# On the Reconstruction of 3-Uniform Hypergraphs from Degree Sequences of Span-Two

Giulia Palma[2] · Andrea Frosini[1] · Simone Rinaldi[2]

## Abstract

A nonnegative integer sequence is $k$-graphic if it is the degree sequence of a $k$-uniform simple hypergraph. The problem of deciding whether a given sequence $\pi$ is 3-graphic has recently been proved to be NP-complete, after years of studies. Thus, it acquires primary relevance to detect classes of degree sequences whose graphicality can be tested in polynomial time in order to restrict the NP-hard core of the problem and design algorithms that can also be useful in different research areas. Several necessary and few sufficient conditions for $\pi$ to be $k$-graphic, with $k \geq 3$, appear in the literature. Frosini et al. defined a polynomial time algorithm to reconstruct $k$-uniform hypergraphs having regular or almost regular degree sequences. Our study fits in this research line providing a combinatorial characterization of span-two sequences, i.e., sequences of the form $\pi = (d, \ldots, d, d-1, \ldots, d-1, d-2, \ldots, d-2)$, $d \geq 2$, which are degree sequences of some 3-uniform hypergraphs. Then, we define a polynomial time algorithm to reconstruct one of the related 3-uniform hypergraphs. Our results are likely to be easily generalized to $k \geq 4$ and to other families of degree sequences having simple characterization, such as gap-free sequences.

**Keywords** 3-Uniform hypergraphs · Degree sequences · Reconstruction problem

## 1 Introduction

The *degree sequence* of a simple hypergraph is the list of its vertex degrees, usually arranged in non-increasing order. Given a nonnegative integer sequence $\pi$, the possibility of an efficient test for the existence of a simple hypergraph having degree sequence $\pi$ remained unsolved for many years (see [1,4]). The degree sequence of a simple $k$-uniform hypergraph, say $k$-hypergraph, is said $k$-graphic, briefly $k$-sequence. Note that very little can be said about the uniqueness of the hypergraphs related to a degree sequence.

In general, 2-sequences do not guarantee such uniqueness but in particular cases. Similarly, the same happens when the degree sequence is regular and almost regular.

The characterization of 2-sequences, i.e., the degree sequences of graphs, was solved by Erdős-Gallai in [7].

Relying on this result, several polynomial time algorithms have been defined to reconstruct the incidence matrix of one of the related graphs.

Assuming that $P \neq$ NP, if $k \geq 3$ an effective characterization does not exist even for the simplest case of 3-hypergraphs (see Deza et al. [6]).

However, necessary and sufficient conditions to detect the $k$-graphicality of an integer sequence $\pi$, with $k \geq 3$, can be found in the literature, and they mainly rely on a result by Dewdney [5], based on a recursive decomposition of $\pi$. This characterization does not yield to an efficient algorithm and the question to determine a more practical characterization remained open. Basing on these results, in [2] the authors presented some general sufficient and testable conditions for $k$-graphicality when $k \geq 3$. Brlek et al. in [4] defined a $P$-time reconstruction algorithm for the case of regular $k$-sequences. Later this result was extended to almost regular $k$-sequences (see [9]). A remarkable fact is that in

✉ Giulia Palma
  giulia.palma2@unisi.it

  Andrea Frosini
  andrea.frosini@unifi.it

  Simone Rinaldi
  rinaldi@unisi.it

1   Dipartimento di Matematica e Informatica, Università di Firenze, Viale Morgagni 65, 50134 Florence, Italy

2   Dipartimento di Ingegneria dell'Informazione e Scienze Matematiche, Università di Siena, Via Roma 56, 53100 Siena, Italy

both cases, all the $k$-sequences satisfy a simple necessary and sufficient condition. In our study, we investigate span-two 3-sequences, i.e., degree sequences where the difference between their maximal and minimal entry is two, and we prove that these conditions are not sufficient any more.

This result allows to further restrict the hard core of the reconstruction problem on 3-uniform hypergraphs, and it moves in the direction of inspecting the potential of reconstruction heuristics that base on the maximum span among the nodes' degrees. We recall that, similarly to what happens to graphs, techniques for random hypergraphs generation, as well as the small world or the scale free properties tend to condense the degree values inside small integers' intervals.

Here, we provide evidence that moving from regular or almost regular sequences to span-two sequences is critical for the reconstruction of one of the related 3-hypergraphs.

Then, we solve the related reconstruction problem, i.e., we define a polynomial time algorithm to reconstruct the incidence matrix of a 3-hypergraph having a span-two 3-sequence.

To this aim, we first identify a family of span-two 3-sequences that we call *basic sequences* and that are used as primitive sub-matrices, say blocks, of the incidence matrix of the hypergraph. At one time, we point out few span-two sequences that are not 3-sequences. Finally, we show how to successfully complete the reconstruction of the hypergraph by vertically concatenating a basic sequence block and some blocks obtained by using as edges the cyclic shifts of 3-dense Lyndon words.

So, in the next section, we give the definitions useful for our study, we point out some relevant previous results and we introduce the reconstruction problem. Section 3 is devoted to the consistency problem concerning 3-hypergraphs having span-two 3-sequences. Then, in Sect. 4, the combinatorial characterization of basic sequences is given, and then reconstruction problem is solved, via a polynomial time algorithm. Then, we show the characterization of span-two degree sequences of a 3-uniform hypergraph. We conclude the article pointing out some open problems concerning the characterization of other families of degree sequences, in Sect. 5.

An advantage of our approach based on combinatorics on words is that all the results are likely to be easily extendable to $k \geq 4$ and to other families of degree sequences having simple characterization, such as gap-free sequences.

This paper is an extended version of [8] presented at the *International Conference on Discrete Geometry and Mathematical Morphology (DGMM)*, May 24-27, 2021, Uppsala (Sweden). The structure is the same, and the same tools are used. The novelty is that in the conference paper only step-two sequences, i.e., of the form $(d^k, (d-2)^m)$ $(k, m > 0)$ have been studied. Now, we extend step-two sequences to span-two sequences. The problem for the span-

two sequences is technically more challenging than for span-two sequences, and we find this being a sufficient additional contribution to be published in a journal paper.

## 2 Definitions, Previous Results and Introduction of the Problem

A *hypergraph* $H$ is defined as a couple $(V, E)$, where $V = (v_1, \ldots, v_n)$ is a finite set of $n$ vertices, and $E$ is the set of hyperedges, i.e., a collection of subsets of $V$, $\{e_1, e_2, \ldots, e_m\}$ where each $e_i$ is non-empty (see [1]). A hypergraph is *simple* if it is loopless, i.e., there are no singletons among its hyperedges, and without parallel hyperedges, i.e., $e \nsubseteq e'$ for any pair $e, e' \in E$. Moreover, a hypergraph is said to be *k-uniform*, briefly *k*-hypergraph, if each hyperedge has cardinality $k$. The *degree* of a vertex $v \in V$ is the number of hyperedges $e \in E$ such that $v \in e$. The *degree sequence* of $H$ is the list of its vertex degrees, usually written in non-increasing order, as $\pi = (d_1, d_2, \ldots, d_n)$, $d_1 \geq d_2 \geq \cdots \geq d_n$. Let us denote by $\sigma(\pi)$ the sum of the elements of $\pi$. When $H$ is $k-$uniform, the sequence $\pi$ is called *k*-graphic. Notice that the case $k = 2$ corresponds to graphs.

In this context, we consider the problem of characterizing 3-sequences and reconstruct one of the related 3-hypergraphs. We recall that each hypergraph $H = (V, E)$, where $|V| = n$, $|E| = m$, can be represented through its *incidence matrix*, i.e., the $m \times n$ binary matrix $A = (a_{i,j})$, with $i = 1 \ldots m$, $j = 1 \ldots n$, such that $a_{i,j} = 1$ if and only if the vertex $v_j$ belongs to the hyperedge $e_i$. Thus, $\sum_{i=1}^{m} = a_{ij} = d_j$ is the degree of the vertex $v_j$, and when $H$ is $k$-uniform we have $\sum_{j=1}^{n} a_{ij} = k$ for each hyperedge $e_i$. So, our study focuses on the characterization of a subset of 3-sequences and the reconstruction of the incidence matrix of their related 3-hypergraphs.

We underline that if a matrix $A$ is the incidence matrix of a hypergraph $H$, then its rows' and columns' sums correspond to the sequence of the cardinalities of the elements of $E$ and to the degree sequence $\pi$, respectively.

Frosini et al. in [9] give a necessary and sufficient condition to test $d$-regular or almost $d$-regular $k$-sequences, i.e., the cases $\pi = (d, \ldots, d)$ or $\pi = (d, \ldots, d, d-1, \ldots, d-1)$, respectively.

