ENUMERATION OF CONNECTED UNIFORM HYPERGRAPHS

TSIRY ANDRIAMAMPIANINA AND VLADY RAVELOMANANA

Abstract. In this paper, we are concerned in counting exactly and asymptotically connected labeled b-uniform hypergraphs ($b \geq 3$). Enumerative results on connected graphs are generalized here to connected uniform hypergraphs. For this purpose, these structures are counted according to the number of vertices and hyperedges. First, we show how to compute step by step the associated exponential generating functions (EGFs) by means of differential equations and provide combinatorial interpretations of the obtained results. Next, we turn on asymptotic enumeration. We establish Wright-like inequalities for hypergraphs and by means of complex analysis, we obtain the asymptotic number of connected b-uniform hypergraphs with n vertices and $(n+\ell)/(b-1)$ hyperedges whenever $\ell = o(n^{1/3}/b^{1/3})$. This latter result confirms a conjecture made by Karoński and Łuczak in [20] about the validity of their formula for excesses in the 'Wright's range'.

Résumé. Dans cet article, nous nous intéressons à l'énumération exacte puis asymptotique des hypergraphes b-uniformes ($b \geq 3$) . Des résultats énumératifs sur les graphes sont ici généralisés pour les hypergraphes b-uniformes. Dans cette optique, ces structures sont énumérées suivant le nombre de sommets et le nombre d'hyperarêtes. Premièrement, nous montrons comment obtenir récursivement leurs fonctions génératrices exponentielles et nous justifions alors les formes des résultats ainsi obtenus par des arguments combinatoires. Ensuite, nous faisons l'énumération asymptotique. Nous établissons des inégalités similaires à celles de Wright pour les hypergraphes, en passant par de l'analyse complexe, nous obtenons l'asymtotique du nombre d'hypergraphes connexes b-uniformes ayant n sommets et $(n+\ell)/(b-1)$ hyperarêtes quand $\ell = o(n^{1/3}/b^{1/3})$. Ce dernier résultat confirme une conjecture de Karoǹski et Łuczak dans [20] pour avoir une formule valide avec des excès dans 'l'intervalle de Wright'.

1. Introduction

In this paper we are concerned with counting exactly and asymptotically members of families of labeled connected b-uniform hypergraphs with a given number of vertices and hyperedges and without multiple hyperedges. A labeled *hypergraph* $\mathcal{H} = (V, E)$ is given by a set V of n vertices with a family E of subsets of V of cardinal ≥ 2 (see Berge [4]). A member of E is called *hyperedge* and $\mathcal{H} = (V, E)$ is said b-*uniform* ($b \geq 2$) iff each member of E contains *exactly* b vertices. Therefore, 2-uniform hypergraphs are simply graphs. Let $\mathcal{H} = (V, E)$ be a hypergraph, uniform or not, then its *excess* is defined as (see [20]):

$$\text{excess}(\mathcal{H}) = \sum_{e \in E} (|e| - 1) - |V| \,. \tag{1}$$

Therefore, if $\mathcal{H} = (V, E)$ is a b-uniform hypergraph then its excess is given by the expression

$$\text{excess}(\mathcal{H}) = |E| \times (b-1) - |V| \,. \tag{2}$$

The notion of excess was first used in [27] where the author obtained substantial enumerative results in the study of connected graphs according to their number of vertices and edges. Wright's results appeared to be very important in the study of random graphs [6, 18, 19]. Later, Bender, Canfield and McKay [2] but also Pittel and Wormald [24] generalized Wright's results and obtained the asymptotic number of connected graphs of any given number of vertices and edges.

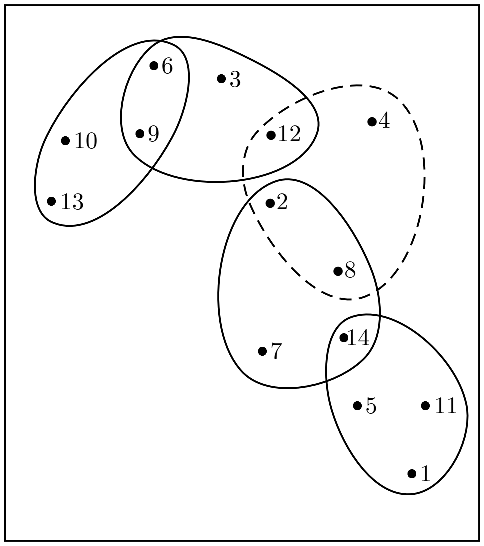
Key words and phrases. b-uniform hypergraphs; enumerative and analytic combinatorics; saddle-point method; generalized Wright's coefficients; random hypergraphs.

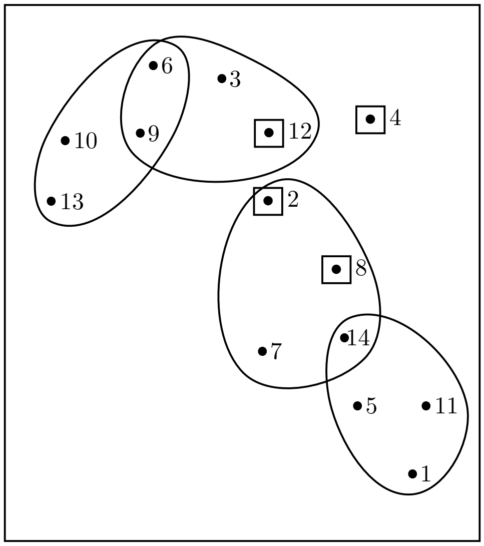
In contrary, much less is known about the number of hypergraphs of a given size. As far as we know, the most important results in these directions are those of Karoński and Łuczak in [20, 21]. In [20], the two authors used 'purely combinatorial arguments' to obtain their results. In this paper, our aim is to obtain analogous results to that of Wright [27, 29]. To do so, our approach is based on generating functions. Following the previously cited works, namely [18], connected hypergraphs with excess -1 are called *hypertrees*, connected hypergraphs with excess 0 are called *unicyclic components* or *unicycles*. Since these structures are labeled, we will use exponential generating functions (EGFs, for short) [15] to encode their number. Then, denote by H_ℓ the EGF of labeled connected b-uniform hypergraphs with excess ℓ. The purpose of this work is to compute the sequence of EGFs $(H_\ell)_{\ell \geq -1}$.

