On Hb-graphs and their Application to General Hypergraph e-adjacency Tensor*

Xavier Ouvrard
CERN, 1 Esplanade des Particules,
1211 MEYRIN, Switzerland
University of Geneva
xavier.ouvrard@cern.ch
Jean-Marie Le Goff
CERN, 1 Esplanade des Particules,
1211 MEYRIN, Switzerland
jean-marie.le.goff@cern.ch
Stéphane Marchand-Maillet
University of Geneva, CUI,
Battelle (Bat A) - Route de Drize, 7
1227 Carouge, Switzerland
stephane.marchand-maillet@unige.ch

Abstract

Working on general hypergraphs requires to properly redefine the concept of adjacency in a way that it captures the information of the hyperedges independently of their size. Coming to represent this information in a tensor imposes to go through a uniformisation process of the hypergraph. Hypergraphs limit the way of achieving it as redundancy is not permitted. Hence, our introduction of hb-graphs, families of multisets on a common universe corresponding to the vertex set, that we present in details in this article, allowing us to

*This work is part of the PhD of X. OUVRARD (CERN/UniGe), co-supervised by Pr. S. MARCHAND-MAILLET (UniGe) and Dr. J.M. LE GOFF (CERN); it has been funded by a doctoral position at CERN. The authors are very grateful to L. TVEDE (CERN) for her careful English review of this article.

have a construction of adequate adjacency tensor that is interpretable in term of m-uniformisation of a general hb-graph. As hypergraphs appear as particular hb-graphs, we deduce two new (e-)adjacency tensors for general hypergraphs. We conclude this article by giving some first results on hypergraph spectral analysis of these tensors and a comparison with the existing tensors for general hypergraphs, before making a final choice.

1 Introduction

Hypergraphs were introduced in [4]. Hypergraphs are defined as a family of nonempty subsets—called hyperedges—of a set of vertices. Elements of a set are unique. Hence elements of a given hyperedge are also unique in a hypergraph. Hypergraphs fit to model collaboration networks—Newman [32, 31]—, co-author networks—Grossman and Ion [16], Taramasco et al. [53]—, chemical reactions —Temkin et al. [55]—, genome—Chauve et al. [7]—, VLSI design—Karypis et al. [21]—and other applications. More generally, hypergraphs perfectly preserve entities grouping information, since they succeed in capturing p-adic relationships. In Berge and Minieka [4], Stell [50] and Bretto [6] hypergraphs are defined differently. In this article, the definition of [6] is used, as it does not impose the union of the hyperedges to cover the vertex set.

Multisets extend sets by allowing duplication of elements. As mentioned in Singh et al. [49], N.G. de Bruijn proposed to Knuth the terminology multiset in replacement of a variety of existing terms, such as bag or weighted set. Multisets are used in database modelling: in Albert [1] relational algebra extension were introduced to manipulate bags—see also Klug [22]—by studying bag algebraic properties. Queries for such bags have been largely studied in a series of articles—see references in Grumbach et al. [17]: bags prevent the costly operation of duplicate search. In Hernich and Kolaitis [19], information integration under bag semantics is studied as well as the tractability of some algorithmic problems: they showed that the GLAV (Global-And-Local-As-View) mapping of two databases problem becomes untractable over such semantic. Multisets are also used in P-computing in the form of labelled multiset, called membrane—see Păun [41] for more details.

Taking advantage of this duplication permissiveness, we construct an extension of hypergraphs called hyper-bag-graphs (hb-graphs for short). There are three main reasons to such an extension. The first one is that multisets are extensively used in databases as they allow the presence of duplicates—Lamperti et al. [24]—removing duplicates (and thus obtaining sets and hypergraphs) being a costly computing operation. The second is that natural hb-graphs —hb-graphs based on multisets with non-negative integer multiplicity values— allow results on the e-adjacency tensor[1], not only for hb-graphs themselves but also, as a special case, for hypergraphs, allowing existing tensors—Banerjee et al. [2], Ouvrard et al. [36], Sun et al. [51]—to have alternatives which have meaningful interpretation to the steps taken during their constructions via the hb-graph uniformisation process. The third reason is linked to the first: allowing vertex multiplicity to be specific to the structure they belong is required by many applications, such as ranking of words by random walks—Bellaachia and Al-Dhelaan [3]—, bag of words for text —Harris [18]—, bag of visual words—Peng et al. [45]—, bag of features for image classification—Nowak et al. [33]—, bag of patterns for finding similarities in time series—Lin et al. [26], Lin and Li [27]—, bag of entities for ranking—Xiong et al. [56]—. We already applied these concepts to define a diffusion by exchange inside hb-graphs to rank not solely vertices but also hb-edges— Ouvrard et al. [39]—.

Section 2 recalls the essential results needed on multisets and then introduces different algebraic representations of natural multisets that will be needed later for the algebraic description of hb-graphs. Section 3 develops a mathematical construction of Hyper-Bag-Graphs (or hb-graphs). Section 4 gives an algebraic description of hb-graphs and draws consequences for the e-adjacency tensor of hypergraphs. Section 5 gives results on the constructed tensors. Section 6 evaluates the constructed tensors and proceeds to a final choice on the hypergraph e-adjacency tensor. Section 7 concludes this article and provides some indications on future work.

2 On multisets

We start by reminding key points on multisets, and introduce afterward two algebraic representations of multisets that will be used in hb-graphs.

[1]We often misuse the word tensor for its hypermatrix representation in a canonical basis in the verbatim of this article, following the abuse made in Qi and Luo [47]. Nonetheless, when writing it mathematically, we make the difference between the tensor $\mathcal{A}$ of dimension n and rank r and its canonical hypermatrix representation, written $\boldsymbol{A} = (a_{i_1 \dots i_r})$ as defined in Subsection 4.2.5. For a full introduction on tensors one can refer for instance to [25]. Only elements of $\mathcal{L}_k^0(V) = \underbrace{V \otimes \dots \otimes V}_{k\text{times}}$ admits the restricted approach that is presented in Section 3.

2.1 Generalities

The concept of multisets is known from ancient times. Knuth [23] mentioned that N.G. de Bruijn coined the word "***multiset***" to designate structures that were previously called "list, bunch, bag, heap, sample, weighted set, collection" in the litterature. In Blizard et al. [5], the concept of multiset is traced back to the very origin of numbers, when numbers were represented by repeating occurrences of symbols. In Knuth [23] and in Patwardhan et al. [40], the first known usage of multisets is attributed to the Indian mathematician Bhaskaracharya around 1150 in the Lilavati manuscript where the permutations of the multiset $\{4, 5, 5, 5, 8\}$ are listed. Multisets were also briefly used by Dedekind in the last two paragraphs of his essay of 1888 "Was sind und was sollen die Zahlen"—Dedekind [12]. In Blizard et al. [5], a full historical approach of multisets is achieved: this approach cites Knuth [23] where the algebra of multisets is put in correspondance with the multiplicative theory of positive integers. Blizard et al. [5] also cite the usage of multisets by Weierstrass for the construction of real numbers. The authors give interesting philosophical aspects of multisets that allow multiplicity without diversity since occurrences of a single element in a multiset cannot be distinguished; this opposes the statements of Frege and Leibniz about the diversity associated with numbers, as exact identity has to be unicity. This point of view can give interesting development on multisets as introduced by Syropoulos [52].

In Blizard et al. [5], the authors start by giving the following naive concept of multiset before building the theory MST of multisets:

"The naive concept of multiset that we now formalize has the following properties: (i) a multiset is a collection of elements in which certain elements may occur more than once; (ii) occurrences of a particular element in a multiset are indistinguishable; (iii) each occurrence of an element in a multiset contributes to the cardinality of the multiset; (iv) the number of occurrences of a particular element in a multiset is a (finite) positive integer; (v) the number of distinguishable (distinct) elements in a multiset need to be finite; and (vi) a multiset is completely determined if we know which elements belong to it and the number of times each element belongs to it."

Nonetheless, the requirement (v) on the finitude is not necessary, and relying on Singh et al. [49], we will give the definitions on multisets without this constraint and this will hold in the remaining of this article. We consider a countable set A of distinct objects and a subset $\mathbb{W} \subseteq \mathbb{R}^+$. We consider m an application from A to $\mathbb{W}$. Then $\mathfrak{A}_m \triangleq (A, m)$[2] is called a

[2]We systematically use fraktur font to designate multisets or mathematical objects

multiset—or **mset** or **bag**—on A. A is called the **ground** or the **universe** of the multiset $\mathfrak{A}_m$, m is called the **multiplicity function** of the multiset $\mathfrak{A}_m$. $\mathfrak{A}_m^\star \triangleq \{x \in A : m(x) \neq 0\}$ is called the **support** of $\mathfrak{A}_m$. The elements of the support of a mset are called its **generators**. A multiset where $\mathbb{W} \subseteq \mathbb{N}$ is called a **natural multiset**.

We write $\mathfrak{M}(A)$ the set of all multisets of universe A. Some extensions of multisets exist where the multiplicity function can have its range in $\mathbb{Z}$—called hybrid set in Loeb [28]. Some other extensions exist like fuzzy multisets—Syropoulos [52]. Most definitions when not restricted to natural multisets can extend to $\mathbb{W} \subseteq \mathbb{R}$.

Several notations of msets exist. One common notation which we will use is: $\mathfrak{A}_m = \{x_i^{m_i} : i \in [\![n]\!] \wedge x_i \in A\}$ with $A = \{x_i : i \in [\![n]\!]\}$ the universe of the mset $\mathfrak{A}_m$ and the notation $\forall i \in [\![n]\!] : m_i = m(x_i)$.

Another useful notation for a natural multiset is the one similar to an unordered list: $\mathfrak{A}_m = \left\{\!\!\left\{ \underbrace{x_1, \ldots, x_1}_{m_1 \text{ times}}, \ldots, \underbrace{x_n, \ldots, x_n}_{m_n \text{ times}} \right\}\!\!\right\}$.

In this last representation, we do not write the elements that are not in the support, also the universe has to be clearly stated[3].

The **m-cardinality** $\#_m\mathfrak{A}_m$ of a mset $\mathfrak{A}_m$ corresponds to the sum of the multiplicities of the elements of its universe: $\#_m\mathfrak{A}_m \triangleq \sum_{x \in A} m(x)$ while the **cardinality** $\#\mathfrak{A}_m$ of a mset $\mathfrak{A}_m$ is the number of elements of its support: $\#\mathfrak{A}_m = |\mathfrak{A}_m^\star|$.

There exists only one multiset of universe A with an empty support that is called the **empty multiset** or also the **trivial multiset** of universe A and written $\emptyset_A$.

Different operations can be defined on the set of all multisets of the same universe. We consider up to the end of this subsection two msets $\mathfrak{A} = (U, m_{\mathfrak{A}})$ and $\mathfrak{B} = (U, m_{\mathfrak{B}})$ on the same universe U. We define different operations such as inclusion, equality, union, intersection, sum,...

The **inclusion** of $\mathfrak{A}$ in $\mathfrak{B}$—written $\mathfrak{A} \subseteq \mathfrak{B}$—if for all $x \in U$:

$$m_{\mathfrak{A}}(x) \leqslant m_{\mathfrak{B}}(x).$$

$\mathfrak{A}$ is then called a **submset** of $\mathfrak{B}$ and said **included** in $\mathfrak{B}$.

If $\mathfrak{A} \subseteq \mathfrak{B}$ and $\mathfrak{B} \subseteq \mathfrak{A}$, then $\mathfrak{A}$ and $\mathfrak{B}$ are said **equal**.

that involves multisets.

[3]When the universe is made clear by the context, we can abusively write: $\mathfrak{A}_m = \{x_i^{m_i} : i \in [\![n]\!] \wedge x_i \in A \wedge m_i \neq 0\}$.

The **union** of $\mathfrak{A}$ and $\mathfrak{B}$ is the mset $\mathfrak{C} \triangleq \mathfrak{A} \cup \mathfrak{B}$ of universe U and of multiplicity function $m_{\mathfrak{C}}$ such that for all $x \in U$:

$$m_{\mathfrak{C}}(x) \triangleq \max\left(m_{\mathfrak{A}}(x), m_{\mathfrak{B}}(x)\right).$$

The **intersection** of $\mathfrak{A}$ and $\mathfrak{B}$ is the mset $\mathfrak{D} \triangleq \mathfrak{A} \cap \mathfrak{B}$ of universe U and of multiplicity function $m_{\mathfrak{D}}$ such that for all $x \in U$:

$$m_{\mathfrak{D}}(x) \triangleq \min\left(m_{\mathfrak{A}}(x), m_{\mathfrak{B}}(x)\right).$$

The **sum** of $\mathfrak{A}$ and $\mathfrak{B}$ is the mset $\mathfrak{S} \triangleq \mathfrak{A} \uplus \mathfrak{B}$ of universe U and of multiplicity function $m_{\mathfrak{S}}$ such that for all $x \in U$:

$$m_{\mathfrak{S}}(x) \triangleq m_{\mathfrak{A}}(x) + m_{\mathfrak{B}}(x).$$

It is worth reminding that $\cup$, $\cap$ and $\uplus$ are commutative and associative laws on msets of the same universe. The empty mset of same universe corresponds to the identity for these operations. $\uplus$ is distributive for $\cup$ and $\cap$. $\cup$ and $\cap$ are distributive one for the other. $\cup$ and $\cap$ are idempotent.

The **difference** of two msets is the mset $\mathfrak{T} \triangleq \mathfrak{A} - \mathfrak{B}$ of universe U and of multiplicity function $m_{\mathfrak{T}}$ such that for all $x \in U$: $m_{\mathfrak{T}}(x) \triangleq m_{\mathfrak{A}}(x) - m_{\mathfrak{A}\cap\mathfrak{B}}(x)$. We cannote that the classical property for sets: $(\mathfrak{A} - \mathfrak{B}) \cap \mathfrak{B} = \emptyset_U$ does not hold for multisets—see Singh et al. [49] for an example.

The **complementation** of $\mathfrak{A}$ referring to a family $\mathfrak{R}$ of multisets $(\mathfrak{A}_i)_{i\in I}$ of universe U such that $\mathfrak{A} \in \mathcal{G}$, is the multiset $\overline{\mathfrak{A}}$ of universe U and of multiplicity function $m_{\overline{\mathfrak{A}}}$ such that for all $x \in U$: $m_{\overline{\mathfrak{A}}}(x) \triangleq \max_{i\in I}\left(m_{\mathfrak{A}_i}(x)\right) - m_{\mathfrak{A}}(x)$.

Finally the **power set** of a multiset $\mathfrak{A}_m$—written $\widetilde{\mathcal{P}}(\mathfrak{A}_m)$—is defined as the set of all submsets of $\mathfrak{A}_m$.

We define the **fusion** of two msets $\mathfrak{A}$ of universe U and $\mathfrak{B}$ of universe V as the multiset $\mathring{\mathfrak{F}} = \mathfrak{A} \oplus \mathfrak{B}$ of universe $U \cup V$ and of multiplicity function $m_{\mathring{\mathfrak{F}}}$ such that for all $x \in U \cup V$: $m_{\mathring{\mathfrak{F}}}(x) \triangleq \begin{cases} m_{\mathfrak{A}}(x) & \text{if } x \in U \backslash V \\ m_{\mathfrak{A}}(x) + m_{\mathfrak{B}}(x) & \text{if } x \in U \cap V \\ m_{\mathfrak{B}}(x) & \text{if } x \in V \backslash U. \end{cases}$

An interesting alternative approach to define multisets is the one given in Syropoulos [52], where a natural multiset $\mathfrak{A}_m = (A, m)$ is viewed as a couple $\langle A_0, \rho\rangle$, where A_0 is the set of instances of elements of A, that includes copies of elements, and ρ is an equivalency relation ρ over A_0:

$$\forall x \in A_0, \forall x' \in A_0 : \ x\rho x' \Leftrightarrow \exists! c \in A : x\rho c \wedge x'\rho c.$$

Two elements of A_0 such that: $x\rho x'$ are said **copies** of one another. 'he unique $c \in A$ is called the original element. x and x' are said copies of

Also A_0/ρ is isomorphic to A and:

$$\forall \overline{x} \in A_0/\rho, \exists! c \in A : |\{x : x \in \overline{x}\}| = m(c) \wedge \forall x \in \overline{x} : x\rho c.$$

The set A_0 is then called a **copy-set** of the multiset $\mathfrak{A}_m$.

We can remark that a copy-set for a given multiset is not unique. Sets ,f equivalency classes of two couples $< A_0, \rho >$ and $< A'_0, \rho' >$ of a given nultiset are isomorphic.

?.2 Algebraic representation of a natural multiset

n this subsection, we propose two algebraic representations of a natural nultiset $\mathfrak{A}_m = (A, m)$ of countable universe $A = \{x_i : i \in [\![n]\!]\}$ and multi- plicity function m.

Vector representation of a natural multiset

A multiset $\mathfrak{A}_m$ can be conveniently represented by a vector of length the cardinality of the universe and where the components of the vector represent the multiplicity of the corresponding element. We assume that the elements of A are given in a fixed order—it is always possible to index these elements by a subset of the positive integer set. Hence $\overrightarrow{\mathfrak{A}_m} \triangleq (m(x_i))_{x_i \in A}$ is called a **vector representation of the multiset** $\mathfrak{A}_m$. This representation requires $|A|$ space and has $|A| - |\mathfrak{A}^\star_m|$ null elements.

The sum of the elements of $\overrightarrow{\mathfrak{A}_m}$ is $\#_m\mathfrak{A}_m$.

This representation will be useful later when considering family of multisets in order to build the incident matrix of a hb-graph.

An alternative representation is obtained using a symmetric hypermatrix. This approach is needed to reach our goal of constructing an e-adjacency tensor for general hb-graphs and hypergraphs.

Hypermatrix representations of a natural multiset

The **unnormalized hypermatrix representation of the natural multiset** $\mathfrak{A}_m = (A, m)$ is the symmetric hypermatrix $\boldsymbol{A_u} \triangleq (a_{u,i_1 \dots i_r})_{i_1,\dots,i_r \in [\![n]\!]}$

of order $r = \#_m\mathfrak{A}_m$ and dimension n such that $a_{u,i_1\ldots i_r} = 1$ if $\forall j \in [\![r]\!]$: $i_j \in [\![n]\!] \wedge x_{i_j} \in \mathfrak{A}_m^\star$. The other elements are null.

Hence the number of non-zero elements in $\boldsymbol{A_u}$ is $\dfrac{r!}{\prod\limits_{x\in\mathfrak{A}_m^\star} m(x)}$ out of the n^r elements of the representation.

The sum of the elements of $\boldsymbol{A_u}$ is then: $\dfrac{r!}{\prod\limits_{x\in\mathfrak{A}_m^\star} m(x)}$.

To normalize $\boldsymbol{A_u}$, we enforce the sum of the elements of the hypermatrix to be the m-rank of the multiset it encodes.

The **normalized hypermatrix representation of the multiset** $\mathfrak{A}_m$ is the symmetric hypermatrix $\boldsymbol{A} \triangleq (a_{i_1\ldots i_r})_{i_1,\ldots,i_r\in[\![n]\!]}$ of order $r = \#_m\mathfrak{A}_m$ and dimension n such that $a_{i_1\ldots i_r} = \dfrac{\prod\limits_{x\in\mathfrak{A}_m^\star} m(x)}{(r-1)!}$ if $\forall j \in [\![r]\!] : i_j \in [\![n]\!] \wedge x_{i_j} \in \mathfrak{A}_m^\star$. The other elements are null.

3 Hb-graphs

In this section, we extend the work on hyper-bag-graphs—hb-graphs for short—introduced in Ouvrard et al. [39]. Hb-graphs extend hypergraphs since families of subsets of a vertex set are replaced by families of msets on a given universe. We start by reviewing the related work that gives a patchwork of elements that are first steps toward hb-graphs, and then expose the concept of hb-graphs as a family of multisets on a given universe; we tackle particularly families of natural multisets on a given universe, where elements are seen as repetition of a given element. [15] express this need in real datasets, where two physical objects can be seen "as the same or equal, if they are indistinguishable, but possibly separate, and identical if they physically coincide". We revisit systematically the common definitions and properties of hypergraphs given in Bretto [6] to extend them to hb-graphs. We give some first applications.

3.1 Related work

Handling structures similar to hypergraphs but having a "hyperedge-based" weighting of vertices occurs quite a few times in the literature, for instance [3], where the authors show that the weighting of vertices at the level of the hyperedges in a hypergraph provides better information retrieval using a modified random walk. They define a vector of weights for each vertex making weights of vertices hyperedge-based.

In [15], the authors define a multiset topology considering a collection of multisets in the power set of a given multiset. The power set of a given multiset is defined as the support of the power mset of an m-set: the power mset of an mset corresponds to the mset of all submsets of that multiset (which implies redundancy). They then study the properties of these multiset topologies. It is a strong background for our work, but multiset topologies include all submset of a given collection in the collection. Multiset topology has to be seen as foundations for potential extension of simplicial complexes.

In [20]—at the same time we were introducing hb-graphs in [38]—the authors consider a hypergraph where the hyperedges are multisets, transforming the initial definition of hypergraphs, to extend the Cheng-Lu model to hypergraphs, to achieve clustering via hypergraph modularity. They use a family of multisets and define the degree of a vertex as the sum of multiplicities and the size of a hyperedge as the sum of the multiplicities of its elements. They obtain good results with their proposed modularity getting a smaller number of hyperedges cut compared to the one achieved with the 2-section of the hypergraph.

In [9], published after [39], a hypergraph with hyperedge-dependent vertex weights is defined by considering a quadruple $\mathcal{H} = (V, E, \omega, \gamma)$ where ω is the edge weight vector, and γ is refined in a weight $\gamma_e(v)$ for every hyperedge $e \in E$. The authors are then using implicitly multisets, but without considering the related algebra. In a recent paper [44], the authors introduce a continuous incident matrix for multimedia retrieval, which is similar to our hb-graph incident matrix.

Finaly, we discovered only recently, due to the polysemy used in the article, that in [42] the authors introduce what they call ([Author's note: PZ-])multi-hypergraphs[4] using multisets, allowing repetitions of vertices in the hyperedges, but where two hyperedges cannot be duplicates. ([Author's note: PZ-])Multi-hypergraphs are a particular case of hb-graphs as we define them in this section. When the hyperedges are all of same cardinality k, the PZ-multi-hypergraph is said k-uniform, and called k-multigraph[5].

[4] Multi-hypergraph is in fact polysemic: multi-hypergraphs originally represent hypergraphs $\mathcal{H} = (V, E)$ where the repetition of hyperedges is authorised in E i.e. E is considered as a family Bretto [6] or a multiset Chazelle and Friedman [8], which is the direct extension of multi-graph.

[5] It is also a polysemy: in Majcher [29], k-multigraphs are multi-graphs—where multiple edges between a couple of vertices can occur—which are also k-graphs—i.e. graphs that are k-regular.

We think that a clear distinction should be made between hypergraphs that are collection of subsets of a vertex set and stuctures that involves either explicitly or implicitly multisets. Multisets require specific operations; moreover multisets break a certain amount of rules, like the Morgan laws, and defining complementation in multisets is not straightforward. As mentioned in [15], the Cantor's theorem is also broken with power set of multisets. In particular, a hyperedge-based weighting of vertices leads to many problems when it comes to intersect or gather hyperedges. Hence, the need of defining properly a structure with a specific vocabulary that allows to make a clear distinction between hypergraphs and this new structure that involves multisets.

3.2 Generalities

Let $V = \{v_i : i \in [\![n]\!]\}$ be a nonempty finite set.

A **hyper-bag-graph** or **hb-graph** for short over V is defined in Ouvrard et al. [39] as a family of msets over a universe V with support a subset of V. The msets are called the **hb-edges** and the elements of V the **vertices**.

We consider for the remainder of the article a hb-graph $\mathfrak{H} = (V, \mathfrak{E})$, with $V = \{v_i : i \in [\![n]\!]\}$ and $\mathfrak{E} = (\mathfrak{e}_j)_{j \in [\![p]\!]}$ the family of its hb-edges.

Each hb-edge $\mathfrak{e}_j \in \mathfrak{E}$ is of universe V and has a multiplicity function associated to it: $m_{\mathfrak{e}_j} : V \to \mathbb{W}$ where $\mathbb{W} \subseteq \mathbb{R}^+$. When the context is clear, the notation m_j is used for $m_{\mathfrak{e}_j}$ and m_{ij} for $m_{\mathfrak{e}_j}(v_i)$.

A hb-graph is said to be with **no repeated hb-edges** if:

$$\forall j_1 \in [\![p]\!], \forall j_2 \in [\![p]\!] : \mathfrak{e}_{j_1} = \mathfrak{e}_{j_2} \Rightarrow j_1 = j_2.$$

A hb-graph where each hb-edge is a natural mset is called a **natural hb-graph**.

The **empty hb-graph** is the hb-graph with an empty set of vertices. A **trivial hb-graph** is a hb-graph with a non empty set of vertices and an empty family of hb-edges.

