On the distribution of the largest connected component size in random graphs with fixed edges set size

Robin Petit

March-April 2017

Contents

1 Introduction

1.1 Definitions

Let's consider $V = \{v_1, \ldots, v_{|V|}\}$ a set of vertices. We denote by $|V|$ the cardinality of the set V. Let's define the function:

$$X : \mathbb{N} \to \mathbb{N} : n \mapsto \frac{n(n-1)}{2}.$$

Definition 1.1. For $(\alpha, \beta) \in \mathbb{N}^2$, if $\beta \geqslant \alpha$, we define:

$$[\![\alpha,\ \beta]\!] := \{n \in \mathbb{N} \text{ s.t. } \alpha \leqslant n \leqslant \beta\}.$$

Definition 1.2. An undirected graph Γ is denoted $\Gamma(V, E)$ for V its vertices set, and E its edges set, with $E = \{e_1, \ldots, e_{|E|}\}$ and $\forall i \in [\![1, |E|]\!] : e_i = \{v_{i1}, v_{i2}\}$ for $1 \leqslant i_1, i_2 \leqslant |V|$ with $i_1 \neq i_2$ (i.e. loops are not tolerated).
Remark. $|E|$ is usually denoted as m, and $|V|$ is sometimes denoted as n. Both these numbers are (non-strictly) positive integers.

Definition 1.3. The set of all the existing graphs having given vertices set V is denoted by $\Gamma(V,\cdot)$. We denote $\Gamma_m(V,\cdot)$ the subset of $\Gamma(V,\cdot)$ such that $|E| = m$.
Remark. We observe that:

$$\Gamma(V,\cdot) = \bigsqcup_{m\in\mathbb{N}} \Gamma_m(V,\cdot).$$

Definition 1.4. For every $n \in \mathbb{N}$, we define $\mathcal{K}_n$ as the *complete graph* of size n.

Lemma 1.5. *For a graph $\Gamma(V,E)$, we have $|E| \leqslant X(|V|)$.*

Proof. We know that $\Gamma(V,E) \leqslant \mathcal{K}_{|V|}$, and $\mathcal{K}_{|V|}$ has exactly $X(|V|)$ edges (vertex v_i is connected to vertices v_{i+1} to $v_{|V|}$, so the number of edges is equal to $\sum_{i=1}^{|V|}(|V|-i) = \sum_{i=0}^{|V|-1} i = X(|V|)$). □

Lemma 1.6. *For given vertices set V and fixed number of edges $m \in \mathbb{N}$, we have:*

$$\left|\Gamma_m(V,\cdot)\right| = \begin{cases} \binom{X(|V|)}{m} & \textit{if } m \leqslant X(|V|) \\ 0 & \textit{else} \end{cases}.$$

Proof. There are $X(|V|)$ edges $e = \{v_i, v_j\}$ which exist in $\mathcal{K}_{|V|}$. A graph $\Gamma(V,E)$ in $\Gamma_m(V,\cdot)$ has m of these $X(|V|)$ edges. The amount of such graphs is then $\binom{X(|V|)}{m}$. □

Corollary 1.7. *For given vertices set V, we have $\left|\Gamma(V,\cdot)\right| = 2^{X(|V|)}$.*

Proof. Since $\Gamma(V,\cdot)$ is given by a disjoint union over m, its cardinality is equal to the sum of the individual cardinalities:

$$\left|\Gamma(V,\cdot)\right| = \sum_{m\in\mathbb{N}} \left|\Gamma_m(V,\cdot)\right| = \sum_{k=0}^{X(|V|)} \left|\Gamma_m(V,\cdot)\right| = \sum_{k=0}^{X(|V|)} \binom{X(|V|)}{m} = 2^{X(|V|)}.$$

□

Definition 1.8. A graph $\Gamma(V,E)$ is said to be connected if for each $v, w \in V$, there exists a path between v and w. We denote by $\chi(V,\cdot)$ the set of all connected graphs having vertices set V. Again, for $m \in \mathbb{N}$, we denote by $\chi_m(V,\cdot) \subset \chi(V,\cdot)$ the set of connected graphs having m edges.
Remark. $\chi(V,\cdot) \subset \Gamma(V,\cdot)$, and:

$$\chi(V,\cdot) = \bigsqcup_{m\in\mathbb{N}} \chi_m(V,\cdot).$$

Lemma 1.9. *For $m < |V|-1$ or $m > X(|V|)$, we have $\chi_m(V,\cdot) = \emptyset$.*

Proof. For $m > X(|V|)$, we know that $\Gamma_m(V,\cdot) = \emptyset$. As $\chi_m(V,\cdot) \subset \Gamma_m(V,\cdot)$, we know that $\chi_m(V,\cdot) = \emptyset$.

