośrednio analogicznych do powyższych twierdzeń twierdzenia dotyczącego spójności grafu $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$. Możemy tylko sformułować proste wnioski.

Wniosek 6.9. Niech $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ będzie zdefiniowany jak w 1.4 oraz $\mathcal{P}_{(n)} = (P_0, P_1, \ldots, P_s)$, wtedy jeśli $\lim_{n\to\infty} \frac{(\sum_{i=2}^{s} i\, P_i)m}{n} - \ln n = \infty$ oraz $P_0 = 0$, to

$$\lim_{n\to\infty} \Pr\{\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)}) \text{ jest spójny}\} = 1.$$

Wniosek 6.10. Niech $\mathcal{P}_{(n)} = (P_s)$ będzie rozkładem jednopunktowym, dla pewnego $s \geqslant 2$, $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ będzie grafem zdefiniowanym w 1.4, wtedy jeśli $m = \frac{n(\ln n + \omega(n))}{s}$, gdzie $\omega(n) \xrightarrow[n\to\infty]{} \infty$, to

$$\lim_{n\to\infty} \Pr\{\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)}) \text{ jest spójny}\} = 1.$$

Zauważmy, że wyniki dotyczące spójności hipergrafów losowych dają odpowiedź na pytanie dotyczące spójności $\mathcal{H}(\mathcal{BG})$ w przypadku, gdy stopnie wierzchołków z $\mathcal{V}$ są ograniczone przez pewną stałą. W przypadku ogólnym umiemy bezpośrednio wyznaczyć liczbę wierzchołków izolowanych ([7]). Możemy jednak podejrzewać, że podobnie jak w przypadku stopni ograniczonych istnienie wierzchołków izolowanych może mieć fundamentalne znaczenie dla spójności grafów $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$. Oszacowanie liczby wierzchołków izolowanych daje także częściową odpowiedź na pytanie dotyczące spójności grafu $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$.

Lemat 6.11. Dla dowolnego $\mathcal{S}$, $\mathcal{S} \subseteq \mathcal{W}, |\mathcal{S}| = k \geqslant 1$, oraz dowolnego $\mathcal{A}$, $\mathcal{A} \subseteq \mathcal{V}$ i $|\mathcal{A}| = t \geqslant 1$, mamy

$$\Pr\left(\Gamma(\mathcal{A}) \subseteq \mathcal{W} \setminus \mathcal{S}\right) = \left(\frac{\mathbb{E}_k(n - X^*)}{(n)_k}\right)^t,$$

gdzie X^* oznacza zmienną losową o rozkładzie $\mathcal{P}_{(n)} = (P_0, P_1, \ldots, P_n)$, czyli zmienną losową określającą stopień wierzchołka $v \in \mathcal{V}$ w grafie $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$.

Dowód. Rozważmy zmienną losową $Y^* = n - X^*$. Tak zdefiniowana zmienna losowa Y^* ma rozkład $\mathcal{P}'_{(n)} = (P_n, P_{n-1}, \ldots, P_0)$. Prawdopodobieństwo, że w grafie $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$dany wierzchołek nie ma sąsiadów w zbiorze $\mathcal{S}$ jest równy prawdopodobieństwu, że zbiór sąsiadów tego samego wierzchołka w grafie $\mathcal{BG}_{m,n}(m, \mathcal{P'}_{(n)})$ zawiera zbiór $\mathcal{S}$, stąd dla

$$\Pr\left(\Gamma_{\mathcal{P}}(\mathcal{A}) \subseteq \mathcal{W} \setminus \mathcal{S}\right) = \Pr\left(\mathcal{S} \subseteq \Gamma_{\mathcal{P'}}(\mathcal{A})\right) =$$

$$\prod_{v \in \mathcal{A}} \left(\sum_{l=0}^{n-k} \sum_{\substack{\mathcal{L} \subseteq \mathcal{W}\setminus\mathcal{S} \\ |\mathcal{L}|=l}} \Pr(\Gamma(v) = \mathcal{S} \cup \mathcal{L}) \right) = \left(\sum_{l=0}^{n-k} \binom{n-k}{l} \frac{P_{n-(l+k)}}{\binom{n}{l+k}} \right)^t =$$