**Theorem 1** *A sequence* $\pi = (d, \ldots, d)$ *(resp.* $\pi = (d, \ldots, d, d-1, \ldots, d-1)$*), is the degree sequence of a $k$-hypergraph $H = (V, E)$, with $|V| = n$ and $|E| = m$, if and only if*

1. $mk = nd$ *(this implies $\sigma(\pi) \equiv_k 0$);*
2. $k \leq n$, $d \leq m$;
3. $d \leq \frac{k}{n}\binom{n}{k}$.

Moreover, given a regular or almost regular integer sequence satisfying the conditions of Theorem 1, they define a *P*-time algorithm to compute one of the associated *k*-hypergraphs by reconstructing its incidence matrix. Such algorithm relies on the notion of Lyndon words and necklaces of fixed density (that we recall hereafter) and uses an already known algorithm for their efficient generation.

A sequence $\pi$ is said to have *span h* if $d_1 - d_n = h$. With $h = 0$ we have regular sequence, while with $h = 1$ we have quasi-regular sequences. Moreover, $\pi$ is said to be *gap-free* if $d_i - d_{i+1} \leq 1$, for all $i = 1, \ldots, n - 1$.

In this paper we investigate 3-sequences having span two (or span-two sequences), i.e., sequences of the form $\pi = (d^g, (d-1)^h, (d-2)^p)$, where the exponential notation $x^i$ indicates the entry $x$ is repeated $i$ times. Then, we show that their characterization and the reconstruction of (one of) the related 3-hypergraphs cannot be obtained as a simple generalization of the results in [8,9].

## 2.1 Necklaces and Lyndon Words

Here, we recall a couple of standard notions of combinatorics of words that will be useful in the sequel. Indeed, we consider each row of a matrix as a binary word, and we group them into equivalence classes according to their cyclic shifts. We consider each row of a binary matrix as a binary finite word $w \in \{0, 1\}^*$, $w = w_1 \ldots w_n$, whose length $n$ is the number of columns of the matrix. Given $i \in \{0, \ldots, n - 1\}$ the *i*th cyclic shift of $w$ is the word $s^i(w) = w_{i+1} \ldots w_n w_1 \ldots w_i$, so that $s^0(w) = w$. We note that applying a cyclic shift to $w$, we obtain a different word, unless the cases $w = (1)^n$ or $w = (0)^n$, of the same length, and having the same number of 1-elements inside. Note that the words repeat after at most $n$ shifts.

Following the notation in [11], the *binary necklace* (briefly *necklace*) of a binary word $w$ is the equivalence class of all its cyclic shifts. We identify a necklace with the lexicographically least representative $w$ in its equivalence class, denoted by $[w]$. The set of all (the words representative of) the necklaces with length $n$ is denoted $N(n)$. For example, $N(4) = \{0000, 0001, 0011, 0101, 0111, 1111\}$. An important class of necklaces are those that are aperiodic. An aperiodic necklace (i.e., one with period greater than or equal to $n$) is called a *Lyndon word*. If $w$ is a Lyndon word, then it cannot be expressed as a power of one if its proper sub-words. Let $L(n)$ denote the set of all Lyndon words with length $n$. For example, $L(4) = \{0001, 0011, 0111\}$. It holds the property that a word $w$ of length $n$ is a Lyndon word if and only if its necklace has $n$ different elements. We denote fixed-density necklaces and Lyndon words in a similar manner by adding the parameter $h$ to represent the density of the word. Thus, the set of necklaces with density $h$ is represented by $N(n, h)$, and the set of Lyndon words with density $h$ is rep-

$$\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix}$$

**Fig. 1** The matrix $M(0111)$. Note that all its columns' and rows' sums equal 3. Since all its rows are different, then it is the incidence matrix of a 3-hypergraph

resented by $L(n, h)$. For example, $N(4, 2) = \{0011, 0101\}$, and $L(4, 2) = \{0011\}$.

It is known from Gilbert et al. [10] that the number of fixed density necklaces and Lyndon words is

$$N(n, h) = \frac{1}{n} \sum_{j \mid \gcd(n,h)} \phi(j) \binom{n|j}{h|j},$$

$$L(n, h) = \frac{1}{n} \sum_{j \mid \gcd(n,h)} \mu(j) \binom{n|j}{h|j},$$

respectively, where the symbols $\phi$ and $\mu$ refer to Euler and Möbius functions. We will be interested in Lyndon words with density 3. In particular, we enlighten the connection between these objects and our problem, denoting by $M(w)$ the matrix obtained by vertically concatenate all the different cyclic shifts of a word $w$. For example, $M(0111)$ is the matrix depicted in Fig. 1.

Since the cardinality of $[w]$ (i.e., the number of rows of $M(w)$), where $w$ is a word of length $n$ and density $h$, is a divisor of $n$, the following trivially holds:

**Proposition 1** *If $w$ is a Lyndon word of length $n$ and density $h$, then the cardinality of $[w]$ is equal to $n$. The rows' and columns' sums of $M(w)$ are all equal to $h$ and it is the incidence matrix of a $h$-hypergraph.*

## 3 The Characterization of 3-Sequences Having Span-Two

Given a generic 3-sequence $\pi = (d_1, \ldots, d_n)$, we define the *complement of* $\pi$, denoted by $\overline{\pi}$, the vector $\overline{\pi} = (d_{\max} - d_1, \ldots, d_{\max} - d_n)$, where, according to Theorem 1, 3., $d_{\max} = \frac{3}{n} \cdot \binom{n}{3}$. A direct consequence of the definition of complement of a degree sequence is the following proposition:

**Proposition 2** *A degree sequence $\pi$ is 3-graphic if and only if its complement is 3-graphic.*

**Proof** Since $\pi$ is 3-graphic, then there exists an incidence matrix $M$ associated with $\pi$. Given the incidence matrix $M_n$ associated with the regular vector $(d_{\max}, d_{\max}, \ldots, d_{\max})$ of length $n$, the reconstructed 3-hypergraph is regular and it is trivially unique since it contains all the possible different

rows. We remove all the rows of the submatrix $M$, and we obtain the incidence matrix of a 3-hypergraph associated with the mirror image of the vector $\overline{\pi}$ (i.e., the vector $\overline{\pi}$ read from right to left). By flipping horizontally the columns of the obtained matrix, we find the incidence matrix of a 3-hypergraph having $\overline{\pi}$ as degree sequence. Hence, $\overline{\pi}$ is 3-graphic.       □

**Example 1** Assuming that $\pi = (3, 3, 1, 1, 1)$ is $3-$graphic, we get that also its complement $\overline{\pi} = (5, 5, 5, 3, 3)$ is $3-$graphic.

In this paper, we characterize the 3-sequences having span two, i.e., of the form

$$\pi = (d^g, (d-1)^h, (d-2)^p), \text{ with } g > 0, h, p > 0.$$

With $h = 0$ we have the *step-two* degree sequences $\pi = (d^g, (d-2)^p)$ (also investigated in [8]).

Let us denote by $\mathcal{P}$, the set of all span-two sequences whose maximum value satisfies the three conditions of Theorem 1. Clearly, span-two 3-sequences are included in $\mathcal{P}$.

We consider the reconstruction problem for span-two sequences of $\mathcal{P}$ of length $n$, starting with the smallest value of $n$. There are no 3-hypergraphs having span-two degree sequences of length $n \leq 4$, indeed an easy exhaustive check reveals it. Concerning span-two sequences in $\mathcal{P}$, whose length is $n \geq 5$, we prove the following result:

**Proposition 3** *Every span-two sequence of the form:*

*(i) $\pi = (2^3, 0^p)$, with $p \geq 2$;*
*(ii) $\pi' = (2, 1, 0^p)$, with $p \geq 2$;*

*is not a 3-graphic sequence.*

**Proof** The instance $\pi = (2^3, 0^p)$ admits a unique matrix which cannot be the incidence matrix of a hypergraph, since it would need two equal rows. Therefore, the only 3-hypergraph that satisfies the degree sequence $\pi$ the one with two equal hyperedges and this goes against the hypothesis that the hypergraph is simple. Concerning $\pi' = (2, 1, 0^p)$, there is no matrix giving it, since we are dealing with 3-hypergraphs.       □

Let us denote by $\mathcal{N}$ the set of sequences of type (*i*) or (*ii*), and their complements. Let us call *basic sequences* the integer sequences $(d^g, (d-1)^h, (d-2)^p) \in \mathcal{P} \setminus \mathcal{N}$, with $2 \leq d \leq 4$, and the sequences $(5^3, 3^p)$, $(5, 4, 3^{p+1})$, with $p \geq 3$. These sequences are called *basic* since their entries have minimal value, and so they constitute the final sequences to be reconstructed in the recursive reconstruction strategy that we are going to define. Their reconstruction (described in Theorem 2) needs to be treated separately from the other

sequences in $\mathcal{P}$. For instance, with $n = 5$ we have the following basic sequences:

$(2, 2, 1, 1, 0)$, $(3, 3, 1, 1, 1)$, $(3, 2, 2, 1, 1)$, $(3, 3, 3, 2, 1)$, $(4, 2, 2, 2, 2)$, $(4, 4, 4, 4, 2)$, $(4, 3, 3, 3, 2)$, $(4, 4, 3, 2, 2)$.