For $m < |V|-1$, let's notice firstly that a cyclic graph needs at least $|V|$ edges: for a cyclic graph, each vertex's degree must be at least 2. We can deduce:

$$2|E| = \sum_{v\in V} \deg(v) \geqslant 2|V|.$$

Thus $|V| \geqslant |E|$ for a cyclic graph. Let's assume (*ad absurdum*) that there exists $\Gamma(V,E) \in \chi_m(V,\cdot)$. Therefore $\Gamma(V,E)$ is acyclic. By definition of $\chi_m(V,\cdot)$, we know that $\Gamma(V,E)$ is connected. But an acyclic and connected graph is a tree, and a tree has exactly $|V|-1 \gneqq m$ edges, which is a contradiction. Therefore, $\chi_m(V,\cdot) = \emptyset$. □

Definition 1.10. Let's define the function:

$$\mu : \mathcal{P}(V) \to [\![1, |V|]\!] : W \mapsto \mu(W) \coloneqq \inf \{i \in [\![1, |V|]\!] \text{ s.t. } v_i \in W\}$$

representing the lowest index of a vertex present in a given subset of $W \subset V$.
Definition 1.11. For every $W \in \mathcal{P}(V)$, we define $\Delta_W : \Gamma(V, \cdot) \to \Gamma(W, \cdot) : \Gamma(V, E) \mapsto \Gamma'(W, E')$ such that:

$$E' = \left\{\{v_i, v_j\} \in E \text{ s.t. } v_i, v_j \in W\right\}.$$

Let's call Δ_W the *restriction to subset* W operator.

1.2 (Largest) Connected Components

Definition 1.12. We define the *connected component of vertex* $v_i \in V$ *in graph* $\Gamma(V, E)$ by the biggest subset (in the sense of inclusion) W of V such that $v_i \in W$ and $\Delta_W(\Gamma(V, E)) \in \chi(W, \cdot)$.

For graph $\Gamma(V, E) \in \Gamma(V, \cdot)$, we define $\left|\mathrm{LCC}(\Gamma(V, E))\right|$ by:

$$\left|\mathrm{LCC}(\Gamma(V, E)\right| \coloneqq \max_{W \in \mathcal{P}(V)} |W| \, \mathbb{I}_{[\Delta_W(\Gamma(V,E) \in \chi(V, \cdot)]}.$$

We then define the *largest connected component of the graph* $\Gamma(V, E)$ as:

$$\mathrm{LCC}(\Gamma(V, E)) \coloneqq \underset{\substack{W \in \mathcal{P}(V) \\ |W| = \left|\mathrm{LCC}(\Gamma(V,E))\right|}}{\arg\min} \mu(W).$$

The set $\Lambda_k^m(V, \cdot)$ is then the set of all graphs $\Gamma(V, E) \in \Gamma(V, \cdot)$, such that $|E| = m$ and $\left|\mathrm{LCC}(\Gamma(V, E))\right| = k$.
Remark. The notations here are consistent since $\left|\mathrm{LCC}(\Gamma(V, E))\right|$ corresponds indeed to the cardinality of $\mathrm{LCC}(\Gamma(V, E))$.

Furthermore, this definition of largest connected component allows to define uniquely the LCC, even though a graph $\Gamma(V, E)$ has several connected component of same size. For example, following graph has two connected component of size 2, i.e. $\{1, 2\}$ (in red) and $\{3, 4\}$ (in blue).

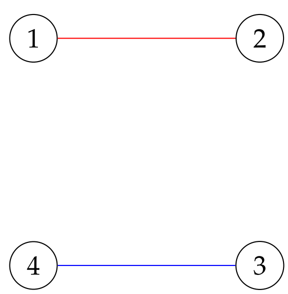

Figure 1: Graph $\Gamma\left([\![1, 4]\!], \left\{\{1, 2\}, \{3, 4\}\right\}\right)$

Nevertheless, the LCC operator yields $\{1, 2\}$ since it minimizes the lowest id of element in connected component (1 for this graph).

Remark. Since $\Lambda_k(V,\cdot) = \bigsqcup_{m=0}^{X(|V|)} \Lambda_k^m(V,\cdot)$ and:

$$\Gamma(V,\cdot) = \bigsqcup_{k=1}^{|V|} \bigsqcup_{m=0}^{X(|V|)} \Lambda_k^m(V,\cdot),$$

we want to know what is $\left|\Lambda_k^m(V,\cdot)\right|$ equal to.