$$= \left(\sum_{r=k}^{n} \binom{n-k}{r-k} \frac{P_{n-r}}{\binom{n}{r}} \right)^t = \left(\frac{1}{(n)_k} \sum_{r=k}^{n} (r)_k P_{n-r} \right)^t = \left(\frac{\mathbb{E}_k(n - X^*)}{(n)_k} \right)^t.$$

□

Twierdzenie 6.12. *Niech μ^a oznacza liczbę wierzchołków izolowanych w grafie $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$, wtedy*

$$\mathbb{E}(\mu^a)_k = (m)_k \sum_{j=0}^{n} \sum_{(j_1, j_2, \ldots, j_k)} \frac{(n)_j}{(n)_{j_1} \ldots (n)_{j_k}} P_{j_1} \ldots P_{j_k} \left[\frac{\mathbb{E}_j(n - X^*)}{(n)_j} \right]^{m-k},$$

gdzie druga suma jest po wszystkich ciągach $(j_1, j_2, \ldots, j_k)$ takich, że $0 \leqslant j_i \leqslant n$ dla $i = 1, 2, \ldots, k$ oraz $j_1 + j_2 + \ldots + j_k = j$. Przyjmujemy dodatkowo, że $(m)_0 = 1$ i $\mathbb{E}(\cdot)_0 = 1$.

Dowód. Niech $\mathcal{S} = \{v_1, v_2, \ldots, v_k\}$ będzie dowolnym k-elementowym podzbiorem zbioru $\mathcal{V}$, wtedy korzystając z lematu 6.11:

$$\Pr(\text{wszystkie wierzchołki w } \mathcal{S} \text{ są izolowane}) =$$

$$= \Pr(\Gamma(v_1), \ldots, \Gamma(v_k) \text{ są parami rozłączne i } \Gamma(\mathcal{V} \setminus \mathcal{S}) \subseteq \mathcal{W} \setminus \bigcup_{i=1}^{k} \Gamma(v_i)) =$$

$$= \sum_{j=0}^{n} \sum_{\substack{(j_1, j_2, \ldots, j_k) \\ j_1 + j_2 + \ldots + j_k = j}} \binom{n}{j_1, j_2, \ldots, j_k, n-j} \frac{P_{j_1}}{\binom{n}{j_1}} \cdots \frac{P_{j_k}}{\binom{n}{j_k}} \left[\frac{\mathbb{E}(n - X^*)_j}{(n)_j} \right]^{m-k} =$$

$$= \sum_{j=0}^{n} \sum_{(j_1, j_2, \ldots, j_k)} \frac{(n)_j}{(n)_{j_1} \ldots (n)_{j_k}} P_{j_1} \ldots P_{j_k} \left[\frac{\mathbb{E}_j(n - X^*)}{(n)_j} \right]^{m-k}$$

Stąd jeśli zdefiniujemy zmienną losową μ^a jak w (3.3) oraz wykorzystamy równość (3.1) otrzymamy:

$$\mathbb{E}(\mu^a)_k = (m)_k \sum_{j=0}^{n} \sum_{(j_1, j_2, \ldots, j_k)} \frac{(n)_j}{(n)_{j_1} \ldots (n)_{j_k}} P_{j_1} \ldots P_{j_k} \left[\frac{\mathbb{E}_j(n - X^*)}{(n)_j} \right]^{m-k},$$

□

Wniosek 6.13. Niech $\mathcal{P}_{(n)} = (P_s(n))$ będzie rozkładem jednopunktowym oraz $ms^2/n - \ln m \to c$, wtedy rozkład zmiennej losowej μ^a w grafie $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ ma rozkład graniczny (zbieżność według dystrybuanty) zmiennej losowej o rozkładzie Poissona z parametrem $\lambda = e^{-c}$.