The outline of this paper is as follows. In the second section, we establish the differential equation satisfied by the EGFs H_ℓ ($\ell \geq -1$). We show how these EGFs can be computed exactly from this combinatorial equation and we retrieve some results that appeared in [25, 20]. In the third section, we give the forms of the expression of H_ℓ. We show that for every $\ell \geq -1$, H_ℓ can be expressed in terms of the EGF of rooted hypertrees and we give combinatorial interpretations of the forms of these EGFs. The next section is devoted to the asymptotic enumeration of uniform hypergraphs. First, we establish Wright-like inequalities for hypergraphs. Next, these inequalities are combined with methods from complex analysis and lead us to the asymptotic number of connected hypergraphs with n vertices and $(n + o(n^{1/3}))/(b-1)$ hyperedges.

2. COMBINATORIAL EQUATIONS SATISFIED BY H_ℓ

2.1. **Hypergraphs surgery.** Let us start with a figure that illustrates, with 4-uniform hypergraph, the main idea from which we deduce the enumeration.





The figure on the left side is a connected 4-uniform hypergraph with 14 vertices and 5 hyperedges one of which is distinguished, namely the dashed hyperedge {2, 4, 8, 12}. The figure on the right side is a 4-uniform hypergraph with also 14 vertices but with only 4 hyperedges. This latter hypergraph is not connected but contains 3 components in which one or more vertices are distinguished (resp. {2, 8}, {4} and {12}). The above figures reflect combinatorial relations between families of connected hypergraphs with on first hand a distinguished hyperedge and on the other hand marked vertices. For instance, we refer the reader to Bergeron, Labelle and Leroux [5] for the use of distinguishing/marking and pointing in combinatorial species. The following lemma describes the relationships between number of edges and excesses in b-uniform connected components with distinguished hyperedge and marked vertices:

Lemma 2.1. *Consider a set $\mathcal{M}$ of connected b-uniform hypergraphs with one or more distinguished vertices. For any couple $(j,\, k)$, denote by m_{jk} the number of connected components, with excess j and with k marked vertices, in $\mathcal{M}$. Then, the hypergraph obtained when creating a (new) hyperedge connecting all the distinguished vertices of all the components in $\mathcal{M}$ is* (**i**) *connected,* (**ii**) *b-uniform and* (**iii**) *has excess ℓ if and only if*

$$\sum_{j,k} k\, m_{jk} = b \qquad \text{and} \qquad \sum_{j,k} (j+k)\, m_{jk} = \ell + 1\,. \tag{3}$$

Proof. It is clear that the created hypergraph is connected and is b-uniform if and only if the total number $(\sum_{j,k} k\, m_{jk})$ of distinguished vertices in the set is equal to b. Let N be the number of (connected) hypergraphs in the considered set $\mathcal{M}$ and let n be the total number of vertices in this set. Let us assign an arbitrary order to the members of the set and let n_i, s_i and k_i be respectively the number of vertices, hyperedges and distinguished vertices in the i-th hypergraph. The excess of the newly created hypergraph is then equals to ℓ if and only if $\sum_{i=1}^{N} s_i(b-1) + (b-1) - n = \ell$. We get $\sum_{i=1}^{N} s_i(b-1) + \sum_{i=1}^{N} k_i - \sum_{i=1}^{N} n_i = \ell + 1$ and $\sum_{i=1}^{N} \left(\{s_i(b-1) - n_i\} + k_i \right) = \ell + 1$. Therefore, $\sum_{j,k} m_{jk}(j+k) = \ell + 1$. ☐

2.2. **Combinatorial equations.** In this paragraph, the previous correspondences are expressed in terms of EGFs. Let us consider the bivariate EGF H_ℓ. We have

$$\begin{aligned} H_\ell(w,z) &= \sum_{s=0}^{\infty} \sum_{n=0}^{\infty} h_\ell(s,n)\, w^s \frac{z^n}{n!} \\ &= \sum_{n=0}^{\infty} h_\ell\Big((n+\ell)/(b-1), n\Big)\, w^{(n+\ell)/(b-1)} \frac{z^n}{n!}\,, \end{aligned} \tag{4}$$

where w (resp. z) is the variable related to the number of hyperedges (resp. labeled vertices). In (4), $h_\ell(s,n)$ denotes the number of connected b-uniform hypergraphs with excess ℓ with s hyperedges and n vertices. Using (2), we note that $h_\ell(s,n) \neq 0$ iff $(n+\ell)/(b-1) \in \mathbb{N}$. The following theorem is inspired by the observations of paragraph 2.1 and gives recursive relation between the EGFs H_ℓ.

Theorem 2.2. *The bivariate EGFs $(H_\ell)_{\ell \geq -1}$ of labeled connected b-uniform hypergraphs satisfy*

$$w \frac{\partial}{\partial w} H_\ell(w,z) = w \sum_{(m_{jk}) \in S_\ell} \left\{ \prod_{j,k} \frac{1}{m_{jk}!} \left(\frac{z^k}{k!} \frac{\partial^k}{\partial z^k} H_j(w,z) \right)^{m_{jk}} \right\} - w \frac{\partial}{\partial w} H_{\ell-b+1}(w,z) \tag{5}$$

where S_ℓ is the following set of matrix:

$$S_\ell = \left\{ (m_{jk})_{\substack{-1 \leq j \leq \ell \\ 1 \leq k \leq b}} \text{ with } (m_{jk} \in \mathbb{N}) \text{ such that } \sum_{j,k} m_{jk}(j+k) = \ell + 1 \text{ and } \sum_{j,k} k m_{jk} = b \right\} \tag{6}$$

and $H_j \equiv 0$ if $j \leq -2$.

Proof. This equation relates, in terms of generating functions, the bijection between two sets of objects described by **a**) and **b**) as follows. **a**) In the left-hand side of (5), we have the EGF of the set of connected hypergraphs with excess ℓ and with a marked hyperedge. **b**) In the right-hand side, there are union of sets of components with one or more distinguished vertices that can be obtained from the removal, in a connected hypergraph of excess ℓ, of a hyperedge. After such removal, in each newly created component, the vertices which belonged to the removed hyperedge are marked. If there is k such distinguished vertices, in terms of EGFs, we then have $\frac{z^k}{k!} \frac{\partial^k}{\partial z^k} H_j(w,z)$. The second member of our equation can be interpreted as the creation of a (future) hyperedge with a total of b distinguished vertices in order to reconnect a set of hypergraphs. In the case where there is only one component, necessarily its excess is $\ell - b - 1$ and there are b of its vertices that are distinguished. These b vertices must not form an already existing hyperedge because we consider here hypergraphs without multiple hyperedges. It is the reason why we have to subtract the term $w \frac{\partial}{\partial w} H_{\ell-b+1}(w,z)$

in the RHS of (5). Observe that by the previous lemma, the definition of the set S_ℓ, *viz.* (6), ensures that the hypergraph obtained by the creation of a hyperedge connecting the marked vertices in the RHS is with excess ℓ and that the hyperedge which is created is formed with b vertices. ☐