Definition 1.13. Let's declare a new random variable $\mathscr{G}(V,m)$, a graph uniformly distributed in $\Gamma_m(V,\cdot)$, thus such that:

$$\forall \Gamma(V,E) \in \Gamma_m(V,\cdot) : \mathbb{P}[\mathscr{G}(V,m) = \Gamma(V,E)] = \frac{1}{\left|\Gamma_m(V,\cdot)\right|} = \frac{1}{\binom{X(|V|)}{m}}.$$

1.3 Objectives

The objective now is to find an expression for $\left|\Lambda_k(V,\cdot)\right|$ since we are looking for:

$$\mathbb{P}\left[\left|\mathrm{LCC}(\mathscr{G}(V,m))\right| = k\right] = \frac{\left|\Lambda_k(V,\cdot)\right|}{\left|\Gamma(V,\cdot)\right|} = \frac{1}{\left|\Gamma(V,E)\right|} \sum_{m=0}^{X(|V|)} \left|\Lambda_k^m(V,\cdot)\right|.$$

Let's denote this value $p_k \coloneqq \mathbb{P}\left[\left|\mathrm{LCC}(\mathscr{G}(V,m))\right| = k\right]$.

2 Preliminary Results

The general idea in order to determine $\left|\Lambda_k^m(V,\cdot)\right|$ is to insert a connected component of size k on vertices set V, and then to tally the configurations placing $m-k$ vertices without making a bigger connected component than the first one.

2.1 $\left|\Lambda_{k=1}(V,\cdot)\right|$

It is trivial to tell $\left|\Lambda_1^m(V,\cdot)\right| = \delta_0^m$[1], i.e. equals one if $m = 0$ and equals zero if $m > 0$: a graph having at least one edge, cannot have a largest connected component of size 1 because if $e = \{v_i, v_j\}$ is an edge in E, then $\{v_i, v_j\} \subset V$ is a connected component of size 2.

2.2 Upper Boundary of m for $\left|\Lambda_k^m(V,\cdot)\right|$

> **Lemma 2.1** (Upper boundary of edges amount for $k = 2$)**.** *For $m > \frac{|V|}{2}$, we have $\Lambda_2^m(V,\cdot) = \emptyset$.*

Proof. To have a largest connected component of size 2, each vertex must have degree 0 or 1. Take $m \in \mathbb{N}$ such that $m > \frac{V}{2}$. Take $\Gamma(V,E)$ such that $|E| = m$, and take $\mathcal{V}_1 \coloneqq \{v \in V \text{ s.t. } \deg(v) \leqslant 1\} \subset V$. Take the restriction $\Gamma'(\mathcal{V}_1, E') = \Delta_{\mathcal{V}_1}(\Gamma(V,E))$.

Since in a graph, the sum of the degree of each vertex is equal to twice the amount of edges, when applied on Γ', it follows that:

$$2\left|E'\right| = \sum_{v\in\mathcal{V}_1} \deg(v) \leqslant \sum_{v\in\mathcal{V}_1} 1 = |\mathcal{V}_1|.$$

[1] δ_i^j is the Kronecker delta operator.

We then deduce that $|E'| \leqslant \frac{|\mathcal{V}_1|}{2} \leqslant \frac{|V|}{2}$. Thus $\mathcal{V}_1$ must be *strictly* included in V, and then there must exist $v \in V$ such that $\deg(v) \geqslant 2$. Thus:

$$\forall m > \frac{|V|}{2} : \forall \Gamma(V,E) \in \Gamma_m(V,\cdot) : \Gamma(V,E) \notin \Lambda_2^m(V,\cdot).$$

□

Lemma 2.2. *For $\Gamma(V,E) \in \Gamma(V,\cdot)$ a graph and $k \in [\![1,|V|]\!]$, if there exists a vertex $v \in V$ such that $\deg(v) = k$, then $\left|\mathrm{LCC}(\Gamma(V,E))\right| \geqslant k+1$.*

Proof. Take $v \in V$ such that $\deg(v) = k$. There exist $\{v_{i_1}, \dots, v_{i_k}\} \subset V$ such that:

$$\forall j \in [\![1,\, k]\!] : \{v, v_{i_j}\} \in E.$$

Thus $\{v, v_{i_1}, \dots, v_{i_k}\}$ is a connected component of size $k+1$. Thus the largest connected component must have size at least that big. □

Proposition 2.3 (Upper boundary of edges amount generalized)**.** *For $k \in [\![1,|V|]\!]$, and $m > \frac{|V|(k-1)}{2}$, we have $\Lambda_k^m(V,\cdot) = \emptyset$.*

Proof. Take $m > \frac{(k-1)|V|}{2}$, and $\Gamma(V,E) \in \Gamma_m(V,\cdot)$. Take $\mathcal{V}_k \coloneqq \{v \in V \text{ s.t. } \deg(v) \leqslant k-1\}$. Let $\Gamma'(\mathcal{V}_k, E')$ be defined by $\Delta_{\mathcal{V}_k}(\Gamma(V,E))$. We know that:

$$2\left|E'\right| = \sum_{v \in \mathcal{V}_k} \deg(v) \leqslant (k-1)|\mathcal{V}_k| \leqslant (k-1)|V|\,.$$