Dowód. Dla $ms^2/n - \ln m \to c$ i ustalonego k, z 6.12, ponieważ

$$(n)_k = n^k \exp(-k^2/2n - k^3/n^2 - O(k^4/n^3 + k/n))$$

mamy:

$$\begin{aligned}\mathbb{E}(\mu^a)_k &= (m)_k \sum_{j=0}^{n} \sum_{(j_1, j_2, \ldots, j_k)} \frac{(n)_j}{(n)_{j_1} \ldots (n)_{j_k}} P_{j_1} \ldots P_{j_k} \left[\frac{\mathbb{E}_j(n - X^*)}{(n)_j}\right]^{m-k} = \\ &= (m)_k \frac{(n)_{ks}}{(n)_s^k} \left[\frac{(n-s)_{ks}}{(n)_{ks}}\right]^{m-k} = (1 - o(1)) m^k \left(1 - \frac{s}{n}\right)^{ksm} e^{-O(mk^2 s^3/n(n-s))}\end{aligned}$$

□

Rozpatrzmy teraz graf $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ ze zdegenerowanym rozkładem $\mathcal{P}_{(n)} = (P_s(n))$, $m = \frac{n(\ln n + \omega(n))}{s}$, gdzie s jest pewną stałą, $\omega(n) \xrightarrow[n\to\infty]{} -\infty$ oraz $\omega(n) = o(\ln n)$. Przy tych założeniach asymptotycznie prawie na pewno graf przecięć $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ nie jest spójny. Z drugiej strony dla $s \geqslant 2$:

$$\lim_{n\to\infty} (ms^2/n - \ln m) = \lim_{n\to\infty} (s \ln n + s\omega(n) - \ln n - \ln(\ln n + \omega(n)) + \ln s) = \infty,$$

Czyli z 6.13 mamy, że

$$\begin{aligned}&\lim_{n\to\infty} \Pr\{\text{W } \mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)}) \text{ istnieje wierzchołek izolowany}\} = \\ &\qquad = \lim_{n\to\infty} \Pr\{\mu^a > 0\} = 1 - \lim_{n\to\infty} \Pr\{\mu^a = 0\} = 0\end{aligned}$$

Analogicznie rozumując możemy pokazać, że asymptotycznie prawie na pewno graf $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ ma wierzchołki izolowane dla $m = \frac{n(\ln n + \omega(n))}{s^2}$. Dlatego dla takich m mamy pewność, że w graf $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ asymptotycznie prawie na pewno jest niespójny.

Wniosek 6.14. Niech $\mathcal{P}_{(n)} = (P_s)$ będzie rozkładem jednopunktowym, dla pewnego $s \geqslant 2$, $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ będzie grafem zdefiniowanym w 1.4, wtedy jeśli $m = \frac{n(\ln n + \omega(n))}{s^2}$, gdzie $\omega(n) \xrightarrow[n\to\infty]{} -\infty$, to

$$\lim_{n\to\infty} \Pr\{\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)}) \text{ jest spójny}\} = 0.$$

Oszacujemy teraz liczbę wierzchołków izolowanych w grafie $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ w przypadku dowolnego rozkładu $\mathcal{P}_{(n)}$ w $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$.

Twierdzenie 6.15. *Niech μ^p oznacza liczbę wierzchołków izolowanych w grafie $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$, wtedy*

$$\mathbb{E}(\mu^p)_k = (m)_k \left(\frac{k}{m}P_1 + \frac{\mathbb{E}(n-X^*)_k}{(n)_k}\right)^m$$

Dowód. Niech $\mathcal{U}$ będzie pewnym k-elementowym podzbiorem zbioru $\mathcal{W}$.