For a general hb-graph, each hb-edge has to be seen as a weighted system of vertices, where the weights of each vertex are hb-edge dependent. In a natural hb-graph the multiplicity function can be viewed as a duplication of the vertices.

In Pearson and Zhang [42], the authors have introduced what they call abusively multi-hypergraphs where the hyperedge set is constituted of multisets of vertices. It corresponds to **natural hb-graphs with no repeated**

-edges, name that we keep, as multi-hypergraphs are hypergraphs where ; subsets of vertices the hyperedges correspond to can be repeated, i.e. ıstitute either a multiset of subsets of vertices—Chazelle and Friedman —or a family of subsets of vertices—Bretto [6].

The **order** of a hb-graph $\mathfrak{H}$—written $O(\mathfrak{H})$—is given by:

$$O(\mathfrak{H}) \triangleq \sum_{i \in [\![n]\!]} \max_{\mathfrak{e} \in \mathfrak{E}} (m_{\mathfrak{e}}(v_i)).$$

The **size** of a hb-graph corresponds to the number of its hb-edges.

If: $\bigcup_{j \in [\![p]\!]} \mathfrak{e}_j^\star = V$ then the hb-graph is said with no isolated vertices. therwise, the elements of $V \backslash \bigcup_{j \in [\![p]\!]} \mathfrak{e}_j^\star$ are called the **isolated vertices**. hey correspond to elements of hb-edges with zero-multiplicity for all hb-lges.

A **hypergraph** is a natural hb-graph where the hb-edge vertices have ıe as multiplicity for any vertex of their support—and by definition, zero ; multiplicity for the ones not in the support.

The **support hypergraph** of a hb-graph $\mathfrak{H} = (V, \mathfrak{E})$ is the hypergraph hose vertices are the ones of the hb-graph and whose hyperedges are the ıpport of the hb-edges. We write $\underline{\mathfrak{H}} \triangleq (V, E)$ with $E \triangleq (\mathfrak{e}^\star)_{\mathfrak{e} \in \mathfrak{E}}$, the ıpport hypergraph of $\mathfrak{H}$.

We can remark that given a hypergraph $\mathcal{H}$, an infinite set $C_{\mathcal{H}}$ of hb-raphs can be generated that all have this hypergraph as support. A hb-dge family is attached to each of the hb-graphs in $C_{\mathcal{H}}$: each hb-edge upport corresponds at least to one hyperedge in $\mathcal{H}$ and reciprocally each yperedge of $\mathcal{H}$ is at least the support of one hb-edge per hb-graph of $C_{\mathcal{H}}$. The unicity of the correspondence is ensured only for hypergraphs and b-graphs without repeated hyperedges.

The **m-range** of a hb-graph—written $r_m(\mathfrak{H})$—is by definition:

$$r_m(\mathfrak{H}) \triangleq \max_{\mathfrak{e} \in \mathfrak{E}} \#_m \mathfrak{e}.$$

The **range** of a hb-graph $\mathfrak{H}$—written $r(\mathfrak{H})$—is the range of its support ıypergraph $\underline{\mathfrak{H}}$.

The **m-co-range** of a hb-graph—written $\mathrm{cr}_m(\mathfrak{H})$—is by definition:

$$\mathrm{cr}_m(\mathfrak{H}) \triangleq \min_{\mathfrak{e} \in \mathfrak{E}} \#_m \mathfrak{e}.$$

The **co-range** of a hb-graph $\mathfrak{H}$—written $\mathrm{cr}(\mathfrak{H})$—is the co-range of its support hypergraph $\underline{\mathfrak{H}}$.

The **global m-cardinality** of a hb-graph $\mathfrak{H}$—written $\#_m\mathfrak{H}$—is the sum of the m-cardinality of its hb-edges.

A hb-graph is said **k-m-uniform** if all its hb-edges have the same m-cardinality k.

A hb-graph is said **k-uniform** if its support hypergraph is k-uniform.

Proposition 1. *A hb-graph $\mathfrak{H}$ is k-m-uniform if and only if:*

$$r_m(\mathfrak{H}) = cr_m(\mathfrak{H}) = k.$$

Proof. Immediate. □

We can still refer for vertices of a hb-graph $\mathfrak{H} = (V, \mathfrak{E})$ to the **degree of** $v \in V$ written $\deg(v) = d(v)$: it corresponds to the degree of the same vertex in the support hypergraph $\underline{\mathfrak{H}}$. The **maximal degree** of a hb-graph $\mathfrak{H}$ is written Δ and corresponds to the maximal degree of the support hypergraph $\underline{\mathfrak{H}}$.

Nonetheless, in hb-graphs, due to the multiplicity function attached to each hb-edge, we can consider another kind of degree. To achieve its proper definition, we define the **hb-star** $\mathfrak{h}(v)$ of a vertex $v \in V$ as the multiset:

$$\mathfrak{h}(v) \triangleq \left\{\mathfrak{e}^{m_{\mathfrak{e}}(v)} : \mathfrak{e} \in \mathfrak{E} \wedge v \in \mathfrak{e}^\star\right\}.$$

The support of $\mathfrak{h}(v)$ is exactly the star of this vertex in the support hypergraph $\underline{\mathfrak{H}}$ and the cardinality of $\mathfrak{h}^\star(v)$ is exactly the degree of v.

The **m-degree of a vertex** $v \in V$ of a hb-graph $\mathfrak{H}$ is then defined as the m-cardinality of the hb-star attached to this vertex:

$$\deg_m(v) \triangleq \#_m\mathfrak{h}(v).$$

We also consider the **maximal m-degree** of a hb-graph $\mathfrak{H}$; we write it:

$$\Delta_m \triangleq \max_{v\in V} \deg_m(v).$$

A hb-graph having all its hb-edges of the same m-degree k is said **m-regular** or **k-m-regular**. A hb-graph is said **regular** if its support hypergraph is regular.

The following proposition is immediate:

Proposition 2. *For any vertex $v \in V$ of a natural hb-graph:*

$$d(v) \leqslant d_m(v) \leqslant \Delta_m.$$

This property is not true for non-natural hb-graphs.

Proposition 3. *If a hb-graph with a nonnegative multiplicity function range is such that: $d(v) = d_m(v)$ for all its vertices, then this hb-graph is a hypergraph.*

We now define the **dual** of a hb-graph $\mathfrak{H} = (V, \mathfrak{E})$ as the hb-graph $\tilde{\mathfrak{H}} \triangleq \left(\tilde{V}, \tilde{\mathfrak{E}}\right)$ such that its set of vertices $\tilde{V} \triangleq \{\tilde{v_j} : j \in [\![p]\!]\}$ is in bijection $f : \mathfrak{E} \to \tilde{V}$ with the family of hb-edges $\mathfrak{E}$ of $\mathfrak{H}$ such that:

$$\forall \tilde{v_j} \in \tilde{V}, \exists! \mathfrak{e}_j \in \mathfrak{E} : \tilde{v_j} = f(e_j).$$

And its family of hb-edges $\tilde{\mathfrak{E}} \triangleq (\tilde{e_i})_{i \in [\![n]\!]}$ is in bijection $g : V \to \tilde{\mathfrak{E}}$ such that $g(v_i) = \tilde{\mathfrak{e}_i}$ for all $i \in [\![n]\!]$—where:

$$\tilde{\mathfrak{e}_i} \triangleq \left\{\tilde{v_j}^{m_{\mathfrak{e}_j}(v_i)} : j \in [\![p]\!] \wedge \tilde{v_j} = f(\mathfrak{e}_j) \wedge v_i \in \mathfrak{e}_j^\star\right\}.$$

In Table 1, we present some dualities between a hb-graph and its dual.

	$\mathfrak{H}$	$\tilde{\mathfrak{H}}$
Vertices	$v_i, i \in [\![n]\!]$	$\tilde{v_j} = f(\mathfrak{e}_j), j \in [\![p]\!]$
Hb-edges	$\mathfrak{e}_j, j \in [\![p]\!]$	$\tilde{\mathfrak{e}_i} = g(v_i), i \in [\![n]\!]$
Multiplicity	$v_i \in \mathfrak{e}_j$ with $m_{\mathfrak{e}_j}(v_i)$	$\tilde{v_j} \in \tilde{\mathfrak{e}_i}$ with $m_{\mathfrak{e}_i}(v_j)$
m-degrees vs m-cardinality	$d_m(v_i)$	$\#_m\tilde{\mathfrak{e}_i}$
	$\#_m\mathfrak{e}_i$	$d_m(\tilde{v_j})$
m-uniformity vs m-regularity	k-m-uniform	k-m-regular
	k-m-regular	k-m-uniform

Table 1: Dualities between a hb-graph and its dual

3.3 Additional concepts for natural hb-graphs

3.3.1 Numbered copy hypergraph of a natural hb-graph

In natural hb-graphs, the hb-edge multiplicity functions have their range which is a subset of $\mathbb{N}$. The vertices in a hb-edge with multiplicity strictly greater than 1 can be seen as copies of the original vertex.

Indexing the copies of the original vertex, makes them seen as "numbered" copies. We consider a vertex v_i belonging to two hb-edges $\mathfrak{e}_{j_1}$ and $\mathfrak{e}_{j_2}$ of the hb-graph $\mathfrak{H} = (V, \mathfrak{E})$, with a multiplicity $m_{\mathfrak{e}_{j_1}}$ in $\mathfrak{e}_{j_1}$ and $m_{\mathfrak{e}_{j_2}}$ in $\mathfrak{e}_{j_2}$. Then $\mathfrak{e}_{j_1} \cap \mathfrak{e}_{j_2}$ holds $\min\left(m_{\mathfrak{e}_{j_1}}(v_i), m_{\mathfrak{e}_{j_2}}(v_i)\right)$ copies: by convention the ones "numbered" from 1 to $\min\left(m_{\mathfrak{e}_{j_1}}, m_{\mathfrak{e}_{j_2}}\right)$. Remaining copies will be in the multiset with the highest multiplicity of v_i.

More generally, we define the **numbered-copy set** of a natural multiset $\mathfrak{A}_m = \{x_i^{m_i} : i \in \llbracket n \rrbracket\}$ as the copy-set $\check{\mathfrak{A}}_m \triangleq \left\{[x_{i\,j}]_{m_i} : i \in \llbracket n \rrbracket\right\}$ where: $[x_{i\,j}]_{m_i}$ is a shortcut to indicate the numbered copies of the original element x_i: $x_{i\,1}$ to $x_{i\,m_i}$ and j is designated as the copy number of the element x_i.

We define the **maximum multiplicity function** of $\mathfrak{H}$ as the function $m : V \to \mathbb{N}$ such that for all $v \in V$: $m(v) \triangleq \max\limits_{\mathfrak{e} \in \mathfrak{E}} m_{\mathfrak{e}}(v)$ and consider the numbered-copy-set $\check{V} \triangleq \left\{[v_{i\,j}]_{m(v_i)} : i \in \llbracket n \rrbracket\right\}$ of the multiset $\left\{v_i^{m(v_i)} : i \in \llbracket n \rrbracket\right\}$

Then each hb-edge $\mathfrak{e}_k = \left\{v_{i_j}^{m_{k\,i_j}} : j \in \llbracket k \rrbracket \wedge i_j \in \llbracket n \rrbracket\right\}$ is associated to a copy-set / equivalency relation $< e_{k\,0}, \rho_k >$ which elements are in $\check{V}$ with the smallest copy numbers possible for any vertex in $\mathfrak{e}_k$. The hypergraph $\mathfrak{H}_0 \triangleq \left(\check{V}, E_0\right)$ where $E_0 \triangleq (e_{k\,0})_{k \in \llbracket p \rrbracket}$ is called the **numbered-copy-hypergraph** of $\mathfrak{H}$.

Proposition 4. *A numbered-copy-hypergraph is unique for a given hb-graph.*

Proof. It is immediate by the way the numbered-copy-hypergraph is built from the hb-graph. □

Allowing duplicates to be numbered prevent ambiguities; nonetheless it has to be seen as a conceptual approach since duplicates are entities that are actually not discernible.

3.3.2 Paths, distance and connected components

More precisely, a **strict m-path** $v_{i_0} \mathfrak{e}_{j_1} v_{i_1} \ldots \mathfrak{e}_{j_s} v_{i_s}$ in a hb-graph $\mathfrak{H}$ from a vertex u to a vertex v is a vertex / hb-edge alternation with s hb-edges $\mathfrak{e}_{j_k}$ such that: $\forall k \in \llbracket s \rrbracket, j_k \in \llbracket p \rrbracket$ and $s+1$ vertices v_{i_k} with $\forall k \in \{0\} \cup \llbracket s \rrbracket$, $i_k \in \llbracket n \rrbracket$ and such that $v_{i_0} = u$, $v_{i_s} = v$, $u \in \mathfrak{e}_{j_1}$ and $v \in \mathfrak{e}_{j_s}$ and that for all $k \in \llbracket s-1 \rrbracket$, $v_{i_k} \in \mathfrak{e}_{j_k} \cap \mathfrak{e}_{j_{k+1}}$.

A **large m-path** $v_{i_0}\mathfrak{e}_{j_1}v_{i_1}\ldots\mathfrak{e}_{j_s}v_{i_s}$ in a hb-graph $\mathfrak{H}$ from a vertex u to a vertex v is a vertex / hb-edge alternation with s hb-edges $\mathfrak{e}_{j_k}$ such that: $\forall k \in \llbracket s \rrbracket, j_k \in \llbracket p \rrbracket$ and $s+1$ vertices v_{i_k} with $\forall k \in \{0\} \cup \llbracket s \rrbracket$, $i_k \in \llbracket n \rrbracket$ and such that $v_{i_0} = u$, $v_{i_s} = v$, $u \in \mathfrak{e}_{j_1}$ and $v \in \mathfrak{e}_{j_s}$ and that for all $k \in \llbracket s-1 \rrbracket$, $v_{i_k} \in \mathfrak{e}_{j_k} \cup \mathfrak{e}_{j_{k+1}}$.

The **length of a m-path** from u to v is the number of hb-edges it traverses; given a path $\mathscr{P}$, we write $l(\mathscr{P})$ its length. It holds that if $\mathscr{P} = v_{i_0}\mathfrak{e}_{j_1}v_{i_1}\ldots\mathfrak{e}_{j_s}v_{i_s}$, we have whatever the path is strict or large: $l(\mathscr{P}) = s$.

In a path $\mathscr{P} = v_{i_0}\mathfrak{e}_{j_1}v_{i_1}\ldots\mathfrak{e}_{j_s}v_{i_s}$, the vertices v_{i_k}, $k \in \llbracket s-1 \rrbracket$ are called the **interior vertices** of the m-path and v_{i_0} and v_{i_s} are called the **extremities** of the m-path.

If the extremities are different copies of the same object, then the m-path is said to be an **almost cycle**. If the extremities designate exactly the same copy of one object, the m-path is said to be a **cycle**.

Proposition 5. *1. For a strict m-path, there are:*

$$\prod_{k \in \llbracket s-1 \rrbracket} m_{\mathfrak{e}_{j_k} \cap \mathfrak{e}_{j_{k+1}}}(v_{i_k})$$

possibilities of choosing the interior vertices along a given m-path $v_{i_0}\mathfrak{e}_{j_1}v_{i_1}\ldots\mathfrak{e}_{j_s}v_{i_s}$ *and:*

$$m_{\mathfrak{e}_{j_1}}(v_{i_0})\left(\prod_{k \in \llbracket s-1 \rrbracket} m_{\mathfrak{e}_{j_k} \cap \mathfrak{e}_{j_{k+1}}}(v_{i_k})\right) m_{\mathfrak{e}_{j_s}}(v_{i_s})$$

possible strict m-paths in between the extremities.

2. For a large m-path, there are:

$$\prod_{k \in \llbracket s-1 \rrbracket} m_{\mathfrak{e}_{j_k} \cup \mathfrak{e}_{j_{k+1}}}(v_{i_k})$$

possibilities of choosing the interior vertices along a given m-path $v_{i_0}\mathfrak{e}_{j_1}v_{i_1}\ldots\mathfrak{e}_{j_s}v_{i_s}$ *and:*

$$m_{\mathfrak{e}_{j_1}}(v_{i_0})\left(\prod_{k \in \llbracket s-1 \rrbracket} m_{\mathfrak{e}_{j_k} \cup \mathfrak{e}_{j_{k+1}}}(v_{i_k})\right) m_{\mathfrak{e}_{j_s}}(v_{i_s})$$

possible large m-paths in between the extremities.

3. As large m-paths between two extremities for a given sequence of interior vertices and hb-edges include strict m-paths, we often refer as ***m-paths*** *for large m-paths.*

4. *When a m-path exists from u to v, it also exists from v to u.*

Proof. All these results come directly of combinatorics over the multisets involved in the different paths. □

An m-path $v_{i_0}\mathfrak{e}_{j_1}v_{i_1}\ldots\mathfrak{e}_{j_s}v_{i_s}$ in a hb-graph corresponds to a unique path in the hb-graph support hypergraph called the **support path**.

Proposition 6. *Every m-path $v_{i_0}\mathfrak{e}_{j_1}v_{i_1}\ldots\mathfrak{e}_{j_s}v_{i_s}$ traversing the same hb-edges and having similar vertices as intermediate and extremity vertices share the same support path.*

Proof. The common support path is then $v_{i_0}\mathfrak{e}^{\star}_{j_1}v_{i_1}\ldots\mathfrak{e}^{\star}_{j_s}v_{i_s}$. □

The notion of distance is similar to the one defined for hypergraphs.

The **geodesic distance** $d(u,v)$ between two vertices u and v of a hb-graph is the length of the shortest m-path between u and v, if it exists, that can be found in the hb-graph. In the case where there is no path between the two vertices, they are said disconnected, and we set: $d(u,v) = +\infty$. A hb-graph is said **connected** if its support hypergraph is connected, disconnected otherwise.

A **connected component** of a hb-graph is a maximal mset of vertices for which there exists a m-path in between every pair of vertices of the mset in the hb-graph.

Proposition 7. *A connected component of a hb-graph is a connected component of one of its copy hypergraph.*

The **diameter** of a hb-graph $\mathfrak{H}$—written diam $(\mathfrak{H})$—is defined as:

$$\operatorname{diam}(\mathfrak{H}) \triangleq \max_{u,v\in V} d(u,v).$$

3.3.3 Adjacency

In Ouvrard et al. [35] we have introduced different concepts of adjacency for hypergraphs. The traditional adjacency is a pairwise relationship. In hb-graphs, hb-edges handle n-adic relationships. Hence, the concept of adjacency in hb-graphs is more than pairwise.

We consider a hb-graph $\mathfrak{H} = (V, \mathfrak{E})$, a positive integer k and k vertices not necessarily distinct belonging to V. We write $\mathfrak{V}_{k,m}$ the mset consisting of these k vertices with multiplicity function m.

The k vertices of $\mathfrak{V}_{k,m}$ are said **k-adjacent** in $\mathfrak{H}$ if it exists $\mathfrak{e} \in \mathfrak{E}$ such that $\mathfrak{V}_{k,m} \subseteq \mathfrak{e}$.

Considering a hb-graph $\mathfrak{H}$ of m-range $\overline{k} = r_m(\mathfrak{H})$, the hb-graph cannot handle any k-adjacency for k strictly greater than $r_m(\mathfrak{H})$. This maximal k-adjacency is called the **$\overline{k}$-adjacency** of $\mathfrak{H}$.

We consider now a hb-edge $\mathfrak{e}$ of a hb-graph $\mathfrak{H}$. The vertices in the support of $\mathfrak{e}$ are said **$e^\star$-adjacent**. For natural hb-graphs, we say that the vertices with nonzero multiplicity in a hb-edge are **e-adjacent**.

We can remark that $e^\star$-adjacency does not support redundancy of vertices while e-adjacency in natural hb-graphs allows it. e-adjacency in natural hb-graphs takes into account the multiplicity of the different vertices, which is not the case of $e^\star$-adjacency.

In non-natural hb-graph, the vertices in the support of hb-edges with multiplicity different from 1 cannot be seen as copies of one another, hence only the $e^\star$-adjacency is valid for this kind of hb-graphs.

Two hb-edges are said **incident** if their support intersection is not empty.

3.3.4 Sum of two hb-graphs

Let $\mathfrak{H}_1 = (V_1, \mathfrak{E}_1)$ and $\mathfrak{H}_2 = (V_2, \mathfrak{E}_2)$ be two hb-graphs.

The **sum of two hb-graphs** $\mathfrak{H}_1$ and $\mathfrak{H}_2$ is the hb-graph, written $\mathfrak{H}_1 + \mathfrak{H}_2$, that has:

- $V_1 \cup V_2$ as vertex set
- $\mathfrak{E}_1 + \mathfrak{E}_2$ as hb-edge family: hb-edges are obtained from the hb-edges of $\mathfrak{E}_1$—respectively $\mathfrak{E}_2$—with same multiplicity for vertices of V_1—respectively V_2—but such that for each hyperedge in $\mathfrak{E}_1$—respectively $\mathfrak{E}_2$—the universe is extended to $V_1 \cup V_2$ and the multiplicity function is extended such that $\forall v \in V_2 \backslash V_1 : m(v) = 0$ (respectively $\forall v \in V_1 \backslash V_2 : m(v) = 0$)
- $\mathfrak{H}_1 + \mathfrak{H}_2 \triangleq (V_1 \cup V_2, \mathfrak{E}_1 + \mathfrak{E}_2)$

This sum is said direct if $\mathfrak{E}_1 + \mathfrak{E}_2$ does not contain any new pair of repeated hb-edge other than the ones already existing in $\mathfrak{E}_1$ and in $\mathfrak{E}_2$, and that $\mathfrak{E}_1$ and $\mathfrak{E}_2$ do not have any common hb-edges. In this case, the sum is written $\mathfrak{H}_1 \oplus \mathfrak{H}_2$ and called the **direct sum of the two hb-graphs** $\mathfrak{H}_1$ and $\mathfrak{H}_2$.

3.4 An example

Example 3.1. *Considering $\mathfrak{H} = (V, \mathfrak{E})$, with $V = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7\}$ and $\mathfrak{E} = \{\mathfrak{e}_1, \mathfrak{e}_2, \mathfrak{e}_3, \mathfrak{e}_4\}$ with: $\mathfrak{e}_1 = \{v_1^2, v_4^2, v_5^1\}$, $\mathfrak{e}_2 = \{v_2^3, v_3^1\}$, $\mathfrak{e}_3 = \{v_3^1, v_5^2\}$, $\mathfrak{e}_4 = \{v_6\}$. It holds:*

	$\mathfrak{e}_1$	$\mathfrak{e}_2$	$\mathfrak{e}_3$	$\mathfrak{e}_4$	$d_m(v_i)$	$\max\{m_{\mathfrak{e}_j}(v_i)\}$
v_1	2	0	0	0	2	2
v_2	0	3	0	0	3	3
v_3	0	1	1	0	2	1
v_4	2	0	0	0	2	1
v_5	1	0	2	0	3	2
v_6	0	0	0	1	1	1
v_7	0	0	0	0	0	0
$\#_m\mathfrak{e}_j$	5	4	3	1		

Therefore the order of $\mathfrak{H}$ is $O(\mathfrak{H}) = 2+3+1+1+2+1+0 = 10$ and its size is $|\mathfrak{E}| = 4$. v_7 is an isolated vertex. $\mathfrak{e}_1$ and $\mathfrak{e}_3$ are incident as well as $\mathfrak{e}_3$ and $\mathfrak{e}_2$. $\mathfrak{e}_4$ is not incident to any hb-edge. v_1, v_4 and v_5 are $e^\star$-adjacent as they hold in $\mathfrak{e}_1^*$, while v_1^2, v_4^1 and v_5^1 are e-adjacent as they hold in $\mathfrak{e}_1$.

The dual of $\mathfrak{H}$ is the hb-graph $\widetilde{\mathfrak{H}} = \left(\widetilde{V}, \widetilde{\mathfrak{E}}\right)$ with $\widetilde{V} = \{\tilde{x_j} : j \in [\![4]\!]\}$ where $\tilde{x_j}$ corresponds to $\mathfrak{e}_j$ of the original hb-graph, with $j \in [\![4]\!]$ and $\tilde{\mathfrak{E}} = \{\tilde{\mathfrak{e}_i} : i \in [\![7]\!]\}$ where $\tilde{\mathfrak{e}_1} = \tilde{\mathfrak{e}_4} = \{\tilde{x_1}^2\}$, $\tilde{\mathfrak{e}_2} = \{\tilde{x_2}^3\}$, $\tilde{\mathfrak{e}_3} = \{\tilde{x_2}^1, \tilde{x_3}^1\}$, $\tilde{\mathfrak{e}_5} = \{\tilde{x_1}^1, \tilde{x_3}^2\}$, $\tilde{\mathfrak{e}_6} = \{\tilde{x_4}^1\}$ and $\tilde{\mathfrak{e}_7} = \emptyset$. The $\tilde{\mathfrak{e}_i}$ corresponds to v_i of the original hb-graph, with $i \in [\![7]\!]$. $\widetilde{\mathfrak{H}}$ has duplicated hb-edges and one empty hb-edge.