Remark 2.3. *We note that it is sufficient to determine the univariate EGFs since the corresponding bivariate EGFs can be deduced from the univariate ones simply using the relation*

$$H_\ell(w,z) = w^{\ell/(b-1)} H_\ell(w^{1/(b-1)} z)\,. \tag{7}$$

Saving the justification of its use for later, let us denote by $T(z)$ the (univariate) EGF corresponding to rooted hypertrees. Since a rooted hypertree is either a root or a root with a non-empty set of rooted (sub)hypertrees, borrowing methods from symbolic combinatorics (cf. [11]), we get

$$T(z) = z\,\exp\left(\frac{T(z)^{(b-1)}}{(b-1)!}\right). \tag{8}$$

Remark 2.4. *Throughout this paper, we use the notation H_ℓ followed by the couple of variables (w,z) to express the bivariate EGF, the notation H_ℓ followed by the variable (z) to express the univariate EGF. Whenever the variable are intentionally omitted, the EGF in used is*
$H_\ell \equiv H_\ell(T(z))$ where $T(z)$ is the EGF of rooted (b-uniform) hypertrees implicitly given by (8).

The EGFs $H_\ell \equiv H_\ell(T(z))$ satisfies the following.

Corollary 2.5. *For excess $\ell = -1$*

$$H_{-1} = T - \frac{(b-1)T^b}{b!}\,, \qquad T \equiv T(z) \tag{9}$$

and for $\ell \geq 0$

$$\begin{aligned} \frac{1}{b-1}\left(\ell H_\ell + T\frac{\mathrm{d}}{\mathrm{d}T}H_\ell\right) &= \sum_{(m_{jk})\in {S_\ell}^*} \left\{\prod_{j,k} \frac{1}{m_{jk}!}\left(\frac{z^k}{k!}\frac{\mathrm{d}^k}{\mathrm{d}z^k}H_j(z)\right)^{m_{jk}}\right\} \\ &\quad -\frac{1}{b-1}\left((\ell-b+1)H_{\ell-b+1}(z) + z\frac{\mathrm{d}}{\mathrm{d}z}H_{\ell-b+1}(z)\right) \end{aligned} \tag{10}$$

where ${S_\ell}^*$ is the same as S_ℓ (see (6)) without the matrix where all coefficients equal zero except for the coefficients $m_{1,1} = b-1$ and $m_{\ell,1} = 1$.

Sketch of proof. Use the fact that

$$w\frac{\partial}{\partial w}H_j(w,z) = \frac{1}{b-1}\left(jH_j(w,z) + z\frac{\partial}{\partial z}H_j(w,z)\right)\,,$$

with (5) and (7) and set $w=1$. For $\ell = -1$, we have $S_{-1} = \{(m_{-11}, m_{-12}, \ldots, m_{-1,b}) = (b,0,0,\ldots,0)\}$. Therefore, we obtain

$$\frac{1}{b-1}\left(-H_{-1} + z\frac{\mathrm{d}}{\mathrm{d}z}H_{-1}(z)\right) = \frac{\left(z\frac{\mathrm{d}}{\mathrm{d}z}H_{-1}(z)\right)^b}{b!}\,,$$

and using the fact that $z\frac{\mathrm{d}}{\mathrm{d}z}H_{-1}(z) = T(z)$, it yields (9). To prove (10), first we note that for $\ell \geq 0$ the range of the matrix in ${S_\ell}^*$ can be rearranged so that the line index ranges from -1 to $\ell - 1$ and the column index ranges from 1 to b. After some algebra, we get

$$\frac{1}{b-1}\left(\ell H_\ell + z\frac{\mathrm{d}}{\mathrm{d}z}H_\ell(z)\right) = J_\ell + \frac{\left(z\frac{\mathrm{d}}{\mathrm{d}z}H_{-1}(z)\right)^{b-1}}{(b-1)!}\left(z\frac{\mathrm{d}}{\mathrm{d}z}H_\ell(z)\right)$$

where J_ℓ is the RHS of (10). Again using $z\frac{\mathrm{d}}{\mathrm{d}z}H_{-1}(z) = T$, we obtain

$$\frac{1}{b-1}\left(\ell H_\ell + \left(z\frac{\mathrm{d}}{\mathrm{d}z}H_\ell(z)\right)\left(1 - \frac{T^{b-1}}{(b-2)!}\right)\right) = J_\ell\,.$$

From (8), we have

$$\frac{\mathrm{d}T}{\mathrm{d}z} = \frac{T}{z\left(1 - \frac{T^{b-1}}{(b-2)!}\right)} \tag{11}$$

and by the chain rule for differentiation we get the desired result. Note also that $\frac{z^k}{k!}\frac{\mathrm{d}^k}{\mathrm{d}z^k}H_j(z)$ can be expressed in terms of T so that (10) is completely a differential equation w.r.t. T. ☐

2.3. **Analytical resolution.** In this section we show how to compute the expression of H_ℓ, $\ell \geq 0$, in terms of the EGF T of rooted hypertrees. We note that the equation (10) for $\ell \geq 0$ allows us to compute recursively the expression of H_j for successive values of j. Thus, for each step, we have to solve a differential equation of order one in the variable T to get the expression of H_j which verifies the condition that ${H_j}_{|T=0} = 0$.