$$\begin{aligned}
&\Pr(\text{wierzchołki ze zbioru } \mathcal{U} \text{ są izolowane}) = \\
&= \Pr\left\{\bigcap_{v\in\mathcal{V}}\left\{\bigcup_{u\in\mathcal{U}}\{\Gamma(v)=u\}\cup\{\Gamma(v)\subseteq\mathcal{W}\setminus\mathcal{U}\} = \right\}\right\} = \\
&= \prod_{v\in\mathcal{V}}\left\{\sum_{u\in\mathcal{U}}\Pr(\Gamma(v)=u) + \Pr(\Gamma(v)\subseteq\mathcal{W}\setminus\mathcal{U})\right\} = \\
&= \left(\frac{k}{m}P_1 + \frac{\mathbb{E}(n-X^*)_k}{(n)_k}\right)^m, \text{ czyli z (3.1)} \\
&\mathbb{E}(\mu^p)_k = (m)_k \left(\frac{k}{m}P_1 + \frac{\mathbb{E}(n-X^*)_k}{(n)_k}\right)^m
\end{aligned}$$

□

Wniosek 6.16. Niech $\mathcal{P}_{(n)} = (P_s)$ będzie rozkładem jednopunktowym oraz $ms/n - \ln n \to c$, wtedy rozkład zmiennej losowej μ^a w grafie $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ ma rozkład graniczny (zbieżność według dystrybuanty) zmiennej losowej o rozkładzie Poissona z parametrem $\lambda = e^{-c}$.

Dowód. Przy $ms/n - \ln n \to c$, ponieważ $(n)_k = n^k \exp(-k^2/2n - k^3/n^2 - O(k^4/n^3 + k/n))$

$$\begin{aligned}
\mathbb{E}(\mu^p)_k &= (m)_k \left(\frac{\mathbb{E}(n-X^*)_k}{(n)_k}\right)^m = (m)_k \left(\frac{(n-s)_k}{(n)_k}\right)^m = \\
&= (m)_k \left(1-\frac{s}{n}\right)^{km} e^{-O\left(\frac{k^2 sm}{n(n-s)}\right)} = \\
&= (1-o(1))m^k \left(1-\frac{s}{n}\right)^{km} \to e^{-ck}
\end{aligned}$$

Zbieżność do rozkładu do rozkładu Poissona wynika bezpośrednio z metody momentów (wniosek 3.3). □

Oczywiście dla s stałego wynik jest identyczny z tym otrzymanym w twierdzeniu 3.4.

Jak wspomnieliśmy wcześniej oprócz spójności grafów $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ interesują nas także własności związane z ich składowymi spójności. Jeśli będziemy znać strukturę największych składowych spójności, będziemy mogli określić, na jak duże grupy dzielą się obiekty i własności podobne w grafach $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$.

O wiele łatwiej niż w przypadku spójności jest sformułować wnioski dotyczące największych składowych spójności w aktywnym losowym grafie przecięć. Możemy sformułować wnioski z twierdzeń 4.1, 4.2 i 4.3.

Z twierdzeń dotyczących przejścia fazowego wynika, że gdy średni stopień wierzchołkowy w hipergrafie $\mathcal{H}(\leqslant s, n, \underline{p})$ przekracza jedynkę, to wielkość największej składowej spójności gwałtownie rośnie. Zauważmy, że średni stopień wierzchołkowy hipergrafu $\mathcal{H}(\mathcal{BG})$ jest równy:

$$\sum_{i=2}^{s}(i-1)\frac{iP_i m}{n} = \alpha$$

Uwaga 6.17. Dla hipergrafów o średnim stopniu $\alpha = 1$ nie możemy zakładać, że średni stopień wierzchołkowy różni się od jedynki. Wynika stąd, że wnioski dotyczące składowych spójności w tych hipergrafach będziemy mogli wyprowadzić tylko dla grafów z rozkładem zdegenerowanym, w których wierzchołek z $\mathcal{V}$ ma co najmniej trzech sąsiadów. W tym przypadku możemy skorzystać z wniosków 2.11 i 5.9.