3.5 Hb-graph representations

Representing hb-graphs can be thought along the two main standards found for hypergraphs —Mäkinen [30]—: the subset standard and the edge standard. But both representations have to be adapted to fit to multisets.

3.5.1 Subset standard

In the subset standard for hypergraphs, a contour line is drawn to surround the vertices of a hyperedge. Each hyperedge is then represented using these contour lines. Depending on how their intersection is represented, we obtain a Venn diagram or an Euler diagram representation of the hypergraph. A Venn diagram systematically represents each possible intersection between hyperedges, while an Euler diagram addresses only the intersections that are needed for the representation. Hence, the Euler diagram is often preferred as this representation scales up a bit better than the Venn diagram; but neither representations scale up to large hypergraphs.

When moving to multisets, a contour line is also drawn around the vertices that are now duplicated. In Radoaca [48], two Venn diagram representations of multisets are proposed for the representation of 2 and 3 multisets: a simplified representation where the parts are not disjoint and a complex representation where the parts are disjoint. Scaling up the number of multisets seems to be hard to achieve. Euler representation of multisets can be based on this work: a simplified and a complex representation can be drawn to depict only the parts needed to be represented. It simplifies the Venn representation and helps to scale up to somewhat larger hb-graphs.

3.5.2 Edge standard

For hypergraphs, there are two main representations in the edge standard: the clique representation and the extra-node representation. The clique representation transforms the hypergraph in its 2-section graph joining every pair of vertices of a hyperedge by an edge, while the extra-node representation is the graph obtained by representing each hyperedge as an extra-vertex and joining it to each vertex of the corresponding hyperedge.

For hb-graphs, the 2-section graph is always representable by considering each hb-edge support: but the quantity of information to display as well as its quality is not optimal in this representation.

Hence, we propose other alternatives based on the extra-node representation. Each hb-edge is represented by an extra-node. For natural hb-graph a first representation called the **extra-node multipartite representation** is achieved by joining each vertex of the hb-edge to the extra node with a number of edges that corresponds to the vertex multiplicity. This representation does not fit for hb-graphs that are not natural; moreover, it is hard to scale up when the values of multiplicities increase.

For hb-graphs with non-negative multiplicity ranges, we propose a second representation based on the extra-node representation of the support hypergraph, but where the thickness of the edges linking the vertices of the hb-edge with its extra-node are proportional either directly to the multiplicity (absolute version) or to the relative multiplicity of the vertex in the hb-edge (unnormalised version).

The **relative multiplicity** of a vertex v_i in a hb-edge $\mathfrak{e}_j$ is defined as:

$$m_{r\,\mathfrak{e}_j}(v_i) \triangleq \frac{m_{\mathfrak{e}_j}(v_i)}{\#_m \mathfrak{e}_j}.$$

We introduce the **normalised relative multiplicity** of a vertex $v_i \in \mathfrak{e}_j$ in the hb-graph as:

$$m_{nr\,\mathfrak{e}_j}(v_i) \triangleq \frac{m_{\mathfrak{e}_j}(v_i)}{\#_m \mathfrak{e}_j} \times \frac{\overline{m}(v_i)}{M}$$

where $\overline{m}(v_i) \triangleq \max_{\mathfrak{e} \in \mathfrak{E}} (m_{\mathfrak{e}}(v_i))$ and $M \triangleq \max_{v \in V} (\overline{m}(v))$. This provides another extra-node representation where the edge thickness is proportional to the normalised relative multiplicity. This representation allows to have a direct view of the importance of each vertex contribution to a hb-edge compared to other hb-edges.

If the hb-graph has some multiplicity function with both negative and positive values, the former representations can be adapted by using different shapes for the edges linking the extra-node to the vertices of the hb-edges.

3.6 Some applications

3.6.1 Prime decomposition representation and elementary operations on hb-graphs

In Corso [11], a natural number network based on common divisors of two vertices is proposed to replace the search for real scale-free networks by the generation of a deterministic network that is also scale-free. Degrees of vertices are studied, as well as the clustering coefficient and the average distance of vertices in the graph. Some topological properties of such networks are tackled in Zhou et al. [57]. In Frahm et al. [13], the authors study the page rank of an integer network built using a directed graph where the vertices are labelled by nonnegative integers; an edge links two vertices m and n if m divides n, m being different of 1 and n, with a weight k corresponding to the maximal k such that m^k divides n, i.e. the valuation of m in n. All these approaches are built using graphs and pairwise relationships.

As already mentioned, multisets can be used for prime decomposition—Blizard et al. [5]. In particular, multisets are intensively used in Tarau [54] to achieve primality decomposition of numbers and to achieve product, division, gcd and lcd of numbers. Using hb-graphs, we can revisit some of the results of Tarau [54] and have a visual representation of simultaneous decomposition of numbers interpretable in terms of elementary operations that transform a hb-graph representation into another one. It should also allow to refine results obtained with graph since multisets handle not solely the n-adicity that could be achieved by sets, but also the hb-edge based weighting of the divisors.

We focus on the prime decomposition of numbers. Considering the set ᴾ of prime numbers, any positive integer n greater or equal to 2 can be lecomposed in a product of prime numbers: p_{i_1} with multiplicity $m_n(i_1)$ o p_{i_n} with multiplicity $m_n(i_n)$. This decomposition is then uniquely decribed by the multiset: $\mathfrak{e}_n = \left\{p_{i_1}^{m_{i_1}}, ..., p_{i_n}^{m_{i_n}}\right\}$. The prime decomposition ıb-graph $\mathfrak{H}_{\mathbb{P}} \triangleq (\mathbb{P}, \mathfrak{E})$—where $\mathfrak{E} \triangleq (\mathfrak{e}_n)_{n \in \mathbb{N} \setminus \{0,1\}}$—is the hb-graph of uniɾerse $\mathbb{P}$ and of hb-edges the prime decomposition of the integers greater han 2. It contains all the possible natural multisets composed of elements ɔf $\mathbb{P}$. To represent this hb-graph, each extra-node is labeled by the corre;ponding number n which decomposition in prime numbers constitutes the nultiset $\mathfrak{e}_n = \left\{p_{i_1}^{m_{i_1}}, ..., p_{i_n}^{m_{i_n}}\right\}$.

We consider a subset A of the hb-edges of $\mathfrak{H}_{\mathbb{P}}$ and write $\mathfrak{H}_A \triangleq (\mathbb{P}, A)$ the ;ub-hb-graph of $\mathfrak{H}_{\mathbb{P}}$ associated to A. We also consider elementary opera;ions on the multipartite extra-node representation of the natural hb-graphs :onstructed to switch between the decomposition involved by the two inte;ers and the prime decomposition of the results. We observed that these ːlementary operations are similar to the elementary operations involved in :he graph edit distance Gao et al. [14], albeit the fact that with extra-nodes ;upplementary operations are possible: deletion of an edge, relabeling of ɪn extra-node, deletion of an extra-node, merging two extra-nodes.

The decomposition in prime of the product mn of two integers m and ɿ is represented by a hb-edge $\mathfrak{e}_{mn}$ of $\mathfrak{H}_{\mathbb{P}}$ which is such that $\mathfrak{e}_{mn} = \mathfrak{e}_m \uplus \mathfrak{e}_n$. Γhe multi-partite extra-node representation of $\mathfrak{H}_{\{\mathfrak{e}_{mn}\}}$ is obtained from the ɔne of $\mathfrak{H}_{\{\mathfrak{e}_m, \mathfrak{e}_n\}}$ by merging the two extra-nodes n and m of the $\mathfrak{e}_n$ and $\mathfrak{e}_m$ representations, while keeping all the existing edges of their respective ːepresentation. It is illustrated with an example in Figure 1.

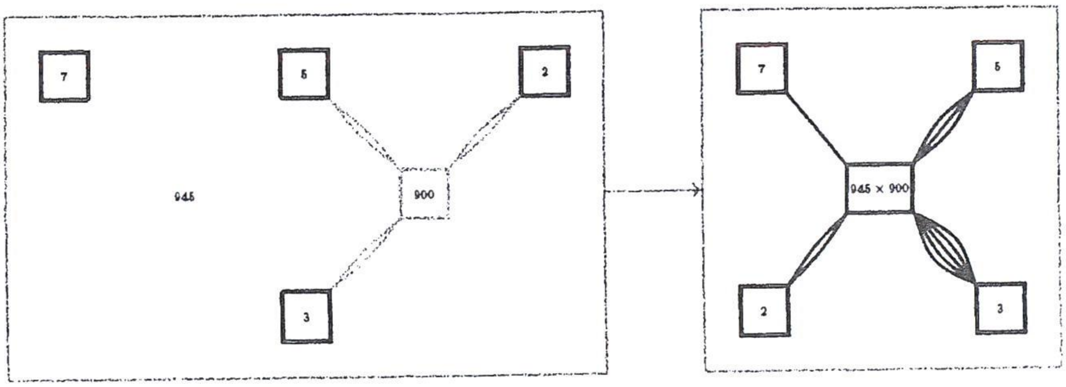


Figure 1: Finding the prime decomposition of mn from the decomposition of m and n with $m = 900$ and $n = 945$.

If we suppose that m divides n, then the decomposition of $n \div m$ in primes is stored in a hb-edge $\mathfrak{e}_{n \div m}$ such that $\mathfrak{e}_{n \div m} = \mathfrak{e}_n \backslash \mathfrak{e}_m$. The multi-partite extra-node representation of $\mathfrak{H}_{\{\mathfrak{e}_{n \div m}\}}$ is obtained from the one of $\mathfrak{H}_{\{\mathfrak{e}_m, \mathfrak{e}_n\}}$ by deleting all edges in the representation of $\mathfrak{e}_m$ and in the same quantities the corresponding edges in the representation of $\mathfrak{e}_n$ and relabeling the extra-node of $\mathfrak{e}_n$ to be the one of $\mathfrak{e}_{n \div m}$.

The decomposition in primes of the greater common divider of two integers m and n is stored in $\mathfrak{e}_{\mathrm{gcd}(m,n)} = \mathfrak{e}_m \cap \mathfrak{e}_n$. The representation of $\mathfrak{H}_{\{\mathfrak{e}_{\mathrm{gcd}(m,n)}\}}$ is obtained from the one of $\mathfrak{H}_{\{\mathfrak{e}_m, \mathfrak{e}_n\}}$ by deleting any edge from one vertex in $\mathfrak{e}_m^*$ to the extra-node representing $\mathfrak{e}_m$ that is greater in quantity than the one linking this vertex to the extra-node representing $\mathfrak{e}_n$ and reciprocally. The final representation is obtained by deleting one of the remaining extra-node vertex and its connected edges and relabeling the other extra-vertex with $\mathrm{gcd}(m,n)$. It is illustrated in Figure 2.

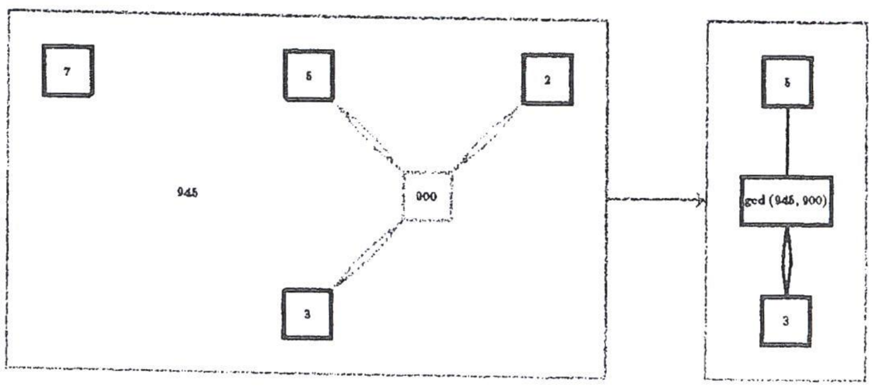


Figure 2: Finding the prime decomposition of $\mathrm{gcd}(m,n)$ from the decomposition of m and n with $m = 900$ and $n = 945$.

The decomposition in primes of the least common multiple of two integers m and n is stored in $\mathfrak{e}_{\mathrm{lcm}(m,n)} = \mathfrak{e}_m \cup \mathfrak{e}_n$. The representation of $\mathfrak{H}_{\{\mathfrak{e}_{\mathrm{lcm}(m,n)}\}}$ is obtained from the one of $\mathfrak{H}_{\{\mathfrak{e}_m, \mathfrak{e}_n\}}$ by deleting any edge from one vertex in $\mathfrak{e}_m^*$ to the extra-node representing $\mathfrak{e}_m$ that is greater in quantity than the one linking this vertex to the extra-node representing $\mathfrak{e}_n$ and reciprocally. The final representation is obtained by deleting one of the remaining extra-node vertex and its connected edges and relabeling the other extra-vertex with $\mathrm{gcd}(m,n)$. It is illustrated in Figure 3.

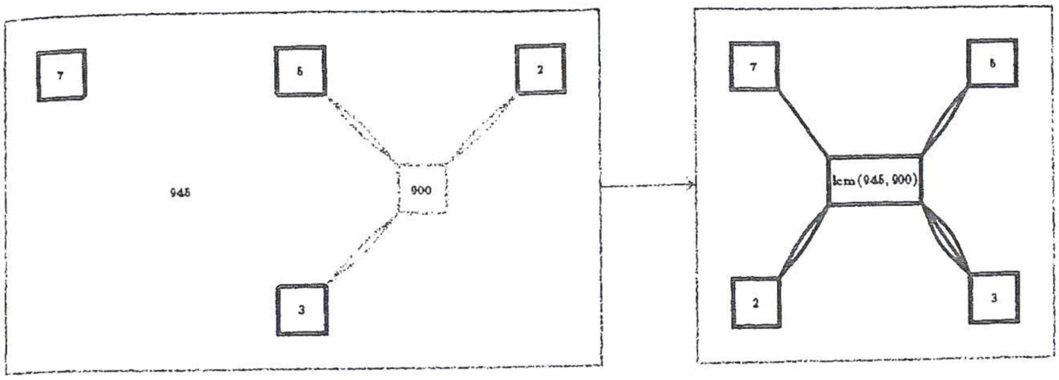


Figure 3: Finding the prime decomposition of lcm (m, n) from the decomposition of m and n with $m = 900$ and $n = 945$.

We can then formulate the property

$$mn = \gcd(m, n) \times \operatorname{lcm}(m, n)$$

by using the hb-graphs $\mathfrak{H}_{\{e_{mn}\}}$ and $\mathfrak{H}_{\{e_{\gcd(m,n)}, e_{\operatorname{lcm}(m,n)}\}}$

It holds: $e_{mn} = e_{\gcd(m,n)} \uplus e_{\operatorname{lcm}(m,n)}$, which can be written: $e_{mn} = (e_m \cap e_n) \uplus (e_m \cup e_n)$, which can be easily observed in the results shown in Figure 1 and Figure 4.

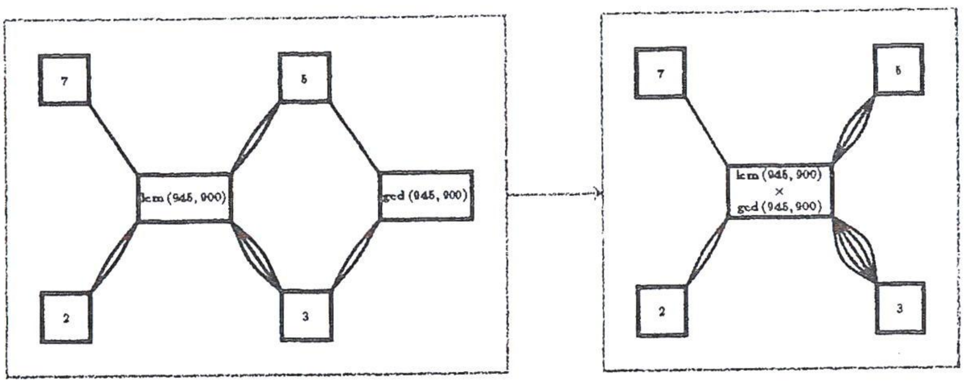


Figure 4: Illustrating the property lcm $(m, n) \times$ gcd $(m, n) = mn$ with $m = 900$ and $n = 945$.

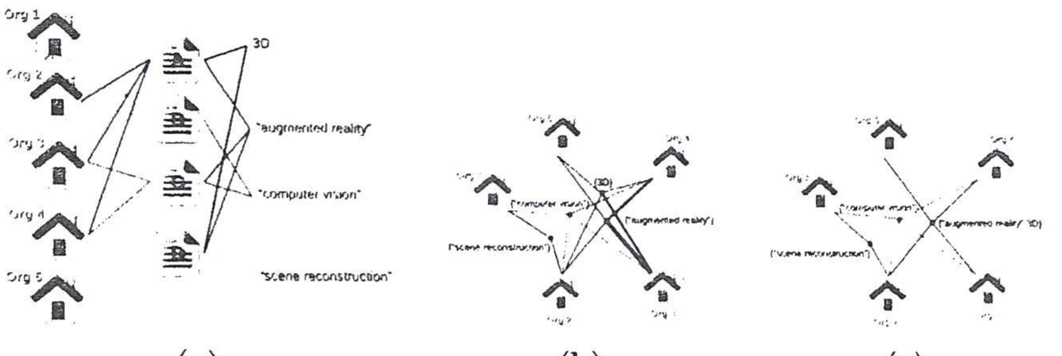
It follows that any connected natural hb-graph can be attached to a number by labeling vertices with prime numbers and multiplying successively the hb-edges by this number: in some way it makes a summary of this hb-graph. Finding the prime labeling such that the number the natural hb-graph represents is minimal is a NP-hard problem.

Reciprocally, being given a number we can use its decomposition to create a bunch of hb-graphs that have this number as overall representation.

3.6.2 Some applications to the modeling and visualisation of textual datasets

In Ouvrard et al. [34], we developed a hypergraph framework for the visualisation of textual datasets. Textual datasets have metadata that describes data instances; these metadata are of various types, and provide different facets of the information space of the textual dataset. For instance, a publication is stored in a database with some metadata giving information not only about it, but also on key features of this publication: it often includes not only authors but also their affiliation that gives access to organisations, some keywords, the abstract, subject categories, and so on. The data instances attached to a publication allow to build a network of those instances. To extract information, we choose one of the types, and build co-occurrences of another type, called facet type. We illustrate this in Figure 5. These co-occurrences can either be modeled as multisets or reduced to the support of these multisets.

So far, we have used hypergraphs in a framework with an extra-node representation—Ouvrard et al. [34]—, which allow to diminish the cognitive load of traditional approaches such as the visualisation of 2-section of hypergraphs —Ouvrard et al. [37]. Using hb-graphs, the representation of the data instances for the visualisation part of the framework can be refined. We illustrate this in Figure 5 with instances of a publication dataset, focusing on keywords and organisations.



(a) (b) (c)

Figure 5: Building co-occurrences: the reference chosen is "keywords", the facet considered is "organisations". From left to right: (a) original network, (b) co-occurrence network using hb-graph, (c) co-occurrence network using the support hypergraph.

In textual analysis, a text is often presented in a first analysis as a ag of words, wherein word multiplicities correspond to the number of ccurrences of the word. A dataset of texts can then be represented by a b-graph where the hb-edges are the bags of words modeling each text, the ertex set being the dictionary of all words occuring in the dataset. The nultiplicity functions can either represent the number of occurrences or the erm frequencies or even the TF-IDF of each term in the dataset.

3.7 Brief conclusion on hb-graphs

Shifting to hb-graphs is not only a change in modeling for visualisation: it also allows to refine any network where hb-edge-based individual weighting of vertices is required.

We have already shown in Ouvrard et al. [39] that diffusion by exchange in hb-graphs provides a fine ranking not only of vertices but also of hb-edges. We think that extending the definitions of hb-graphs to support negative multiplicities or even complex values could open the door to a wider variety of applications.

We can foresee many applications of hb-graphs; they give the best of several worlds: multisets, sets, graphs, and, also, as we will see in the next section, algebraic and polynomial approaches. Natural hb-graphs support the duplication of elements, while general hb-graphs allow the weighting of elements in a refined manner with respect to hypergraphs.

4 Algebraic representation of a hb-graph

In this section, we consider only hb-graphs with no repeated hb-edge.

4.1 Incidence matrix of a hb-graph

A multiset is well defined by its universe and its multiplicity function. It can be represented by the multiset vector representation.

Hb-edges of a given hb-graph have the same universe. Let n and p be two positive integers and a non-empty hb-graph $\mathfrak{H} = (V, \mathfrak{E})$ with vertex set $V = \{v_i : i \in [\![n]\!]\}$ and $\mathfrak{E} = (\mathfrak{e}_j)_{j\in[\![p]\!]}$. We define the **incidence matrix** of the hb-graph $\mathfrak{H}$ as the matrix H :

$$H \triangleq [m_j(v_i)]_{\substack{i\in[\![n]\!]\\ j\in[\![p]\!]}}.$$

This incidence matrix is intensively used in Ouvrard et al. [39] to formalize the diffusion by exchanges in hb-graphs.

Proposition 8. *Any nonnegative matrix with real coefficients is the incident matrix of an hb-graph.*

Any nonnegative matrix with integer coefficients is the incident matrix of a natural hb-graph.

Using a matrix to store the information implies a 2-adic relationships: between vertex and hb-edge. The multi-adic relationship involved by hb-edges has to be reconstructed from the columns of the incident matrix. The same occurs with HH^T which is sometimes taken in the litterature as an adjacency matrix: it is pairwise relationships between vertices.

4.2 e-adjacency tensor of a natural hb-graph

To build the e-adjacency tensor $\mathcal{A}(\mathfrak{H})$ of a general natural hb-graph $\mathfrak{H} = (V, \mathfrak{E})$ without repeated hb-edge—with vertex set $V = \{v_i : i \in [\![n]\!]\}$ and hb-edge set $\mathfrak{E} = (\mathfrak{e}_j)_{j \in [\![p]\!]}$—we take an approach similar to Ouvrard et al. [36, 35] using the strong link between cubical symmetric tensors and homogeneous polynomials.

4.2.1 Related work

In [43], that we mentioned in the former section, the ([Author's note]: $\boldsymbol{k}$)-**adjacency tensor** for k-PZ-multigraphs—i.e. k-m-uniform natural hb-graph with no repeated hb-edge—is defined as following:

Definition 1. *Let $\dddot{\mathcal{H}}_k = (V, E)$ be a k-PZ-multigraph on a finite set of vertices $V = \{v_i : i \in [\![n]\!]\}$ and a set of edges $E = (e_j)_{j \in [\![p]\!]}$.*

*The ([Author's note]: **k)-adjacency tensor** of a k-PZ-multigraph is the symmetric tensor $\mathcal{A}_{\dddot{\mathcal{H}}_m} \in \mathcal{T}_{k,n}$ of CHR $\mathbf{A}_{\dddot{\mathcal{H}}_k} = (a_{i_1,\dots,i_k})_{i_1,\dots,i_k \in [\![n]\!]}$ such that:*

$$a_{i_1,\dots,i_k} = \begin{cases} \dfrac{m_{j_1}! \dots m_{j_s}!}{(m-1)!}, & \textit{if } \{\{v_{i_1}, \dots, v_{i_m}\}\} = \{v_{j_1}^{m_{j_1}}, \dots, v_{j_s}^{m_{j_s}}\} \in E \\ 0, & \textit{otherwise.} \end{cases}$$

The author then study some spectral properties of k-PZ-multigraph.

As we need further refinements and interpretability of our process in term of m-uniformisation and polynomial homogeneisation, and having developed it independently, we start after having given the expectations of such a tensor to first introduce a tensor for elementary hb-graphs, before retrieving a tensor for k-m-uniform natural hb-graph with no repeated hb-edge that corresponds to the ([Author's note]: $\boldsymbol{k}$)-**adjacency tensor** of a k-PZ-multigraph and handling the case of general hb-graphs.

4.2.2 Expectations for the e-adjacency tensor

We formulate the expectations for the e-adjacency tensor of a natural hb-graph we want to construct. For general hypergraphs, we have insisted on the interpretability of the construction using a hypergraph uniformisation process: it has imposed the filling of the hyperedges with additional and two-by-two different vertices since hyperedges cannot have duplicated elements.

As natural hb-graphs allow "naturally" vertices to be duplicated, we can think on different ways of filling the hb-edges with additional vertices. We have chosen three different ways: the straightforward approach, the silo approach and the layered approach—as it was already done for the latter for general hypergraphs in Ouvrard et al. [35, 36].

We first give some expected properties of such a tensor, some of them being more qualitative than quantifiable.

Expectation 4.1. *The e-adjacency tensor should be nonnegative, symmetric and its generation should be as simple as possible.*

The motivation behind is that nonnegative symmetric hypermatrices have nice properties: they can be described with a small number of values. A hb-edge can be described with only one tuple of indices and their corresponding coefficient, the other coefficients of the tensor being obtained by permuting the indices of the first one while the same value is kept. Moreover, in the spectral theory, symmetric nonnegative tensors ensure interesting properties as their spectral radius is positive for a nonzero tensor—Qi and Luo [47]—; furthermore there is at most one H-eigenvalue that corresponds to the spectral radius with a positive Perron H-eigenvector.