Lemma 2.6. *Let us define θ as*

$$\theta = 1 - \frac{T^{b-1}}{(b-2)!}\,. \tag{12}$$

For all $j \geq -1$ and for all $k \geq 0$, there is a function f_{jk} such that

$$\frac{\mathrm{d}^k}{\mathrm{d}z^k}H_j(z) = \frac{f_{jk}(\theta)}{z^k T^j}\,. \tag{13}$$

Denoting $f_j \equiv f_{j0}$, in particular we have

$$H_j = \frac{f_j(\theta)}{T^j}\,. \tag{14}$$

Proof. From (11) and by the chain rule for differential we deduce (13):

$$f_{j,k+1}\left(\theta\right) = -(b-1)\frac{{f_{jk}}'(\theta)}{\theta} + (b-1){f_{jk}}'(\theta) - j\frac{f_{jk}(\theta)}{\theta} - kf_{jk}(\theta)\,. \tag{15}$$

The change of variable given by (12) allow us to deduce that $f_\ell(\theta)$ satisfies:
(16)

$$\frac{\mathrm{d}\left(f_\ell(\theta)\right)}{-(b-2)!} = \left(\sum_{(m_{jk})\in {S_\ell}^*}\prod_{j,k}\frac{1}{m_{jk}!}\left(\frac{f_{jk}(\theta)}{k!}\right)^{m_{jk}} - \frac{1}{b-1}\left((\ell-b+1)f_{\ell-b+1,0}(\theta) + f_{\ell-b+1,1}(\theta)\right)\right)\mathrm{d}\theta$$

☐

3. On the form of the EGFs H_ℓ

In order to establish the forms of the EGFs H_ℓ, we introduce some definitions.
Definitions. The *degree* of a vertex v is the number of the hyperedges that contain v.
A *special hyperedge* is one that contains 3 or more vertices of degree at least 2.
A *special vertex* is either a vertex that belongs to a special hyperedge or a vertex of degree ≥ 3.
A *pendant hyperedge* is one where there are $(b-1)$ vertices of degree 1. In the following, we call *path* a sequence of hyperedges. A path is also characterized by a starting vertex that belongs to the first hyperedge and by an ending vertex that belongs to the last hyperedge, and the sequence of hyperedges defining a path is such that each hyperedge contains exactly $(b-2)$ vertices of degree 1 in the hypergraph and where any pair of successive hyperedges share exactly one vertex that is not the starting nor the ending vertex of the path. We distinguish four kind of paths:

- α*-path*: a path that starts from and ends to the same special vertex, there are at least 2 hyperedges in an α-path and if there are exactly 2 hyperedges in an α-path then it is said to be *minimal.*
- β*-path*: a path that connects 2 special vertices such that if any hyperedge in the path is broken, these 2 special vertices become disconnected, there is at least 1 hyperedge in such a path; a single

hyperedge β-path is said to be *minimal*.

- γ-*path*: a path that joins 2 special vertices such that these vertices remain connected even if this path is broken, there is at least 1 hyperedge in such a path; a single hyperedge γ-path is said to be *minimal*.

A *basic hypergraph* is an unlabeled hypergraph that can be obtained from a labeled hypergraph by the following procedure:

- Discard the labels.
- Remove recursively pendant hyperedges.
- Shrink paths to a minimal special path of the same kind.

Thus, basic hypergraph has the same excess as any hypergraph from which it may be obtained: in the procedure we have described each time a hyperedge is removed (this happens when shrinking a path), it is just as if we have removed $(b-1)$ vertices. Furthermore, basic hypergraph has, for a given kind of paths, as much number of this kind of paths as in any (original) labeled hypergraph from which it may be obtained.

Let us enumerate the number of hypergraphs from which a fixed basic hypergraph with excess ℓ can be obtained. Let J be the EGF of such hypergraphs. Let m be the number of vertices in the basic hypergraph and respectively c_α, c_β and c_γ be the number of α-, β- and γ-paths, then

$$J = \frac{1}{g}\frac{T^m}{\theta^p} \tag{17}$$

where g is the number of automorphisms (e.g. [14]) of the basic hypergraph and $p = c_\alpha + c_\beta + c_\gamma$ is the number of α-, β- or γ-paths. The proof of this relation is immediate since “original” hypergraphs are obtained by rooting m rooted hypertrees in the basic hypergraph and by re-inserting p “chains” of eventually zero length in α- β- and γ-paths of the basic hypergraph. Thus, each hypergraph is obtained g times because of the number of choices where the m rooted hypertrees can be fixed. Furthermore, with s denoting the number of hyperedges in the considered basic hypergraph, there is a positive rational λ such that:

$$J = \frac{1}{g}\frac{T^{s(b-1)-\ell}}{\theta^p} = \lambda\frac{(1-\theta)^s}{T^\ell\theta^p}\,, \tag{18}$$

and necessary $s \geq \left\lfloor \frac{\ell+1}{b-1} + 1 \right\rfloor$.

Lemma 3.1. *For any basic hypergraph with excess ℓ, the total number of α-, β- and γ-paths verifies*

$$c_\alpha + c_\beta + c_\gamma \leq 3\ell$$

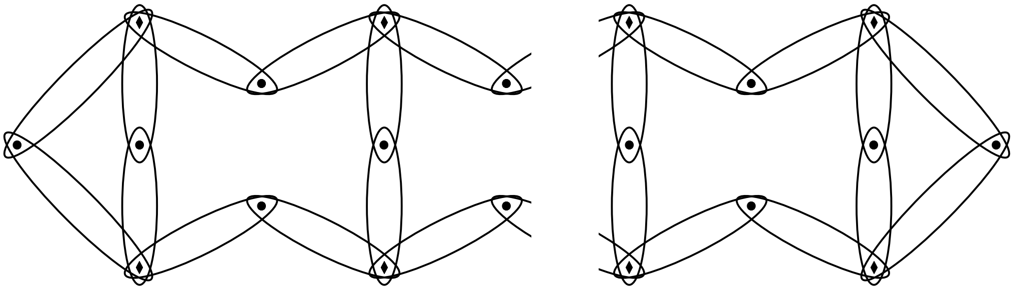
Proof. Let B_0 be the hypergraph induced by the special vertices in the basic hypergraph, let m_0 be the number of special vertices and r_0 be the number of special hyperedges then

$$m_0 + c_\alpha + 2(b-2)c_\alpha + (b-2)c_\beta + (b-2)c_\gamma + \ell = (b-1)\left(r_0 + 2c_\alpha + c_\beta + c_\gamma\right)\,. \tag{19}$$

Thus, $m_0 - r_0(b-1) + \ell = c_\alpha + c_\beta + c_\gamma$ and

$$-\text{excess}(B_0) + \ell = p\,. \tag{20}$$

Therefore, to determine the maximum of the number p over the basic hypergraph, it is sufficient to determine the minimal value of excess(B_0). excess(B_0) has minimal value if B_0 is a forest where hypertrees are either a single vertex (of degree 3 in the basic hypergraph) or a hyperedge (with exactly 3 vertices of degree 2 in the basic hypergraph). Therefore, if there is a hypergraph with hypergraph induced by the special vertices satisfying the above condition, we can deduce the maximum of the number p. The construction of such hypergraph is depicted in the following figure:



where only the vertices of degree at least 2 are represented. The basic hypergraph that can be obtained from the hypergraph in the figure above is only with special vertices of degree 3 and the hypergraph induced by these vertices consists of exactly $2 \times \ell$ isolated vertices (because the excess of the hypergraph is ℓ). Thus, $p \leq 3\ell$. □