Twierdzenie 6.18. *Niech $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$będzie zdefiniowany jak w 1.4,$\mathcal{P}_{(n)} = (P_0, P_1, \ldots, P_s)$ oraz dla pewnej stałej $\alpha > 0$ mamy:*

$$m = \alpha\frac{n}{\sum_{i=2}^{s}(i-1)iP_i}, \text{ wtedy}$$

(i) *jeśli $\alpha < 1$ to asymptotycznie prawie na pewno największa składowa spójności w grafie w grafie $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ jest wielkości $\theta(s \ln n)$;*
(ii) *jeśli $\alpha > 1$ to asymptotycznie prawie na pewno największa składowa spójności w grafie w grafie $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ jest wielkości $\theta(n/s)$.*

Dowód. Własności "największa składowa zawiera co najmniej k wierzchołków" oraz "największa składowa zawiera co najwyżej k wierzchołków" są własnościami monotonicznymi, dlatego możemy korzystać z twierdzeń dotyczących asymptotycznej równoważności modeli hipergrafów losowych.

Niech zatem $\underline{p}$ spełnia równość:

$$\sum_{i=2}^{s}(i-1)\binom{n-1}{i-1}p_i = \alpha$$

Podstawmy za $m_i = \binom{n}{i}p_i = N_i p_i$ dla $i = 2, 3, \ldots, s$, wtedy

$$\sum_{i=2}^{s}(i-1)\frac{im_i}{n} = \alpha.$$

Czyli $m_i = O(n)$, zatem dla dowolnego $p_i' = \frac{m_i}{N_i} + O\left(\sqrt{\frac{m_i(N_i - m_i))}{N_i^3}}\right) = \frac{m_i}{N_i} + O(\frac{\sqrt{m_i}}{N_i})$ mamy:

$$\sum_{i=2}^{s}(i-1)\binom{n-1}{i-1}p_i' = \alpha + o(1), \text{ dla } i = 2, 3, \dots, s$$

Zauważmy, że jeśli $0 < \alpha < 1$ oraz $\alpha'(n) = \alpha + o(1)$, to istnieje ϵ i takie $n_0(\epsilon)$, że dla każdego $n \geqslant n_0$ mamy $0 < \alpha - \epsilon < \alpha'(n) < \alpha + \epsilon < 1$ (podobnie w przypadku $\alpha > 1$). Czyli z 2.8 mamy, że jeśli pewna własność monotoniczna zachodzi asymptotycznie prawie na pewno dla dowolnego hipergrafu $\mathcal{H}(\leqslant s, n, \underline{p})$ o średnim stopniu α dla dowolnego $0 < \alpha < 1$, to zachodzi asymptotycznie prawie na pewno dla dowolnego $\mathcal{H}(\leqslant s, n, \underline{m})$ o średnim stopniu α. Podobnie dla hipergrafu $\mathcal{H}(\leqslant s, n, \underline{m})$ takiego, że

$$\sum_{i=2}^{s}(i-1)\frac{im_i}{n} = \alpha$$

mamy $m_i = O(n)$ i $\sum_{i=2}^{s} m_i = m = O(n)$, czyli oczywiście dla $i = 2, 3 \dots, s$ mamy $\frac{m_i^2}{N_i} + \sqrt{m} = o(m^{2/3})$ oraz

$$\sum_{i=2}^{s}(i-1)\left(\frac{im_i}{n} + \frac{im_i^{2/3}}{n}\right) = \alpha + o(1).$$

Zatem z twierdzenia 2.8 i 5.11 jeśli własności "największa składowa zawiera co najmniej k wierzchołków" oraz "największa składowa zawiera co najwyżej k wierzchołków" zachodzą dla grafu $\mathcal{H}(\leqslant s, n, \underline{p})$ o średnim stopniu $\alpha < 1$ ($\alpha > 1$), to zachodzą dla hipergrafu $\mathcal{H}(\mathcal{BG})$ o średnim stopniu $\alpha < 1$ ($\alpha > 1$). Zatem jeśli dla $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ z rozkładem $\mathcal{P}_{(n)} = (P_0, P_1, \dots, P_s)$ $\mathcal{H}(\mathcal{BG})$ ma średni stopień $\alpha < 1$ ($\alpha > 1$),to twierdzenie dla grafu $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ wynika z 6.4, 4.1 i 4.3. □