We observe that, according to the definition, the sequence $(5, 5, 5, 3, 3)$ (resp. $(5, 4, 3, 3, 3)$) is not a basic sequence. Indeed, it is the complement of $(3, 3, 1, 1, 1)$ (resp. $(3, 3, 3, 2, 1)$).

Let us recall the following (standard) partial order on integer sequences $\pi$ and $\pi'$ having the same length $n$: we say that $\pi \preceq \pi'$ if and only if for each $1 \leq i \leq n$, it holds $\pi(i) \leq \pi'(i)$.

We underline that the following theorem extends Proposition 2 of [8].

**Theorem 2** *Each basic sequence is 3-graphic and it can be reconstructed using the shifts of three Lyndon words at most.*

**Proof** We obtain the result by defining, for each type of basic sequence $\pi$, a class of 3-hypergraphs satisfying it. We consider separately the cases of the basic sequences:

1. $(2^g, 1^h, 0^p)$, with $g > 0$, $h \geq 0$, $p > 0$ and such that $g + h + p > 4$, and $|g - h| \equiv_3 0$;
2. $(3^g, 2^h, 1^p)$, with $g > 0$, $h \geq 0$, $p > 0$ and such that $g + h + p > 4$, and $|p - h| \equiv_3 0$;
3. $(4^g, 3^h, 2^p)$, with $g > 0$, $h \geq 0$, $p > 0$ and such that $g + h + p > 4$, and $|g - p| \equiv_3 0$;
4. $(5^3, 3^p)$, $(5, 4, 3^{p'})$ with $p \geq 3$, $p' \geq 2$.

Let us study the 4 cases separately.

1. Let us define the following algorithm:

---

**Algorithm 1** CompBasic

---

1: **Input:** an integer sequence $\pi$ of length $n$.
2: **Output:** An integer matrix $M$ having column sums $\pi$.

3: $n = g + h + p$; $w = (1^3, 0^{n-3})$; $M = \emptyset$;
4: let $\pi_M$ be the column sums vector of $M$;
5: **for** $i = 1 : n - 3$ **do**
6:     $M' = M \ominus s^{i-1}(w)$;    % $\ominus$ being the vertical concatenation of two matrices;
7:                         % $s^{i-1}(w)$ being the $(i - 1)$th cyclic shift of $w$;
8:     **if** $\pi_{M'} \preceq \pi$ **then**
9:        $M = M'$;
10:    **end if**
11: **end for**
12: **return** $M = M \ominus (\pi \setminus \pi_M)$.

---

      □

We proceed in proving,

**Property 1** Let $\pi = (2^g, 1^h, 0^p)$, with $g > 0$, $h, p > 0$, and $\sigma(\pi) \equiv_3 0$ (the sum of the elements of $\pi$ being a multiple of three). The run of *CompBasic*$(\pi)$ returns the incidence matrix $M$ of a 3-hypergraph having degree sequence $\pi$.

**Proof** We start by observing that on input $\pi = (2^g, 1^h, 0^p)$, at each run $i$ of the for-loop, with $i > 1$, the sequence $\pi_M(2 : n)$ is weakly decreasing. This statement can be proved by induction:
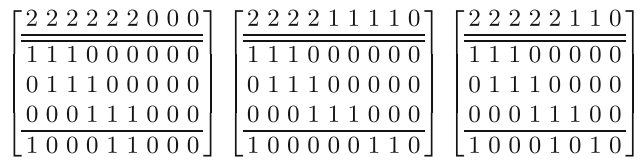
*Base:* when $i = 2$, it holds $\pi_M = (1, 2, 2, 1, 0^{n-4})$;
*Step:* by inductive hypothesis we have that the sequence $\pi_M$ is weakly decreasing up to the $i$th iteration of the for-loop. Iteration $i + 1$ acts on the following possible configurations (position $i + 1$ of $\pi_{M'}$ is underlined):

$(i)$ $\pi_M = (1, 2, \ldots, \underline{2}, x, 0^{n-i-2})$, with $x \in \{0, 1, 2\}$. The vertical concatenation $M' = M \ominus s^i(w)$ produces $\pi_{M'} = (1, 2, \ldots, \underline{3}, x+1, 1, 0^{n-i-3}) \not\leq \pi$, so no updates are produced on $M$ and $\pi_M$ maintains the weakly decreasing property;

$(ii)$ $\pi_M = (1, 2, \ldots, \underline{1}, x, 0^{n-i-2})$, with $x \in \{0, 1\}$. The vertical concatenation $M' = M \ominus s^i(w)$ produces $\pi_{M'} = (1, 2, \ldots, \underline{2}, x+1, 1, 0^{n-i-3})$, that, according to the upper bound imposed by $\pi$, can be accepted, updating the matrix $M$ and still preserving the weakly decreasing property. Note that, if $M'$ is not accepted, then $\pi(i + 2) = x$, and so $\pi(i + 1) = \pi_{M'}(i + 1) + 1$. Consequently, $i + 1$ is a position where $\pi$ and $\pi_{M'}$ differs;

$(iii)$ $\pi_M = (1, 2, \ldots, \underline{0}, 0, 0^{n-i-2})$. The vertical concatenation $M' = M \ominus s^i(w)$ produces $\pi_{M'} = (1, 2, \ldots, \underline{1}, 1, 1, 0^{n-i-3})$, that, according to the upper bound imposed by $\pi$, can be accepted, updating the matrix $M$ and still preserving the weakly decreasing property.

Finally, we observe that, at the end of the for-loop, $\pi_M$ and $\pi$ does not differ on three consecutive elements $i$, $i + 1$, $i + 2$: if so, the $i$th iteration of the for-loop would have filled the gap.

As a consequence of the two above properties, it holds that at the end of the for-loop $\pi_M$ and $\pi$ may differ on the positions $\{1, g, n-1, n\}$. Since, by construction, the number of positions where they can differ is a multiple of three, then it can be zero or three only. So, the final row $\pi \setminus \pi_M$ added to $M$ is either void or has three elements 1, so the output of *CompBasic*$(\pi)$ is the incidence matrix of a 3-hypergraph as desired. □

An application of Algorithm 1 for the reconstruction of sequences of type $(2^g, 1^h, 0^p)$ is shown in Fig. 2:

$$\begin{bmatrix} 2 & 2 & 2 & 2 & 2 & 2 & 0 & 0 & 0 \\ \hline 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}$$

$$\pi = (2^6, 0^3).$$

$$\begin{bmatrix} 2 & 2 & 2 & 2 & 1 & 1 & 1 & 1 & 0 \\ \hline 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{bmatrix}$$

$$\pi = (2^4, 1^4, 0).$$

$$\begin{bmatrix} 2 & 2 & 2 & 2 & 2 & 1 & 1 & 0 \\ \hline 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

$$\pi = (2^5, 1^2, 0).$$

**Fig. 2** An application of Algorithm 3 for the reconstruction of sequences of type $(2^g, 1^h, 0^p)$

2. Now we consider a sequence $\pi = (3^g, 2^h, 1^p)$, with $g > 0$, $h \geq 0$, $p > 0$ and such that $g + h + p > 4$, and $|p - h| \equiv_3 0$. Let us define *CompBasic2* to be the variant of *CompBasic* where line 3 is updated with the following lines to initialize the matrix $M$

3.1: $n = g + h + p$; $w = (1^3, 0^{n-3})$;
3.2: **if** $g = 1$ **then**;
3.3:     $M = (1, 0^{n-3}, 1, 1)$;
3.4: **else**
3.5:     $M = (1, 1, 0^{n-3}, 0, 1)$;
3.6: **end if**

Again, we proceed in proving

**Property 2** Let $\pi = (3^g, 2^h, 1^p)$, with $g > 0$, $h, p > 0$, and $\sigma(\pi) \equiv_3 0$. The run of *CompBasic2*$(\pi)$ returns the incidence matrix $M$ of a 3-hypergraph having degree sequence $\pi$.

**Proof** Also in this case, on input $\pi = (3^g, 2^h, 1^p)$, at each run $i$ of the for-loop, with $i > 1$, the sequence $\pi_M(2 : n - 2)$ is weakly decreasing.