We deduce that $|E'| \leqslant \frac{(k-1)|V|}{2} < m = |E|$. Thus $|E| \gneqq |E'|$, and this implies that there exists $v \in V$ such that $\deg(v) \geqslant k$. By previous lemma, largest connected component size must be at least $k+1$. □

Remark. We can understand this upper boundary as $m > \frac{|V|(k-1)}{2} = \frac{|V|}{k}\frac{k(k-1)}{2} = \frac{|V|}{k} \cdot X(k)$. So in order to have a LCC of size k, edges can be distributed to make $\left\lfloor \frac{|V|}{k} \right\rfloor$ complete graphs having each $X(k)$ edges. The maximum amount of edges is then given by $\frac{|V|(k-1)}{2}$.

2.3 $\left|\Lambda_{k=2}(V,\cdot)\right|$

Example of size 2 is a bit more complicated:

$$\forall m \in \left[\!\left[1, \left\lfloor \frac{|V|}{2} \right\rfloor\right]\!\right] : \left|\Lambda_2^m(V,\cdot)\right| = \begin{cases} \frac{1}{m!} \prod_{k=0}^{m-1} \binom{|V|-2k}{2} & \text{if } m \leqslant \frac{|V|}{2} \\ 0 & \text{else} \end{cases}.$$

Proof. For $m > \frac{|V|}{2}$, result is shown in Lemma 2.1. The part $\prod_{k=0}^{m-1} \binom{|V|-2k}{2}$ corresponds to the choice of m edges without making a connected component of size $\geqslant 3$.

$\binom{|V|-2\cdot 0}{2}$ is the choice of the first edge (two vertices) among $|V|$ vertices, $\binom{|V|-2}{2}$ is the choice of the second edge (two vertices) among the $|V|-2$ vertices left, etc. At step ℓ, only $|V| - 2(\ell-1)$ vertices are available because two are selected per step, and a selected vertex cannot be used again, otherwise its degree would be $\geqslant 2$, and then the largest component size would be $\geqslant 3$.

The $\frac{1}{m!}$ comes from the fact that the order the edges are selected doesn't matter (so for each choice of m edges, there are $m!$ permutations of these). □

Remark. This can also be expressed as:

$$\left|\Lambda_2^m(V,\cdot)\right| = \frac{1}{m!}\frac{|V|!}{2^m\left(|V|-2m\right)!},$$

by simplification of the product.

3 Processing on Examples — Cardinality of $\Lambda_k^m(V,\cdot)$

$$\begin{aligned}
\left|\Lambda_3^0(V,\cdot)\right| = \left|\Lambda_3^1(V,\cdot)\right| &= 0\\
\left|\Lambda_3^2(V,\cdot)\right| &= \binom{|V|}{3}\binom{3}{2}\\
\left|\Lambda_3^3(V,\cdot)\right| &= \binom{|V|}{3}\binom{3}{3} + \binom{|V|}{3}\binom{3}{2}\binom{|V|-3}{2}\\
\left|\Lambda_3^4(V,\cdot)\right| &= \binom{|V|}{3}\binom{3}{3}\binom{|V|-3}{2} + \binom{|V|}{3}\binom{3}{2}\binom{|V|-3}{3}\binom{3}{2} + \binom{|V|}{3}\binom{3}{2}\binom{|V|-3}{2}\binom{|V|-5}{2}.
\end{aligned}$$

Definition 3.1. Let's denote equally $\left|\Lambda_k^m(n)\right| = \Lambda_k^m(n) \equiv \left|\Lambda_k^m(V,\cdot)\right|$ for V such that $|V| = n$.

This notation allows to lighten the expressions.
Remark. Current explorations tend to a formula looking something like:

$$\left|\Lambda_k^m(|V|)\right| = \binom{|V|}{k}\sum_{\ell=k-1}^{\min\left(m,X(k)\right)}\left|\Lambda_k^\ell(k)\right|\sum_{p=1}^{k}\left|\Lambda_p^{m-\ell}(|V|-k)\right|\beta_{p\,\ell}(m,k,|V|),$$

with $\beta_{p\,\ell}(m,k,|V|)$, a coefficient.

The idea behind this formula is explained in introduction of Section 2: to find the amount of graphs having n vertices, m edges and a largest connected component of size k, let's place a connected component of size k somewhere in the graph (so choose k in $|V|$ vertices), and then multiply this by the amount of possible graphs of largest connected component of size $p \in \{1,\dots,k\}$ (so lower or equal to k).