Twierdzenie 6.19. *Niech $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$będzie zdefiniowany jak w 1.4,$\mathcal{P}_{(n)} = (P_0, P_1, \dots, P_s)$ oraz dla pewnej stałej $\alpha > 0$ mamy:*

$$m = \alpha\frac{n}{\sum_{i=2}^{s}(i-1)iP_i}, \text{ wtedy}$$

(i) *jeśli $\alpha < 1$, to asymptotycznie prawie na pewno największa składowa spójności w grafie $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ jest wielkości $\theta(\ln n)$;*
(ii) *jeśli $\alpha > 1$, to asymptotycznie prawie na pewno największa składowa spójności w grafie $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ jest wielkości $\theta(n)$.*

Dowód. Dowolnej składowej spójności w grafie $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ odpowiada pewna składowa spójności w $\mathcal{H}(\mathcal{BG})$.

Rozpatrzmy przypadek $\alpha < 1$. Niech S ($O(|S|) = \ln n$) będzie składową spójności o indeksie cyklicznym k w $\mathcal{H}(\mathcal{BG})$. Odpowiada jej dokładnie jedna składowa

S^a w grafie $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ generowana przez krawędzie ze składowej S. Udowodnimy, że $|S^a| = \theta(|S|)$ asymptycznie prawie na pewno. Musimy policzyć ile krawędzi co najmniej dwuelementowych zawiera S, ile z nich odpowiada kilku wierzchołkom w $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ oraz ile krawędzi jednoelementowych ma S, gdyż wielkości te będą decydować o mocy składowej S^a.

$\sum_{i=2}^{s}(i-1)T_i = |S|+k-1$, gdzie T_i oznacza liczbę krawędzi o i wierzchołkach w S oraz $s = O(1)$. Stąd liczba krawędzi o co najmniej dwóch wierzchołkach w składowej S jest równa $\sum_{i=2}^{s} T_i = \theta(|S| + k)$. Dla pewnej stałej C indeks cykliczny składowych w $\mathcal{H}(\mathcal{BG})$ jest mniejszy od C z prawdopodobieństwem dążącym do 1(lemat 4.8), czyli asymptotycznie prawie na pewno $\sum_{i=2}^{s} T_i = \theta(|S|)$. Z 5.7 skoro $m^2/N_i = O(1)$, więc liczba krawędzi wielokrotnych o co najmniej dwóch wierzchołkach w rodzinie wyznaczonej przez wierzchołki z $\mathcal{V}$ asymptotycznie prawie na pewno wynosi co najwyżej $\omega(n)$ dla dowolnej funkcji rosnącej $\omega(n)$. Możemy zatem założyć, że $\omega(n) = o(\ln n)$. Wtedy krawędzie, które odpowiadają więcej niż jednemu wierzchołkowi w $\mathcal{V}$ mogą wygenerować co najwyżej $o(\ln n)$ dodatkowych wierzchołków w składowej S^a. W składowej S są jeszcze krawędzie jednoelementowe, które odpowiadają pewnym wierzchołkom z S^a. Średnio jest ich $mP_i|S|/n = O(|S|)$ i powstają one niezależnie, zatem jeśli przez X oznaczymy zmienną losową określającą liczbę krawędzi jednoelementowych w S, to $\mathbb{E}X = O(|S|)$ oraz $Var(X) \leqslant \mathbb{E}X$. Z nierówności Czebyszewa mamy zatem, że liczba tych wierzchołków jest równa $O(|S|)$ asymptotycznie prawie na pewno.