This statement can be proved by induction:

*Base:* when $i = 2$, it holds either $\pi_M = (2, 1^5, 0^{n-6})$ or $\pi_M = (2, 2, 1^4, 0^{n-6})$ or $\pi_M = (2, 3, 2, 1, 0^{n-4})$;
*Step:* by inductive hypothesis we have that the sequence $\pi_M$ is weakly decreasing up to the $i$th iteration of the for-loop. Iteration $i + 1$ acts on the following possible configurations (position $i + 1$ of $\pi_{M'}$ is underlined):

$i)$ $\pi_M = (2, 3, \ldots, 3, \underline{2}, 1, 0^{n-i-4}, 1, 1)$. The vertical concatenation $M' = M \ominus s^i(w)$ produces $\pi_{M'} = (2, 3, \ldots, \underline{3}, 2, 1, 0^{n-i-5}, 1, 1)$ that, according to the upper bound imposed by $\pi$, can be accepted, updating the matrix $M$ and still preserving the weakly decreasing property;

$ii)$ $\pi_M = (2, 3, \ldots, 3, \underline{1}, x, 0^{n-i-4}, 1, 1)$, with $x \in \{1, 0\}$. By construction, this case cannot be obtained;

$iii)$ the remaining cases, where the element $\pi_M(i) < 3$, has been already considered in Property 1.

To reach the hypothesis, we observe that, at the end of the for-loop, $\pi_M$ and $\pi$ does not differ on three consecutive
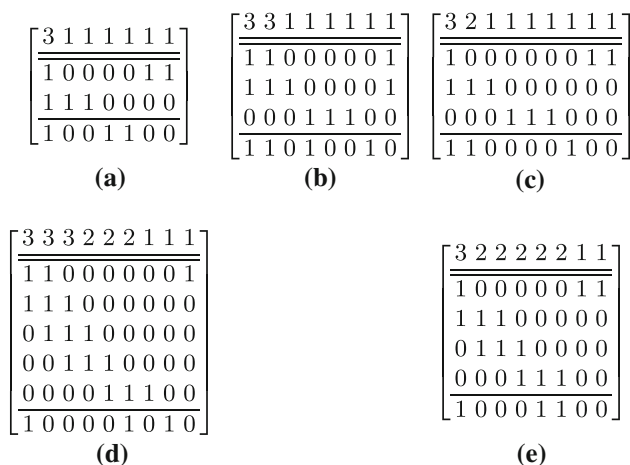
$$\begin{bmatrix} 3 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ \hline 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ \hline 1 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

**(a)**

$$\begin{bmatrix} 3 & 3 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

**(b)**

$$\begin{bmatrix} 3 & 2 & 1 & 1 & 1 & 1 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

**(c)**

$$\begin{bmatrix} 3 & 3 & 3 & 2 & 2 & 2 & 1 & 1 & 1 \\ \hline 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}$$

**(d)**

$$\begin{bmatrix} 3 & 2 & 2 & 2 & 2 & 2 & 1 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \end{bmatrix}$$

**(e)**

**Fig. 3** The application of *CompBasic2* to the sequences of type $(3^g, 2^h, 1^p)$ with parameters **a**: $g = 1$, $h = 0$, $p = 6$; **b**: $g = 2$, $h = 0$, $p = 6$; **c**: $g = 1$, $h = 1$, and $p = 7$; **d**: $g = 3$, $h = 3$, $p = 3$; **e**: $g = 1$, $h = 5$, $p = 2$

elements $i$, $i + 1$, $i + 2$: If so, the $i$th iteration of the for-loop would have fill the gap.

As a consequence of the two above properties, it holds that at the end of the for-loop $\pi_M$ and $\pi$ may differ either in the positions $\{1, g + h, n - 2, n - 1\}$ or in the positions $\{1, g + h, n - 3, n - 2\}$, according to the value $g = 1$ or $g > 1$, respectively. Since, by construction, the number of positions where they can differ is a multiple of three, then it can be zero or three only. So, the final row $\pi \backslash \pi_M$ added to $M$ is either void or has three elements 1, so the output of *CompBasic($\pi$)* is the incidence matrix of a 3-hypergraph as desired. □

Some runs of *CompBasic2* for the reconstruction of sequences of type $(3^g, 2^h, 1^p)$, with different values of the parameters $g$, $h$ and $p$, are shown in Fig. 3:

3. Now, we consider the class of sequences $\pi = (4^g, 3^h, 2^p)$, with $g > 0$, $h \geq 0$, $p > 0$ and such that $n = g + h + p \geq 6$, and $|g - p| \equiv_3 0$. Let us define the variant *CompBasic3* of *CompBasic* where the following changes are provided: Let $M$ be a $m \times n$ matrix and $r$ an $n$ length vector, we indicate by $M \backslash r$ the matrix obtained by deleting row $r$ from $M$, if present

5: **for** $i = 1 : n$ **do**
12.1: **if** $(\pi \backslash \pi_M) = \{(2, 1, 0^{n-4}), (2, 0^{n-2}, 1), (2, 2, 1, 1, 0^{n-4})\}$
12.2: $M = (M \backslash s^{n-4}(w))$ or $(M = (M \backslash s^{n-3}(w))$;
12.3: **end if**
12.4: **if** *Check (M)*;
12.5: **return** *Special (M)*
12.6: **else**
12.7: **return** $M = M \ominus RecMin(\pi \backslash \pi_M)$;

12.8: **end if**

The procedures *Check (M)* and *Special(M)* deal with few special cases that arise after the partial reconstruction of $M$ in the *for loop* starting in line 5. We describe in words their simple behaviors: *Check (M)* returns 1 if $\pi' = (\pi \backslash \pi_M)$ is one of the vectors

$$\{\pi_1 = (2, 0^{n-1}, 1), \pi_2 = (2, 1, 0^{n-1}), \pi_3 = (2, 2, 1, 1, 0^{n-4})\}.$$

Those cases are treated apart by the procedure *Special(M)*: It removes from $M$ the rows

$$(0, 1^3, 0^{n-4}), (0^2, 1^3, 0^{n-5}), (0^3, 1^3, 0^{n-6})$$

in $\pi_1$, $\pi_2$, and $\pi_3$, respectively. The vector $\pi'$ is updated accordingly.

Such a row deletion is always possible when the lengths of $\pi_1$ and $\pi_2$ are greater than six, and the length of $\pi_3$ is greater than seven, as witnessed by Fig. 4.

Then, *Special* returns a matrix $A$ satisfying the updated $\pi'$ and involving two Lyndon words different from $(1^3, 0^n)$. All the possible cases are shown in Fig. 4. Finally, the matrix $A$ is vertically concatenated with $M$ to obtain the final output of *CompBasic3($\pi$)*.

So, the procedure *RecMin* acts on the remaining vectors $\pi' = (\pi \backslash \pi_M)$, i.e., whose elements are in $\{0, 1, 2\}$, all but the last one in decreasing order, $\sigma(\pi') \equiv_3 0$, $\sigma(\pi') > 3$, the element 2 occurs two times at most and they are different from $\pi_3$. We denote this class by $\mathcal{C}$.

---

**Algorithm 2** RecMin

1: **Input:** an integer sequence $\pi' \in \mathcal{C}$.
2: **Output:** A binary matrix $A$ having column sums $\pi'$.

3: $n = length(\pi')$; $m = \frac{\sigma(\pi')}{3}$; $A(1 : m, 1 : n) = 0$;
4: $A(1, i) = 1$; $A(1, j) = 1$; $A(1, j) = 1$; where $i$, $j$, $k$ are the indexes in $\pi'$ of the element(s) 2 and the greatest index(s) of the element(s) 1, if at least one element 2 is present in $\pi'$;
5: let $i_{max}$ be the maximum index of the elements 1 in $\pi' \backslash \pi_A$
6: **for** $j = 1 : \frac{i_{max}}{3}$ **do**
7: $A(j+1, j) = 1$; $A(j+1, j + \frac{i_{max}}{3}) = 1$; $A(j+1, j + 2\frac{i_{max}}{3}) = 1$;
8: **end for**
9: **return** $A$.

---

We prove

**Property 3** Let $\pi = (4^g, 3^h, 2^p)$, with $g > 0$, $h \geq 0$, $p > 0$ and such that $n = g + h + p \geq 6$, and $|g - p| \equiv_3 0$. The run of *CompBasic3 ($\pi$)* returns the incidence matrix $M$ of a 3-hypergraph having degree sequence $\pi$.