Idea of proof to be extended later on. In order to prove the equality of the cardinalities, let's find a bijective function Ω between $\Lambda_k^m(V,\cdot)$ and a set like:

$$\mathfrak{Q}_k^m(V) := \bigsqcup_{\substack{W\in\mathcal{P}(V)\\|W|=k}}\ \bigsqcup_{\ell=k-1}^{\min\left(m,X(k)\right)} \Lambda_k^\ell(W,\cdot)\times\left(\bigsqcup_{p=1}^{k}\Lambda_p^{m-\ell}(V\setminus W,\cdot)\right).$$

□

Lemma 3.2. *The sets $\chi_\ell(V,\cdot)$ and $\Lambda_{|V|}^\ell(V,\cdot)$ are equal.*

Proof. A graph $\Gamma(V,E)$ is connected if and only if its largest connected component contains all its vertices, i.e. $\mathrm{LCC}(\Gamma(V,E)) = V$.

This is equivalent to say that $\left|LCC(\Gamma(V,E))\right| = |V|$ since $\forall W \in \mathcal{P}(V) : |W| = |V| \Rightarrow V = W$:

$$\forall W \in \mathcal{P}(V) : \left|\left\{\widetilde{W} \in \mathcal{P}(V) \text{ s.t. } |W| = \left|\widetilde{W}\right|\right\}\right| = \binom{|V|}{|W|},$$

and $\binom{|V|}{|V|} = 1$, thus $\left\{W \in \mathcal{P}(V) \text{ s.t. } |W| = |V|\right\} = \{V\}$. □

3.1 Decomposing set $\Lambda_k(V)$

Definition 3.3. For $k \in \mathbb{N}$, and $\alpha \in \mathbb{N}$, we define:

$$\Lambda_{k,\alpha}(V,\cdot) := \left\{\Gamma(V,E) \in \Lambda_k(V,\cdot) \text{ s.t. } \left|\left\{W \in \mathcal{P}(V) \text{ s.t. } \Delta_W(\Gamma(V,E)) \in \chi(W) \text{ and } |W| = \left|LCC(\Gamma(V,E))\right|\right\}\right| = \alpha\right\},$$

the class of all graphs in $\Lambda_k(V,\cdot)$ having exactly α connected components of maximum size.

Again, for $m \in [\![1, X(|V|)]\!]$, we define $\Lambda^m_{k,\alpha}(V,\cdot)$ by $\Lambda^m_k(V,\cdot) \cap \Lambda_{k,\alpha}(V,\cdot)$.

Remark. Even though several connected components of maximum size do exist in a graph, *the one* LCC is still defined unambiguously!

Lemma 3.4.

1. *For $k > |V|$ or $k = 0$, we have: $\forall \alpha \in \mathbb{N} : \Lambda_{k,\alpha}(V,\cdot) = \emptyset$.*
2. *For $k \in [\![1, |V|]\!]$ and $\alpha > \left\lfloor \frac{|V|}{k} \right\rfloor$, we have $\Lambda_{k,\alpha}(V,\cdot) = \emptyset$.*

Proof.

1. For $k > |V|$ or $k = 0$, it is obvious that: $\Lambda_k(V,\cdot) = \emptyset$ (and then $\Lambda_{k,\alpha}(V,\cdot)$).
2. Take such k and α. Assume (*ad absurdum*) that there exists $\Gamma(V,E) \in \Lambda_{k,\alpha}(V,\cdot)$. We have then $L_1, \dots, L_\alpha \in \mathcal{P}(V)$ such that $\forall i \in [\![1, \alpha]\!] : |L_i| = k$. Also, since the L_i's are connected component, they are disjoint, i.e. $\forall (i,j) \in [\![1, \alpha]\!]^2 : i \neq j \Rightarrow L_i \cap L_j = \emptyset$.

   Thus $\bigcup_{i=1}^{\alpha} L_i \subset V$, and $\sum_{i=1}^{\alpha} |L_i| \leqslant |V|$. But:

   $$\sum_{i=1}^{\alpha} |L_i| = \alpha k > \left\lfloor \frac{|V|}{k} \right\rfloor k \geqslant \frac{|V|}{k} k = |V|,$$

   which leads a contradiction: $|V| > |V|$. We deduce that $\Lambda_{k,\alpha}(V,\cdot) = \emptyset$.

□

Corollary 3.5.

$$\forall k < |V| : \Lambda_k(V,\cdot) = \bigsqcup_{m=k-1}^{X(|V|)} \bigsqcup_{\alpha=1}^{\left\lfloor \frac{|V|}{k} \right\rfloor} \Lambda^m_{k,\alpha}(V,\cdot).$$

Proof. Unions are trivially disjointed.

Now show the equality. The right-hand side is trivially included in $\Lambda_k(V,\cdot)$ (by definition of $\Lambda^m_{k,\alpha}(V,\cdot)$).