Expectation 4.2. *The tensor should be globally invariant to vertex permutation in the original hb-graph.*

By globally invariant we mean that a permutation of rows on each face of the hypermatrix follows the same permutation than the one involved in the vertex permutation. We do not expect the special vertices added for the filling of the hb-edges to follow the same rule.

Expectation 4.3. *The e-adjacency tensor should allow the unique reconstruction of the hb-graph it is originated from.*

The e-adjacency tensor should describe the hb-graph in a unique way up to a permutation of indices, so that no two hb-graphs have the same e-adjacency tensor unless they are isomorphic. This is a strong requirement as it enforces the addition of special vertices even for k-m-uniform hb-graphs, where the $\bar{k}$-adjacency corresponds to the k-adjacency. Hence, the special vertices will be systematically generated and added to the final tensor in order to meet this expectation.

Expectation 4.4. *Given the choice of two representations, the one that can be described with the least elements possible should be chosen. Then the sparsest e-adjacency tensor should be chosen.*

It forces the hypermatrix to be easily describable before ensuring the lowest sparsity possible. The fact that the hypermatrix is symmetric will help.

Expectation 4.5. *The e-adjacency tensor should allow direct retrieval of the vertex degrees.*

It is a requirement made for all k-adjacency tensors of uniform hypergraphs. It is also the case for the e-adjacency tensors of Sun et al. [51], Banerjee et al. [2] and for the first e-adjacency tensor we built for general hypergraphs.

4.2.3 Elementary hb-graph

A hb-graph that has only one non repeated hb-edge in its hb-edge family is called an **elementary hb-graph.**

Claim 4.1. *Let* $\mathfrak{H} = (V, \mathfrak{E})$ *be a hb-graph with no repeated hb-edge. Then:*

$$\mathfrak{H} = \bigoplus_{\mathfrak{e} \in \mathfrak{E}} \mathfrak{H}_{\mathfrak{e}}$$

where $\mathfrak{H}_{\mathfrak{e}} = (V, (\mathfrak{e}))$ *is the elementary hb-graph associated to the hb-edge* $\mathfrak{e}$.

Proof. Let $\mathfrak{e}_1 \in \mathfrak{E}$ and $\mathfrak{e}_2 \in \mathfrak{E}$. As $\mathfrak{H}$ is with no repeated hb-edge, $\mathfrak{e}_1 + \mathfrak{e}_2$ does not contain any new pairs of repeated elements. Thus $\mathfrak{H}_{\mathfrak{e}_1} + \mathfrak{H}_{\mathfrak{e}_2}$ is a direct sum and can be written $\mathfrak{H}_{\mathfrak{e}_1} \oplus \mathfrak{H}_{\mathfrak{e}_2}$.

A straightforward iteration over elements of $\mathfrak{e} \in \mathfrak{E}$ leads to the result. □

We need first to define the $\overline{k}$-adjacency hypermatrices for an elementary hb-graph and for a m-uniform hb-graph.

4.2.4 Normalised $\overline{k}$-adjacency tensor of an elementary hb-graph

We consider an elementary hb-graph $\mathfrak{H}_\mathfrak{e} = (V, (\mathfrak{e}))$ where $V = \{v_i : i \in [\![n]\!]\}$ and $\mathfrak{e}$ is a multiset of universe V and multiplicity function m. The support of $\mathfrak{e}$ is $\mathfrak{e}^\star = \{v_{j_1}, \dots, v_{j_k}\}$ by considering, without loss of generality: $1 \leqslant j_1 < \dots < j_k \leqslant n$.

$\mathfrak{e}$ is the multiset: $\mathfrak{e} = \left\{ v_{j_1}^{m_{j_1}}, \dots, v_{j_k}^{m_{j_k}} \right\}$ where $m_j = m(v_j)$.

The normalised hypermatrix representation of $\mathfrak{e}$, written $\boldsymbol{Q}_\mathfrak{e}$, describes uniquely the mset $\mathfrak{e}$. Thus the elementary hb-graph $\mathfrak{H}_\mathfrak{e}$ is also uniquely described by $\boldsymbol{Q}_\mathfrak{e}$ as $\mathfrak{e}$ is its unique hb-edge. $\boldsymbol{Q}_\mathfrak{e}$ is of rank $r = \#_m\mathfrak{e} = \sum_{j=1}^{k} m_j$ and dimension n.

Hence, the definition:

Definition 2. *Let $\mathfrak{H} = (V, (\mathfrak{e}))$ be an elementary hb-graph with $V = \{v_i : i \in [\![n]\!]\}$ and $\mathfrak{e}$ the multiset $\left\{ v_{j_1}^{m_{j_1}}, \dots, v_{j_k}^{m_{j_k}} \right\}$ of m-range $r = \#_m\mathfrak{e}$, universe V and multiplicity function m.*

*The **normalised $\overline{k}$-adjacency hypermatrix of an elementary hb-graph** $\mathfrak{H}_\mathfrak{e}$ is the normalised representation of the multiset $\mathfrak{e}$, i.e. the symmetric hypermatrix $\boldsymbol{Q}_\mathfrak{e} \triangleq (q_{j_1 \dots j_r})$ of rank r and dimension n where the only nonzero elements are:*

$$q_{\sigma(j_1)^{m_{i\sigma(j_1)}} \dots \sigma(j_k)^{m_{i\sigma(j_k)}}} = \frac{m_{ij_1}! \dots m_{ij_k}!}{(r-1)!}$$

where $\sigma \in \mathcal{S}_{[\![r]\!]}$[6].

In a elementary hb-graph, the $\overline{k}$-adjacency corresponds to $\#_m\mathfrak{e}$-adjacency. This hypermatrix encodes the $\overline{k}$-adjacency of the elementary hb-graph.

4.2.5 Hb-graph polynomial

Homogeneous polynomial associated to a hypermatrix: With an approach similar to Ouvrard et al. [36] where full details are given, let $e_1, \dots, e_n$ be the canonical basis of $\mathbb{R}^n$.

[6] $\mathcal{S}_{[\![r]\!]}$ designates the set of permutations on $[\![r]\!]$

$(e_{i_1} \otimes \ldots \otimes e_{i_k})_{i_1,\ldots,i_k \in [\![n]\!]}$ is a basis of $\mathcal{L}_k^0(\mathbb{K}^n)$, where $\otimes$ is the Segre outer product.

A covariant tensor $\mathcal{Q} \in \mathcal{L}_k^0(\mathbb{K}^n)$ is associated to an hypermatrix $\boldsymbol{Q} = (q_{i_1 \ldots i_r})_{i_1,\ldots,i_r \in [\![n]\!]}$ by writting $\mathcal{Q}$ as:

$$\mathcal{Q} = \sum_{i_1,\ldots,i_r \in [\![n]\!]} q_{i_1 \ldots i_r} e_{i_1} \otimes \ldots \otimes e_{i_r}$$

$\boldsymbol{Q}$ is called the **canonical hypermatrix representation** (CHR for short) of $\mathcal{Q}$.

Considering n variables z_i attached to the n vertices v_i and $z = \sum_{i \in [\![n]\!]} z_i e_i$, the multilinear matrix product $(z, \ldots, z).\boldsymbol{Q} = (z)_{[r]} . \boldsymbol{Q}$ is a polynomial $P(z_0)$[7]:

$$P(z_0) = \sum_{i_1,\ldots,i_r \in [\![n]\!]} q_{i_1 \ldots i_r} z_{i_1} \ldots z_{i_r}$$

of degree r.

Elementary hb-graph polynomial: Let $\mathfrak{H}_{\mathfrak{e}} = (V, (\mathfrak{e}))$ be a hb-graph with $V = \{v_i : i \in [\![n]\!]\}$ and $\mathfrak{e}$ the multiset $\left\{ v_{j_1}^{m_{j_1}}, \ldots, v_{j_k}^{m_{j_k}} \right\}$ of m-range $r = \#_m \mathfrak{e}$, universe V and multiplicity function m.

Using the normalised $\overline{k}$-adjacency hypermatrix $\boldsymbol{Q}_{\mathfrak{e}} = (q_{i_1 \ldots i_r})_{i_1,\ldots,i_r \in [\![n]\!]}$, which is symmetric, we can write the reduced version of its attached homogeneous polynomial $P_{\mathfrak{e}}$:

$$\begin{aligned} P_{\mathfrak{e}}(z_0) &= \frac{r!}{m_{j_1}! \ldots m_{j_k}!} q_{j_1^{m_{j_1}} \ldots j_k^{m_{j_k}}} z_{j_1}^{m_{j_1}} \ldots z_{j_k}^{m_{j_k}} \\ &= \#_m \mathfrak{e} \, z_{j_1}^{m_{j_1}} \ldots z_{j_k}^{m_{j_k}}. \end{aligned}$$

Hb-graph polynomial: Considering a hb-graph $\mathfrak{H} = (V, \mathfrak{E})$ with no-repeated hb-edge, with $V = \{v_i : i \in [\![n]\!]\}$ and $\mathfrak{E} = (\mathfrak{e}_i)_{i \in [\![p]\!]}$.

[7]Where: $z_0 = (z_1, \ldots, z_n)$

This hb-graph is summarized by a polynomial of degree $r_{\mathfrak{H}} = \max\limits_{\mathfrak{e} \in \mathfrak{E}} \#_m (\mathfrak{e})$:

$$\begin{aligned} P(z_0) &\triangleq \sum_{i \in \llbracket p \rrbracket} c_{e_i} P_{e_i}(z_0) \\ &= \sum_{i \in \llbracket p \rrbracket} c_{e_i} \frac{r_i!}{m_{ij_1}! \dots m_{ijk_i}!} q_{j_1^{m_{ij1}} \dots j_{k_i}^{m_{ijk_i}}} z_{j_1}^{m_{ij_1}} \dots z_{j_{k_i}}^{m_{ijk_i}} \\ &= \sum_{i \in \llbracket p \rrbracket} c_{e_i} \#_m \mathfrak{e}_i \, z_{j_1}^{m_{ij_1}} \dots z_{j_{k_i}}^{m_{ijk_i}} \end{aligned}$$

where c_{e_i} is a technical coefficient. $P(z_0)$ is called the **hb-graph polynomial**. The choice of c_{e_i} will be further made in order to retrieve the m-degree of the vertices from the e-adjacency tensor.

4.2.6 $\overline{k}$-adjacency hypermatrix of a m-uniform natural hb-graph

We now extend to m-uniform hb-graphs the $\overline{k}$-adjacency hypermatrix obtained in the case of an elementary hb-graph.

In the case of a r-m-uniform natural hb-graph with no repeated hb-edge, each hb-edge has the same m-cardinality r. Hence the $\overline{k}$-adjacency of a r-m-uniform hb-graph corresponds to r-adjacency where r is the m-rank of the hb-graph. The $\overline{k}$-adjacency tensor of the hb-graph has rank r and dimension n. The elements of the $\overline{k}$-adjacency hypermatrix are:

$$a_{i_1 \dots i_r}$$

with $i_1, \dots, i_r \in \llbracket n \rrbracket$.

The associated hb-graph polynomial is homogeneous of degree r.

We obtain the definition of the $\overline{k}$-adjacency tensor of a r-m-uniform hb-graph by summing the $\overline{k}$-adjacency tensor attached to each hyperedge with a coefficient c_i equals to 1 for each hyperedge.

Definition 3. *Let $\mathfrak{H} = (V, \mathfrak{E})$ be a hb-graph. $V = \{v_i : i \in \llbracket n \rrbracket\}$.*

*The **$\overline{k}$-adjacency hypermatrix of a r-m-uniform hb-graph** $\mathfrak{H} = (V, \mathfrak{E})$ is the hypermatrix $\mathbf{A}_{\mathfrak{H}} = (a_{i_1 \dots i_r})_{i_1, \dots, i_r \in \llbracket n \rrbracket}$ defined by:*

$$\mathbf{A}_{\mathfrak{H}} \triangleq \sum_{i \in \llbracket p \rrbracket} \mathbf{Q}_{e_i}$$

where $\mathbf{Q}_{e_i}$ is the $\overline{k}$-adjacency hypermatrix of the elementary hb-graph associated to the hb-edge $\mathfrak{e}_i = \left\{ v_{j_1}^{m_{ij_1}}, \dots, v_{j_{k_i}}^{m_{ijk_i}} \right\} \in \mathfrak{E}$.

The only non-zero elements of $\mathbf{Q}_{\mathfrak{e}_i}$ *are the elements with indices obtained by permutation of the multiset* $\left\{j_1^{m_{ij_1}}, \dots, j_{k_i}^{m_{ij_{k_i}}}\right\}$ *and are all equal to:*

$$\frac{m_{ij_1}! \dots m_{ij_{k_i}}!}{(r-1)!}.$$

This definition corresponds to the definition given by Pearson and Zhang [43], which is a symmetrized version of the one given in Pearson and Zhang [42].

We can remark that when a r-m-uniform hb-graph has 1 as vertex multiplicity for any vertices in each hb-edge support of all hb-edges, then this hb-graph is a r-uniform hypergraph: in this case, we retrieve the result of the degree-normalized tensor defined in Cooper and Dutle [10].

Claim 4.2. *The m-degree of a vertex* v_j *in a* r*-m-uniform hb-graph* $\mathfrak{H}$ *of the* $\overline{k}$*-adjacency hypermatrix is:*

$$\deg_m(v_j) = \sum_{j_2,\dots,j_r \in [\![n]\!]} a_{jj_2\dots j_r}.$$

Proof. $\sum\limits_{j_2,\dots,j_r \in [\![n]\!]} a_{jj_2\dots j_r}$ has non-zero terms only for the corresponding hb-edges $\mathfrak{e}_i$ that have v_j in it. Such a hb-edge contains v_j and is described by $\mathfrak{e}_i = \left\{v_j^{m_{ij}}, v_{l_2}^{m_{il_2}}, \dots, v_{l_{k_i}}^{m_{il_{k_i}}}\right\}$. It means that the multiset $\{\{j_2, \dots, j_r\}\}$ corresponds exactly to the multiset $\left\{j^{m_{ij}-1}, l_2^{m_{il_2}}, \dots, l_{k_i}^{m_{il_{k_i}}}\right\}$. For each $\mathfrak{e}_i$ such that $v_j \in \mathfrak{e}_i$, there are $\dfrac{(r-1)!}{(m_{ij}-1)!m_{il_2}! \dots m_{il_{k_i}}!}$ possible permutations of the indices j_2 to j_l and $a_{jj_2\dots j_r} = \dfrac{m_{ij}!m_{il_2}! \dots m_{il_{k_i}}!}{(r-1)!}$.

Also: $\sum\limits_{j_2,\dots,j_r \in [\![n]\!]} a_{jj_2\dots j_r} = \sum\limits_{i \in [\![p]\!] \,:\, v_j \in \mathfrak{e}_i} m_{ij} = \deg_m(v_j)$. □

4.2.7 Elementary operations on hb-graphs

In Ouvrard et al. [36], we describe two elementary operations that are used in the hypergraph uniformisation process. We describe here two similar operations and some additional operations for hb-graphs.

Operation 4.1. *Let* $\mathfrak{H} = (V, \mathfrak{E})$ *be a hb-graph.*
Let w_1 *be a constant weighted function on hb-edges with constant value* 1.
The weighted hb-graph $\mathfrak{H}_1 \triangleq (V, \mathfrak{E}, w_1)$ *is called the* ***canonical weighted hb-graph*** *of* $\mathfrak{H}$.
The application $\phi_{cw} : \mathfrak{H} \mapsto \mathfrak{H}_1$ *is called the* ***canonical weighting operation.***

Operation 4.2. *Let* $\mathfrak{H}_1 = (V, \mathfrak{E}, w_1)$ *be a canonical weighted hb-graph.*
Let $c \in \mathbb{R}^{++}$. *Let* w_c *be a constant weighted function on hb-edges with constant value* c.
The weighted hb-graph $\mathfrak{H}_c \triangleq (V, \mathfrak{E}, w_c)$ *is called the* ***c-dilated hb-graph*** *of* $\mathfrak{H}$.
The application $\phi_{c\text{-}d} : \mathfrak{H}_1 \mapsto \mathfrak{H}_c$ *is called the* ***c-dilatation operation.***

Operation 4.3. *Let* $\mathfrak{H}_w = (V, \mathfrak{E}, w)$ *be a weighted hb-graph. Let* $y \notin V$ *be a new vertex.*
The ***y-complemented hb-graph*** *of* $\mathfrak{H}_w$ *is the hb-graph* $\tilde{\mathfrak{H}}_{\tilde{w}} \triangleq \left(\tilde{V}, \tilde{\mathfrak{E}}, \tilde{w}\right)$ *where:*

- $\tilde{V} \triangleq V \cup \{y\}$;
- $\tilde{\mathfrak{E}} \triangleq (\xi(\mathfrak{e}))_{\mathfrak{e} \in \mathfrak{E}}$ *where the map* $\xi : \mathfrak{E} \to \mathcal{M}\left(\tilde{V}\right)$ *is such that for all* $\mathfrak{e} \in \mathfrak{E}$, $\xi(\mathfrak{e}) \triangleq \left\{x^{m_{\xi(\mathfrak{e})}(x)} : x \in \tilde{V}\right\} \in \mathcal{M}\left(\tilde{V}\right)$ *with:*

$$m_{\xi(\mathfrak{e})}(x) \triangleq \begin{cases} m_{\mathfrak{e}}(x) & \text{if } x \in \mathfrak{e}^\star \\ r_{\mathfrak{H}} - \#_m \mathfrak{e} & \text{if } x = y \end{cases};$$

- *the weight function* $\tilde{w}$ *is such that* $\forall \mathfrak{e} \in \mathfrak{E}$: $\tilde{w}(\xi(\mathfrak{e})) \triangleq w(\mathfrak{e})$.

The application $\phi_{y\text{-}c} : \mathfrak{H}_w \mapsto \tilde{\mathfrak{H}}_{\tilde{w}}$ *is called the* ***y-complemented operation.***

Operation 4.4. *Let* $\mathfrak{H}_w = (V, \mathfrak{E}, w)$ *be a weighted hb-graph. Let* $y \notin V$ *be a new vertex. Let* $\alpha \in \mathbb{R}^{++}$.
The y^α***-vertex-increased hb-graph*** *of* $\mathfrak{H}_w$ *is the hb-graph* $\mathfrak{H}^+_{w^+} \triangleq (V^+, \mathfrak{E}^+, w^+)$ *where:*

- $V^+ \triangleq V \cup \{y\}$*;*
- $\mathfrak{E}^+ \triangleq (\phi(\mathfrak{e}))_{\mathfrak{e}\in\mathfrak{E}}$ *where the map* $\phi : \mathfrak{E} \to \mathcal{M}(V^+)$ *such that for all* $\mathfrak{e} \in \mathfrak{E}$, $\phi(\mathfrak{e}) \triangleq \left\{x^{m_{\phi(\mathfrak{e})}(x)} : x \in V^+\right\} \in \mathcal{M}(V^+)$ *with:*

$$m_{\phi(\mathfrak{e})}(x) \triangleq \begin{cases} m_{\mathfrak{e}}(x) & \text{if } x \in \mathfrak{e}^\star \\ \alpha & \text{if } x = y \end{cases};$$

- *the weight function is* w^+ *is such that* $\forall \mathfrak{e} \in \mathfrak{E}$*:* $w^+(\phi(\mathfrak{e})) \triangleq w(\mathfrak{e})$.

The application $\phi_{y^\alpha\text{-}v} : \mathfrak{H}_w \mapsto \mathfrak{H}^+_{w^+}$ *is called the* y^α***-vertex-increasing operation***.

Operation 4.5. *The* ***merged hb-graph*** $\widehat{\mathfrak{H}_{\widehat{w}}} \triangleq \left(\widehat{V}, \widehat{\mathfrak{E}}, \widehat{w}\right)$ *of a family* $(\mathfrak{H}_i)_{i\in I}$ *of weighted hb-graphs with* $\forall i \in I : \mathfrak{H}_i = (V_i, \mathfrak{E}_i, w_i)$ *is the weighted hb-graph where:*

- $\widehat{V} \triangleq \bigcup_{i\in I} V_i$*;*
- $\widehat{\mathfrak{E}} \triangleq (\psi(\mathfrak{e}))_{\mathfrak{e}\in\sum_{i\in I}\mathfrak{E}_i}$[a] *where the map* $\psi : \sum_{i\in I}\mathfrak{E}_i \to \mathcal{M}\left(\widehat{V}\right)$ *such that for all* $\mathfrak{e} \in \sum_{i\in I}\mathfrak{E}_i$, $\psi(\mathfrak{e}) \in \mathcal{M}\left(\widehat{V}\right)$ *and is the multiset* $\left\{x^{m_{\psi(\mathfrak{e})}(x)} : x \in \widehat{V}\right\}$, *with:*

$$m_{\psi(\mathfrak{e})}(x) \triangleq \begin{cases} m_{\mathfrak{e}}(x) & \text{if } x \in \mathfrak{e}^\star \\ 0 & \text{otherwise} \end{cases};$$

- $\forall \mathfrak{e} \in \mathfrak{E}_i$, $\widehat{w}(\mathfrak{e}) \triangleq w_i(\mathfrak{e})$.

The application $\phi_m : (\mathfrak{H}_i)_{i\in I} \mapsto \widehat{\mathfrak{H}}$ *is called the* ***merging operation***.

[a] $\sum_{i\in I}\mathfrak{E}_i$ is the family obtained with all elements of each family $\mathfrak{E}_i$.

Operation 4.6. *Decomposing a hb-graph $\mathfrak{H} = (V, \mathfrak{E})$ into a family of hb-graphs $(\mathfrak{H}_i)_{i \in I}$, where $\mathfrak{H}_i = (V, \mathfrak{E}_i)$ such that $\mathfrak{H} = \bigoplus_{i \in I} \mathfrak{H}_i$ is called a* ***decomposition operation*** *$\phi_d : \mathfrak{H} \mapsto (\mathfrak{H}_i)_{i \in I}$.*

The direct sum of two hb-graphs appears as a merging operation in one way and as a decomposition operation in the opposite way. For a given hb-graph, different decomposition operations exist. Nonetheless, the decomposition in elementary hb-graphs is unique as well as the decomposition in m-uniform hb-graphs representing the different levels of m-uniformity in that hb-graph.

We now focus on the preservation of e-adjacency through these different operations, as it is fundamental to ensure the soundness of the constructed hypermatrix.

Definition 4. *Let $\mathfrak{H} = (V, \mathfrak{E})$ and $\mathfrak{H}' = (V', \mathfrak{E}')$ be two hb-graphs.*

Let $\phi : \mathfrak{H} \mapsto \mathfrak{H}'$ be an elementary operation between a hb-graph and another one.

ϕ is said ***preserving e-adjacency*** *if vertices of V' that are e-adjacent in $\mathfrak{H}'$ and also in V are e-adjacent in $\mathfrak{H}$.*

ϕ is said ***preserving exactly e-adjacency*** *if vertices that are e-adjacent in $\mathfrak{H}'$ are e-adjacent in $\mathfrak{H}$ and reciprocally.*

We can extend these definitions to $\psi : (\mathfrak{H}_i)_{i \in I} \mapsto \mathfrak{H}'$.

Definition 5. *Let $(\mathfrak{H}_i)_{i \in I}$ be a family of hb-graphs with $\forall i \in I, \mathfrak{H}_i = (V_i, \mathfrak{E}_i)$ and $\mathfrak{H}' = (V', \mathfrak{E}')$ a hb-graph.*

Let $\psi : (\mathfrak{H}_i)_{i \in I} \mapsto \mathfrak{H}'$ be an elementary operation between a family of hb-graphs and a hb-graph.

ψ is said ***preserving e-adjacency*** *if vertices that are e-adjacent in $\mathfrak{H}'$ and also in $V = \bigcup_{i \in I} V_i$ are e-adjacent vertices in exactly one of the $\mathfrak{H}_i, i \in I$.*

ψ is said ***preserving exactly e-adjacency*** *if vertices that are e-adjacent in $\mathfrak{H}'$ are e-adjacent in exactly one of the $\mathfrak{H}_i, i \in I$ and reciprocally.*

We can extend these definitions to $\nu : \mathfrak{H} \mapsto (\mathfrak{H}_i)_{i \in I}$.

Definition 6. *Let $(\mathfrak{H}_i)_{i \in I}$ be a family of hb-graphs with $\forall i \in I, \mathfrak{H}_i = (V_i, \mathfrak{E}_i)$ and $\mathfrak{H} = (V, \mathfrak{E})$ a hb-graph.*

Let $\nu : \mathfrak{H} \mapsto (\mathfrak{H}_i)_{i \in I}$ be an elementary operation between a hb-graph and a family of hb-graphs.