Lemma 3.2. *$H_\ell = \frac{f_\ell(\theta)}{T^\ell}$ with f_ℓ a polynomial of maximum degree $\left\lfloor \frac{\ell+1}{b-1} + 1 \right\rfloor$.*

Proof. A matrix in ${S_\ell}^*$ corresponds to constructions as the one we have described in lemma 2.1. After having assigned an arbitrary order to the marked hypergraphs used in a such construction, let:

- the i-th hypergraph be of excess ℓ_i such that $\ell_i + 1 = q_i(b-1) + r_i$
- $\ell + 1 = q(b-1) + r$

with $q,\ q_i \geq 0$ and $r,\ r_i < b-1$.
As in the proof of lemma 2.1, we get here $\ell + 1 = q(b-1) + r = \sum_{i=1}^{N}(q_i(b-1) + r_i - 1 + k_i)$ since $\sum_{i=1}^{N}(r_i - 1 + k_i) \geq 0$, we deduce that $q \geq \sum_{i=1}^{N} q_i$. So,

$$q + 1 \geq \sum_{i=1}^{N} q_i + 1\,.$$

The lemma follows, since the summation in the right-hand side maximizes the degree of θ in $f_\ell(\theta)$. □

Using lemmas (3.1) and (3.2) with combinatorial identities , we obtain the following theorem and its corollary about the forms of the EGFs H_ℓ.

Theorem 3.3. *The EGF of connected b-uniform hypergraphs with excess ℓ can be put into the form*

$$H_\ell = \frac{(1-\theta)^{\left\lfloor \frac{\ell+1}{b-1}+1 \right\rfloor}}{T^\ell} \sum_{p=0}^{3\,\ell} A_{\ell\,p} \left(\frac{1-\theta}{\theta}\right)^p$$

with the coefficients $A_{\ell\,p}$ being rational.

Corollary 3.4. *The EGF of connected b-uniform hypergraphs with excess $\ell \geq 1$ can be rewritten as*

$$H_\ell = \frac{1}{T^\ell} \sum_{j=-3\,\ell}^{\left\lfloor \frac{\ell+1}{b-1}+1 \right\rfloor} c_j(\ell, b)\,\theta^j\,,$$

where $c_j(\ell, b) \in \mathbb{Q}$.

The proofs of theorem 3.3 and corollary 3.4 are omitted in this extended abstract.

4. Asymptotic results

4.1. **Wright-like inequalities for hypergraphs.** In order to compute the asymptotic number of connected ℓ-excess hypergraphs of a given size, we need the following result which gives the first two terms of H_ℓ. Let us recall that $\theta = 1 - T^{b-1}/(b-2)!$.

Lemma 4.1. *Developing the two first coefficients of the partial fraction form of* H_ℓ*, we get for* $\ell \geq 1$

$$T(z)^\ell H_\ell(z) = \frac{\lambda_\ell (b-1)^{2\ell}}{3\,\ell\,\theta(z)^{3\ell}} - \frac{(\kappa_\ell - \nu_\ell(b-2))(b-1)^{2\ell-1}}{(3\,\ell-1)\,\theta(z)^{3\ell-1}} + \sum_{j=-3\ell+2}^{\left\lfloor \frac{\ell+1}{b-1}+1 \right\rfloor} c_j(\ell,b)\,\theta(z)^j \,. \tag{21}$$

In (21), $(\lambda_\ell)_{\ell\in\mathbb{N}}$ *is defined recursively by* $\lambda_0 = \frac{1}{2}$ *and*

$$\lambda_\ell \;=\; \frac{1}{2}\lambda_{\ell-1}(3\ell-1) \;+ \frac{1}{2}\sum_{t=0}^{\ell-1}\lambda_t\lambda_{\ell-1-t}\,, \qquad (\ell \geq 1)\,. \tag{22}$$

Similarly, define $(\nu_\ell)_{\ell\geq 1}$*,* $(\mu_\ell)_{\ell\geq 0}$ *and* $(\kappa_\ell)_{\ell\geq 1}$ *as follows:* $\nu_1 = \frac{5}{12}$ *and*

$$\begin{aligned} \nu_\ell \;&=\; \frac{1}{2}\lambda_{\ell-1} + \frac{1}{6}(3\ell-4)(3\ell-2)\lambda_{\ell-2} + \frac{1}{2}\sum_{t=0}^{\ell-2}(3t+2)\lambda_t\lambda_{\ell-2-t} \\ &+\; \frac{1}{6}\sum_{s=0}^{\ell-2}\sum_{t=0}^{\ell-2-s}\lambda_s\lambda_t\lambda_{\ell-2-s-t}\,(\ell\geq 2)\,. \end{aligned} \tag{23}$$

$$\kappa_\ell \;=\; \frac{1}{2}\left((3\ell-2)\mu_{\ell-1} + (3b\ell-b-2\ell)\,\lambda_{\ell-1}\right) + \sum_{t=0}^{\ell-1}\mu_t\lambda_{\ell-1-t}\,. \tag{24}$$

$\mu_0 = b-1$ *and for* $\ell \geq 1$*,* μ_ℓ *is given by*

$$\mu_\ell \;=\; \kappa_\ell - \nu_\ell(b-2) + \lambda_\ell(b-\frac{2}{3})\,, \qquad (\ell\geq 1)\,. \tag{25}$$

Sketch of proof. Use the differential equation (10) given in corollary 2.5 with corollary 3.4, mainly focusing on the ‘first two terms’ of H_ℓ after a bit of standard algebra we get (21).

We are now ready to state similar inequalities such as those obtained by Wright in [29]. If A and B are two formal power series such that for all $n \geq 0$ we have $[z^n]\,A(z) \leq [z^n]\,B(z)$ then we denote this relation $A \preceq B$ (or $A(z) \preceq B(z)$).

Lemma 4.2. *For any* $\ell \geq 1$*,* H_ℓ *satisfies*

$$\frac{\lambda_\ell(b-1)^{2\ell}}{3\,\ell\,T(z)^\ell\,\theta(z)^{3\ell}} - \frac{(\kappa_\ell - \nu_\ell(b-2))(b-1)^{2\ell-1}}{(3\,\ell-1)\,T(z)^\ell\,\theta(z)^{3\ell-1}} \preceq H_\ell(z) \preceq \frac{\lambda_\ell(b-1)^{2\ell}}{3\,\ell\,T(z)^\ell\,\theta(z)^{3\ell}}\,, \tag{26}$$

where $(\lambda_\ell)_{\ell\in\mathbb{N}}$*,* $(\kappa_\ell)_{\ell\in\mathbb{N}^\star}$ *and* $(\nu_\ell)_{\ell\in\mathbb{N}^\star}$ *are defined as in lemma 4.1.*

The proof of this lemma will be provided in the full paper.