**Proof** Since the use of the shifts of $w = (1^3, 0^{n-3})$ in $M$ prevents each column sum to overcame 3, from Proposition 2,

$$M_1 = \begin{bmatrix} 4\,2\,2\,2\,2\,2\,2\,2 \\ \hline 1\,1\,1\,0\,0\,0\,0\,0 \\ 0\,\underline{1}\,\underline{1}\,1\,0\,0\,0\,0 \\ 0\,0\,0\,1\,1\,1\,0\,0 \\ 0\,0\,0\,0\,1\,1\,1\,0 \\ 1\,0\,0\,0\,0\,0\,1\,1 \\ 2\,0\,0\,0\,0\,0\,0\,1 \end{bmatrix} \quad M_2 = \begin{bmatrix} 4\,3\,3\,2\,2\,2\,2 \\ \hline 1\,1\,1\,0\,0\,0\,0 \\ 0\,1\,1\,1\,0\,0\,0 \\ 0\,0\,\underline{1}\,\underline{1}\,1\,0\,0 \\ 0\,0\,0\,0\,1\,1\,1 \\ 1\,0\,0\,0\,0\,1\,1 \\ 2\,1\,0\,0\,0\,0\,0 \end{bmatrix} \quad M_3 = \begin{bmatrix} 4\,4\,4\,4\,2\,2\,2\,2 \\ \hline 1\,1\,1\,0\,0\,0\,0\,0 \\ 0\,1\,1\,1\,0\,0\,0\,0 \\ 0\,0\,1\,1\,1\,0\,0\,0 \\ 0\,0\,0\,\underline{1}\,\underline{1}\,1\,0\,0 \\ 0\,0\,0\,0\,0\,1\,1\,1 \\ 1\,0\,0\,0\,0\,0\,1\,1 \\ 2\,2\,1\,1\,0\,0\,0\,0 \end{bmatrix}$$

$$A_1 = \begin{bmatrix} 2\,1\,1\,1\,0\,0\,0\,1 \\ \hline 1\,0\,1\,0\,0\,0\,0\,1 \\ 1\,1\,0\,1\,0\,0\,0\,0 \end{bmatrix} \quad A_2 = \begin{bmatrix} 2\,1\,1\,1\,1\,0\,0 \\ \hline 1\,0\,1\,0\,1\,0\,0 \\ 1\,1\,0\,1\,0\,0\,0 \end{bmatrix} \quad A_3 = \begin{bmatrix} 2\,2\,1\,2\,1\,1\,0\,0 \\ \hline 1\,1\,0\,1\,0\,0\,0\,0 \\ 1\,0\,1\,0\,1\,0\,0\,0 \\ 0\,1\,0\,1\,0\,1\,0\,0 \end{bmatrix}$$

**Fig. 4** Three vectors $\pi$, i.e., $\{(4, 2^7), (4, 3^2, 2^4), (4^4, 2^4)\}$ that originate the cases $\pi_1$, $\pi_2$, and $\pi_3$. The procedure *Special* acts on them first deleting from the correspondent matrices $M_i$, $i = 1, \ldots, 3$ the rows whose elements are underlined. Then, the updated vectors $\pi_1$, $\pi_2$, and $\pi_3$ are obtained and reconstructed by the matrices $A_1$, $A_2$, and $A_3$, respectively. The upper and the lower parts of the matrices have to be vertically arranged to obtain the incidence matrices of three 3-hypergraphs. Each solution contains three Lyndon words at most

we obtain that the for loop starting in line 5 produces a weakly decreasing sub-sequence $\pi_M(3 : n)$. Furthermore, since no elements 1 are present in $\pi$, then $\pi_M(g + 1 : n - 1) = \pi$ (recall, yet from Proposition 2 that one of the three positions where $\pi_M$ and $\pi$ may differ is where the sequence of elements 2 in $\pi$ ends, changing into the sequence of elements 1. No differences are present when the sequence of elements 3 changes into that of elements 2).

Since $\pi(n) = 2$, the extension of the *for loop* variable $i$ in line 5 until $n$ adds to $M$ either the row $(1, 1, 0^{n-3}, 1)$ or the row $(1, 0^{n-3}, 1, 1)$ or both, according to $p \equiv_3 1$, $p \equiv_3 2$ or $p \equiv_3 0$ (see Fig. 4 for some examples).

An easy check reveals that for all the three of the above additions the sequence $\pi'(1 : n - 1) = \pi(1 : n - 1) \backslash \pi_M(1 : n - 1)$ is weakly decreasing, but for the case $\pi = (4, 2^{3k+1})$, with $k \geq 2$, where $\pi' = (2, 0^{3k}, 1)$. Furthermore, by construction, $\pi'$ may contain at most two elements 2 starting from its first position, and it holds $\sigma(\pi') \equiv_3 0$.

So, if $\pi'$ turns out to be $\pi_1$, $\pi_2$ or $\pi_3$, then *Special* $(M)$ successfully completes the reconstruction of a matrix $M$ that is the incidence matrix of a 3-hypergraph and it uses three Lyndon words at most. Otherwise $\pi \backslash \pi_M$ is a sequence in $\mathcal{C}$, so *RecMin* creates a new matrix $A$ by adding (line 4 of *RecMin*) a first row that allows, by construction, $\pi' \backslash \pi_A$ to be a zero vector, but for a single sequence of elements 1 of length greater than three. Then, *RecMin* completes the reconstruction of $A$ by adding rows that are shifts of the same Lyndon word (lines $6 - 8$ in *RecMin*). So, also in this case, the final output of *RecBasic3* $(\pi)$, i.e., the vertical concatenation of $M$ and $A$, produces the incidence matrix of a 3-hypergraph by using three Lyndon words at most. $\square$

4. The following cases end our analysis: $(5^3, 3^p)$, $(5, 4, 3^{p'})$ with $p \geq 3$, $p' \geq 2$.

$$\begin{bmatrix} 5\,5\,5\,3\,3\,3 \\ \hline 1\,1\,1\,0\,0\,0 \\ 0\,1\,1\,1\,0\,0 \\ 0\,0\,1\,1\,1\,0 \\ 1\,0\,0\,0\,1\,1 \\ 1\,1\,0\,0\,0\,1 \\ \hline 1\,1\,0\,1\,0\,0 \\ 0\,1\,1\,0\,1\,0 \\ 1\,0\,1\,0\,0\,1 \end{bmatrix} \qquad \begin{bmatrix} 5\,4\,3\,3\,3\,3\,3 \\ \hline 1\,1\,1\,0\,0\,0\,0 \\ 0\,1\,1\,1\,0\,0\,0 \\ 0\,0\,1\,1\,1\,0\,0 \\ 0\,0\,0\,1\,1\,1\,0 \\ 1\,0\,0\,0\,0\,1\,1 \\ 1\,1\,0\,0\,0\,0\,1 \\ 1\,1\,0\,0\,0\,1\,0 \\ 1\,0\,0\,0\,1\,0\,1 \end{bmatrix}$$

**Fig. 5** The reconstruction of $\pi = (5, 5, 5, 3, 3, 3)$ on the left, and $\pi = (5, 4, 3, 3, 3, 3, 3)$ on the right

(1) Concerning $\pi = (5^3, 3^p)$, $p \geq 3$, the 3-hypergraph can be reconstructed using the matrix obtained from all cyclic shifts of the word $1^3 0^{n-3}$, except $0^{n-3} 1^3$. Now, we have to reconstruct the matrix associated with the sequence $(2^3, 0^{n-6}, 1^3)$. This can be done using the Lyndon words $110^{n-5} 10^2$, $0110^{n-5} 10$, and $1010^{n-4} 1$. So, we need three Lyndon words, except for the case where $\pi = (5^3, 3^3)$, where two Lyndon words are needed. See Fig. 5 (left).

(2) Concerning the case $\pi = (5, 4, 3^{p'})$, $p' \geq 2$, the hypergraph can be reconstructed using the matrix obtained from all cyclic shifts of the word $1^3 0^{n-3}$, except $0^{n-3} 1^3$. Now, we have to reconstruct the matrix associated with the sequence $(2, 1, 0^{n-6}, 1^3)$. This can be done using the Lyndon word $110^{p'-2} 10$, and its shift $10^{p'-2} 101$. See Fig. 5, (right).

All the basic sequences have been considered, and for each of them, we have provided a reconstruction strategy that involves three Lyndon words at most, so reaching the thesis. $\square$

Observe that the reconstruction of a basic sequence requires two Lyndon words at most, except for the cases:

(i) $(4^{3g'}, 2^{3p'})$ (with minimal length seven);
(ii) $(4^g, 3^h, 2^p)$ (with minimal length five);
(iii) $(5^3, 3^p)$, $p > 3$ (with minimal length seven).

We point out that the sequences of length 5 are: $(4, 3, 3, 3, 2)$, $(4, 4, 3, 2, 2)$, and those of length 6 are: $(4, 4, 3, 3, 2, 2)$, $(4, 4, 4, 4, 3, 2)$, $(4, 3, 2, 2, 2, 2)$, $(4, 3, 3, 3, 3, 2)$. These six sequences can be reconstructed with two Lyndon words, as special cases of Theorem 2:

So we can state the following:

**Corollary 1** *Each basic sequence having length less than seven can be reconstructed using two Lyndon words at most.*

The procedure that reconstructs the incidence matrix associated with a basic element $\pi$ and that relies on the proof of Theorem 2 is denoted by **RecBasic**$(\pi)$.