*ν is said **preserving** e-adjacency if vertices that are e-adjacent in one of the $\mathfrak{H}_i, i \in I$ and also in V are e-adjacent in $\mathfrak{H}$.*

*ν is said **preserving exactly** e-adjacency if vertices that are e-adjacent in one of the $\mathfrak{H}_i, i \in I$ are e-adjacent in $\mathfrak{H}$ and reciprocally.*

Claim 4.3. *Let $\mathfrak{H} = (V, \mathfrak{E})$ be a hb-graph.*

The canonical weighting operation, the c-dilatation operation, the merging operation, and the decomposition operation preserve exactly e-adjacency.

The y-complemented operation and the y^{α}-vertex-increasing operation preserve e-adjacency.

Proof. Immediate. □

Claim 4.4. *The composition of two operations which preserve (respectively exactly) e-adjacency preserves (respectively exactly) e-adjacency.*

The composition of two operations where one preserves exactly e-adjacency and the other preserves e-adjacency preserves e-adjacency.

Proof. Immediate. □

4.2.8 Processes involved for building the e-adjacency tensor

In a general natural hb-graph $\mathfrak{H}$, hb-edges do not have the same m-cardinality: the rank of the $\overline{k}$-adjacency tensor of the elementary hb-graph associated to each hb-edge depends on the m-cardinality of the hb-edge. As a consequence, the hb-graph polynomial is no more homogeneous. Nonetheless, techniques to homogenize such a polynomial are well known.

We introduce here the hb-graph m-uniformisation process (Hm-UP for short) which transforms a given hb-graph of m-range $r_{\mathfrak{H}}$ into a $r_{\mathfrak{H}}$-m-uniform hb-graph written $\overline{\mathcal{H}}$: this uniformisation can be mapped to the homogenization of the attached polynomial of the original hb-graph, called the polynomial homogenization process (PHP).

The Hm-UP can be achieved by different means of filling the not-at-the-level hb-edges so they reach a m-range of $r_{\mathfrak{H}}$:

- **straightforward m-uniformisation** levels directly all hb-edges by adding a Null vertex Y_1 with a multiplicity being the difference between the hb-graph m-rank and the hb-edge m-cardinality. It is achieved by considering the Y_1-complemented hb-graph of $\mathfrak{H}$.

silo m-uniformisation processes each of the m-uniform sub-hb-graphs obtained by gathering all hb-edges of a given m-cardinality r in a single sub-hb-graph, which is then $Y_r^{r_{\mathfrak{H}}-r}$-vertex-increased. A single $r_{\mathfrak{H}}$-m-uniformized hb-graph is then obtained by merging them.

- **layered m-uniformisation** processes m-uniform sub-hb-graphs of increasing m-cardinality by successively adding a vertex and merging it to the sub-hb-graph of the above layer. The layered homogenization process applied to hypergraphs was explained with full details in Ouvrard et al. [36]; it involves two-phase step iterations based on successive $\{Y_k^1\}$-vertex-increased hb-graphs and merging with the dilated weighted hb-graph of the next layer.

4.2.9 On the choice of the technical coefficient $c_{\mathfrak{e}_i}$

To comply to the expectations, the technical coefficient $c_{\mathfrak{e}_i}$ has to be chosen such that by using the elements of the e-adjacency hypermatrix $\boldsymbol{A} = (a_{i_1 \dots i_r})_{i_1,\dots,i_r \in [\![n]\!]}$, one can retrieve:

1. the m-degree of the vertices: $\sum\limits_{i_2,\dots,i_r \in [\![n]\!]} a_{i i_2 \dots i_r} = \deg_m(v_i)$.
2. the number of hb-edges $|\mathfrak{E}|$.

Similarly to Ouvrard et al. [35], we consider a hb-graph $\mathfrak{H} = (V, \mathfrak{E})$ that we decompose in a family of r-m-uniform hb-graphs $(\mathfrak{H}_r)_{r \in [\![r_{\mathfrak{H}}]\!]}$.

We consider $\mathcal{R}$ the equivalency relation defined on the family of hb-edges $\mathfrak{E}$ of $\mathfrak{H}$: $\mathfrak{e} \mathcal{R} \mathfrak{e}' \Leftrightarrow \#_m \mathfrak{e} = \#_m \mathfrak{e}'$.

$\mathfrak{E}/\mathcal{R}$ is the set of classes of hb-edges of same m-cardinality. The elements of $\mathfrak{E}/\mathcal{R}$ are the sets: $\mathfrak{E}_r = \{\mathfrak{e} \in \mathfrak{E} : \#_m \mathfrak{e} = r\}$.

Considering $R = \{r : \mathfrak{E}_r \in \mathfrak{E}/\mathcal{R}\}$, we set $\mathfrak{E}_r = \emptyset$ for all $r \in [\![r_{\mathfrak{H}}]\!] \backslash R$.

For all $r \in [\![r_{\mathfrak{H}}]\!]$, $\mathfrak{H}_r = (V, \mathfrak{E}_r)$ is r-m-uniform.

It holds: $\mathfrak{E} = \bigcup\limits_{r \in [\![r_{\mathfrak{H}}]\!]} \mathfrak{E}_r$ and $\mathfrak{E}_{r_1} \cap \mathfrak{E}_{r_2} = \emptyset$ for all $r_1 \neq r_2$, hence $(\mathfrak{E}_r)_{r \in [\![r_{\mathfrak{H}}]\!]}$ constitutes a partition of $\mathfrak{E}$ which is unique from the way it has been defined.

Hence:

$$\mathfrak{H} = \bigoplus_{r \in [\![r_{\mathfrak{H}}]\!]} \mathfrak{H}_r.$$

Each of these r-m-uniform hb-graph $\mathfrak{H}_r$, where the $\overline{k}$-adjacency is achieved by r-adjacency, can be associated to a $\overline{k}$-adjacency tensor $\mathcal{A}_r$ viewed as a hypermatrix $\boldsymbol{A}_{\mathfrak{H}_r} = \left(a_{(r) i_1 \dots i_r}\right)$ of order r, hyper-cubic and symmetric of dimension $|V| = n$.

We write $\left(a_{i_1 \ldots i_{r_{\mathfrak{H}}}}\right)_{i_1,\ldots,i_{r_{\mathfrak{H}}} \in [\![n_1]\!]}$ the e-adjacency hypermatrix associated to $\mathfrak{H}$ where $n_1 = n + n_A$, n_A corresponds to the number of different special vertices added in the hb-edges. n_A depends on the way the hypermatrix is built:

- $n_A = 1$ for the straightforward process;
- $n_A = r_{\mathfrak{H}} - 1$ for the two silo and layered processes.

For a given $r \in [\![r_{\mathfrak{H}}]\!]$, the number of hb-edges in $\mathfrak{H}_r$ is given by summing the elements of $\boldsymbol{A}_{\mathfrak{H}_r}$:

$$\begin{aligned}\sum_{i_1,\ldots,i_r \in [\![n]\!]} a_{(r)i_1 \ldots i_r} &= \sum_{i=1}^{n} \sum_{i_2,\ldots,i_r \in [\![n]\!]} a_{(r)ii_2 \ldots i_r} \\ &= \sum_{i=1}^{n} \deg_m(v_i) \\ &= r\,|\mathfrak{E}_r|\end{aligned}$$

In the m-uniformized hb-graph of $\mathfrak{H}$, the number of hb-edges can also be calculated using:

$$\sum_{i_1,\ldots,i_{r_{\mathfrak{H}}} \in [\![n_1]\!]} a_{i_1 \ldots i_{r_{\mathfrak{H}}}} = r_{\mathfrak{H}}\,|\mathfrak{E}|.$$

As

$$\begin{aligned}|\mathfrak{E}| &= \sum_{r=1}^{r_{\mathfrak{H}}} |\mathfrak{E}_r| \\ &= \sum_{r=1}^{r_{\mathfrak{H}}} \frac{1}{r} \sum_{i_1,\ldots,i_r \in [\![n]\!]} a_{(r)i_1 \ldots i_r}\end{aligned}$$

It follows:

$$\sum_{i_1,\ldots,i_{r_{\mathfrak{H}}} \in [\![n_1]\!]} a_{i_1 \ldots i_{r_{\mathfrak{H}}}} = \sum_{r=1}^{r_{\mathfrak{H}}} \frac{r_{\mathfrak{H}}}{r} \sum_{i_1,\ldots,i_r \in [\![n]\!]} a_{(r)i_1 \ldots i_r}.$$

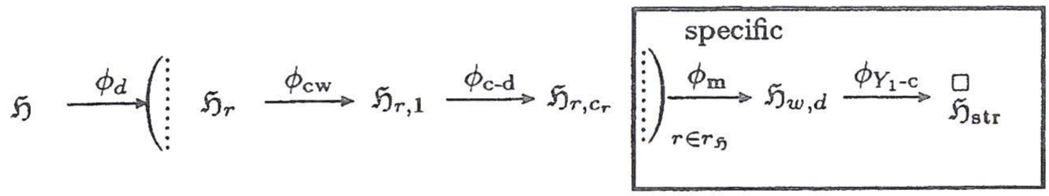


Figure 6: Operations on the original hb-graph to m-uniformize it in the straightforward approach. Parenthesis with vertical dots indicate parallel operations.

Also, choosing for all $i \in [\![p]\!]$: $c_{\mathfrak{e}_i} = \frac{r_{\mathfrak{H}}}{r}$ where $r = \#_m \mathfrak{e}_i$, we write for all $r \in [\![r_{\mathfrak{H}}]\!]$:

$$c_r \triangleq \frac{r_{\mathfrak{H}}}{r}.$$

c_r is the technical coefficient for the corresponding layer of level r of the hb-graph $\mathfrak{H}$.

Hence, the Hm-UP is initiated by applying the canonical weighting to each m-uniform hb-graph $\mathfrak{H}_r$ that transforms it into $\mathfrak{H}_{r,1}$. Then the c_r-dilatation operation is applied to each weighted m-uniform hb-graph $\mathfrak{H}_{r,1}$ to obtain its c_r-dilated hb-graph $\mathfrak{H}_{r,c_r}$.

4.2.10 Straightforward approach

Straightforward m-uniformisation: We first decompose $\mathfrak{H} = \bigoplus_{r \in [\![r_{\mathfrak{H}}]\!]} \mathfrak{H}_r$ as seen in sub-section 4.2.9.

We then transform each $\mathfrak{H}_r$, $r \in [\![r_{\mathfrak{H}}]\!]$ into a canonical weighted hb-graph $\mathfrak{H}_{r,1}$ that is dilated with the help of the dilatation coefficient c_r to obtain the c_r-dilated hb-graph $\mathfrak{H}_{r,c_r}$.

This family $(\mathfrak{H}_{r,c_r})$ is then merged into the hb-graph: $\mathfrak{H}_{w,d} = \bigoplus_{r \in [\![r_{\mathfrak{H}}]\!]} \mathfrak{H}_{r,c_r}$. To get a m-uniform hb-graph we finally generate a vertex $Y_1 \notin V$ and apply to $\mathfrak{H}_{w,d}$ the Y_1-complemented operation to obtain $\tilde{\mathfrak{H}}_{w,d} = \overset{\square}{\mathfrak{H}}_{\text{sil}}$ the Y_1-complemented hb-graph of $\mathfrak{H}_{w,d}$. $\overset{\square}{\mathfrak{H}}_{\text{str}}$ is called the **straightforward m-uniformized hb-graph** of $\mathfrak{H}$.

The different steps are summarized in Figure 6.

Claim 4.5. *The transformation* $\phi_s : \mathfrak{H} \mapsto \overset{\square}{\mathfrak{H}}_{str}$ *preserves e-adjacency.*

Proof. $\phi_{\text{s}} = \phi_{y_1\text{-c}} \circ \phi_{\text{m}} \circ \left(\vdots \phi_{\text{c-d}} \circ \phi_{\text{cw}} \vdots \right) \circ \phi_{\text{d}}$.[8]

All these operations either preserve e-adjacency or preserve exactly e-adjacency. Also, by composition, ϕ_{s} preserves e-adjacency. □

Straightforward homogenization: In order to homogenize the hb-graph polynomial we introduce an additional variable y_1 that corresponds to the additional vertex Y_1 used during the Hm-UP.

The normalised $\overline{k}$-adjacency hypermatrix of the elementary hb-graph corresponding to the hb-edge $e_i = \left\{ v_{j_1}^{m_{ij_1}}, \ldots, v_{j_{k_i}}^{m_{ij_{k_i}}} \right\}$ is $\boldsymbol{Q}_{e_i}$ of rank $\rho_i = \#_m e_i$ and dimension n. The corresponding reduced polynomial is $P_{e_i}(\boldsymbol{z_0}) = \rho_i z_{j_1}^{m_{ij_1}} \ldots z_{j_{k_i}}^{m_{ij_{k_i}}}$.

To transform this polynomial of degree ρ_i into a polynomial of degree $r_{\mathfrak{H}}$ we have to multiply it by $y_1^{m_{i\,n+1}}$ where $m_{i\,n+1} = r_{\mathfrak{H}} - \rho_i$. It corresponds to adding the vertex Y_1 with multiplicity $m_{i\,n+1}$.

The term $P_{e_i}(\boldsymbol{z_0})$ with the attached tensor $\mathcal{P}_{e_i}$ of rank ρ_i and dimension n is transformed in:

$$R_{e_i}(\boldsymbol{z_1}) = P_{e_i}(\boldsymbol{z_0})\, y_1^{m_{i\,n+1}} = \rho_i z_{j_1}^{m_{ij_1}} \ldots z_{j_{k_i}}^{m_{ij_{k_i}}} y_1^{m_{i\,n+1}}$$

[9] with the attached tensor $\mathcal{R}_{e_i}$ of rank $r_{\mathfrak{H}}$ and dimension $n+1$.

The CHR of the tensor $\mathcal{R}_{e_i}$ is the hypermatrix $\boldsymbol{R}_{e_i} = \left(r_{i_1 \ldots i_{r_{\mathfrak{H}}}} \right)$. The elements that are non-zero in $\boldsymbol{R}_{e_i}$ have all the same value:

$$\rho_{\text{str},e_i} = \rho_i \frac{m_{ij_1}! \ldots m_{ij_{k_i}}! m_{i\,n+1}!}{r_{\mathfrak{H}}!}.$$

The indices of the non-zero elements of $\boldsymbol{R}_{e_i}$ are obtained by permutation of the elements of the multiset:

$$\left\{ j_1^{m_{ij_1}}, \ldots, j_{k_i}^{m_{ij_{k_i}}}, (n+1)^{m_{i\,n+1}} \right\}.$$

The number of possible permutations is:

$$\frac{r_{\mathfrak{H}}!}{m_{ij_1}! \ldots m_{ij_{k_i}}! m_{i\,n+1}!}.$$

[8] $\left(\vdots \quad \vdots \right)_{\ldots \in \ldots}$ indicates parallel operations on each member of the family as specified in index of the right parenthesis.

[9] $\boldsymbol{z_k} = (z_1, \ldots, z_n, y_1, \ldots, y_k)$

The hb-graph polynomial $P(z_0) = \sum\limits_{i \in [\![p]\!]} c_i P_{\mathfrak{e}_i}(z_0)$ is transformed into a homogeneous polynomial:

$$R(z_1) = \sum_{i \in [\![p]\!]} c_i R_{\mathfrak{e}_i}(z_1) = \sum_{i \in [\![p]\!]} c_i z_{j_1}^{m_{ij_1}} \dots z_{j_{k_i}}^{m_{ij_{k_i}}} y_1^{m_{i\,n+1}}$$

□

representing the straightforward m-uniformized hb-graph $\mathfrak{H}_{\text{str}}$ of $\mathfrak{H}$ with attached hypermatrix $\boldsymbol{R} = \sum\limits_{i=1}^{p} c_{\mathfrak{e}_i} \boldsymbol{R}_{\mathfrak{e}_i}$ where $c_{\mathfrak{e}_i} = \frac{r_{\mathfrak{H}}}{\rho_i} = \frac{r_{\mathfrak{H}}}{\#_m \mathfrak{e}_i}$. This provides a direct homogenization of the whole hb-graph polynomial.

Definition 7. *The **straightforward e-adjacency tensor** $\mathcal{A}_{str,\mathfrak{H}}$ of a hb-graph $\mathfrak{H} = (V, \mathfrak{E})$ is the tensor of CHR $\boldsymbol{A}_{str,\mathfrak{H}}$ defined by:*

$$\boldsymbol{A}_{str,\mathfrak{H}} \triangleq \sum_{i \in [\![p]\!]} c_{\mathfrak{e}_i} \boldsymbol{R}_{\mathfrak{e}_i}.$$

where for $\mathfrak{e}_i = \left\{ v_{j_1}^{m_{ij_1}}, \dots, v_{j_{k_i}}^{m_{ij_{k_i}}} \right\} \in \mathfrak{E}$, $c_{\mathfrak{e}_i} = \frac{r_{\mathfrak{H}}}{\#_m \mathfrak{e}_i}$ is the dilatation coefficient and $\boldsymbol{R}_{\mathfrak{e}_i} = \left(r_{i_1 \dots i_{r_{\mathfrak{H}}}} \right)$ is the hypermatrix whose elements have only two possible values: 0 and:

$$\rho_{str,\mathfrak{e}_i} = \frac{m_{ij_1}! \dots m_{ij_{k_i}}! m_{i\,n+1}!}{r_{\mathfrak{H}}!} \#_m \mathfrak{e}_i$$

—with $m_{i\,n+1} = r_{\mathfrak{H}} - \#_m \mathfrak{e}_i$. The indices of the non-zero elements of $\boldsymbol{R}_{\mathfrak{e}_i}$ are obtained by permutation of the elements of the multiset:

$$\left\{ j_1^{m_{ij_1}}, \dots, j_{k_i}^{m_{ij_{k_i}}}, (n+1)^{m_{i\,n+1}} \right\}.$$

Remark 4.1. *In practice, writing $\boldsymbol{A}_{str,\mathfrak{H}} = \left(a_{l_1 \cdots l_{r_{\mathfrak{H}}}} \right)$, the element of $\boldsymbol{A}_{str,\mathfrak{H}}$ of indices $l_1, \cdots, l_{r_{\mathfrak{H}}}$ such that:*

$$\{\{l_1, \cdots, l_{r_{\mathfrak{H}}}\}\} = \left\{ j_1^{m_{ij_1}}, \dots, j_{k_i}^{m_{ij_{k_i}}}, [n+1]^{m_{i\,n+1}} \right\},$$

corresponding to a hb-edge $\mathfrak{e}_i = \left\{ v_{j_1}^{m_{ij_1}}, \dots, v_{j_{k_i}}^{m_{ij_{k_i}}} \right\}$ of the original hb-graph $\mathfrak{H}$, is:

$$a_{l_1 \dots l_{r_{\mathfrak{H}}}} = \frac{m_{ij_1}! \dots m_{ij_{k_i}}! m_{i\,n+1}!}{(r_{\mathfrak{H}} - 1)!}$$

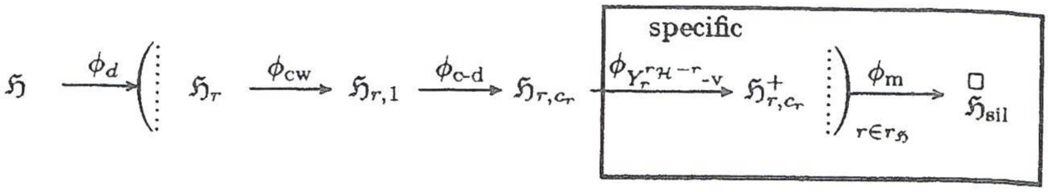


Figure 7: Operations on the original hb-graph to m-uniformize it in the silo approach. Parenthesis with vertical dots indicate parallel operations.

4.2.11 Silo approach

Silo m-uniformisation: The first steps are similar to the straightforward approach.

The hb-graph $\mathfrak{H}$ is decomposed in layers $\mathfrak{H} = \bigoplus_{r \in [\![r_\mathfrak{H}]\!]} \mathfrak{H}_r$ as described in sub-section 4.2.9. Each $\mathfrak{H}_r, r \in [\![r_\mathfrak{H}]\!]$ is canonically weighted and c_r-dilated to obtain $\mathfrak{H}_{r,c_r}$.

We generate $r_\mathfrak{H} - 1$ new vertices $Y_i \notin V$, $i \in [\![r_\mathfrak{H} - 1]\!]$.

We then apply to each $\mathfrak{H}_{r,c_r}, r \in [\![r_\mathfrak{H} - 1]\!]$ the $Y_r^{r_\mathfrak{H}-r}$-vertex-increasing operation to obtain $\mathfrak{H}^+_{r,c_r}$ the $Y_r^{r_\mathfrak{H}-r}$-complemented hb-graph for each $\mathfrak{H}_{r,c_r}, r \in [\![r_\mathfrak{H} - 1]\!]$. The family $\left(\mathfrak{H}^+_{r,c_r}\right)_{r \in r_\mathfrak{H}}$ is then merged using the merging operation to obtain the $r_\mathfrak{H}$-m-uniform hb-graph $\widehat{\mathfrak{H}_{\widehat{w}}} = \overset{\square}{\mathfrak{H}}_{sil}$. $\overset{\square}{\mathfrak{H}}_{sil}$ is called the **silo m-uniformized hb-graph** of $\mathfrak{H}$.

The different steps are summarized in Figure 7.

Claim 4.6. *The transformation $\phi_s : \mathfrak{H} \mapsto \overset{\square}{\mathfrak{H}}_{sil}$ preserves e-adjacency.*

Proof. $\phi_s = \phi_m \circ \left(\vdots \phi_{y_r^{r_\mathfrak{H}-r}\text{-v}} \circ \phi_{c\text{-}d} \circ \phi_{cw} \vdots \right) \circ \phi_d$.

The operations involved in ϕ_s either preserve e-adjacency or preserve exactly e-adjacency: also, by composition, ϕ_s preserves e-adjacency. ☐

Silo homogenization: In this homogenization process we suppose that the hb-edges are sorted by m-cardinality.

During the silo uniformisation, we added $r_\mathfrak{H} - 1$ vertices Y_1 to $Y_{r_\mathfrak{H}-1}$ into the universe, i.e. the vertex set. These vertices correspond to $r_\mathfrak{H} - 1$ additional variables—respectively y_1 to $y_{r_\mathfrak{H}-1}$—that we introduce to homogenize the hb-graph polynomial.

The normalised $\bar{k}$-adjacency hypermatrix of the elementary hb-graph corresponding to the hb-edge $\mathfrak{e}_i = \left\{ v_{j_1}^{m_{ij_1}}, \ldots, v_{j_{k_i}}^{m_{ij_{k_i}}} \right\}$ is $\boldsymbol{Q}_{\mathfrak{e}_i}$ of rank $\rho_i = \#_m \mathfrak{e}_i$ and dimension n. The corresponding reduced polynomial is $P_{\mathfrak{e}_i}(z_0) = \rho_i z_{j_1}^{m_{ij_1}} \ldots z_{j_{k_i}}^{m_{ij_{k_i}}}$ of P has for degree the m-cardinality of the hb-edge $\mathfrak{e}_i$, i.e. $\#_m \mathfrak{e}_i$. To transform it into a polynomial of degree $r_{\mathfrak{H}}$, we use the additional variable $y_{\#_m \mathfrak{e}_i}$ with multiplicity $m_{i\,\#_m \mathfrak{e}_i} = r_{\mathfrak{H}} - \#_m \mathfrak{e}_i$.

The term $P_{\mathfrak{e}_i}(z_0)$ with attached tensor $\mathcal{P}_{\mathfrak{e}_i}$ of rank $\#_m \mathfrak{e}_i$ and dimension n is transformed in $R_{\mathfrak{e}_i}(z_{\#_m \mathfrak{e}_i}) = P_{\mathfrak{e}_i}(z_0)\, y_{\#_m \mathfrak{e}_i}^{m_{i\,n+\#_m \mathfrak{e}_i}}$ with attached tensor $\mathcal{R}_{\mathfrak{e}_i}$ of rank $r_{\mathfrak{H}}$ and dimension $n+1$.

The CHR of the tensor $\mathcal{R}_{\mathfrak{e}_i}$ is the hypermatrix $\boldsymbol{R}_{\mathfrak{e}_i} = \left(r_{i_1 \ldots i_{r_{\mathfrak{H}}}} \right)$. All the non-zero elements of $\boldsymbol{R}_{\mathfrak{e}_i}$ have the same value:

$$\rho_{\text{sil},\mathfrak{e}_i} = \rho_i \frac{m_{ij_1}! \ldots m_{ij_{k_i}}! m_{i\,n+\#_m \mathfrak{e}_i}!}{r_{\mathfrak{H}}!}$$

—with $m_{i\,n+\#_m \mathfrak{e}_i} = r_{\mathfrak{H}} - \#_m \mathfrak{e}_i$. The indices of the non-zero elements of $\boldsymbol{R}_{\mathfrak{e}_i}$ are obtained by permutation of the multiset:

$$\left\{ j_1^{m_{ij_1}}, \ldots, j_{k_i}^{m_{ij_{k_i}}}, [n + \#_m \mathfrak{e}_i]^{m_{i\,n+\#_m \mathfrak{e}_i}} \right\}.$$

P is transformed into a homogeneous polynomial

$$R(z_{r_{\mathfrak{H}}-1}) = \sum_{i \in [\![p]\!]} c_i R_{\mathfrak{e}_i}(z_{\#_m \mathfrak{e}_i}) = \sum_{i \in [\![p]\!]} c_i z_{j_1}^{m_{ij_1}} \ldots z_{j_{k_i}}^{m_{ij_{k_i}}} y_{\#_m \mathfrak{e}_i}^{m_{i\,n+\#_m \mathfrak{e}_i}}$$

□

representing the silo m-uniformized hb-graph $\mathfrak{H}_{\text{sil}}$ of $\mathfrak{H}$ with attached hypermatrix $\boldsymbol{R} = \sum_{i \in [\![p]\!]} c_{\mathfrak{e}_i} \boldsymbol{R}_{\mathfrak{e}_i}$ where: $c_{\mathfrak{e}_i} = \frac{r_{\mathfrak{H}}}{\rho_i} = \frac{r_{\mathfrak{H}}}{\#_m \mathfrak{e}_i}$.