The following lemma gives the order of magnitude of the two first coefficients of the partial fraction form of H_ℓ.

Lemma 4.3. *We have*

$$\lambda_\ell = 3\left(\frac{3}{2}\right)^\ell \frac{\ell!}{2\pi}\left(1 + O\left(\frac{1}{\ell}\right)\right)\,, \tag{27}$$

$$\Big|\,\kappa_\ell - \nu_\ell(b-2)\,\Big| = O\left(\ell\lambda_\ell\right)\,. \tag{28}$$

Proof. To prove (27), it suffices to remark that $\lambda_\ell = 3\,\ell b_\ell$ where the sequence (b_ℓ) corresponds to the Wright’s coefficients defined in [27, eq. (3.2)]. Therefore, by the proof of Lambert Meertens reported in [2] (see also Voblyĭ [26]), (27) holds. The remaining proof of (28) is technical and is omitted in this extended abstract. □

4.2. **A lemma from contour integration.** In order to get rid of the asymptotic behavior of the coefficients of $H_\ell(z)$, we need a last intermediate step. Define $h_n(a,\beta)$ as follows

$$\frac{1}{T(z)^a\,\left(1-\frac{T(z)^{b-1}}{(b-2)!}\right)^{3a+\beta}} = \sum_{n\geq 0} h_n(a,\beta)\frac{z^n}{n!}\,. \tag{29}$$

The following lemma is an application of the saddle point method [8, 11] which is well suited to cope with our analysis :

Lemma 4.4. *Let* $a \equiv a(n)$ *be such that* $a(b-1)\to 0$ *but* $\frac{a(b-1)n}{\ln n^2}\to\infty$ *and let* β *be a fixed number. Then* $h_n(an,\beta)$ *defined in (29) satisfies*

$$\begin{aligned} h_n(an,\beta) &= \frac{n!}{\sqrt{2\pi n\Big(b-1\Big)}\,\Big((b-1)!\Big)^{\frac{an+n}{b-1}}}\Big(1-(b-1)u_0\Big)^{(1-\beta)} \\ &\times\ \exp\left(n\Phi(u_0)\right)\left(1+O\Big(\sqrt{a(b-1)}\Big)+O\Big(\frac{1}{\sqrt{a(b-1)n}}\Big)\right)\,, \end{aligned} \tag{30}$$

where

$$\begin{aligned} \Phi(u) &= u-\left(\frac{a+1}{b-1}\right)\ln u - 3\,a\ln\left(1-(b-1)u\right) \\ u_0 &= \frac{3/2\,ab-a+1-1/2\,\sqrt{\Delta}}{b-1} \qquad \textit{with}\ \ \Delta = 9\,a^2b^2-12\,a^2b+12\,ab+4\,a^2-12\,a. \end{aligned} \tag{31}$$

Proof. Cauchy's integral formula gives

$$h_n(an,\beta) = n!\,[z^n]\,\frac{1}{T(z)^{an}\,\left(1-\frac{T(z)^{b-1}}{(b-2)!}\right)^{3a\,n+\beta}} = \frac{n!}{2\pi i}\oint \frac{1}{\left(T(z)\right)^{a\,n}\,\left(1-\frac{T(z)^{(b-1)}}{(b-2)!}\right)^{(3a\,n+\beta)}}\,\frac{dz}{z^{n+1}}\,. \tag{32}$$

Note that the radius of convergence of the series $T(z)$ is given by $\sqrt[(b-1)]{(b-2)!}\,\exp\left(-1/(b-1)\right)$. We make the substitution $u = T(z)^{(b-1)}/(b-1)!$ and get successively

$$\begin{aligned} T(z) &= \sqrt[(b-1)]{(b-1)!\,u}\,, \quad z = \sqrt[(b-1)]{(b-1)!\,u}\,e^{-u} \quad \text{and} \\ dz &= \left(\frac{1}{(b-1)\,u}-1\right)\,\Big((b-1)!\,u\Big)^{\frac{1}{(b-1)}}\,e^{-u}\,du\,. \end{aligned} \tag{33}$$

From (32), we then obtain

$$h_n(a\,n,\beta) = \frac{n!}{2\pi i\Big((b-1)!\Big)^{(a\,n+n)/(b-1)}}\oint \frac{\left(1-(b-1)u\right)^{1-\beta}}{(b-1)\,u}\,\exp\left(n\Phi(u)\right)du\,, \tag{34}$$

where $\Phi(u) = u-\left(\frac{a+1}{b-1}\right)\ln u - 3\,a\ln\left(1-(b-1)u\right)$. The big power in the integrand, *viz.* $\exp\left(n\Phi(u)\right)$, suggests us to use the saddle point method. Investigating the roots of $\Phi'(u)=0$, we find two saddle points, $u_0 = \frac{3/2\,ab-a+1-1/2\,\sqrt{\Delta}}{b-1}$ and $u_1 = \frac{3/2\,ab-a+1+1/2\,\sqrt{\Delta}}{b-1}$ with $\Delta = 9\,a^2b^2-12\,a^2b+12\,ab+4\,a^2-12\,a$ Moreover, we have $\Phi''(u) = \frac{a+1}{(b-1)u^2}+3\,\frac{a(-b+1)^2}{(1-(b-1)u)^2}$ so that for $u\notin\{0,1/(b-1)\}$, $\Phi''(u)>0$. The main point of the application of the saddle point method here is that $\Phi^{'}(u_0)=0$ and $\Phi^{''}(u_0)>0$, hence $n\Phi(u_0\exp{(i\tau)})$ is well approximated by $n\Phi(u_0)-n{u_0}^2\Phi^{''}(u_0)\frac{\tau^2}{2}$ in the vicinity of $\tau=0$. If we integrate (34) around a circle passing vertically through $u=u_0$ in the z-plane, we obtain

$$h_n(an,\beta) = \frac{n!}{2\pi\,\Big((b-1)!\Big)^{(an+n)/(b-1)}}\int_{-\pi}^{\pi}\frac{\Big(1-(b-1)u_0e^{i\tau}\Big)^{1-\beta}}{(b-1)}\,\exp\Big(n\Phi(u_0e^{i\tau})\Big)\,d\tau \tag{35}$$

where

$$\Phi(u_0 e^{i\tau}) = u_0 \cos\tau + i u_0 \sin\tau - \frac{a+1}{b-1}\ln u_0 - i\frac{a+1}{b-1}\tau - 3a\ln(1-(b-1)u_0 e^{i\tau})\ . \tag{36}$$