## 4 Reconstructing a 3-Hypergraph from an Element of $\mathcal{P}$

We recall that Sawada et al. [11] presented a constant amortized time (CAT) algorithm *FastFixedContent* for the exhaustive generation of necklaces $N(n, h)$ of fixed length and density. Moreover, Sawada [12] shows that a slight modification of it, here denoted *GenLyndon*$(n, h)$, can be used for the CAT generation of the Lyndon words $L(n, h)$. This latter constructs a generating tree of the words, and since the tree has height $h$, the computational cost of generating $k$ words of $L(n, h)$ is $O(k \cdot h \cdot n)$.

Let us put together the previous algorithms and define the procedure $RecP(\pi)$ to reconstruct the 3-hypergraph of a span-two degree sequence $\pi$, if it exists. The pseudo-code of the procedure is provided:

---

**Algorithm 3** $RecP(\pi)$

---

**Input:** The sequence $\pi = (d^g, (d-1)^h, (d-2)^p) \in \mathcal{P} \setminus \mathcal{N}$ of length $n \geq 5$.
**Output:** An incidence matrix $M$ of a $3-$hypergraph whose degree sequence is $\pi$.
We solve $RecP(\pi')$, where $\pi'$ is the minimum between $\pi$ and $\overline{\pi}$ in the lexicographical order.
**Step 1:** We initialize $f = 0$, and the vector $D = \pi'$.
**Step 2:** While $D$ is not a basic sequence, do $D = D - (3^n)$ and $f = f + 1$.
**Step 3:** $B = RecBasic(D)$, using the Lyndon words $\ell_1, \ldots, \ell_r$, where $r = 2$ or $3$ depending on $D$.
If $f + r > L(n, 3)$, then give **Failure**.
**Step 4:** By applying *GenLyndon*$(n, 3)$, generate the sequence of Lyndon words, and let $w_1, \ldots, w_q$, be the set of these Lyndon words after removing $\ell_1, \ldots, \ell_r$.
**Step 5:** Using $w_1, \ldots w_f$, we generate the matrix of the necklaces $M = M(w_1) \ominus \ldots \ominus M(w_f)$. **Return** $M = M \ominus B$.

---

**Theorem 3** *Any sequence $\pi \in \mathcal{P} \setminus \mathcal{N}$ can be reconstructed by* RecP.

**Proof** The proof consists in showing that, for any $\pi \in \mathcal{P} \setminus \mathcal{N}$ of length $n > 4$ there is a sufficient number of Lyndon words required by $RecP(\pi)$.

We need to consider the cases $n = 5$ and $n = 6$ separately. Concerning these two cases, we know from Corollary 1 that all basic sequences can be reconstructed using at most two Lyndon words, so we will just need to prove that all the sequences $\pi \in \mathcal{P} \setminus \mathcal{N}$ which are not basic can be reconstructed.

(1) For $n = 5$, we have $d_{\max} = 6$ and $L(5, 3) = 2$. There are only four sequences in $\mathcal{P} \setminus \mathcal{N}$: $\pi_1 = (3^2, 1^3)$ and its complement $\overline{\pi_1} = (5^3, 3^2)$, and $\pi_2 = (4, 2^4)$ and its complement $\overline{\pi_2} = (4^4, 2)$. $\pi_1$ and $\pi_2$ are basic sequences, and they can be reconstructed with two Lyndon words, so also their complements can be reconstructed.

(2) For $n = 6$, we have $d_{\max} = 10$ and $L(6, 3) = 3$. Any sequence $\pi = (d^g, (d-1)^h, (d-2)^p)$ with $d \leq d_{\max}$, such that $\pi \leq \overline{\pi}$ has $d \leq 7$. The cases $d = 2, 3$, and $4$ concerns basic sequences, otherwise it holds $d \leq 7$, so $\pi - (3^n)$ is a basic sequence, then three Lyndon words are sufficient for the reconstruction.

Now, let $\pi \in \mathcal{P}$ be a generic sequence of length $n > 6$, such that $\pi$ is less than $\overline{\pi}$ in the lexicographic order, and denote by $l(B)$ the number of words needed to reconstruct the basic sequence $B = B(\pi)$ associated with $\pi$. The number of Lyndon words of length $n$ and density 3 is:

$$L(n, 3) = \begin{cases} \frac{1}{n}\binom{n}{3} & \text{if } n \text{ is not a multiple of 3,} \\ \frac{1}{n}[\binom{n}{3} - \frac{n}{3}] & \text{otherwise.} \end{cases}$$

We show that the inequality of Step 3 of *RecP* is always satisfied on input $\pi$, i.e., there are enough Lyndon words to reconstruct 3-hypergraph related to $\pi$. Precisely, a simple computation shows that the following inequality holds:

$$\frac{\lceil \frac{d_{\max}}{2} \cdot \frac{3}{n}\binom{n}{3} \rceil - d_B}{3} + l(B) \leq L(n, 3)$$

where:

– $\lceil \frac{d_{\max}}{2} \cdot \frac{3}{n}\binom{n}{3} \rceil$ is the maximal degree of a span-two degree sequence of length $n$ in $\mathcal{P} \setminus \mathcal{N}$ up to complement;
– $d_B$ is the greatest degree of a basic sequence (hence either $d_B = 2$, or $d_B = 3$, or $d_B = 4$, or $d_B = 5$).

From the previous inequality we get:

**Fig. 6** The reconstruction of the sequences of length 5

$$
\begin{bmatrix} 4\,4\,3\,2\,2 \\ \hline 1\,1\,1\,0\,0 \\ \hline 1\,1\,0\,1\,1 \\ 1\,1\,0\,1\,0 \\ 1\,0\,1\,1\,0 \\ 0\,1\,1\,0\,1 \end{bmatrix}
\begin{bmatrix} 4\,3\,3\,3\,2 \\ \hline 1\,1\,1\,0\,0 \\ \hline 0\,1\,1\,1\,0 \\ 1\,1\,0\,0\,1 \\ 1\,0\,0\,1\,1 \\ 1\,0\,1\,1\,0 \end{bmatrix}
\begin{bmatrix} 4\,4\,3\,3\,2\,2 \\ \hline 1\,1\,1\,0\,0\,0 \\ 0\,1\,1\,1\,0\,0 \\ 0\,0\,1\,1\,1\,0 \\ 1\,1\,0\,0\,0\,1 \\ 1\,0\,0\,0\,1\,1 \\ 1\,1\,0\,1\,0\,0 \end{bmatrix}
\begin{bmatrix} 4\,4\,4\,4\,3\,2 \\ \hline \hline 1\,1\,1\,0\,0\,0 \\ 0\,1\,1\,1\,0\,0 \\ 0\,0\,1\,1\,1\,0 \\ 1\,1\,0\,0\,0\,1 \\ 1\,0\,0\,0\,1\,1 \\ 1\,0\,1\,1\,0\,0 \\ 0\,1\,0\,1\,1\,0 \end{bmatrix}
\begin{bmatrix} 4\,3\,3\,3\,3\,2 \\ \hline 1\,1\,1\,0\,0\,0 \\ 0\,1\,1\,1\,0\,0 \\ 0\,0\,1\,1\,1\,0 \\ 1\,1\,0\,0\,0\,1 \\ 1\,0\,0\,0\,1\,1 \\ 1\,0\,0\,1\,1\,0 \end{bmatrix}
\begin{bmatrix} 4\,3\,2\,2\,2\,2 \\ \hline 1\,1\,1\,0\,0\,0 \\ 0\,1\,1\,1\,0\,0 \\ 1\,1\,0\,0\,0\,1 \\ 1\,0\,0\,0\,1\,1 \\ 1\,0\,0\,1\,1\,0 \end{bmatrix}
$$

- if $n$ is a multiple of 3:

$$
6l(B) - 2d_B \leq \frac{(n-1)(n-2)}{2}
$$

that is always satisfied starting from $d_B = 4$ and $n = 7$ on.

- if $n$ is not a multiple of 3, we get:

$$
6l(B) - 2d_B + 2 \leq \frac{(n-1)(n-2)}{2} \, .
$$

Again the inequality is always satisfied starting from $d_B = 4$ and $n = 9$ on.

$\square$

**Remark 1** Observe that **Failure** occurs if the number of Lyndon words of $L(n, 3)$ is not sufficient to reconstruct the matrix, and the above statement prevents this to occur.

**Remark 2** Theorem 3 clarifies the insertion of $\pi = (5^3, 3^p)$ and $\pi = (5, 4, 3^{p+1})$, with $p \geq 3$ among the basic sequences: the action of *RecP* on them removes a block $(3^n)$, and produces the sequences $(2^3, 0^p)$ and $(2, 1, 0^{p+1})$, which are not the degree sequences of a 3-hypergraph, as stated in Proposition 3.