Now take $\Gamma(V,E) \in \Lambda_k(V,\cdot)$. We know that $\Lambda^{|E|}_k(V,\cdot)$ with $|E| \leqslant X(|V|)$. As well, we know that the amount of connected components of size $\left|LCC(\Gamma(V,E))\right| = k$ is at least 1 (because $\Gamma(V,E) \in \Lambda_k(V,\cdot)$), and lower or equal to $\left\lfloor \frac{|V|}{k} \right\rfloor$ by previous Lemma. □

3.2 The Set $\mathfrak{Q}^m_{k,\alpha}(V)$

3.2.1 Special case of $\alpha = 1$

From now on, let's write:

$$\mathfrak{Q}^m_{k,1}(V) \coloneqq \bigsqcup_{\substack{W\in\mathcal{P}(V)\\|W|=k}} \bigsqcup_{\ell=k-1}^{\min\left(m,X(k)\right)} \chi_\ell(W,\cdot) \times \left(\bigsqcup_{p=1}^{k-1} \Lambda^{m-\ell}_p(V,\cdot)\right).$$

Proposition 3.6. *For* $k \in [\![1,|V|]\!]$ *and* $m \in \left[\!\left[1, X(|V|)\right]\!\right]$, *we have:*

$$\Lambda^m_{k,1}(V,\cdot) \cong \mathfrak{Q}^m_{k,1}(V),$$

i.e. there exists a bijection between these two sets.

Proof. Consider the function defined by:

$$\Omega : \Lambda^m_{k,1}(V,\cdot) \to \mathfrak{Q}^m_{k,1}(V) : \Gamma(V,E) \mapsto \left(\Delta_{LCC(\Gamma(V,E))}(\Gamma(V,E)), \Delta_{V\setminus LCC(\Gamma(V,E))}(\Gamma(V,E))\right).$$

We know that $\Omega(\Gamma(V,E))$ gives a partition of $\Gamma(V,E)$ because there doesn't exist $e = \{v_{i_1}, v_{i_2}\} \in E$ such that $v_{i_1} \in LCC(\Gamma(V,E))$ and $v_{i_2} \in V \setminus LCC(\Gamma(V,E))$ by definition of connected component. Thus if:

$$\left(\Gamma_1(LCC(\Gamma(V,E)), E_{LCC}), \Gamma_2(V\setminus LCC(\Gamma(V,E)), E_{V\setminus LCC})\right) = \Omega(\Gamma(V,E)),$$

we have $E_{LCC} \sqcup E_{V\setminus LCC} = E$.

Let's prove that Ω is injective. Take $\Gamma_1(V,E_1)$ and $\Gamma_2(V,E_2)$ in $\Lambda^m_k(V,\cdot)$. Assume that $\Omega(\Gamma_1(V,E_1)) = \Omega(\Gamma_2(V,E_2))$. This implies that $LCC(\Gamma_1(V,E_1)) = LCC(\Gamma_2(V,E_2))$ (and thus that their complementary in V are equal as well) and that the restriction of $\Gamma_1(V,E_1)$ and $\Gamma_2(V,E_2)$ to their common LCC are equal. Their restriction to the complementary of the common LCC in V are equal as well.

Thus $\Gamma_1(V,E_1)$ and $\Gamma_2(V,E_2)$ have the same partition and are then equal (i.e. $E_1 = E_2$).

Now, let's prove that Ω is surjective. For $W \in \mathcal{P}(V)$, take:

$$\left(\Gamma_1(W,E_1), \Gamma_2(V\setminus W, E_2)\right) \in \mathfrak{Q}^m_{k,1}(V).$$

We deduce that if $\ell \coloneqq |E_1|$, we have $|E_2| = m - \ell$. Also, $\left|LCC(\Gamma_2(V\setminus W,E_2))\right| \lneqq k$ and $\Gamma_1(W,E_1) \in \chi_\ell(W)$.

For $\Gamma(V,E) \coloneqq \Gamma(V\setminus W \sqcup W, E_1 \sqcup E_2)$, we have indeed that $\Gamma(V,E) \in \Lambda^m_k(V,\cdot)$ and $\Omega(\Gamma(V,E)) = (\Gamma_1,\Gamma_2)$ because $\left|LCC(\Gamma(V,E))\right| = |W| = k$ since W is the only connected component of size k. □

3.2.2 General Case

Definition 3.7. For $k, \alpha \in \mathbb{N}^*$, let's define:

$$\mathcal{P}_{k,\alpha}(V) \coloneqq \left\{ \{W_1,\ldots,W_\alpha\} \in \mathcal{P}\left(\mathcal{P}(V)\right) \text{ s.t. } \left\{ \begin{array}{l} \forall i \in [\![1,\alpha]\!] : |W_i| = k \\ \forall (i,j) \in [\![1,\alpha]\!]^2 : i \neq j \Leftrightarrow W_i \cap W_j = \emptyset \end{array} \right. \right\},$$

thus $\mathcal{P}_{k,\alpha}(V)$ is the set of all sets containing α subsets of V which are disjointed and of size k.