Denoting by $\mathfrak{Re}(z)$ the real part of z, if $f(\tau) = \mathfrak{Re}(\Phi(u_0 e^{i\tau}))$ we have

$$f(\tau) = u_0 \cos\tau - \frac{a+1}{b-1}\ln u_0 - 3a\ln u_0 - 3a\ln(b-1) - \frac{3a}{2}\ln\left(1 + \frac{1}{(b-1)^2 u_0^2} - \frac{2\cos\tau}{(b-1)u_0}\right). \tag{37}$$

It comes

$$f'(\tau) = \frac{d}{d\tau}\mathfrak{Re}(h(u_0 e^{i\tau})) = -u_0 \sin\tau - \frac{3a\sin\tau}{u_0(b-1) + \frac{1}{(b-1)u_0} - 2\cos\tau}\ . \tag{38}$$

Therefore, if $\tau = 0$ $f^{'}(\tau) = 0$. Also, $f(\tau)$ is a symmetric function of τ and in $[-\pi, -\tau_0] \cup [\tau_0, \pi]$, for any given $\tau_0 \in (0,\, \pi)$, and $f(\tau)$ takes its maximum value for $\tau = \tau_0$. Since $|\exp(\Phi(u))| = \exp(\mathfrak{Re}(\Phi(u)))$, when splitting the integral in (35) into three parts, viz. "$\int_{-\pi}^{-\tau_0} + \int_{-\tau_0}^{\tau_0} + \int_{\tau_0}^{\pi}$", we know that it suffices to integrate from $-\tau_0$ to τ_0, for a convenient value of τ_0, because the others can be bounded by the magnitude of the integrand at τ_0. In fact, we have

$$\Phi(u_0 e^{i\theta}) = \Phi(u_0) + \sum_{p\geq 2} \phi_p (e^{i\theta} - 1)^p \tag{39}$$

where $\phi_p = \frac{{u_0}^p}{p!}\Phi^{(p)}(u_0)$. We easily compute $\Phi^{(p)}(u_0) = (-1)^p (p-1)!\Big(\frac{a+1}{(b-1){u_0}^p} + \frac{3a(1-b)^p}{(1-(b-1)u_0)^p}\Big)$, for $p \geq 2$. Whenever $ab \to 0$, we have

$$(b-1)u_0 = 1 - \sqrt{3\,(b-1)\,a} + (3/2\,b - 1)\,a + O\left(b^{3/2} a^{3/2}\right)\ . \tag{40}$$

Therefore, we obtain after a bit of algebra

$$|\phi_p| \leq O\left(\frac{2^p}{a^{\frac{p}{2}-1}(b-1)^{\frac{p}{2}}}\right)\ , \qquad \text{as } a(b-1) \to 0\ . \tag{41}$$

On the other hand,

$$|e^{i\tau} - 1| = \sqrt{2(1-\cos\tau)} < \tau\ , \qquad \tau > 0\ . \tag{42}$$

Thus, the summation in (39) can be bounded for values of τ and a such that $\tau \to 0$, $ab \to 0$ ($a \to 0$) but $\frac{\tau}{\sqrt{a}} \to 0$ and we have

$$\Big|\sum_{p\geq 4} \phi_p (e^{i\tau} - 1)^p\Big| \leq \sum_{p\geq 4} |\phi_p \tau^p| \leq \sum_{p\geq 4} O\Big(\frac{2^p \tau^p}{a^{\frac{p}{2}-1}(b-1)^{\frac{p}{2}}}\Big) = O\Big(\frac{\tau^4}{a(b-1)}\Big)\ . \tag{43}$$

It follows that for $\tau \to 0$, $a(b-1) \to 0$ and $\frac{\tau}{\sqrt{a(b-1)}} \to 0$, $\Phi(u_0 e^{i\tau})$ can be rewritten as

$$\begin{aligned}\Phi(u_0 e^{i\tau}) &= \Phi(u_0) - \frac{1}{(b-1)}\left(1 - \frac{\sqrt{a}}{\sqrt{3(b-1)}}\,\frac{3b-4}{2} + \frac{(9b^2 - 12b + 4)}{12(b-1)}a\right)\tau^2 \\ &- \frac{i}{(b-1)}\left(1 - \frac{(3b-4)\sqrt{a}}{2\sqrt{3(b-1)}} + \frac{(9b^2-12b+4)}{12(b-1)}\,a\right)\tau^3 + O\left(\frac{\tau^4}{a(b-1)}\right)\ .\end{aligned} \tag{44}$$

Therefore, if $a(b-1) \to 0$ but $\frac{a(b-1)n}{(\ln n)^2} \to \infty$, if we let $\tau_0 = \frac{\ln n}{\sqrt{n\, u_0^2 \Phi''(u_0)}}$ (with $u_0^2\Phi''(u_0) = \frac{2}{b-1} + O(\sqrt{a(b-1)})$) we can remark (as already said) that it suffices to integrate (35) from $-\tau_0$ to τ_0, using the magnitude of the integrand at τ_0 to bound the resulting error. In fact,

$$\begin{aligned}&\Big|(1-(b-1)u_0 e^{i\tau_0})^{(1-\beta)} \exp\left(n\Phi(u_0 e^{i\tau_0}) - n u_0 + \frac{n(a+1)}{(b-1)}\ln u_0 + 3an\ln\left(1-(b-1)u_0\right)\right)\Big| = \\ &\Big|1-(b-1)u_0 e^{i\tau_0}\Big|^{(1-\beta)} \exp\Big(-\frac{n}{2}u_0^2\Phi''(u_0)\,{\tau_0}^2 + O\Big(n\frac{{\tau_0}^4}{a(b-1)}\Big)\Big) = O\Big(e^{-\frac{(\ln n)^2}{2}}\Big)\ .\end{aligned} \tag{45}$$

The rest of the proof is now standard application of the saddle point method (see for instance De Bruijn [8, Chapters 5 & 6]) and is omitted in this extended abstract. After a bit of algebra, one gets the formula (30). ☐

4.3. **Asymptotic number of connected hypergraphs.** We are now ready to state the main result of this section.

