$(k-2)$-linear connected components in hypergraphs of rank k

Florian Galliot — Université Grenoble Alpes, France
Sylvain Gravier — Université Grenoble Alpes, France
Isabelle Sivignon — Université Grenoble Alpes, Grenoble INP, France

Abstract

We define a *q-linear path* in a hypergraph $\mathcal{H}$ as a sequence $(e_1, \ldots, e_L)$ of edges of $\mathcal{H}$ such that $|e_i \cap e_{i+1}| \in [\![1, q]\!]$ and $e_i \cap e_j = \varnothing$ if $|i - j| > 1$. In this paper, we study the connected components associated to these paths when $q = k - 2$ where k is the rank of $\mathcal{H}$. If $k = 3$ then $q = 1$ which coincides with the well-known notion of *linear path* or *loose path*. We describe the structure of the connected components, using an algorithmic proof which shows that the connected components can be computed in polynomial time. A consequence of our algorithmic result is that tractable cases for the NP-complete problem of "Paths Avoiding Forbidden Pairs" in a graph can be deduced from the recognition of a special type of line graph of a hypergraph.

1 Introduction

There are many possible definitions for a path between two vertices in an undirected hypergraph $\mathcal{H}$. For instance, a *linear path* (or *loose path*) is a sequence of edges such that any two consecutive edges intersect on exactly one vertex and any two non-consecutive edges do not intersect. Linear paths in 3-uniform hypergraphs are our main focus. Although the existence of these paths is the subject of numerous extremal results in the literature (see [1] for instance), there has been no previous study of the connected components associated to these paths, be it in terms of algorithmic computation or description of their structure. We carry out this study in this paper.

For this, we develop methods that actually generalize to hypergraphs of rank $k \geq 4$ when replacing linearity with a notion of $(k-2)$-*linearity*: we ask for any two consecutive edges to intersect on at most $k - 2$ vertices, instead of exactly one vertex. These paths are called $(k-2)$-*linear paths*, and each vertex of $\mathcal{H}$ has its own associated $(k-2)$-*linear connected component*. Our first main result describes the structure of the subhypergraph of $\mathcal{H}$ induced by a $(k-2)$-linear connected component. The proof is algorithmic and provides us with our second main result: an algorithm that computes a $(k-2)$-linear connected component in polynomial time $O(m^2k)$, where m is the number of edges of $\mathcal{H}$. Note that this time remains polynomial even if k is part of the input.

Finally, we present an application of our algorithmic result to the "Paths Avoiding Forbidden Pairs" decision problem (known as PAFP). We define a bicolored version of the line graph of a hypergraph, and we show that PAFP is in polynomial time on any class of bicolored graphs G for which we can efficiently find a hypergraph whose bicolored line graph is G. This is interesting since PAFP is NP-complete in general [2], with only one known tractable case of significance [3].

2 The $(k-2)$-linear connectivity problem

Definition 2.1. Let $q \geq 1$. A *q-linear path* in $\mathcal{H}$ is a sequence $\overrightarrow{P} = (e_1, \ldots, e_L)$ of edges of $\mathcal{H}$ such that for all $1 \leq i < j \leq L$: $|e_i \cap e_j| \in [\![1, q]\!]$ if $j = i + 1$ and $e_i \cap e_j = \varnothing$ otherwise.

Definition 2.2. Let $q \geq 1$ be an integer and let $X, Y \subseteq V(\mathcal{H})$ be nonempty such that $|X \cap Y| \leq q$. A q-linear path $\overrightarrow{P} = (e_1, \ldots, e_L)$ in $\mathcal{H}$ is said to be *from X to Y* if:

- If $X \cap Y \neq \varnothing$, then $L = 0$.
- If $X \cap Y = \varnothing$, then $L \geq 1$ and:
  (i) $X \cap e_1 \neq \varnothing$, and if $L \geq 2$ then $X \cap e_i = \varnothing$ for all $2 \leq i \leq L$.
  (ii) $Y \cap e_L \neq \varnothing$, and if $L \geq 2$ then $Y \cap e_i = \varnothing$ for all $1 \leq i \leq L-1$.