Remark. We can tell:

$$\left|\mathcal{P}_{k,\alpha}(V)\right| = \frac{1}{\alpha!}\prod_{\beta=0}^{\alpha-1}\binom{|V|-k\beta}{k} = \frac{1}{\alpha!}\frac{|V|!}{(k!)^{\alpha}(|V|-k\alpha)!}.$$

Definition 3.8. For $k \in [\![1, |V|]\!]$, $m \in [\![0, X(|V|)]\!]$, and $\alpha \in \left[\!\left[2, \left\lfloor \frac{|V|}{k} \right\rfloor\right]\!\right]$, let's define:

$$\begin{aligned}\mathfrak{Q}^{m}_{k,\alpha}(V) &\coloneqq \bigsqcup_{\substack{\{W_1,\dots,W_\alpha\}\in\mathcal{P}_{k,\alpha}(V)\\ \mu(W_1)<\dots<\mu(W_\alpha)}} \bigsqcup_{\substack{(i_1,\dots,i_\alpha)\in[\![k-1,X(k)]\!]^{\alpha}\\ \text{s.t. } \sum_{j=1}^{\alpha} i_j \leqslant \min\left(m,\alpha X(k)\right)}} \left[\left(\prod_{j=1}^{\alpha}\chi_{i_j}(W_j,\cdot)\right)\times\left(\bigsqcup_{p=1}^{k-1}\Lambda_p^{m-\sum_{j=1}^{\alpha} i_j}\left(V\setminus\bigsqcup_{j=1}^{\alpha}W_j,\cdot\right)\right)\right]\\ &= \bigsqcup_{\substack{\{W_1,\dots,W_\alpha\}\in\mathcal{P}_{k,\alpha}(V)\\ \mu(W_1)<\dots<\mu(W_\alpha)}} \bigsqcup_{\Sigma=\alpha(k-1)}^{\min\left(m,\alpha X(k)\right)} \bigsqcup_{\substack{(i_1,\dots,i_\alpha)\in[\![k-1,X(k)]\!]^{\alpha}\\ \text{s.t. } \sum_{j=1}^{\alpha} i_j = \Sigma}} \left[\left(\prod_{j=1}^{\alpha}\chi_{i_j}(W_j,\cdot)\right)\times\left(\bigsqcup_{p=1}^{k-1}\Lambda_p^{m-\Sigma}\left(V\setminus\bigsqcup_{j=1}^{\alpha}W_j,\cdot\right)\right)\right]\end{aligned}$$

Theorem 3.9. *For $(k,m) \in [\![1, |V|]\!] \times [\![0, X(|V|)]\!]$ and $\alpha \in \left[\!\left[1, \left\lfloor \frac{|V|}{k} \right\rfloor\right]\!\right]$, we have:*

$$\Lambda^{m}_{k,\alpha}(V,\cdot) \cong \mathfrak{Q}^{m}_{k,\alpha}(V).$$

Proof. For $\alpha = 1$, the theorem is proven by Proposition 3.6.[2]

Now, for $\alpha \geqslant 2$, we have the function:

$$\begin{aligned}&\Omega_\alpha : \Lambda^{m}_{k,\alpha}(V,\cdot) \to \mathfrak{Q}^{m}_{k,\alpha}(V) :\\ &\quad \Gamma(V,E) \mapsto \left(\Delta_{W_1}(\Gamma(V,E)),\dots,\Delta_{W_\alpha}(\Gamma(V,E)),\Delta_{V\setminus\bigsqcup_{j=1}^{\alpha}W_j}(\Gamma(V,E))\right),\end{aligned}$$

for $W_1,\dots,W_\alpha$ the subsets of V two by two disjoints, such that $\forall i \in [\![1, \alpha]\!] : |W_i| = k$, and that:

$$\forall i \in [\![2, \alpha]\!] : \mu(W_{i-1}) \lneqq \mu(W_i).$$

We know that $W_1,\dots,W_\alpha$ are the only connected components of size k because $\Gamma(V,E) \in \Lambda^{m}_{k,\alpha}(V,\cdot)$. And also, values of $\mu(W_j)$ can't be equal for different indices by definition of connected components. This implies that function Ω_α is properly defined.

Now, prove that Ω_α is bijective.