The next theorem is an immediate consequence of Theorem 3 and Remark 2.

**Theorem 4** *A sequence* $\pi \in \mathcal{P} \backslash \mathcal{N}$ *is 3-graphic if and only if* $\pi$ *can be reconstructed using the Algorithm* RecP.

The validity of $RecP(\pi)$ is a simple consequence of Theorem 4. Clearly, the obtained matrix is the incidence matrix of a 3-hypergraph having degree sequence $\pi$, indeed by construction, all the rows are distinct. Moreover, the algorithm always terminates since, at each iteration, it adds as many rows as possible to the final solution. Concerning the complexity analysis, we need to generate $O(m)$ different Lyndon words and shifts each of them $O(n)$ times. Thus, since the algorithm *GenLyndon*$(n,3)$ requires $O(f \cdot h \cdot n)$, that is $O(3 \cdot f \cdot n)$ steps to generate $f$ words of $L(n, 3)$, the whole process takes polynomial time.

As a consequence of Theorem 4 we have a simple characterization of the span-two sequences which are 3-graphic.
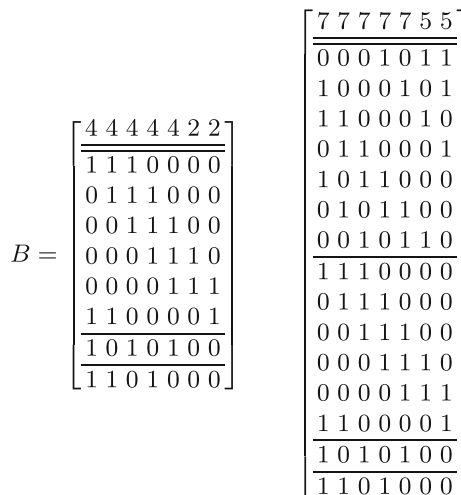
$$
B = \begin{bmatrix} 4\,4\,4\,4\,2\,2 \\ \hline \hline 1\,1\,1\,0\,0\,0\,0 \\ 0\,1\,1\,1\,0\,0\,0 \\ 0\,0\,1\,1\,1\,0\,0 \\ 0\,0\,0\,1\,1\,1\,0 \\ 0\,0\,0\,0\,1\,1\,1 \\ 1\,1\,0\,0\,0\,0\,1 \\ \hline 1\,0\,1\,0\,1\,0\,0 \\ \hline 1\,1\,0\,1\,0\,0\,0 \end{bmatrix}
\qquad
\begin{bmatrix} 7\,7\,7\,7\,5\,5 \\ \hline \hline 0\,0\,0\,1\,0\,1\,1 \\ 1\,0\,0\,0\,1\,0\,1 \\ 1\,1\,0\,0\,0\,1\,0 \\ 0\,1\,1\,0\,0\,0\,1 \\ 1\,0\,1\,1\,0\,0\,0 \\ 0\,1\,0\,1\,1\,0\,0 \\ 0\,0\,1\,0\,1\,1\,0 \\ \hline 1\,1\,1\,0\,0\,0\,0 \\ 0\,1\,1\,1\,0\,0\,0 \\ 0\,0\,1\,1\,1\,0\,0 \\ 0\,0\,0\,1\,1\,1\,0 \\ 0\,0\,0\,0\,1\,1\,1 \\ 1\,1\,0\,0\,0\,0\,1 \\ \hline 1\,0\,1\,0\,1\,0\,0 \\ \hline 1\,1\,0\,1\,0\,0\,0 \end{bmatrix}
$$

**Fig. 7** On the left, the matrix $B = RecBasic((7^5, 5^2))$; on the right the reconstruction of $\pi = (7^5, 5^2)$

**Corollary 2** *The span-two 3-graphic sequences are exactly all the sequences in* $\mathcal{P} \backslash \mathcal{N}$.

### 4.1 Two Examples of Application of *RecP*($\pi$)

We illustrate the action of *RecP* with two examples, the second one involving a gap-free sequence.

**Example 2** Let us consider the span-two sequence $\pi = (7^5, 5^2)$ having length $n = 7$, $d_{\max} = 15$. Moreover, $L(7, 3) = 5$ and, in particular, by applying *GenLyndon*$(7, 3)$ we compute the Lyndon words $\ell_1 = 1110000$, $\ell_2 = 1101000$, $\ell_3 = 1010100$, $\ell_4 = 1100100$, $\ell_5 = 1100010$.

- In Step 1, we initialize $f = 0$, and $D = \min\{(7^5, 5^2), (10^2, 8^5)\} = (7^5, 5^2)$.
- From Step 2, we get the basic sequence $D = (7^5, 5^2) - (3^7) = (4^5, 2^2)$, and $f = 1$.
- In Step 3, we apply the procedure $B = RecBasic(D)$, using the Lyndon words $\ell_1, \ell_2, \ell_3$. See Fig. 7, (left).

Then $r = 3$. Since $f + r = 4 < L(7, 3) = 5$, *RecP* does not give **Failure**.

From Step 4, removing $\ell_1, \ell_2, \ell_3$ from the set of all available Lyndon words, we get $\ell_4$ and $\ell_5$.

$$B = \begin{bmatrix} 5 & 4 & 3 & 3 & 3 & 3 & 3 \\ \hline 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ \hline 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 8 & 7 & 6 & 6 & 6 & 6 & 6 \\ \hline 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ \hline 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ \hline 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}$$

**Fig. 8** On the left, the matrix $B = RecBasic((8, 7, 6^5))$; on the right the reconstruction of $\pi = (8, 7, 6^5)$

In Step 5, to reconstruct $(3^7)$, we use, for example, all the cyclic shifts of $\ell_5$. Thus, vertically concatenating the obtained matrices, we get the solution depicted in Fig. 7 (right).

**Example 3** Let us consider the span-two sequence $\pi = (8, 7, 6^5)$ having length $n = 7$, again $d_{\max} = 15$ and we have the five Lyndon words $\ell_1, \ell_2, \ell_3, \ell_4, \ell_5$.

- In Step 1, we initialize $f = 0$, and $D = \min\{(8, 7, 6^5), (9^5, 8, 7)\} = (8, 7, 6^5)$.
- From Step 2, we get the basic sequence $D = (8, 7, 6^5) - (3^7) = (5, 4, 3^5)$, and $f = 1$.
- In Step 3, we apply the procedure $B = RecBasic(D)$, using the Lyndon words $\ell_1, \ell_2$. See Fig. 8 (left).

Then $r = 2$. Since $f + r = 3 < L(7, 3) = 5$, $RecP$ does not give **Failure**.

- From Step 4, removing $\ell_1, \ell_2$, we get $\ell_3, \ell_4$ and $\ell_5$.
- In Step 5, to reconstruct $(3^7)$, we use, for example, all the cyclic shifts of $\ell_5$. Thus, vertically arrange the obtained matrices, we get the matrix depicted in Fig. 8 (right).

## 5 Future Developments and Conclusions

In this paper we defined a polynomial time algorithm that reconstructs the incidence matrix of a 3-uniform hypergraph, when such a hypergraph exists, from its degree sequence in case it is a span-two sequence.

The novelty of our approach concerns the use of notions from combinatorics on words, in particular Lyndon words,

to enhance the reconstruction strategy. The main result concerns the characterization of span-two sequences which are 3-graphic, and states that actually, very few sequences among those satisfying the conditions (1), (2), and (3) of Theorem 1 are not 3-graphic, i.e., those belonging to the set $\mathcal{N}$.

*Degree sequences having span $h > 2$:* we believe that an analogous rather simple characterization of sequences which are not 3-graphic cannot be obtained also for generic span-$h$ sequences, $h > 2$. For instance, with $h = 3$, the set $\mathcal{N}_3$ of span-three sequences satisfying the conditions 1., 2., and 3. of Theorem 1 that are not 3-graphic contains:

$$(3, 1^3, 0), (4, 1^5), (4, 2, 1^3), (4, 2, 1^2),$$
$$(4, 3, 3, 1, 1), (4, 2, 1^2), (4^3, 2, 1)$$

and

$$(3, 2, 1, 0^p), (3^2, 2, 1, 0^p), (3^r, 0^p) \text{ with}$$
$$p \geq 1, \text{ and } r = 1, 2, 3,$$

and their complements. Again, we stress that few nonzero elements of the sequences prevent them from being 3-sequences. Therefore, if the number of nonzero elements overcomes 6, there exists only 10 sequences in $\mathcal{N}_3$.