Definition 8. *The silo e-adjacency tensor $\mathcal{A}_{sil,\mathfrak{H}}$ of a hb-graph $\mathfrak{H} = (V, \mathfrak{E})$ is the tensor of CHR $\boldsymbol{A}_{sil,\mathfrak{H}} \triangleq \left(a_{i_1 \ldots i_{r_{\mathfrak{H}}}} \right)_{i_1, \ldots, i_{r_{\mathfrak{H}}} \in [\![n]\!]}$ defined by:*

$$\boldsymbol{A}_{sil,\mathfrak{H}} \triangleq \sum_{i \in [\![p]\!]} c_{\mathfrak{e}_i} \boldsymbol{R}_{\mathfrak{e}_i}$$

and where for $\mathfrak{e}_i = \left\{ v_{j_1}^{m_{ij_1}}, \ldots, v_{j_{k_i}}^{m_{ij_{k_i}}} \right\} \in \mathfrak{E}$, $c_{\mathfrak{e}_i} = \frac{r_{\mathfrak{H}}}{\#_m \mathfrak{e}_i}$ is the dilatation coefficient and $\boldsymbol{R}_{\mathfrak{e}_i} = \left(r_{i_1 \ldots i_{r_{\mathfrak{H}}}} \right)$ is the hypermatrix whose elements have only two possible values, 0 and:

$$\rho_{sil,\mathfrak{e}_i} = \frac{m_{ij_1}! \ldots m_{ij_{k_i}}! m_{i\,n+\#_m \mathfrak{e}_i}!}{r_{\mathfrak{H}}!} \#_m \mathfrak{e}_i$$

—*with:* $m_{i\,n+\#_m e_i} = r_{\mathfrak{H}} - \#_m e_i$. *The indices of the non-zero elements of* $\boldsymbol{R}_{e_i}$ *are obtained by permutation of the elements of the multiset:*

$$\left\{ j_1^{m_{ij_1}}, \dots j_{k_i}^{m_{ij_{k_i}}}, [n+\#_m e_i]^{m_{i\,n+\#_m e_i}} \right\}.$$

Remark 4.2. *In this case,*

$$\boldsymbol{A}_{sil,\mathfrak{H}} = \sum_{r \in [\![r_{\mathfrak{H}}]\!]} c_r \sum_{e_i \in \{e : \#_m e = r\}} \boldsymbol{R}_{e_i}$$

where $c_r = \dfrac{r_{\mathfrak{H}}}{r}$.

Remark 4.3. *In practice, writing* $\boldsymbol{A}_{sil,\mathfrak{H}} = \left(a_{l_1 \cdots l_{r_{\mathfrak{H}}}} \right)$, *the element of* $\boldsymbol{A}_{sil,\mathfrak{H}}$ *of indices* $l_1, \cdots, l_{r_{\mathfrak{H}}}$ *such that:*

$$\{\{l_1, \dots, l_{r_{\mathfrak{H}}}\}\} = \left\{ j_1^{m_{ij_1}}, \dots, j_{k_i}^{m_{ij_{k_i}}}, [n+\#_m e_i]^{m_{i\,n+\#_m e_i}} \right\},$$

corresponding to a hb-edge $e_i = \left\{ v_{j_1}^{m_{ij_1}}, \dots, v_{j_{k_i}}^{m_{ij_{k_i}}} \right\}$ *of the original hb-graph* $\mathfrak{H}$, *is:*

$$a_{l_1 \dots l_{r_{\mathfrak{H}}}} = \frac{m_{ij_1}! \dots m_{ij_{k_i}}! m_{i\,n+\#_m e_i}!}{(r_{\mathfrak{H}} - 1)!}$$

4.2.12 Layered approach

Layered uniformisation: The first steps are similar to the straightforward approach.

The hb-graph $\mathfrak{H}$ is decomposed in layers $\mathfrak{H} = \bigoplus_{r \in [\![r_{\mathfrak{H}}]\!]} \mathfrak{H}_r$ as described in sub-section 4.2.9. Each $\mathfrak{H}_r, r \in [\![r_{\mathfrak{H}}]\!]$ is canonically weighted and c_r-dilated to obtain $\mathfrak{H}_{r,c_r}$.

We generate $r_{\mathfrak{H}} - 1$ new vertices $Y_i \notin V$, $i \in [\![r_{\mathfrak{H}} - 1]\!]$ and write $V_s = \{Y_i : i \in [\![r_{\mathfrak{H}} - 1]\!]\}$

A two-phase step iteration is considered: the inflation phase (IP) and the merging phase (MP). At step $k = 0$, $\mathcal{K}_0 = \mathfrak{H}_{1,c_1}$ and no further action is made but increasing k by 1. At step $k > 0$, the input is the k-m-uniform weighted hb-graph $\mathcal{K}_k$ obtained from the previous iteration. In the IP, $\mathcal{K}_k$ is transformed into $\mathcal{K}_k^+$ the Y_k^1-vertex-increased hb-graph, which is $(k+1)$-m-uniform.

The MP merges the hypergraphs $\mathcal{K}_k^+$ and $\mathfrak{H}_{k+1,c_{k+1}}$ into a single $(k+1)$-m-uniform hb-graph $\widehat{\mathcal{K}_{\widehat{w}}}$.

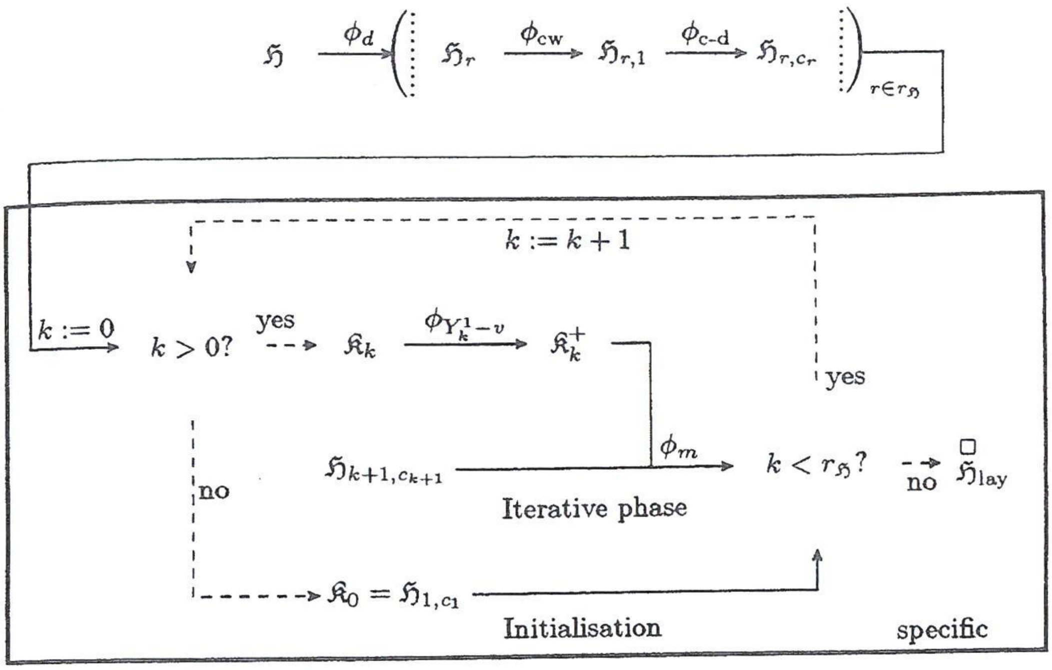


Figure 8: Operations on the original hb-graph to m-uniformize it in the layered approach. Parenthesis with vertical dots indicate parallel operations.

We iterate while $k < r_{\mathfrak{H}}$, increasing in between each step k by 1. When k reaches $r_{\mathfrak{H}}$, we stop iterating and the last $\widehat{\mathcal{K}_{\widehat{w}}}$ obtained, written $\overset{\square}{\mathfrak{H}}_{\text{lay}}$ is called the layered m-uniformized hb-graph of $\mathfrak{H}$.

The different steps are summarized in Figure 8.

Claim 4.7. *The transformation* $\phi_s : \mathfrak{H} \mapsto \overset{\square}{\mathfrak{H}}_{lay}$ *preserves e-adjacency.*

Proof. $\phi_s = \psi \circ \left(\vdots \phi_{\text{c-d}} \circ \phi_{\text{cw}} \vdots \right) \circ \phi_{\text{d}}$, where ψ is called the iterative layered operation that converts the family obtained by $\left(\vdots \phi_{\text{c-d}} \circ \phi_{\text{cw}} \vdots \right) \circ \phi_{\text{d}}$ and transform it into the V_S-layered m-uniform hb-graph of $\mathfrak{H}$.

All the operations $\phi_{\text{c-d}}$, $\phi_{\text{c-w}}$ and ϕ_{d} either preserve e-adjacency or preserve exactly e-adjacency, and so forth by composition.

The iterative layered operation preserves e-adjacency as the operations involved are preserving e-adjacency and that the family of hb-graphs at the input contains hb-edge families that are totally distinct.

Also by composition ϕ_s preserves e-adjacency. □

Layered homogenization: The idea is to sort the hb-edges as in the silo homogenization and consider as well $r_{\mathfrak{H}} - 1$ additional vertices Y_1 to $Y_{r_{\mathfrak{H}}-1}$ into the universe, corresponding to $r_{\mathfrak{H}} - 1$ additional variables respectively y_1 to $y_{r_{\mathfrak{H}}-1}$.

But in this case, these vertices are added successively to each hb-edge to fill the hb-edges so they reach all the same m-cardinality $r_{\mathfrak{H}}$: a hb-edge of initial cardinality $\#_m \mathfrak{e}$ will be filled with elements $Y_{\#_m \mathfrak{e}}$ to $Y_{r_{\mathfrak{H}}-1}$. It corresponds to adding the k-m-uniform sub-hb-graph $\mathfrak{H}_k$ with the $k+1$-m-uniform sub-hb-graph $\mathfrak{H}_{k+1}$ by filling the hb-edge of $\mathfrak{H}_k$ with the additional vertex Y_k to get a homogenized $k+1$-m-uniform sub-hb-graph of the homogenized hb-graph $\overline{\mathfrak{H}}$.

The normalised $\overline{k}$-adjacency hypermatrix of the elementary hb-graph corresponding to the hb-edge $\mathfrak{e}_i = \left\{ v_{j_1}^{m_{ij_1}}, \dots, v_{j_{k_i}}^{m_{ij_{k_i}}} \right\}$ is $\boldsymbol{Q}_{\mathfrak{e}_i}$ of rank $\rho_i = \#_m \mathfrak{e}_i$ and dimension n. The corresponding reduced polynomial is:

$$P_{\mathfrak{e}_i}(z_0) = \rho_i z_{j_1}^{m_{ij_1}} \dots z_{j_{k_i}}^{m_{ij_{k_i}}}$$

of degree $\#_m \mathfrak{e}_i$.

All the hb-edges of same m-cardinality m belongs to the same layer of level m. To transform the hb-edge of m-cardinality $\#_m \mathfrak{e}_i + 1$ we fill it with the element $Y_{\#_m \mathfrak{e}_i}$.

In this case, the polynomial $P_{\mathfrak{e}_i}(z_0)$ is transformed into:

$$R_{(1)\mathfrak{e}_i}\left(z_{\#_m \mathfrak{e}_i}\right) = P_{\mathfrak{e}_i}(z_0)\, y_{\#_m \mathfrak{e}_i}^1$$

of degree $\#_m \mathfrak{e}_i + 1$.

Iterating over the layers, the polynomial $P_{\mathfrak{e}_i}(z_0)$ is transformed into the polynomial:

$$R_{(r_{\mathfrak{H}} - \#_m \mathfrak{e}_i)\mathfrak{e}_i}\left(z_{r_{\mathfrak{H}}-1}\right) = P_{\mathfrak{e}_i}(z_0)\, y_{\#_m \mathfrak{e}_i}^1 \dots y_{r_{\mathfrak{H}}-1}^1$$

of degree $r_{\mathfrak{H}}$.

The polynomial $P_{\mathfrak{e}_i}(z_0)$ with attached tensor $\mathcal{P}_{\mathfrak{e}_i}$ of rank $\#_m \mathfrak{e}_i$ and dimension n is transformed in:

$$R_{(r_{\mathfrak{H}} - \#_m \mathfrak{e}_i)\mathfrak{e}_i}\left(z_{r_{\mathfrak{H}}-1}\right) = R_{\mathfrak{e}_i}(z_0)\, y_{\#_m \mathfrak{e}_i}^1 \dots y_{r_{\mathfrak{H}}-1}^1$$

with attached tensor $\mathcal{R}_{(r_{\mathfrak{H}} - \#_m \mathfrak{e}_i)\mathfrak{e}_i}$ of rank $r_{\mathfrak{H}}$ and dimension $n + r_{\mathfrak{H}} - 1$.

The CHR of the tensor $\mathcal{R}_{(r_{\mathfrak{H}} - \#_m \mathfrak{e}_i)\mathfrak{e}_i}$ is the hypermatrix $\boldsymbol{R}_{(r_{\mathfrak{H}} - \#_m \mathfrak{e}_i)\mathfrak{e}_i} = \left(r_{(r_{\mathfrak{H}} - \#_m \mathfrak{e}_i)\, i_1 \dots i_{r_{\mathfrak{H}}}} \right)$. The elements of $\boldsymbol{R}_{(r_{\mathfrak{H}} - \#_m \mathfrak{e}_i)\mathfrak{e}_i}$ have only two possible values, 0 and:

$$\rho_{\text{lay},(r_{\mathfrak{H}} - \#_m \mathfrak{e}_i)\mathfrak{e}_i} = \rho_i \frac{m_{ij_1}! \dots m_{ij_{k_i}}!}{r_{\mathfrak{H}}!}.$$

The indices of the non-zero elements of $\boldsymbol{R}_{(r_{\mathfrak{H}}-\#_m \mathfrak{e}_i)\mathfrak{e}_i}$ are obtained by permutation of the elements of the multiset:

$$\left\{ j_1^{m_{ij_1}}, \ldots, j_{k_i}^{m_{ij_{k_i}}}, [n+\#_m \mathfrak{e}_i]^1, \ldots, [n+r_{\mathfrak{H}}-1]^1 \right\}.$$

And P is transformed in a homogeneous polynomial:

$$\begin{aligned} R\left(z_{r_{\mathfrak{H}}-1}\right) &= \sum_{i\in[\![p]\!]} c_i R_{(r_{\mathfrak{H}}-\#_m \mathfrak{e}_i)\mathfrak{e}_i}\left(z_{r_{\mathfrak{H}}-1}\right) \\ &= \sum_{i\in[\![p]\!]} c_i \rho_i z_{j_1}^{m_{ij_1}} \ldots z_{j_{k_i}}^{m_{ij_{k_i}}} y_{\#_m \mathfrak{e}_i}^1 \ldots y_{r_{\mathfrak{H}}-1}^1 \end{aligned}$$

$\square$

representing the layered m-uniformized hb-graph $\mathfrak{H}_{\text{lay}}$ with attached hypermatrix:

$$\boldsymbol{R} = \sum_{i\in[\![p]\!]} c_{\mathfrak{e}_i} \boldsymbol{R}_{(r_{\mathfrak{H}}-\#_m \mathfrak{e}_i)\mathfrak{e}_i},$$

where:

$$c_{\mathfrak{e}_i} = \frac{r_{\mathfrak{H}}}{\rho_i} = \frac{r_{\mathfrak{H}}}{\#_m \mathfrak{e}_i}.$$

Definition 9. *The* ***layered e-adjacency tensor*** $\mathcal{A}_{lay,\mathfrak{H}}$ *of a hb-graph* $\mathfrak{H} = (V, \mathfrak{E})$ *is the tensor of CHR* $\boldsymbol{A}_{lay,\mathfrak{H}} \triangleq \left(a_{i_1 \ldots i_{r_{\mathfrak{H}}}}\right)_{i_1,\ldots,i_{r_{\mathfrak{H}}} \in [\![n]\!]}$ *defined by:*

$$\boldsymbol{A}_{lay,\mathfrak{H}} \triangleq \sum_{i\in[\![p]\!]} c_{\mathfrak{e}_i} \boldsymbol{R}_{(r_{\mathfrak{H}}-\#_m \mathfrak{e}_i)\mathfrak{e}_i}$$

where for $\mathfrak{e}_i = \left\{ v_{j_1}^{m_{ij_1}}, \ldots, v_{j_{k_i}}^{m_{ij_{k_i}}} \right\} \in \mathfrak{E}$, $c_{\mathfrak{e}_i} = \frac{r_{\mathfrak{H}}}{\#_m \mathfrak{e}_i}$ *is the dilatation coefficient and* $\boldsymbol{R}_{(r_{\mathfrak{H}}-\#_m \mathfrak{e}_i)\mathfrak{e}_i} = \left(r_{(r_{\mathfrak{H}}-\#_m \mathfrak{e}_i)i_1 \ldots i_{r_{\mathfrak{H}}}}\right)$ *is the hypermatrix whose elements have only two possible values 0, and:*

$$\rho_{lay,(r_{\mathfrak{H}}-\#_m \mathfrak{e}_i)\mathfrak{e}_i} = \frac{m_{ij_1}! \ldots m_{ij_{k_i}}!}{r_{\mathfrak{H}}!} \#_m \mathfrak{e}_i.$$

The indices of the non-zero elements of $\boldsymbol{R}_{(r_{\mathfrak{H}}-\#_m \mathfrak{e}_i)\mathfrak{e}_i}$ *are obtained by permutation of the elements of the multiset:*

$$\left\{ j_1^{m_{ij_1}}, \ldots, j_{k_i}^{m_{ij_{k_i}}}, [n+\#_m \mathfrak{e}_i]^1, \ldots, [n+r_{\mathfrak{H}}-1]^1 \right\}.$$

Remark 4.4. $\mathcal{A}_{lay,\mathfrak{H}}$ *can also be written:*

$$\mathcal{A}_{lay,\mathfrak{H}} = \sum_{r\in[\![r_{\mathfrak{H}}]\!]} c_r \sum_{\mathfrak{e}_i \in \{\mathfrak{e}:\#_m \mathfrak{e} = r\}} \boldsymbol{R}_{\mathfrak{e}_i},$$

where $c_r = \frac{r_{\mathfrak{H}}}{r}$.

Remark 4.5. *In practice, writing* $A_{lay,\mathfrak{H}} = \left(a_{l_1 \cdots l_{r_{\mathfrak{H}}}}\right)$, *the element of* $A_{lay,\mathfrak{H}}$ *of indices* $l_1, \cdots, l_{r_{\mathfrak{H}}}$ *such that:*

$$\{\{l_1, \cdots, l_{r_{\mathfrak{H}}}\}\} = \left\{j_1^{m_{ij_1}}, \ldots, j_{k_i}^{m_{ij_{k_i}}}, [n + \#_m \mathfrak{e}_i]^1, \ldots, [n + r_{\mathfrak{H}} - 1]^1\right\},$$

corresponding to a hb-edge $\mathfrak{e}_i = \left\{v_{j_1}^{m_{ij_1}}, \ldots, v_{j_{k_i}}^{m_{ij_{k_i}}}\right\}$ *of the original hb-graph* $\mathfrak{H}$, *is:*

$$a_{l_1 \ldots l_{r_{\mathfrak{H}}}} = \frac{m_{ij_1}! \ldots m_{ij_{k_i}}!}{(r_{\mathfrak{H}} - 1)!}$$

5 Results on the constructed tensors

We remind that each of the tensors obtained is of rank $r_{\mathfrak{H}}$ and of dimension $n + n_{\mathcal{A}}$ where $n_{\mathcal{A}}$ is:

- in the straightforward approach: $n_{\mathcal{A}} = 1$.
- in the silo approach and the layered approach: $n_{\mathcal{A}} = r_{\mathfrak{H}} - 1$.

5.1 Fulfillment of the expectations

We revisit the expectations and prove they are all met.

Expectation 5.1. *The e-adjacency tensor should be nonnegative, symmetric and its generation should be as simple as possible.*

Proof. The tensors that have been built are all nonnegative and symmetric. Their generation mostly depends on the content of the hb-edges. Only the layered approach has a less simple generation process: it is nonetheless retained as it is the only possible approach that keeps hypergraph interpretability via Hm-UP without needing hb-graphs for general hypergraphs. □

Expectation 5.2. *The tensor should be globally invariant to vertex permutation in the original hb-graph.*

Proof. Let $\mathfrak{H} = (V, \mathfrak{E})$ be a hb-graph with vertex set $V = \{v_i : i \in [\![n]\!]\}$ and $\mathfrak{E} = (\mathfrak{e}_j)_{j \in [\![p]\!]}$. We do the proof only for the straightforward tensor, since the other proofs are similar.

Let $\pi \in \mathcal{S}_n$ be a permutation, that corresponds to a relabeling of the vertices from the universe of the original hb-graph. Applying this relabeling to the content of the hb-edge $\mathfrak{e}_j = \left\{ v_{i_1}^{m_{j\,i_1}}, \ldots, v_{i_j}^{m_{j\,i_j}} \right\}$ transforms it into $\mathfrak{e}_j = \left\{ v_{\pi(i_1)}^{m_{j\,i_1}}, \ldots, v_{\pi(i_j)}^{m_{j\,i_j}} \right\}$. The original hb-edge was stored in the elements of the form: $a_{i_1^{m_{j\,i_1}} \ldots i_j^{m_{j\,i_j}} (n+1)^{r_{\mathfrak{H}} - \#_m \mathfrak{e}_j}}$. It follows that the relabeled hb-graph will have its elements stored in $a_{\pi(i_1)^{m_{j\,i_1}} \ldots \pi(i_j)^{m_{j\,i_j}} (n+1)^{r_{\mathfrak{H}} - \#_m \mathfrak{e}_j}}$. It corresponds to a permutation of the elements in the tensor corresponding to the relabeling: the two tensors differ only by a reshuffling of the indices corresponding to π, which means that the tensor constructed is globally invariant to a relabeling in the original hb-graph. □

Expectation 5.3. *The e-adjacency tensor should allow the unique reconstruction of the hb-graph it is originated from.*

Proof. In the way the elements have been constructed, there is a one-to-one mapping between the hb-edge and the ordered indices of the coefficient in the tensor. From these indices, the hb-graph can be reconstructed with no ambiguity. □

Expectation 5.4. *Given the choice of two representations, the one that can be described with the least elements possible should be chosen. Then the sparsest e-adjacency tensor should be chosen.*

Proof. The straightforward tensor requires only one additional element to capture the information. The three representations are anyway economic as only one element each time needs to be described in the hypermatrix for a given hb-edge, the others being obtained by permutation of its indices. □

Expectation 5.5. *The e-adjacency tensor should allow direct retrieval of the vertex degrees.*

Proof. The proof will be given in the next subsection. □

5.2 Information on hb-graph

5.2.1 m-degree of vertices

We built the different tensors so that the retrieval of the vertex m-degree is possible; the null vertex(-ices) added during the Hm-UP give(s) additional information on the structure of the hb-graph.

Claim 5.1. *Let us consider for $j \in [\![n]\!]$ a vertex $v_j \in V$.*

Then for each of the e-adjacency tensors built, it holds:

$$\sum_{j_2,\dots,j_{r_{\mathfrak{H}}} \in [\![n+n_{\mathcal{A}}]\!]} a_{jj_2\dots j_{r_{\mathfrak{H}}}} = \sum_{i:v_j \in \mathfrak{e}_i} m_{ij} = deg_m(v_j)$$

Proof. For $j \in [\![n]\!]$:

$\sum_{j_2,\dots,j_{r_{\mathfrak{H}}} \in [\![n+n_{\mathcal{A}}]\!]} a_{jj_2\dots j_{r_{\mathfrak{H}}}}$ has non-zero terms only for the corresponding hb-edges of the original hb-graph $\mathfrak{e}_i$ containing v_j. Such a hb-edge is described by $\mathfrak{e}_i = \left\{v_j^{m_{ij}}, v_{l_2}^{m_{il_2}}, \dots, v_{l_k}^{m_{il_k}}\right\}$. This means that the multiset $\{\{j_2, \dots, j_{r_{\mathfrak{H}}}\}\}$ corresponds exactly to the multiset $\left\{j^{m_{ij}-1}, l_2^{m_{il_2}}, \dots, l_k^{m_{il_k}}\right\}$.