Injective Take $\Gamma_1(V,E_1), \Gamma_2(V,E_2) \in \Lambda^{m}_{k,\alpha}(V,\cdot)$. Let's assume that:

$$\Omega_\alpha(\Gamma_1(V,E)) = \Omega_\alpha(\Gamma_2(V,E_2)).$$

We can deduce that Γ_1 and Γ_2 have the same connected components, and that their restrictions to these connected components are equal as well. By similar argument than in proof of Proposition 3.6, we find that Γ_1 and Γ_2 must be equal.

[2]This Proposition is not exactly expressed the same way because $\mathfrak{Q}^{m}_{k,1}(V)$ does not consider a set $\{W_1\} \in \mathcal{P}\left(\mathcal{P}(V)\right)$ but only a subset $W \in \mathcal{P}(V)$. Yet, the proof is equivalent.

Surjective Take:

$$\big(\Gamma_1(W_1, E_1), \dots, \Gamma_\alpha(W_\alpha, E_\alpha), \Gamma(W, E)\big) \in \mathfrak{Q}^m_{k,\alpha}(V),$$

and proof that there exists a graph $\hat{\Gamma}(V, \hat{E}) \in \Lambda^m_{k,\alpha}(V, \cdot)$ such that:

$$\Omega_\alpha(\hat{\Gamma}(V, \hat{E})) = \big(\Gamma_1(W_1, E_1), \dots, \Gamma_\alpha(W_\alpha, E_\alpha), \Gamma(V, E)\big)\,.$$

We know that $\{W_1, \dots, W_\alpha\} \in \mathcal{P}_{k,\alpha}(V)$ by definition of $\mathfrak{Q}^m_{k,\alpha}(V)$, and that:

$$V = W \sqcup \bigsqcup_{j=1}^{\alpha} W_j.$$

As well, by definition of $\mathfrak{Q}^m_{k,\alpha}(V)$, we know that if $\hat{E} \coloneqq E \sqcup \bigsqcup_{j=1}^{\alpha} E_j$, then $\left|\hat{E}\right| = m$. So $\hat{\Gamma}(V, \hat{E})$, the graph created by *assembling* the different components Γ_1 to Γ_α and Γ, is indeed in $\Lambda^m_{k,\alpha}$ because:

$$\forall j \in [\![1,\, \alpha]\!] : \left|\mathrm{LCC}(\Gamma_j(W_j, E_j)\right| = k \qquad \text{and } \left|\mathrm{LCC}(\Gamma(V, E))\right| \lneqq k$$

by definition of $\mathfrak{Q}^m_{k,\alpha}(V)$.

Finally, $\Omega_\alpha(\hat{\Gamma}(V, \hat{E}))$ yields indeed $(\Gamma_1, \dots, \Gamma_\alpha, \Gamma)$ since connected components are ordered according to the μ function defined in Definition 1.10 □

Remark. The problem of the largest connected component size has been reduced to a problem of connected graphs and recursive combinatorics.

Recursive values like these ones can be computed pretty efficiently thanks to dynamic programming.

4 Conclusion

We have then proved that the cardinality of set $\Lambda^m_k(V, \cdot)$ is equal to:

$$\begin{aligned}
\left|\Lambda^m_k(V, \cdot)\right| &= \left|\bigsqcup_{\alpha=1}^{\left\lfloor \frac{|V|}{k} \right\rfloor} \Lambda^m_{k,\alpha}(V, \cdot)\right| = \sum_{\alpha=1}^{\left\lfloor \frac{|V|}{k} \right\rfloor} \left|\Lambda^m_{k,\alpha}(V, \cdot)\right| = \sum_{\alpha=1}^{\left\lfloor \frac{|V|}{k} \right\rfloor} \left|\mathfrak{Q}^m_{k,\alpha}(V)\right| \\
&= \sum_{\alpha=1}^{\left\lfloor \frac{|V|}{k} \right\rfloor} \left|\mathcal{P}_{k,\alpha}(V)\right| \sum_{\Sigma=\alpha(k-1)}^{\min(m, \alpha X(k))} \sum_{\substack{(i_1, \dots, i_\alpha) \in [\![k-1, X(k)]\!]^\alpha \\ \sum_{j=1}^{\alpha} i_j = \Sigma}} \left(\prod_{j=1}^{\alpha} \left|\chi_{i_j}(k)\right| \times \sum_{p=1}^{k-1} \left|\Lambda_p^{m-\Sigma}(|V| - k\alpha)\right| \right),
\end{aligned}$$

from which we eventually deduce:

$$\forall k \in [\![1, |V|]\!] : \mathbb{P}\left[\left|\mathrm{LCC}(\mathscr{G}(V, m))\right| = k\right] = \frac{\left|\Lambda^m_k(V, \cdot)\right|}{\left|\Gamma(V, \cdot)\right|}.$$