Unfortunately, it seems that this characterization does not immediately lead to an algorithm for the reconstruction of 3-sequences having span $h$, when $h > 2$. This is due to the fact that the characterization of basic sequences for $h > 2$ is quite hard, and so it is the formulation of an analogous of Theorem 2. To better realize this statement, we investigate the case $h = 3$. An exhaustive search computes the following the span-3 basic sequences whose sums are multiple of three:

$\diamond$ $(3^p, 2^q, 1^r, 0^s)$: having length greater than five, and $p, s \geq 1, q, r \geq 0$;
$\diamond$ $(4^p, 3^q, 2^r, 1^s)$ and $(5^p, 4^q, 3^r, 2^s)$: with $p, s \geq 1$, $q, r \geq 0$ and length greater than five;
$\diamond$ $(6, 6, 3^q)$: with $q \geq 4$;
$\diamond$ $(6, 6, 6, 3^q)$: with $q \geq 3$;
$\diamond$ $(6, 5, 4, 3^s)$ and $(6, 6, 5, 4, 3^s)$: with $s > 1$.
$\diamond$ $(7, 4^{3q+2})$: with $q \geq 1$;

excluding the sequences in $\mathcal{N}_3$.

If we want to use a reconstruction strategy similar to that in **RecP**, it is necessary to provide a statement analogous to that of Theorem 2, i.e., we have to prove that any basic sequence can be reconstructed using a number of Lyndon words bounded by a constant.

This task turns out to be extremely hard when the relations between the run lengths of the entries in the sequences assume the form of a linear Diophantine equation with more than two variables, as for the gap-free sequences. As an example,

concerning the sequences $(4^p, 3^q, 2^r, 1^s)$, the coefficients $p, s \geq 1, r \geq 0$ have to satisfy the equation $4p + 2r + s = 3z$, whose solutions lie on $3D$ hyper-planes.

It is worth mentioning that, in the specific case of gap-free sequences, some different approaches (see [2,3,9]) may help the reconstruction in the case of span $h > 2$. Again, these ad hoc strategies do not admit a generalization as the degree sequences span increases, underlining the non-polynomiality of the general characterization problem.

*Strategy failure.* The strategy used in the definition of *RecP* roughly consists in removing blocks of $(3^n)$ from the sequence $\pi$ in input, until we reach a basic sequence, i.e., a sequence from which we cannot remove any other block. Then, we prove that this basic sequence can be reconstructed using a minimal number of Lyndon words and that the overall number of Lyndon words which are needed is less than or equal to $L(n, 3)$. We would like to provide an example of the failure of this strategy.

**Example 4** Let us consider the sequence $\pi = (13, 11, 9, 8, 8, 7, 6)$ of length $n = 7$. In this case, $d_{max} = 21$, $L(7, 3) = 7$, and the words are listed in example 3.

The sequence is 3-graphic, indeed one possible 3-hypergraph solution has the incidence matrix in Fig. 9 (observe that it contains all the 7 Lyndon words). A generalization of our algorithm would work as follows:

(1) The complement of $\pi$ is $(15, 14, 13, 13, 12, 10, 7)$ so we reconstruct $\pi$;
(2) We use two Lyndon words to remove two blocks $(3^7)$ obtaining $\pi' = (7, 5, 3, 2, 2, 1, 1, 0)$, which is a basic sequence;
(3) The sequence $\pi'$ can only be reconstructed using six Lyndon words, as follows:

So, the application of our algorithm would require eight Lyndon words, so giving failure as output, whilst $\pi$ can be actually reconstructed.

*k-uniform hypergraphs.* The algorithm *RecP* we have defined is tuned for 3-uniform hypergraphs. A possible direction for further research follows its generalization to obtain a polynomial algorithm for the reconstruction of generic $k$-uniform hypergraphs having a step-two degree sequence. We recall that the identification of the degree sequences of $k$-uniform hypergraphs, $k \geq 3$, is an NP-hard problem. Therefore, the proposed studies aim at limiting its NP-hard core, being fully aware of the impossibility of a good general characterization. However, finding a compact nice looking characterization would be of great interest in order to design algorithms for real-life applications.

$$
\begin{bmatrix}
13 & 11 & 9 & 8 & 8 & 7 & 6 \\
\hline
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 1 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 1 & 1
\end{bmatrix}
\qquad
\begin{bmatrix}
7 & 5 & 3 & 2 & 2 & 1 & 1 & 0 \\
\hline
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\
1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 1 & 1 & 0 & 0 & 0 & 0
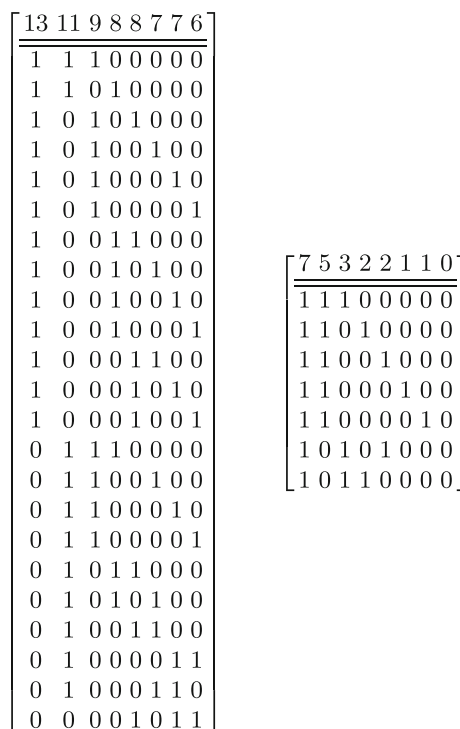\end{bmatrix}
$$

**Fig. 9** The reconstruction of the incidence matrices of two 3-hypergraphs having degree sequences $\pi = (13, 11, 9, 8, 8, 7, 6)$ and $\pi'(7, 5, 3, 2, 2, 1, 1, 0)$

## References

1. Berge, C.: Hyperpgraphs. North-Holland, Amsterdam (1989)
2. Behrens, S., Erbes, C., Ferrara, M., Hartke, S.G., Reiniger, B., Spinoza, H.: New results on degree sequences of uniform hypergraphs. Electron. J. Comb. **20**, P14 (2013)
3. Brlek, S., Frosini, A.: A tomographical interpretation of a sufficient condition on $h$-graphical sequences. In: Normand, N., Guédon, J., Autrusseau, F. (eds.) Discrete Geometry for Computer Imagery. DGCI 2016, pp. 95–104. Springer, Cham (2016)
4. Colbourn, C.J., Kocay, W.L., Stinson, D.R.: Some NP-complete problems for hypergraph degree sequences. Discrete Appl. Math. **14**, 239–254 (1986)
5. Dewdney, A.K.: Degree sequences in complexes and hypergraphs. Proc. Am. Math. Soc. **53**(2), 535–540 (1975)

6. Deza, A., Levin, A., Meesum, S.M., Onn, S.: Optimization over degree sequences. SIAM J. Discrete Math. **32**(3), 2067–2079 (2018)

7. Erdös, P., Gallai, T.: Graphs with prescribed degrees of vertices (Hungarian). Mat. Lapok **11**, 264–274 (1960)

8. Frosini, A., Palma, G., Rinaldi, S.: On the reconstruction of 3-uniform hypergraphs from step-two degree sequences. In: Lindblad, J., Malmberg, F., Sladoje, N. (eds.) Discrete Geometry and Mathematical Morphology. DGMM 2021. Lecture Notes in Computer Science, vol. 12708, pp. 338–347. Springer, Cham (2021)

9. Frosini, A., Picouleau, C., Rinaldi, S.: On the degree sequences of uniform hypergraphs. In: Gonzalez-Diaz, R., Jimenez, M.J., Medrano, B. (eds.) Discrete Geometry for Computer Imagery. DGCI 2013. Lecture Notes in Computer Science, vol. 7749, pp. 300–310. Springer, Berlin (2013)

10. Gilbert, E.N., Riordan, J.: Symmetry types of periodic sequences. Ill. J. Math. **5**(4), 657–665 (1961)

11. Ruskey, F., Sawada, J.: An efficient algorithm for generating necklaces with fixed density. SIAM J. Comput. **29**, 671–684 (1999)

12. Sawada, J.: A fast algorithm to generate necklaces with fixed content. Theor. Comput. Sci. **301**, 477–489 (2003)

**Giulia Palma** received degree in Applied Mathematics from Università degli Studi di Siena, in 2019. She is currently pursuing Ph.D. degree in Computer Science with Università degli Studi di Pisa. Her research interests include combinatorics, discrete tomography and game theory.



**Andrea Frosini** received degree in Mathematics from Università degli Studi di Siena, in 1996. He is currently a researcher in Computer Science, Department of Mathematics and Computer Science, University of Firenze.



**Simone Rinaldi** received degree in Mathematics from Università degli Studi di Siena. He is currently a professor at University of Siena. His research interests include combinatorics, discrete tomography and theory of formal languages.