Ostatecznie jeśli $\ln n = O(|S|)$, to $|S^a| = \theta(\sum_{i=1}^{s} T_i) = \theta(|S|)$. Więc liczba wierzchołków odpowiadających sobie składowych spójności w $\mathcal{H}(\mathcal{BG})$ i w $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ rzędowo powinna być taka sama (dla $\alpha \leqslant 1$). Czyli (i) wynika z dowodu twierdzenia 6.18.

Dla danego $\alpha > 1$ niech S będzie największą składową spójności w $\mathcal{H}(\mathcal{BG})$. Dowolna krawędź grafu $\mathcal{H}(\mathcal{BG})$ ma co najwyżej s wierzchołków, więc odpowiadająca składowej S w hipergrafie $\mathcal{H}(\mathcal{BG})$ składowa spójności w grafie $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ ma co najmniej $|S|/s$ wierzchołków. Ponieważ z dowodu twierdzenia 6.18 $S = \theta(n)$ asymptotycznie prawie na pewno oraz $m = O(n)$, więc oczywiście największa składowa w $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ ma moc $\theta(n)$ z prawdopodobieństwem dążącym do 1 przy $n \to \infty$. □

Twierdzenie 6.20. *Niech $\mathcal{P}_{(n)} = (P_s)$ będzie rozkładem jednopunktowym, dla pewnego $s \geqslant 3$, $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ będzie grafem zdefiniowanym w 1.4, $m = \alpha\frac{n}{s(s-1)}$ dla pewnej stałej $\alpha > 0$, wtedy:*

(i) *jeśli $\alpha < 1$, to największa składowa spójności w grafie $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ ma $\theta(s \ln n)$ wierzchołków;*

(ii) *jeśli $\alpha = 1$, to największa składowa spójności grafu $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ ma mniej niż $\omega_1(n)n^{2/3}$ i więcej niż $n^{2/3}/\omega(n)$ wierzchołków, gdzie $\omega(n)$ jest dowolną funkcją dążącą do nieskończoności z n i $\omega_1(n) = \theta(\ln n)$;*

(iii) *jeśli $\alpha > 1$, to największa składowa spójności w grafie $\mathcal{IG}^{pasywny}(m, \mathcal{P}_{(n)})$ ma $\theta(n/s)$ wierzchołków* *z prawdopodobieństwem dążącym do* 1 *przy* $n \to \infty$.

Dowód. (i) oraz (iii) są bezpośrednim wnioskiem z 6.18. Natomiast w przypadku (ii) zauważmy, że 2.11 i 5.9 dla grafu dwudzielnego $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ z rozkładem zdegenerowanym $\mathcal{P}_{(n)} = (P_s)$, jeśli pewna własność monotoniczna zachodzi asymptotycznie prawie na pewno dla $\mathcal{H}(s, n, p)$, przy $p = m/\binom{n}{s}$ oraz $m^2/\binom{n}{s} = o(1)$, to ta własność zachodzi asymptotycznie prawie na pewno dla $\mathcal{H}(s, n, m)$ i $\mathcal{H}(\mathcal{BG})$. Czyli dla $s \geqslant 3$, skoro $m = O(n)$, (ii) jest wnioskiem z twierdzenia 4.2 i faktu 6.4. □

Twierdzenie 6.21. *Niech* $\mathcal{P}_{(n)} = (P_s)$ *będzie rozkładem jednopunktowym, dla pewnego* $s \geqslant 3$ *,* $\mathcal{BG}_{m,n}(m, \mathcal{P}_{(n)})$ *będzie grafem zdefiniowanym w 1.4,* $m = \alpha \frac{n}{s(s-1)}$ *dla pewnej stałej* $\alpha > 0$ *, wtedy:*

(i) *jeśli* $\alpha < 1$, *to największa składowa spójności w grafie* $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ *ma* $\theta(\ln n)$ *wierzchołków;*

(ii) *jeśli* $\alpha = 1$, *to największa składowa spójności w grafie* $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ *ma mniej niż* $\omega_1(n)n^{2/3}$ *i więcej niż* $n^{2/3}/\omega(n)$ *wierzchołków, gdzie* $\omega(n)$ *jest dowolną funkcją dążącą do nieskończoności z* n *i* $O(\omega_1(n)) = \ln n$*;*

(iii) *jeśli* $\alpha > 1$, *to największa składowa spójności w grafie* $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ *ma* $\theta(n)$ *wierzchołków*

z prawdopodobieństwem dążącym do 1 *przy* $n \to \infty$.