In the straightforward approach, for each $\mathfrak{e}_i$ such that $v_j \in \mathfrak{e}_i$, there are:

$$\frac{(r_{\mathfrak{H}}-1)!}{(m_{ij}-1)!m_{il_2}!\dots m_{il_k}!m_{in+1}!}$$

possible permutations of the indices j_2 to $j_{r_{\mathfrak{H}}}$ and

$$a_{jj_2\dots j_{r_{\mathfrak{H}}}} = \frac{m_{ij}!m_{il_2}!\dots m_{il_k}!m_{in+1}!}{(r_{\mathfrak{H}}-1)!}.$$

In the silo approach, for each $\mathfrak{e}_i$ such that $v_j \in \mathfrak{e}_i$, there are

$$\frac{(r_{\mathfrak{H}}-1)!}{(m_{ij}-1)!m_{il_2}!\dots m_{il_k}!m_{in+\#_m\mathfrak{e}_i}!}$$

possible permutations of the indices j_2 to $j_{r_{\mathfrak{H}}}$ and

$$a_{jj_2\dots j_{r_{\mathfrak{H}}}} = \frac{m_{ij}!m_{il_2}!\dots m_{il_k}!m_{in+\#_m\mathfrak{e}_i}!}{(r_{\mathfrak{H}}-1)!}.$$

In the layered approach, for each $\mathfrak{e}_i$ such that $v_j \in \mathfrak{e}_i$, there are:

$$\frac{(r_{\mathfrak{H}}-1)!}{(m_{ij}-1)!m_{il_2}!\dots m_{il_k}!}$$

ossible permutations of the indices j_2 to $j_{r_{\mathfrak{H}}}$ which have all the same value qual to:

$$a_{jj_2\ldots j_{r_{\mathfrak{H}}}} = \frac{m_{ij}!m_{il_2}!\ldots m_{il_k}!}{(r_{\mathfrak{H}}-1)!}.$$

Also, whatever the approach taken:

$$\sum_{j_2,\ldots,j_{r_{\mathfrak{H}}}\in[\![n]\!]} a_{jj_2\ldots j_{r_{\mathfrak{H}}}} = \sum_{i:v_j\in\mathfrak{e}_i} m_{ij} = \deg_m(v_j).$$

□

5.2.2 Additional vertex information

The additional vertices carry information on the hb-graph hb-edges that lepends on the approach taken.

Claim 5.2. *The layered e-adjacency tensor allows the retrieval of the distribution of the hb-edges.*

Proof. For $j \in [\![n_A]\!]$:

$\sum\limits_{j_2,\ldots,j_{r_{\mathfrak{H}}}\in[\![n+n_A]\!]} a_{n+jj_2\ldots j_{r_{\mathfrak{H}}}}$ has non-zero terms only for the corresponding hb-edges of the m-uniformized hb-graph $\overline{\mathfrak{e}_i}$ containing v_j. Such a hb-edge s described by[10]:

$$\overline{\mathfrak{e}_i} = \left\{v_k^{m_{ik}} : 1 \leqslant k \leqslant n+n_A\right\}.$$

This means that the multiset:

$$\{\{j_2,\ldots,j_{r_{\mathfrak{H}}}\}\}$$

corresponds exactly to the multiset:

$$\left\{(n+j)^{m_{in+j}-1}\right\} + \left\{k^{m_{ik}} : 1 \leqslant k \leqslant n+n_A, k \neq j\right\}.$$

The number of possible permutations of elements in this multiset is:

$$\frac{(r_{\mathfrak{H}}-1)!}{(m_{in+j}-1)!\prod\limits_{k\in[\![n]\!]} m_{ik}! \prod\limits_{\substack{k\in[\![n+1;n+n_A]\!]\\ k\neq j}} m_{ik}!}$$

[10]With the convention, that for $j \in [\![n_A]\!]$: $v_{n+j} = y_j$

and the elements corresponding to one hb-edge are all equal to:

$$\frac{\prod\limits_{k\in[\![n_{\mathcal{A}}]\!]} m_{i\,k}!}{(r_{\mathfrak{H}}-1)!}.$$

Thus:

$$\sum_{j_2,\ldots,j_{r_{\mathfrak{H}}}\in[\![n+n_{\mathcal{A}}]\!]} a_{n+j\,j_2\ldots j_{r_{\mathfrak{H}}}} = \sum_{j_2,\ldots,j_{r_{\mathfrak{H}}}\in[\![n]\!]} m_{i\,n+j} = \deg_m(Y_j).$$

The interpretation differs between the different approaches.

For the silo approach: There is one added vertex in each hb-edge. The silo of hb-edges of m-cardinality m_s ($m_s \in [\![r_{\mathfrak{H}}-1]\!]$) is associated to the null vertex Y_{m_s}. The multiplicity of Y_{m_s} in each hb-edge of the silo is $r_{\mathfrak{H}}-m_s$.

Hence:

$$\frac{\deg_m(Y_j)}{r_{\mathfrak{H}}-m_s} = |\{\mathfrak{e} : \#_m\mathfrak{e} = m_s\}|.$$

The number of hb-edges in the silo m_s is then deduced by the following formula:

$$|\{\mathfrak{e} : \#_m\mathfrak{e} = m_s\}| = |\mathfrak{E}| - \sum_{m_s\in[\![r_{\mathfrak{H}}-1]\!]} \frac{\deg_m(Y_j)}{r_{\mathfrak{H}}-m_s}.$$

For the layered approach: The vertex Y_j corresponds to the layer of level j added to each hb-edge with m-cardinality less or equal to j with a multiplicity of 1.

Also:

$$\deg_m(Y_j) = |\{\mathfrak{e} : \#_m\mathfrak{e} \leqslant j\}|.$$

Hence, for $j \in [\![2; r_{\mathfrak{H}}-1]\!]$:

$$|\{\mathfrak{e} : \#_m\mathfrak{e} = j\}| = \deg_m(Y_j) - \deg_m(Y_{j-1})$$

and:

$$|\{\mathfrak{e} : \#_m\mathfrak{e} = 1\}| = \deg_m(Y_1)$$

$$|\{\mathfrak{e} : \#_m\mathfrak{e} = r_{\mathfrak{H}}\}| = |\mathfrak{E}| - \deg_m(Y_{r_{\mathfrak{H}}-1})$$

For the straightforward approach: In a hb-edge of m-cardinality $j \in [\![r_{\mathfrak{H}} - 1]\!]$, the vertex Y_1 is added with multiplicity $r_{\mathfrak{H}} - j$. The number of hb-edges with m-cardinality j can be retrieved by considering the elements of $A_{\text{str},\mathfrak{H}}$ of index $(n+1)i_1 \dots i_{r_{\mathfrak{H}}-1}$ where $1 \leqslant i_1 \leqslant \dots \leqslant i_j \leqslant n$ and $i_{j+1} = \dots = i_{r_{\mathfrak{H}}-1} = n+1$ and the elements with indices obtained by permutation.

It follows for $j \in [\![r_{\mathfrak{H}} - 1]\!]$:

$$\begin{aligned} |\{\mathfrak{e} : \#_m \mathfrak{e} = j\}| &= |\{\mathfrak{e} \colon Y_1 \in \mathfrak{e} \wedge m_{\mathfrak{e}}(Y_1) = r_{\mathfrak{H}} - j\}| \\ &= \sum_{\substack{i_1, \dots, i_{r_{\mathfrak{H}}-1} \in [\![n+1]\!] \\ |\{i_k = n+1\}| = r_{\mathfrak{H}} - j - 1}} a_{n+1 i_1 \dots i_{r_{\mathfrak{H}}-1}} \end{aligned}$$

The terms $a_{n+1 i_1 \dots i_{r_{\mathfrak{H}}-1}}$ of this sum are non-zero only for the corresponding hb-edges $\overline{\mathfrak{e}}$ of the m-uniformized hb-graph having Y_1 with multiplicity $r_{\mathfrak{H}} - j$ in it. Such a hb-edge is described by:

$$\overline{\mathfrak{e}_i} = \{v_k^{m_{i\,k}} : 1 \leqslant k \leqslant n\} + \left\{Y_1^{r_{\mathfrak{H}} - j}\right\}.$$

It means that the multiset:

$$\{\{i_1, \dots, i_{r_{\mathfrak{H}}-1}\}\}$$

corresponds exactly to the multiset:

$$\{k^{m_{i\,k}} : k \in [\![n]\!]\} + \{n + 1^{r_{\mathfrak{H}} - j - 1}\}.$$

The number of possible permutations in this multiset is:

$$\frac{(r_{\mathfrak{H}} - 1)!}{\prod\limits_{k \in [\![n]\!]} m_{i\,k}! \, (r_{\mathfrak{H}} - j - 1)!}$$

and the elements corresponding to one hb-edge are all equal to:

$$\frac{\prod\limits_{k \in [\![n]\!]} m_{i\,k}! \times (r_{\mathfrak{H}} - j)!}{(r_{\mathfrak{H}} - 1)!}.$$

Hence:

$$\frac{1}{r_{\mathfrak{H}} - j} \sum_{\substack{i_2, \dots, i_{r_{\mathfrak{H}}} \in [\![n+1]\!] \\ |\{i_k = n+1 : 2 \leqslant k \leqslant r_{\mathfrak{H}}\}| = r_{\mathfrak{H}} - j - 1}} a_{n+1 i_2 \dots i_{r_{\mathfrak{H}}}} = |\{\mathfrak{e} : \#_m \mathfrak{e} = j\}$$

The number of hb-edges of m-cardinality $r_{\mathfrak{H}}$ can be retrieved by:

$$|\{\mathfrak{e} : \#_m \mathfrak{e} = r_{\mathfrak{H}}\}| = |\mathfrak{E}| - \sum_{j \in \llbracket r_{\mathfrak{H}} - 1 \rrbracket} |\{\mathfrak{e} : \#_m \mathfrak{e} = j\}| .$$

□

5.3 Some first results on spectral analysis

Let $\mathfrak{H} = (V, \mathfrak{E})$ be a general hb-graph of e-adjacency tensor $\mathcal{A}_{\mathfrak{H}}$ of CHR $\boldsymbol{A}_{\mathfrak{H}} = \left(a_{i_1 \dots i_{k_{\max}}}\right)$ of order $k_{\max}$ and dimension $n + n_{\mathcal{A}}$.

We write $d_{m,i} = \deg_m(v_i)$ if $i \in \llbracket n \rrbracket$ and $d_{m,n+i} = \deg_m(N_i)$ if $i = n + j,\ j \in \llbracket n_{\mathcal{A}} \rrbracket$.

In the e-adjacency hypermatrix $\boldsymbol{A}_{\mathfrak{H}}$, the diagonal entries are no longer equal to zero. As all elements of $\boldsymbol{A}_{\mathfrak{H}}$ are non-negative real numbers and as we have shown in the previous subsection that:

$$\sum_{i_2, \dots, i_m \in \llbracket n + n_{\mathcal{A}} \rrbracket} a_{i i_2 \dots i_{r_{\mathfrak{H}}}} = \begin{cases} d_{m,i} & \text{if } i \in \llbracket n \rrbracket \\ d_{m,n+j} & \text{if } i = n + j,\ j \in \llbracket n_{\mathcal{A}} \rrbracket . \end{cases}$$

It follows:

Claim 5.3. *The e-adjacency tensor $\mathcal{A}_{\mathfrak{H}}$ of a general hb-graph $\mathfrak{H} = (V, \mathfrak{E})$ has eigenvalues λ of its CHR $\boldsymbol{A}_{\mathfrak{H}}$ such that:*

$$|\lambda| \leqslant \max\left(\Delta_m, \Delta_m^{\star}\right) \tag{1}$$

where $\Delta_m = \max\limits_{i \in \llbracket n \rrbracket} \left(d_{m,i}\right)$ *and* $\Delta_m^{\star} = \max\limits_{i \in \llbracket n_{\mathcal{A}} \rrbracket} \left(d_{m,n+i}\right)$

Proof. From:

$$\forall i \in \llbracket 1, n \rrbracket ,\ \left(\mathcal{A} x^{m-1}\right)_i = \lambda x_i^{m-1}, \tag{2}$$

since $a_{i i_2 \dots i_m}$ are non-negative real numbers, it holds for all λ that:

$$|\lambda - a_{i \dots i}| \leqslant \sum_{\substack{i_2, \dots, i_m \in \llbracket n + n_{\mathcal{A}} \rrbracket \\ \delta_{i i_2 \dots i_m = 0}}} a_{i i_2 \dots i_m} \tag{3}$$

Considering the triangular inequality:

$$|\lambda| \leqslant |\lambda - a_{i \dots i}| + |a_{i \dots i}| \tag{4}$$

Combining (3) and (4) yield:

$$|\lambda| \leqslant \sum_{\substack{i_2,\ldots,i_m \in [\![n+n_{\mathcal{A}}]\!] \\ \delta_{ii_2\ldots i_m}=0}} a_{ii_2\ldots i_m} + |a_{i\ldots i}| \,. \quad (5)$$

But, irrespective of the approach taken, if $\{i^{r_{\mathfrak{H}}}\}$ is an hb-edge of the hb-graph then:

$$|a_{i\ldots i}| = r_{\mathfrak{H}}$$

otherwise:

$$|a_{i\ldots i}| = 0$$

and thus writing $\Delta_m = \max_{i \in [\![n]\!]} (\deg_m(v_i))$ and $\Delta_m^\star = \max_{i \in [\![n_{\mathcal{A}}]\!]} (\deg_m(N_i))$ and using (5) and $d_{m,i} = \sum_{\substack{i_2,\ldots,i_m \in [\![n+n_{\mathcal{A}}]\!] \\ \delta_{ii_2\ldots i_m}=0}} a_{ii_2\ldots i_m} + a_{i\ldots i}$ yield:

$$|\lambda| \leqslant \max(\Delta_m, \Delta_m^\star) \,.$$

□

Remark 5.1. *In the straightforward approach:*

$$\begin{aligned} \Delta_m^\star &= \deg_m(N_1) \\ &= \sum_{j \in [\![r_{\mathfrak{H}}-1]\!]} (r_{\mathfrak{H}} - j) \left|\{\mathfrak{e} : \#_m\mathfrak{e} = j\}\right| \end{aligned}$$

In the silo approach:

$$\begin{aligned} \Delta_m^\star &= \max_{j \in [\![r_{\mathfrak{H}}-1]\!]} (\deg_m(N_j)) \\ &= \max_{j \in [\![r_{\mathfrak{H}}-1]\!]} \left((r_{\mathfrak{H}} - j) \left|\{\mathfrak{e} : \#_m\mathfrak{e} = j\}\right|\right) \end{aligned}$$

In the layered approach:

$$\begin{aligned} \Delta_m^\star &= \max_{j \in [\![r_{\mathfrak{H}}-1]\!]} (\deg_m(N_j)) \\ &= \max_{j \in [\![r_{\mathfrak{H}}-1]\!]} \left(\left|\{\mathfrak{e} : \#_m\mathfrak{e} \leqslant j\}\right|\right) \\ &= \left|\{\mathfrak{e} : \#_m\mathfrak{e} \leqslant r_{\mathfrak{H}} - 1\}\right| \end{aligned}$$

The values of Δ_m are independent of the approach taken.

5.4 Categorisation of the constructed tensors

5.4.1 On classification of hypermatrices

Most of the definitions and results of this subsubsection are taken directly from Qi and Luo [47], which is the first book on tensor spectral analysis.

We consider a tensor $\mathcal{A}$ and its CHR of order m and dimension n: $\boldsymbol{A} = (a_{i_1 \dots i_m})_{i_1,\dots,i_m \in [\![n]\!]}$. The set of such hypermatrices is written $T_{m,n}$. The subset of $T_{m,n}$ of hypermatrices with only nonnegative coefficients is written $N_{m,n}$.

A hypermatrix $\boldsymbol{A} = (a_{i_1 \dots i_m})_{i_1,\dots,i_m \in [\![n]\!]} \in T_{m,n}$ is said **reducible** if it exists a nonempty proper subset $J \subsetneq [\![n]\!]$ such that: $\forall i_1 \in J, \forall i_2, \dots, i_m \in [\![n]\!] \setminus J : a_{i_1 \dots i_m} = 0$. A hypermatrix that is not reducible is said **irreducible**. The notion of reducibility has to be seen as a possible way of reducing the dimensionality of the problem.

Irreducible nonnegative hypermatrices have plenty of nice properties, particularly the Perron-Frobenius theorem for irreducible nonnegative hypermatrices which is one of the two declinations of the extension of the Perron-Frobenius theorem for irreducible nonnegative matrices.

The Perron-Frobenius theorem for irreducible nonnegative matrices states that the eigenvector associated to the spectral radius of a nonnegative matrix is non negative, and moreover, if this matrix is irreducible then this eigenvector is positive and its eigenvalue is the unique one associated with a nonnegative eigenvector.

The Perron-Frobenius theorem has a lot of applications in probability with stochastic matrices and is the basis for algorithms such as PageRank, ensuring that the convergence is feasible—Pillai et al. [46].

But hypermatrices manipulated in hypergraph theory are reducible. In this case, this first extension of the theorem of Perron-Frobenius cannot be used.

Weak irreducibility of non negative hypermatrices has been introduced to help to solve this problem. To define weak irreducibility, an **associated graph** $\mathcal{G}(\mathcal{A})$ is built out of the nonnegative hypercubic CHR $\boldsymbol{A}$ which represents $\mathcal{A}$ by considering as vertex set $[\![n]\!]$ and building the edges by considering an edge from i to j if it exists $a_{ii_2 \dots i_m} \neq 0$ such that $j \in \{i_2, \dots, i_m\}$.

A directed graph is **strongly connected** if for any ordered pair of vertices (i, j) of the graph, it exists a directed path from i to j.

A tensor $\mathcal{A}$ is said **weakly irreducible** if its associated graph is strongly onnected. A tensor that is not weakly irreducible is said **weakly reducible**.

We now consider a tensor $\mathcal{A}$ of nonnegative CHR $A = (a_{i_1 \dots i_m}) \in N_{m,n}$.

The **representative vector of** $\mathcal{A}$ is the vector $u(\mathcal{A})$ of coordinates $u_i(\mathcal{A}) = \sum\limits_{i_2,\dots,i_m \in [\![n]\!]} a_{ii_2 \dots i_n}$. $\mathcal{A}$ is called a **strictly nonnegative tensor** if its representative vector is positive.

Let now J be a proper nonempty subset of $[\![n]\!]$. The tensor $\mathcal{A}_J$ of CHR $A_J = (\alpha_{i_1 \dots i_m}) \in N_{m,|J|}$ such that $\alpha_{i_1 \dots i_m} = a_{i_1 \dots i_m}$ if $i_1, \dots, i_m \in J$ is called the **principal subtensor of** $\mathcal{A}$ **associated to** J.

$\mathcal{A}$ is said a **nontrivially nonnegative tensor** if it exists a principal subtensor of $\mathcal{A}$ that is a strictly nonnegative tensor.

The following proposition will be helpfull in the spectral analysis of hb-graphs.

Proposition 9. *A nonnegative tensor has a positive eigenvalue if and only if it is nontrivially nonnegative.*

As a consequence, a nontrivially nonnegative tensor has its spectral radius positive.

Qi and Luo [47] give a procedure to check easily if a nonnegative tensor is nontrivially negative or not.

5.4.2 Classification of the tensors built

The hypermatrices constructed in the three approaches are symmetric and nonnegative. This ensures that these hypermatrices have their spectral radius $\rho(\mathcal{A})$ which is an H^+-eigenvalue of $\mathcal{A}$, which means that $\rho(\mathcal{A})$ has a nonnegative eigenvector—all the components of the vector are nonnegative—associated to it.

A uniform hypergraph tensor is shown in Pearson and Zhang [42] to be weakly irreducible if and only if the hypergraph is connected. In the e-adjacency tensors this result does not hold.

Nonetheless, we can claim the following results:

Claim 5.4. *Let $\mathfrak{H} = (V, \mathfrak{E})$ be a non-m-uniform hb-graph with:* $\bigcup\limits_{\mathfrak{e} \in \mathfrak{E}} \mathfrak{e}^\star = V$.

If $\mathfrak{H}$ is connected then its straightforward e-adjacency tensor is weakly irreducible.

Proof. This proof combines arguments of Pearson and Zhang [42], with arguments of Qi and Luo [47] on weak irreducibility of the adjacency tensor of a uniform hypergraph, and some specific arguments related to the e-adjacency tensors we use for hb-graphs.

Let $\mathfrak{H} = (V, \mathfrak{E})$ be a non-m-uniform hb-graph with: $\bigcup_{e \in \mathfrak{E}} e^{\star} = V$.

Suppose that the straightforward e-adjacency tensor $\mathcal{A}_{\text{str},\mathfrak{H}}$ of CHR $\mathcal{A}_{\text{str},\mathfrak{H}} \in T_{r_{\mathfrak{H}},n+1}$ is weakly reducible, it means that the associated graph $\mathcal{G}\left(\mathcal{A}_{\text{str},\mathfrak{H}}\right)$ is not strongly connected and hence its matrix representation $A_{\mathcal{G}\left(\mathcal{A}_{\text{str},\mathfrak{H}}\right)} = (\alpha_{ij})$ is reducible.

As $\mathcal{A}_{\text{str},\mathfrak{H}}$ is symmetric, its associated graph $\mathcal{G}\left(\mathcal{A}_{\text{str},\mathfrak{H}}\right)$ is bidirectional and $A_{\mathcal{G}\left(\mathcal{A}_{\text{str},\mathfrak{H}}\right)} = (\alpha_{ij})$ is symmetric.

It means that it exists a nonempty proper subset J of $[\![n+1]\!]$ such that $\forall i \in J, \forall j \in [\![n+1]\!] \backslash J : \alpha_{ij} = 0$.

As $\mathfrak{H}$ is not m-uniform, J cannot be reduced to the singleton $\{n+1\}$ since the special vertex has to be linked to vertices of the hb-edges of m-cardinality lower than the maximal m-range.

For symmetric reasons, J cannot be $[\![n]\!]$, otherwise it would mean that the special vertex $n+1$ is isolated, which is not possible as the hb-graph is not m-uniform.

Thus it exists at least one $i_1 \in J$, which represents an original vertex of $\mathfrak{H}$ such that $a_{i_1 i_2 \dots i_{r_{\mathfrak{H}}}} = 0$ when at least one of the indices i_2, ..., $i_{r_{\mathfrak{H}}}$ is in $[\![n+1]\!] \backslash J$. In these $r_{\mathfrak{H}} - 1$ indices, at least one corresponds to an original vertex of $\mathfrak{H}$.

It indicates in this case that the group of original vertices of $\mathfrak{H}$ represented in J are disconnected from the original vertices that are in $[\![n+1]\!] \backslash J$. Hence the hb-graph is disconnected. □

By: $\bigcup_{e \in \mathfrak{E}} e^{\star} = V$, we require that there are no unused vertices in the universe.

If the hb-graph is m-uniform and connected, the straightforward e-adjacency tensor $\mathcal{A}_{\text{str},\mathfrak{H}}$ is weakly reducible as the additional vertex is not used and hence is isolated in the associated graph of $\mathcal{A}_{\text{str},\mathfrak{H}}$. In this case, one can use $\mathcal{A}_{\text{str},\mathfrak{H}}|_{[\![n]\!]}$ the principal sub-tensor of $\mathcal{A}_{\text{str},\mathfrak{H}}$ related to $[\![n]\!]$ which is weakly irreducible. As $\mathcal{A}_{\text{str},\mathfrak{H}}|_{[\![n]\!]}$ is weakly irreducible it is strictly nonnegative and the $\mathcal{A}_{\text{str},\mathfrak{H}}$ is a non-trivially nonnegative tensor, which means by using the theorem from Qi and Luo [47] that $\mathcal{A}_{\text{str},\mathfrak{H}}$ has a positive eigenvalue and hence $\rho\left(\mathcal{A}_{\text{str},\mathfrak{H}}\right) > 0$.

Nonnegative tensor weak irreducibility is a desirable property as it ensures that the tensor has a unique positive Perron vector (up to a multiplicative constant) associated to its spectral positive radius $\rho(\mathcal{A})$, which ensures the convergence of algorithms to find it.

To ensure weak irreducibility, we could transform the straightforward e-adjacency tensor so its associated graph is always strongly connected. It is sufficient to add the special vertex to each of the hyperedges: in this case it will force the associated graph to be connected, albeit the spectral radius upper-bound will be increased to the maximum between the maximal m-degree and the number of hb-edges.

Moreover:

Claim 5.5. *The three e-adjacency tensors built for hb-graphs are non-trivially nonnegative tensors when the hb-graph is connected and that the union of the support of hb-edges covers the vertex set.*

Proof. We already know that for a connected hb-graph which is not m-uniform, the straightforward e-adjacency tensor is weakly irreducible, hence non-trivially nonnegative. For a m-uniform hb-graph, the principal sub-tensor of $\mathcal{A}_{\text{str},\mathfrak{H}}$ composed of its n first indices is weakly irreducible, hence strictly non negative. The proof is similar to the one of Pearson and Zhang [42] and Qi and Luo [47] for hypergraphs, so we omit it.