Definition 2.3. Let $x \in V(\mathcal{H})$. The *q-linear connected component of x* in $\mathcal{H}$ is defined as:

$$LCC^q_{\mathcal{H}}(x) = \{y \in V(\mathcal{H}) \text{ such that there exists a } q\text{-linear path from } x \text{ to } y \text{ in } \mathcal{H}\}.$$

We will always assume that $\mathcal{H}$ is k-uniform. Indeed, if $\mathcal{H}$ is of rank k, then let $\mathcal{H}_0$ be the k-uniform hypergraph obtained from $\mathcal{H}$ by adding $k - |e|$ new vertices to each edge e: it is easy to see that there exists a q-linear path from x to y in $\mathcal{H}$ if and only if there exists one in $\mathcal{H}_0$.
We are interested in the case $q = k-2$. See Figure 1 (all figures will illustrate the main case of interest $k = 3$, where the notion of $(k-2)$-linear path coincides with that of a linear path). Let $x^* \in V(\mathcal{H})$: we want to understand the structure of $\mathcal{H}[LCC^{k-2}_{\mathcal{H}}(x^*)]$ (which denotes the subhypergraph of $\mathcal{H}$ induced by $LCC^{k-2}_{\mathcal{H}}(x^*)$) and to design an efficient algorithm to compute it. The idea for the algorithm is to build $\mathcal{H}[LCC^{k-2}_{\mathcal{H}}(x^*)]$ edge by edge, deciding for each examined edge whether it should be accepted as part of the component or (maybe temporarily) rejected.

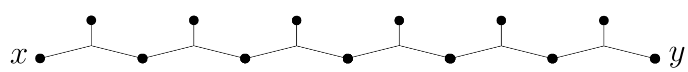

Figure 1: A $(k-2)$-linear path from x to y in the case $k = 3$. Each "claw" joining three vertices represents an edge.

3 Extendable paths, islands and archipelagos

Assume $k = 3$ for now, to alleviate notations. Consider the algorithm in the middle of its execution: part of the component has been constructed, i.e. edges have already been accepted (whose vertices we deem "accepted" as well). The difficulty is we cannot instantly accept any edge $e = \{a, b, u\}$ such that a and b have already been accepted, as illustrated in Figure 2. In this example, even though we know a linear path from x^* to $\{a, b\}$, we cannot prolong it with the edge e, because the resulting path from x^* to u would not be linear. To accept the edge e, we would need a and b to be *separated*, i.e. we would need a known linear path from x^* to $\{a, b\}$ **that does not contain both** a **and** b: only then could we prolong that path with the edge e to get a linear path from x^* to u.

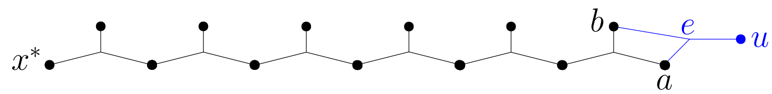

Figure 2: There is no linear path from x^* to u in this picture.

Let $\{a, b\}$ be a pair of accepted vertices that are not separated. Suppose that we are assessing an edge $e' = \{c, d, u\}$ as in Figure 3, with c and d already accepted. To fill the gap in Figure 3 and accept e', we need a linear path from $\{a, b\}$ to $\{c, d\}$ that can be prolonged on both ends, i.e. that does not contain both a and b and does not contain both c and d. Such a path will be deemed $(\{a, b\}, \{c, d\})$-extendable. This principle extends to general $k \geq 3$ as follows:

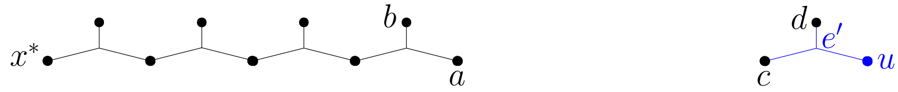

Figure 3: We need an $(\{a,b\},\{c,d\})$-extendable path to link x^* to u.

Definition 3.1. Let $X, Y \subseteq V(\mathcal{H})$ such that $1 \leq |X|, |Y| \leq k-1$ and $|X \cap Y| \leq k-2$. An (X,Y)-*extendable path* in $\mathcal{H}$ is a $(k-2)$-linear path $\overrightarrow{P} = (e_1, \ldots, e_L)$ from X to Y in $\mathcal{H}$ with the additional property if $L \geq 1$ that $|e_1 \cap X| \leq k-2$ and $|e_L \cap Y| \leq k-2$.