Dowód. (i) oraz (iii) wynikają z 6.19. Dla (ii) przypomnijmy, że w dowodzie twierdzenia 6.19 pokazaliśmy, że wielkość odpowiadających sobie składowych w $\mathcal{IG}^{aktywny}(m, \mathcal{P}_{(n)})$ i $\mathcal{H}(\mathcal{BG})$ rzędowo jest taka sama. Czyli (ii) wynika z dowodu twierdzenia 6.20.

□

Oznaczenia

$[n] = \{1, 2, \ldots, n\}$ $N_i = \binom{n}{i}$
$(n)_k = n!/(n-k)!$
$b_n = o(a_n)$, gdy $b_n/a_n \to 0$ przy $n \to \infty$
$b_n = O(a_n)$, gdy istnieje taka stała C, że $|b_n| \leqslant Ca_n$
$b_n = \theta(a_n)$, gdy $b_n = O(a_n)$ i $a_n = O(b_n)$

Bibliografia

[1] B. Bollobás, *Random Graphs,* Academic Press, London 1985

[2] B. Bollobás. *The evolution of random graphs,* Trans. Am. Math Soc, 286, 257-274 (1984)

[3] K. L. Chung, *A Course in Probability Theory,* Harcourt, Brace & World,inc.(1968)

[4] P. Erdős, A. Rényi, *On random graphs I,* Publ. Math. Debrecen 6,290-297 (1959)

[5] P. Erdős, A. Rényi, *On the evolution of random graphs,* Publ. Math. Inst. Hungar. Acad. Sci. 5, 17-61, (1960)

[6] P. Erdős, J. Spencer, *Probabilistic methods in combinatorics,* Akadémia Kiadó, Budapest 1974

[7] E. Godehardt, J. Jaworski, *Two models of Random Intersection Graphs for classification* w : M. Schwaiger, O. Opitz eds.: Exploratory data analysis in empirical research, Springer, Berlin-Heidelberg-New York, 67-81 (2002)

[8] J. Jakubiowski, R. Sztencel, *Wstęp do teorii prawdopodobieństwa,* Script, Warszawa 2001

[9] S. Janson, T. Łuczak, A. Ruciński, *Random Graphs,* ,Wiley, New York, 2000

[10] M. Karoński i T. Łuczak, *Random Hypergraphs,* Combinatorics, Paul Erdős is Eighty (Tom2), Keszthely(Węgry)(1993), Budapeszt(1996), 283-293

[11] W. Kordecki, *On the connectedness of random s-graphs,* Discussiones Mathematicae 4(1981),51-58

[12] W. Kordecki, *On the connectedness of random hypergraphs*, Commentaciones Mathematicae 25(1985),265-283

[13] T. Łuczak, *On the equivalence of two basic models of random graphs,* w Random Graphs '87 , Proceedings, Poznań, 1987, John Wiley & Sons, Chichester, eds. M. Karoński, J. Jaworski & A. Ruciński, 151-158.(14)

[14] Z. Palka, A. Ruciński, *Niekonstruktywne metody matematyki dyskretnej,* Wydawnictwo Naukowo-Techniczne, Warszawa 1996

[15] A.Rényi, *Probality theory,* Akadémia Kiadó, Budapest (1970)

[16] J. Schmidt-Pruzan, E. Shamir, *Component structure in the evolution of random hypergraphs,* Combinatorica 5(1)(1985) 81-94

[17] R.L.Graham, J.Nešetřil, *The mathematics of Paul Erdős,* Springer-Verlang, Berlin Heidelberg (1997)