We have already explained in Ouvrard et al. [38] how a hb-graph $\mathfrak{H} = (V, \mathfrak{E})$ can be decomposed in layers by considering $\mathfrak{E}_k = \{\mathfrak{e} \in \mathfrak{E}, \sharp_m \mathfrak{e} = k\}$ and the hb-graphs $\mathfrak{H}_k = (V, \mathfrak{E}_k)$ to obtain $\mathfrak{H} = \bigoplus_{k \in \llbracket r_{\mathfrak{H}} \rrbracket} \mathfrak{H}_k$.

For the silo e-adjacency tensor, the principal sub-tensor to be considered is the one obtained using $J = \llbracket n \rrbracket \cup \{n + k : 1 \leqslant k \leqslant r_{\mathfrak{H}} - 1 \wedge \mathfrak{E}_k \neq \emptyset\}$. $\mathcal{A}_{\text{sil},\mathfrak{H}}|_J$ is then weakly irreducible, hence strictly non negative. The tensor $\mathcal{A}_{\text{sil},\mathfrak{H}}$ is therefore a non-trivially nonnegative tensor.

For the layered e-adjacency tensor $\mathcal{A}_{\text{lay},\mathfrak{H}}$, we consider the principal sub-tensor $\mathcal{A}_{\text{lay},\mathfrak{H}}|_K$ where $K = \llbracket n \rrbracket \cup \{n + k : k \llbracket k_{\min}, r_{\mathfrak{H}} - 1 \rrbracket\}$ with $k_{\min} = \min\{k : \mathfrak{E}_k \neq \emptyset\}$. This principal sub-tensor is weakly irreducible as the hb-graph is connected, hence strictly non negative. Hence, the tensor $\mathcal{A}_{\text{lay},\mathfrak{H}}$ is non-trivially nonnegative. □

As a consequence the spectral radii of those tensors are positive.

5.4.3 A remark on the tensors built

The uniformisation process used to built the three tensors uses the fact that some additional vertices are added to the hyperedges; if it does not change the number of hb-edges and the degree of the vertices that are in the original hb-graph, it has an impact on the connectivity of the uniform hb-graph compared to the original one.

Nonetheless, to address this problem, it is always possible to consider each connected component of the original hb-graph separately and to build a tensor for each of this connected component. In this case we do not change the connectivity.

So we can always assume that we address only connected hb-graphs.

6 Evaluation and final choice

6.1 Evaluation

We have put together some key features of the e-adjacency tensors proposed in this article: the straightforward approach tensor $\mathcal{A}_{\text{str},\mathfrak{H}}$, the silo approach tensor $\mathcal{A}_{\text{sil},\mathfrak{H}}$ and $\mathcal{A}_{\text{lay},\mathfrak{H}}$ for the layered approach.

The CHR of the constructed tensors all have the same order $r_{\mathfrak{H}}$. $\boldsymbol{A}_{\text{sil},\mathfrak{H}}$ and $\boldsymbol{A}_{\text{lay},\mathfrak{H}}$ dimensions are $r_{\mathfrak{H}} - 2$ bigger than $\boldsymbol{A}_{\text{str},\mathfrak{H}}$ ($n-2$ in the worst case). $\boldsymbol{A}_{\text{str},\mathfrak{H}}$ has a total number of elements $\dfrac{(n+1)^{r_{\mathfrak{H}}}}{(n+r_{\mathfrak{H}}-1)^{r_{\mathfrak{H}}}}$ times smaller than the two other tensors.

Elements of $\boldsymbol{A}_{\text{str},\mathfrak{H}}$—respectively $\boldsymbol{A}_{\text{sil},\mathfrak{H}}$—are repeated $\dfrac{1}{n_j!}$—respectively $\dfrac{1}{n_{j\,k}!}$—times less than elements of $\boldsymbol{A}_{\text{lay},\mathfrak{H}}$. The total number of non-null elements filled for a given hb-graph in $\boldsymbol{A}_{\text{str},\mathfrak{H}}$ and $\boldsymbol{A}_{\text{sil},\mathfrak{H}}$ are the same and is smaller than the total number of non-null elements in $\boldsymbol{A}_{\text{lay},\mathfrak{H}}$.

Whatever the approach taken, the tensors are symmetric: a unique value is needed to have full description of an hb-edge; moreover, this value depends only on the hb-edge composition.

All tensors are symmetric and allow the reconstruction of the hb-graph from these elements.

Nodes degree can be retrieved as it has been shown previously. Additional information on hb-edges is easier to retrieve with the silo and the layered approach.

6.2 First choice

Insofar as the straightforward tensor is weakly irreducible for non m-uniform connected hb-graph, and as it is a sufficient desirable property to choose it, even if it is at the price of less practicability to retrieve information on hb-edges, we take $\mathcal{A}_{\text{str},\mathfrak{H}}$ for definition of the e-adjacency tensor of the hb-graph. The preservation of the information on the shape of the hb-edges through the added null vertex allows to retrieve information on the hb-edge cardinality.

6.3 Hypergraphs and hb-graphs

Hypergraphs are particular case of hb-graphs and hence the e-adjacency tensor defined for hb-graphs can be used for hypergraphs. Since the multiplicity function for vertices of a hyperedge seen as hb-edge has its values in $\{0, 1\}$, the e-adjacency tensor elements differ only by a factorial due to the cardinality of the hyperedge.

We retain as definition for the e-adjacency tensor of a hypergraph:

Definition 10. *The e-adjacency tensor of a hypergraph* $\mathcal{H} = (V, E)$ *having maximal cardinality of its hyperedges* $r_{\mathcal{H}} = k_{\max}$ *is the tensor* $\mathcal{A}_{\mathcal{H}}$ *of CHR* $\boldsymbol{A}_{\mathcal{H}} \triangleq \left(a_{i_1 \dots i_{r_{\mathcal{H}}}}\right)_{1 \leqslant i_1, \dots, i_{r_{\mathcal{H}}} \leqslant n}$ *defined by:*

$$\boldsymbol{A}_{\mathcal{H}} \triangleq \sum_{i \in \llbracket p \rrbracket} c_{e_i} \boldsymbol{R}_{e_i}$$

and where for $e_i = \left\{v_{j_1}, \dots, v_{j_{k_i}}\right\} \in E$, $c_{e_i} = \frac{k_{\max}}{k_i}$ *is the dilatation coefficient and* $\boldsymbol{R}_{e_i} = \left(r_{i_1 \dots i_{r_{\mathcal{H}}}}\right)$ *is the associated tensor to* e_i*, having all non-zero elements of same value. The non-zero elements of* $\boldsymbol{R}_{e_i}$ *are:*

$$r_{j_1 \dots j_{k_i} (n+1)^{k_{\max} - k_i}} = \frac{(k_{\max} - k_i)!}{k_{\max}!}$$

and all the ones whose indices are obtained by permutation of:

$$j_1 \dots j_{k_i} \, (n+1)^{k_{\max} - k_i}.$$

As in Ouvrard et al. [36], we compare the e-adjacency tensors obtained by Banerjee et al. [2] and by Sun et al. [51] with the one chosen in this article. The results are presented in Table 3.

	$\mathcal{A}_{str,\mathfrak{H}}$	$\mathcal{A}_{sil,\mathfrak{H}}$	$\mathcal{A}_{lay,\mathfrak{H}}$
Hypermatrix representation	$\mathbf{A}_{str,\mathfrak{H}}$	$\mathbf{A}_{sil,\mathfrak{H}}$	$\mathbf{A}_{lay,\mathfrak{H}}$
Order	$r_{\mathfrak{H}}$	$r_{\mathfrak{H}}$	$r_{\mathfrak{H}}$
Dimension	$n+1$	$n+r_{\mathfrak{H}}-1$	$n+r_{\mathfrak{H}}-1$
Total number of elements	$(n+1)^{r_{\mathfrak{H}}}$	$(n+r_{\mathfrak{H}}-1)^{r_{\mathfrak{H}}}$	$(n+r_{\mathfrak{H}}-1)^{r_{\mathfrak{H}}}$
Total number of elements potentially used by the way the tensor is build	$(n+1)^{r_{\mathfrak{H}}}$	$(n+r_{\mathfrak{H}}-1)^{r_{\mathfrak{H}}}$	$(n+r_{\mathfrak{H}}-1)^{r_{\mathfrak{H}}}$
Number of repeated elements per hb-edge $\mathfrak{e}_j = \left\{ v_{i_1}^{m_{ji_1}}, \dots, v_{i_j}^{m_{ji_j}} \right\}$	$\frac{r_{\mathfrak{H}}!}{m_{ji_1}!\dots m_{ji_j}!n_j!}$ with $n_j = r_{\mathfrak{H}} - \#_m\mathfrak{e}_j$	$\frac{r_{\mathfrak{H}}!}{m_{ji_1}!\dots m_{ji_j}!n_{jk}!}$ with $n_{jk} = r_{\mathfrak{H}} - \#_m\mathfrak{e}_j$	$\frac{r_{\mathfrak{H}}!}{m_{ji_1}!\dots m_{ji_j}!}$
Number of elements to be filled per hb-edge of size s before permutation	Constant 1	Constant 1	Constant 1
Number of elements to be described to derived the tensor by permutation of indices	$\|\mathfrak{E}\|$	$\|\mathfrak{E}\|$	$\|\mathfrak{E}\|$
Value of elements corresponding to a hb-edge	Dependent of hb-edge composition $\frac{m_{ji_1}!\dots m_{ji_j}!n_j!}{(r_{\mathfrak{H}}-1)!}$	Dependent of hb-edge composition $\frac{m_{ji_1}!\dots m_{ji_j}!n_{jk}!}{(r_{\mathfrak{H}}-1)!}$	Dependent of hb-edge composition $\frac{m_{ji_1}!\dots m_{ji_j}!}{(r_{\mathfrak{H}}-1)!}$
Symmetric	Yes	Yes	Yes
Reconstructivity	Straightforward: delete special vertices	Straightforward: delete special vertices	Straightforward: delete special vertices
Nodes degree	Yes	Yes	Yes
Information on hb-edges	Yes, but not straightforward	Yes	Yes
Spectral analysis	Special vertex increases the amplitude of the bounds	Special vertices increase the amplitude of the bounds	Special vertices increase the amplitude of the bounds
Interpretability of the tensor in term of hb-graph	Yes	Yes	Yes

Table 2: Evaluation of the hb-graph e-adjacency tensor depending on construction

$\mathcal{A}_{str,\mathfrak{H}}$ refers to the e-adjacency tensor built with the straightforward approach;

$\mathcal{A}_{sil,\mathfrak{H}}$ refers to the e-adjacency tensor built with the silo approach;

$\mathcal{A}_{lay,\mathfrak{H}}$ refers to the e-adjacency tensor built with the layered approach.

	$\mathcal{B}_{\mathcal{H}}$	$\mathcal{S}_{\mathcal{H}}$	$\mathcal{A}_{\mathcal{H}}$
Hypermatrix representation	$B_{\mathcal{H}}$	$S_{\mathcal{H}}$	$A_{\mathcal{H}}$
Order	$k_{\max}$	$k_{\max}$	$k_{\max}$
Dimension	n	n	$n + k_{\max} - 1$
Total number of elements	$n^{k_{\max}}$	$n^{k_{\max}}$	$(n + k_{\max} - 1)^{k_{\max}}$
Total number of elements potentially used by the way the tensor is build	$n^{k_{\max}}$	$n^{k_{\max}}$	$(n + k_{\max} - 1)^{k_{\max}}$
Number of non-zero elements for a given hypergraph	$\sum\limits_{s=1}^{k_{\max}} \alpha_s \lvert E_s \rvert$ with $\alpha_s = p_s(k_{\max}) \frac{k_{\max}!}{k_1!...k_s!}$	$\sum\limits_{s=1}^{k_{\max}} s! \lvert E_s \rvert$	$\sum\limits_{s=1}^{k_{\max}} \alpha_s \lvert E_s \rvert$ with $\alpha_s = \frac{k_{\max}!}{k_1!...k_s!n_s!}$ with $n_s = k_{\max} - s$
Number of repeated elements per hyperedge of size s	$\frac{k_{\max}!}{k_1!...k_s!}$	s	$\frac{k_{\max}!}{k_1!...k_s!n_s!}$ with $n_s = k_{\max} - s$
Number of elements to be filled per hyperedge of size s before permutation	Varying $p_s(k_{\max})$	Varying s if prefix is considered as non-permuting part	Constant 1
Number of elements to be described to derive the tensor by permutation of indices	$\sum\limits_{s=1}^{k_{\max}} p_s(k_{\max}) \lvert E_s \rvert$	$\sum\limits_{s=1}^{k_{\max}} s \lvert E_s \rvert$	$\lvert E \rvert$
Value of elements of a hyperedge	Dependent of hyperedge composition $\frac{s}{\alpha_s}$	Dependent of hyperedge composition $\frac{1}{(s-1)!}$	Dependent of hyperedge size $\frac{(k_{\max} - s)!}{s(k_{\max} - 1)!}$
Symmetric	Yes	No	Yes
Reconstructivity	Need computation of duplicated vertices	Need computation of duplicated vertices	Straightforward: delete special vertices
Nodes degree	Yes	Yes	Yes
Hyperedge cardinality	Not straightforward	Not straightforward	Yes
Spectral analysis	Yes	Yes	Special vertices increase the amplitude of the bounds
Interpretability of the tensor in term of hypergraph / hb-graph	No / No	No / No	No / Yes

Table 3: Evaluation of the hypergraph e-adjacency tensor $\mathcal{B}_{\mathcal{H}}$ designates the adjacency tensor defined in Banerjee et al. [2]. $\mathcal{S}_{\mathcal{H}}$ designates the adjacency tensor defined in Sun et al. [51]. $\mathcal{A}_{\mathcal{H}}$ refers to the e-adjacency tensor as defined in this article.

7 Conclusion

Extending the concept of hypergraphs to support multisets and introducing hb-graphs has allowed us to define a systematic approach to build the e-adjacency tensor of a hb-graph. In return, as hypergraphs appear as particular case of hb-graphs, the e-adjacency tensors are applicable to general hypergraphs. Hb-graphs are a good modeling framework for many real problems and already allow some nice refinements of existing work.

The tensor constructed in Banerjee et al. [2] appears as a transformation of the hypergraph $\mathcal{H} = (V, E)$ into a weighted hb-graph $\mathcal{H}_{\mathcal{B}} = (V, E', w_e)$: with the same vertex set but with hb-edges obtained from the hyperedges of the original hypergraph such that for a given hyperedge all the hb-edges having this hyperedge as support are considered with multiplicities of vertices such that it reaches $k_{\max}$.

We still have to focus on the analysis of the behavior of our constructed e-adjacency tensor regarding the diffusion process, particularly concerning the m-uniformation process. The fact that information on hb-edges for hb-graphs and therefore for hyperedge in hypergraphs are stored in the e-adjacency tensor should provide a nice explicitation on the role of the variety of m-cardinality of hb-edges.

References

[1] Albert, J. (1991). Algebraic properties of bag data types. In *VLDB*, volume 91, pages 211–219. Citeseer.

[2] Banerjee, A., Char, A., and Mondal, B. (2017). Spectra of general hypergraphs. *Linear Algebra and its Applications*, 518:14–30.

[3] Bellaachia, A. and Al-Dhelaan, M. (2013). Random walks in hypergraph. In *Proceedings of the 2013 International Conference on Applied Mathematics and Computational Methods, Venice Italy*, pages 187–194.

[4] Berge, C. and Minieka, E. (1973). *Graphs and hypergraphs*, volume 7. North-Holland publishing company Amsterdam.

[5] Blizard, W. D. et al. (1988). Multiset theory. *Notre Dame Journal of formal logic*, 30(1):36–66.

[6] Bretto, A. (2013). Hypergraph theory. *An introduction. Mathematical Engineering. Cham: Springer.*

[7] Chauve, C., Patterson, M., and Rajaraman, A. (2013). Hypergraph covering problems motivated by genome assembly questions. In *International Workshop on Combinatorial Algorithms*, pages 428–432. Springer.

[8] Chazelle, B. and Friedman, J. (1988). A deterministic view of random sampling and its use in geometry. In *Foundations of Computer Science, 1988., 29th Annual Symposium on*, pages 539–549. IEEE.

[9] Chitra, U. and Raphael, B. J. (2019). Random walks on hypergraphs with edge-dependent vertex weights. *arXiv preprint arXiv:1905.08287*.

[10] Cooper, J. and Dutle, A. (2012). Spectra of uniform hypergraphs. *Linear Algebra and its Applications*, 436(9):3268–3292.

[11] Corso, G. (2004). Families and clustering in a natural numbers network. *Physical Review E*, 69(3):036106.

[12] Dedekind, R. (1965). Was sind und was sollen die zahlen? In *Was sind und was sollen die Zahlen?. Stetigkeit und Irrationale Zahlen*, pages 1–47. Springer.

[13] Frahm, K. M., Chepelianskii, A., and Shepelyansky, D. (2012). Pagerank of integers. *Journal of Physics A: Mathematical and Theoretical*, 45(40):405101.

[14] Gao, X., Xiao, B., Tao, D., and Li, X. (2010). A survey of graph edit distance. *Pattern Analysis and applications*, 13(1):113–129.

[15] Girish, K. and Jacob, J. S. (2012). On multiset topologies. *Theory and Applications of Mathematics & Computer Science*, 2(1):37–52.

[16] Grossman, J. W. and Ion, P. D. (1995). On a portion of the well-known collaboration graph. *Congressus Numerantium*, pages 129–132.

[17] Grumbach, S., Libkin, L., Milo, T., and Wong, L. (1996). Query languages for bags: expressive power and complexity. *ACM SIGACT News*, 27(2):30–44.

[18] Harris, Z. S. (1954). Distributional structure. *Word*, 10(2-3):146–162.

[19] Hernich, A. and Kolaitis, P. G. (2017). Foundations of information integration under bag semantics. In *Logic in Computer Science (LICS), 2017 32nd Annual ACM/IEEE Symposium on*, pages 1–12. IEEE.

[20] Kaminski, B., Poulin, V., Pralat, P., Szufel, P., and Theberge, F. (2018). Clustering via hypergraph modularity. *arXiv preprint arXiv:1810.04816*.

[21] Karypis, G., Aggarwal, R., Kumar, V., and Shekhar, S. (1999). Multilevel hypergraph partitioning: applications in vlsi domain. *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 7(1):69–79.

[22] Klug, A. (1982). Equivalence of relational algebra and relational calculus query languages having aggregate functions. *Journal of the ACM (JACM)*, 29(3):699–717.

[23] Knuth, D. (1968). The art of computer programming 1: Fundamental algorithms 2: Seminumerical algorithms 3: Sorting and searching 4: Generating all trees. history of combinatorial generation. *MA: Addison-Wesley*, 30.

[24] Lamperti, G., Melchiori, M., and Zanella, M. (2000). On multisets in database systems. In *Workshop on Membrane Computing*, pages 147–215. Springer.

[25] Lee, J. M. (2013). Smooth manifolds. In *Introduction to Smooth Manifolds*, pages 1–31. Springer.

[26] Lin, J., Khade, R., and Li, Y. (2012). Rotation-invariant similarity in time series using bag-of-patterns representation. *Journal of Intelligent Information Systems*, 39(2):287–315.

[27] Lin, J. and Li, Y. (2009). Finding structural similarity in time series data using bag-of-patterns representation. In *International Conference on Scientific and Statistical Database Management*, pages 461–477. Springer.

[28] Loeb, D. (1992). Sets with a negative number of elements. *Advances in Mathematics*, 91(1):64–74.

[29] Majcher, Z. (1987). Alternating cycles and realizations of a degree sequence. *Commentationes Mathematicae Universitatis Carolinae*, 28(3):467–480.

[30] Mäkinen, E. (1990). How to draw a hypergraph. *International Journal of Computer Mathematics*, 34(3-4):177–185.

[31] Newman, M. E. (2001a). Scientific collaboration networks. i. network construction and fundamental results. *Physical review E*, 64(1):016131.

[32] Newman, M. E. (2001b). Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Physical review E*, 64(1):016132.

[33] Nowak, E., Jurie, F., and Triggs, B. (2006). Sampling strategies for bag-of-features image classification. In *European conference on computer vision*, pages 490–503. Springer.

[34] Ouvrard, X., Le Goff, J., and Marchand-Maillet, S. (2018a). Hypergraph modeling and visualisation of complex co-occurence networks. *Electronic Notes in Discrete Mathematics*, 70:65–70. TCDM 2018, University of Derby.

[35] Ouvrard, X., Le Goff, J., and Marchand-Maillet, S. (2018b). On adjacency and e-adjacency in general hypergraphs: Towards a new e-adjacency tensor. *Electronic Notes in Discrete Mathematics*, 70:71–76. TCDM 2018, University of Derby.

[36] Ouvrard, X., Le Goff, J.-M., and Marchand-Maillet, S. (2017a). Adjacency and tensor representation in general hypergraphs part 1: e-adjacency tensor uniformisation using homogeneous polynomials. *arXiv preprint arXiv:1712.08189*.

[37] Ouvrard, X., Le Goff, J.-M., and Marchand-Maillet, S. (2017b). Networks of collaborations: Hypergraph modeling and visualisation. *arXiv preprint arXiv:1707.00115*.

[38] Ouvrard, X., Le Goff, J.-M., and Marchand-Maillet, S. (2018c). Adjacency and tensor representation in general hypergraphs. part 2: Multisets, hb-graphs and related e-adjacency tensors. *arXiv preprint arXiv:1805.11952*.

[39] Ouvrard, X., Le Goff, J.-M., and Marchand-Maillet, S. (2018d). Diffusion by exchanges in hb-graphs: Highlighting complex relationships. *CBMI Proceedings*.

[40] Patwardhan, K. S., Naimpally, S. A., and Singh, S. L. (2001). Lilavati of bhaskaracarya. a treatise of mathematics of vedic tradition.

[41] Păun, G. (2006). Introduction to membrane computing. In *Applications of Membrane Computing*, pages 1–42. Springer.

[42] Pearson, K. J. and Zhang, T. (2014). On spectral hypergraph theory of the adjacency tensor. *Graphs and Combinatorics*, 30(5):1233–1248.

[43] Pearson, K. J. and Zhang, T. (2015). The laplacian tensor of a multi-hypergraph. *Discrete Mathematics*, 338(6):972–982.

[44] Pedronette, D. C. G., Valem, L. P., Almeida, J., and Torres, R. d. S. (2019). Multimedia retrieval through unsupervised hypergraph-based manifold ranking. *IEEE Transactions on Image Processing*.

[45] Peng, X., Wang, L., Wang, X., and Qiao, Y. (2016). Bag of visual words and fusion methods for action recognition: Comprehensive study and good practice. *Computer Vision and Image Understanding*, 150:109–125.

[46] Pillai, S. U., Suel, T., and Cha, S. (2005). The perron-frobenius theorem: some of its applications. *IEEE Signal Processing Magazine*, 22(2):62–75.

[47] Qi, L. and Luo, Z. (2017). *Tensor analysis: spectral theory and special tensors*, volume 151. SIAM.

[48] Radoaca, A. (2015). Simple venn diagrams for multisets. In *Symbolic and Numeric Algorithms for Scientific Computing (SYNASC), 2015 17th International Symposium on*, pages 181–184. IEEE.

[49] Singh, D., Ibrahim, A., Yohanna, T., and Singh, J. (2007). An overview of the applications of multisets. *Novi Sad Journal of Mathematics*, 37(3):73–92.

[50] Stell, J. (2012). Relations on hypergraphs. *Relational and Algebraic Methods in Computer Science*, pages 326–341.

[51] Sun, L., Zhou, J., and Bu, C. (2018). Spectral properties of general hypergraphs. *Linear Algebra and its Applications*.

[52] Syropoulos, A. (2000). Mathematics of multisets. In *Workshop on Membrane Computing*, pages 347–358. Springer.

[53] Taramasco, C., Cointet, J.-P., and Roth, C. (2010). Academic team formation as evolving hypergraphs. *Scientometrics*, 85(3):721–740.

[54] Tarau, P. (2014). Towards a generic view of primality through multiset decompositions of natural numbers. *Theoretical Computer Science*, 537:105–124.

[55] Temkin, O. N., Zeigarnik, A. V., and Bonchev, D. (1996). *Chemical reaction networks: a graph-theoretical approach*. CRC Press.

[56] Xiong, C., Callan, J., and Liu, T.-Y. (2016). Bag-of-entities representation for ranking. In *Proceedings of the 2016 ACM International Conference on the Theory of Information Retrieval*, pages 181–184. ACM.

[57] Zhou, T., Wang, B.-H., Hui, P., and Chan, K. (2006). Topological properties of integer networks. *Physica A: Statistical Mechanics and its Applications*, 367:613–618.