We now define islands and archipelagos. An island is basically a pleasant structure inside of which we have all the extendable paths we need. If all subsets of size $k-1$ are separated, then we have a single island with entry $\{x^*\}$. If $k-1$ new vertices are discovered at once, then they form the entry of a new island. All in all, we get several islands, and the only way to reach an island from x^* is through an edge that contains the entire entry of that island: we call this an x^*-archipelago.

Definition 3.2. Let $\mathcal{I}$ be a subhypergraph of $\mathcal{H}$ and $\varepsilon \subset V(\mathcal{I})$ such that $1 \leq |\varepsilon| \leq k-1$. We say $\mathcal{I}$ is an *island with entry* ε if, for all $X \subset V(\mathcal{I})$ satisfying $1 \leq |X| \leq k-1$ (and $X \neq \varepsilon$ if $|\varepsilon| = k-1$), there exists an (ε, X)-extendable path in $\mathcal{I}$. See Figure 4 for examples of islands when $k = 3$.

Definition 3.3. An x^*-*archipelago* is a subhypergraph $\mathcal{A}$ of $\mathcal{H}$ in which there exist subhypergraphs $\mathcal{I}_1, \ldots, \mathcal{I}_N$ that are pairwise-disjoint islands with respective entries $\varepsilon_1, \ldots, \varepsilon_N$ such that:

(i) $\varepsilon_1 = \{x^*\}$.
(ii) $|\varepsilon_i| = k-1$ for all $2 \leq i \leq N$.
(iii) $V(\mathcal{A}) = V(\mathcal{I}_1) \cup \ldots \cup V(\mathcal{I}_N)$.
(iv) Each edge in $E(\mathcal{A}) \setminus (E(\mathcal{I}_1) \cup \ldots \cup E(\mathcal{I}_N))$ is (i,j)-*crossing* for some $i \neq j$ i.e. is of the form $\{x\} \cup \varepsilon_j$ where $x \in V(\mathcal{I}_i)$, and the digraph G defined by $V(G) = \{\mathcal{I}_1, \ldots, \mathcal{I}_N\}$ and $E(G) = \{(\mathcal{I}_i, \mathcal{I}_j), \text{there exists an } (i,j)\text{-crossing edge}\}$ contains a spanning arborescence rooted at $\mathcal{I}_1$.

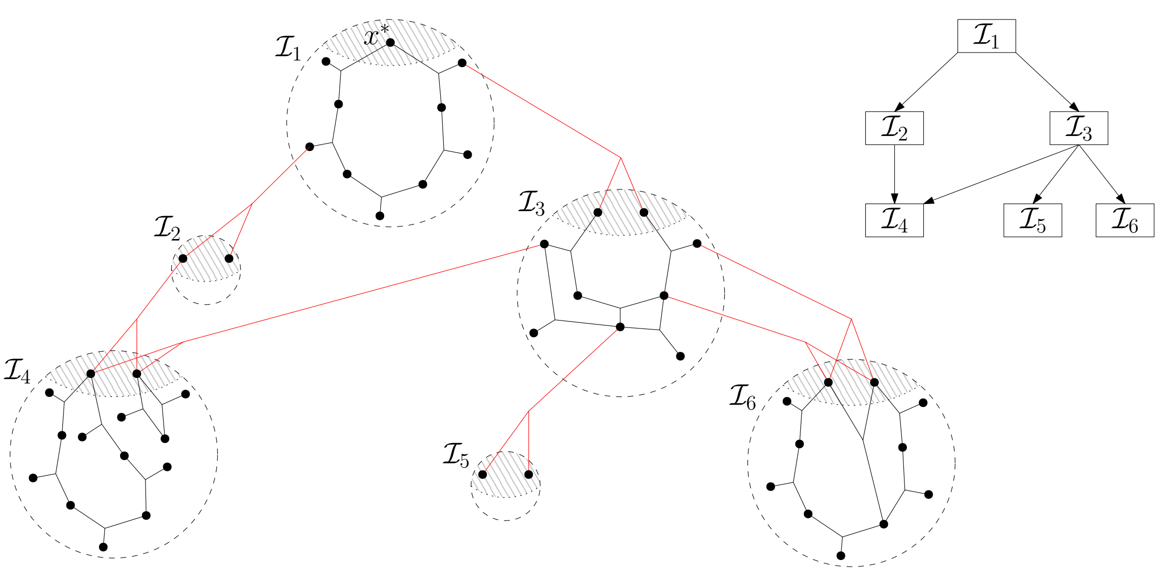

Figure 4: Example of an x^*-archipelago when $k = 3$, with the digraph from item (iv) on the right.