Theorem 4.5. *For* $\ell \equiv \ell(n)$ *such that* $\ell = o\left(\sqrt[3]{\frac{n}{b}}\right)$ *as* $n \to \infty$, *the number of connected* b*-uniform hypergraphs built with* n *vertices and having excess* ℓ *satisfies*

$$(46)\quad \sqrt{\frac{3}{2\,\pi}}\quad \frac{\Big(b-1\Big)^{\frac{\ell}{2}} \quad e^{\frac{\ell}{2}} \quad n^{n+\frac{3\,\ell}{2}-\frac{1}{2}}}{12^{\frac{\ell}{2}}\,\ell^{\frac{\ell}{2}}\,\Big((b-2)!\Big)^{\frac{n+\ell}{b-1}}} \exp\left(\frac{n}{b-1}-n\right)\left(1+O\left(\frac{1}{\sqrt{\ell}}\right)+O\left(\sqrt{\frac{b\,\ell^3}{n}}\right)\right).$$

We urge the reader to compare the methods and results obtained by Karoński and Łuczak in [20] with ours. In particular, the authors of [20] obtained results concerning various kinds of hypergraphs (smooth hypergraphs, clean hypergraphs, etc.). Unlike their results, where the excesses are of order $o(\log n/\log\log n)$, the theorem above states that the three variables n, ℓ and b can tend together to infinity but (46) remains valid whenever $\ell = o\left(\sqrt[3]{\frac{n}{b}}\right)$. Note also that by setting $b = 2$ in (46), we retrieve Wright's formula for graphs [29]. We remark also that the powerful methods developed in [2] and in [24] can be used to extend the validity of our asymptotic result.

Acknowledgments

The authors would like to thank warmly Gerard Duchamp (LIPN – University of Paris-Nord) and Jean-Gabriel Luque (IGM – Marne la Vallée) for providing proofs of the identities used to prove theorem 3.3 and its corollary. They would like to thank also Michal Karoński for useful discussions about [20] during the AofA meeting in Gdansk (2000).

References

[1] Bender E. A. (1974). Asymptotic Methods in Enumeration, *SIAM Review*, 16:485–515.

[2] Bender, E. A., Canfield, E. R. and McKay B. D. (1990). The asymptotic number of labeled connected graphs with a given number of vertices and edges. *Random Structures and Algorithms*, 1:127–169.

[3] Bender, E. A., Canfield, E. R. and McKay B. D. (1992). Asymptotic properties of labeled connected graphs. *Random Structures and Algorithms*, 3:183–202.

[4] Berge, C. (1976). *Graphs and hypergraphs* North-Holland Mathematical Library.

[5] Bergeron, F., Labelle, G. and Leroux, P. (1998). *Combinatorial Species and Tree-like Structures* Encyclopedia of Mathematics and its Applications, Vol. 67, Cambridge Univ. Press.

[6] Bollobás, B. (1985) *Random Graphs* Academic Press, London.

[7] Bouwkamp and De Bruijn *Hypergraphs and a functional equation* Journal of Combinatorial Theory, Ser.

[8] De Bruijn, N. G. (1981). *Asymptotic Methods in Analysis.* Dover, New-York.

[9] Cayley, A. (1889). A Theorem on Trees. *Quart. J. Math. Oxford Ser.*, 23:376–378.

[10] Flajolet, P., Knuth, D. E. and Pittel B. (1989). The first cycles in an evolving graph. *Discrete Math.*, 75:167–215.

[11] Flajolet, P. and Sedgewick, R. *Analytic Combinatorics.* To appear (chapters are available as Inria research reports). See *http : //algo.inria.fr/flajolet/Publications/books.html.*

[12] Flajolet, P. and Sedgewick, R. *The average case analysis of algorithms: Counting and generating functions*

[13] Goulden, I. P. and Jackson, D. M. (1983). *Combinatorial Enumeration.* Wiley, New York.

[14] Harary, F. and Palmer, E. (1973). *Graphical Enumeration.* Academic Press, New-York and London.

[15] Herbert S. Wilf *generatingfunctionology*

[16] Janson, S. (1993). Multicyclic components in random graphs process. *Random Structures and Algorithms*, 4:71–84.

[17] Janson, S. (2000). Growth of components in random graphs. *Random Structures and Algorithms*, 17:343-356.

[18] Janson, S., Knuth, D. E., Łuczak, T. and Pittel B. (1993). The birth of the giant component. *Random Structures and Algorithms*, 4:233–358.

[19] Janson, S., Łuczak, T. and Rucinski A. (2000). *Random Graphs.* John Wiley, New York.

[20] Karoński, M. and Łuczak, T. (1997). The number of sparsely edges connected uniform hypergraphs. *Discrete Math.* 171:153–168.

[21] Karoński, M. and Łuczak, T. and (2002). The phase transition in a random hypergraph. *J. Comput. Appl. Math.* 142:125–135.

[22] Knuth, D. E. and Pittel, B. (1989). A recurrence related to trees. *Proc. Am. Math. Soc.*, 105:335–349.

[23] Rényi, A. (1959). On connected graphs I. *Publ. Math. Inst. Hungarian Acad. Sci.* 4:385–388.

[24] Pittel, B. and Wormald N. (2003). Counting connected graphs inside-out. *J. Comb. Th. Series B*, to appear.

[25] Selivanov, B. I. (1972). Perechislenie odnorodnykh hipergrafov c protstoĭ ciklovoĭ strukturoĭ, Kombinatornyĭ Analiz 2, pages 60–67.

[26] Voblyĭ, V. A. O koeffitsientakh Raĭta i Stepanova-Raĭta *Matematicheskie Zametki* 42 (1987), 854–862. English translation, V. A. Voblyi, "Wright and Stepanov-Wright coefficients," *Mathematical Notes* **42** (1987), 969–974.

[27] Wright, E. M. (1977). The Number of Connected Sparsely Edged Graphs. *Journal of Graph Theory*, 1:317–330.

[28] Wright, E. M. (1978). The Number of Connected Sparsely Edged Graphs. II. Smooth graphs and blocks *Journal of Graph Theory*, 2:299–305.

[29] Wright, E. M. (1980). The Number of Connected Sparsely Edged Graphs. III. Asymptotic results *Journal of Graph Theory*, 4:393–407.

E-mail address: `{tsr,vlad}@lipn.univ-paris13.fr`

Tsiry ANDRIAMAMPIANINA, LIPN – UMR CNRS 7030, Université de Paris-Nord, F 93430 Villetaneuse, France.

Vlady RAVELOMANANA, LIPN – UMR CNRS 7030, Université de Paris-Nord, F 93430 Villetaneuse, France.