4 Main results

Theorem 4.1. *$\mathcal{H}[LCC_{\mathcal{H}}^{k-2}(x^*)]$ is the unique maximal x^*-archipelago in $\mathcal{H}$.*

Theorem 4.2. *$LCC_{\mathcal{H}}^{k-2}(x^*)$ can be computed in $O(m^2k)$ time where $m = |E(\mathcal{H})|$.*

These are our two main results about $(k-2)$-linear connected components: the first is structural, the second is algorithmic. The core of the proof is to show that, if $\mathcal{A}$ is an x^*-archipelago, then any edge intersecting both $V(\mathcal{A})$ and $V(\mathcal{H}) \setminus V(\mathcal{A})$ that is not of the form $\varepsilon_i \cup \{u\}$ can be added to $\mathcal{A}$ while preserving the archipelago structure. The addition of such an edge can require to redefine the islands: we may create a new one, enlarge an existing one, or merge several existing ones.

5 Application to the "Paths Avoiding Forbidden Pairs" problem

PAFP takes as input a *bicolored graph* G (i.e. a graph with red and blue edges) and two vertices x, y of G, and returns "yes" if and only if there exists a blue induced path from x to y in G.
Let $k \geq 3$. We define a bicolored version of the line graph of a k-uniform hypergraph:

Definition 5.1. Let $\mathcal{H}$ be a k-uniform hypergraph. The *bicolored line graph* of $\mathcal{H}$ is the bicolored graph G such that $V(G) = E(\mathcal{H})$ and, for all distinct $e_1, e_2 \in V(G)$, there is a blue (resp. red) edge between e_1 and e_2 in G if and only if $1 \leq |e_1 \cap e_2| \leq k-2$ (resp. if and only if $|e_1 \cap e_2| = k-1$).

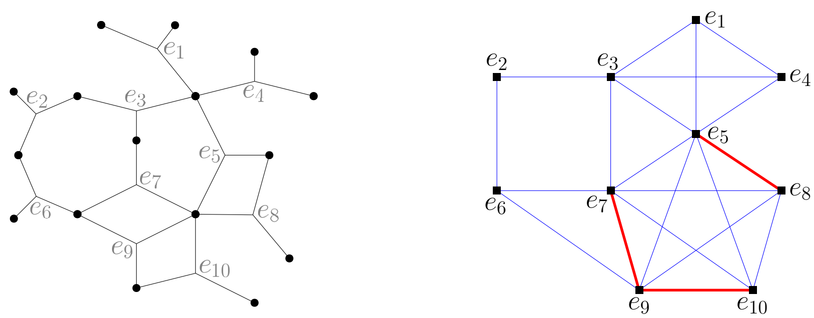

Figure 5: A 3-uniform hypergraph and its bicolored line graph. Red edges are represented in bold.

If G is the bicolored line graph of some k-uniform hypergraph $\mathcal{H}$, then there exists a blue induced path between two vertices of G if and only if there exists a $(k-2)$-linear path in $\mathcal{H}$ between the corresponding (hyper)edges. By Theorem 4.2, PAFP is thus tractable on any class of bicolored graphs G that are bicolored line graphs for which we can compute a pre-image in polynomial time.

References

[1] E. Jackowska. The 3-color Ramsey number for a 3-uniform loose path of length 3. *Australas. J. Comb.*, 63 (2), 2015, pp.314-320.

[2] H. N. Gabow, S. N. Maheswari, L. J. Osterweil. On two problems in the generation of program test paths. *IEEE Trans. Softw. Eng.*, 2 (3), 1976, pp.227–231.

[3] H. Yinnone. On paths avoiding forbidden pairs of vertices in a graph. *Discrete Appl. Math.*, 74 (1), 1997, pp.85-92.