Information Processing Letters 111 (2011) 231–234

Contents lists available at ScienceDirect

Information Processing Letters

www.elsevier.com/locate/ipl

Computing simple-path convex hulls in hypergraphs

Francesco M. Malvestuto *, Mauro Mezzini, Marina Moscarini

Computer Science Dept., Sapienza University of Rome, Italy

ARTICLE INFO

Article history:
Received 12 May 2010
Received in revised form 30 November 2010
Accepted 30 November 2010
Available online 3 December 2010
Communicated by J. Torán

Keywords:
Graph algorithms
Simple-path convexity
Simple-path convex hull
Totally balanced hypergraphs

ABSTRACT

In a connected hypergraph a vertex set X is *simple-path convex* (*sp-convex*, for short) if either $|X| \leqslant 1$ or X contains every vertex on every simple path between two vertices in X (Faber and Jamison, 1986 [7]), and the *sp-convex hull* of a vertex set X is the minimal superset of X that is *sp*-convex. In this paper, we give a polynomial algorithm to compute *sp*-convex hulls in an arbitrary hypergraph.

1. Introduction

While several convexity notions exist for graphs (e.g., *g*-convexity [7], *m*-convexity [5,7], *ap*-convexity [4], *tp*-convexity [3], Steiner convexity [2,10]), fewer convexity notions have been defined explicitly for hypergraphs. The first hypergraph convexity that has been introduced is *simple-path convexity* (*sp-convexity*, for short) [7], which is a generalization of *ap*-convexity. Recently [8], *m*-convexity has been generalized to hypergraphs and another hypergraph convexity, which is stronger than *m*-convexity and is called *c-convexity*, has been introduced; moreover, efficient algorithms to compute *m*-convex and *c*-convex hulls have been given [8]. On the other hand, no result on the complexity of the problem of computing the *sp*-convex hull of a vertex set exists except for the case that the family of *sp*-convex sets is a convex geometry, in which case an efficient algorithm can be easily derived from well-known properties of totally balanced hypergraphs [1,7]. In this paper we state a characterization of *sp*-convex sets, which leads to solve the *sp*-convex hull problem in an arbitrary hypergraph in $O(n^3ms)$ time where n is the number of its vertices, m is the number of its edges and s is the sum of the cardinalities of its edges.

The rest of the paper is organized as follows. Section 2 contains basic notions on hypergraphs and simple-path convexity. In Section 3 we present an *sp*-convex hull algorithm for totally balanced hypergraphs. In Section 4 we first state a characterization of *sp*-convex sets in an arbitrary hypergraph and, then, give our *sp*-convex hull algorithm.

2. Definitions

In this section we recall some hypergraph-theoretic definitions from [6].

A *hypergraph* is a (possibly empty) set $\boldsymbol{H}$ of nonempty sets; the elements of $\boldsymbol{H}$ are called the (*hyper*)*edges* of $\boldsymbol{H}$ and their union the *vertex set* of $\boldsymbol{H}$, denoted by $V(\boldsymbol{H})$. The *degree* of a vertex of $\boldsymbol{H}$ is the number of edges containing it.

A hypergraph is *trivial* if it has only one edge, and *nontrivial* otherwise. A *partial hypergraph* of hypergraph $\boldsymbol{H}$ is a nonempty subset of $\boldsymbol{H}$.

The *subhypergraph* of $\boldsymbol{H}$ *induced* by a nonempty subset X of $V(\boldsymbol{H})$ is the hypergraph $\{A \cap X\colon\ A \in \boldsymbol{H}\} \backslash \{\varnothing\}$.

A *path* between two vertices a and b of $\boldsymbol{H}$ is a sequence $\pi = (a_0, A_1, a_1, \ldots, A_k, a_k), k \geqslant 0,$ where $a_0 = a,\ a_k = b,$ and if $k \geqslant 1$ the a_i's are pairwise distinct vertices of $\boldsymbol{H}$, the

* Corresponding author.
E-mail address: malvestuto@di.uniroma1.it (F.M. Malvestuto).

0020-0190/$ – see front matter
doi:10.1016/j.ipl.2010.11.026

A_i's are pairwise distinct edges of $\boldsymbol{H}$, and $\{a_{i-1}, a_i\} \subseteq A_i$ for $1 \leqslant i \leqslant k$; by $V(\pi)$ and $\boldsymbol{H}(\pi)$ we denote the set of vertices and edges on the path π, respectively, that is, $V(\pi) = \{a_0, a_1, \ldots, a_k\}$ and $\boldsymbol{H}(\pi) = \{A_1, \ldots, A_k\}$. If $\boldsymbol{H}$ is a graph (i.e., every edge has cardinality less than 3), then path $\pi = (a_0, A_1, a_1, \ldots, A_k, a_k)$ will be written simply as $(a_0, a_1, \ldots, a_k)$ and is *chordless* if no two non-consecutive vertices are adjacent in $\boldsymbol{H}$.

Two vertices a and b of a hypergraph are *connected* if there exists a path between a and b. A hypergraph is *connected* if every two vertices are connected. The *connected components* of a hypergraph are its maximal connected partial hypergraphs.

A path π in $\boldsymbol{H}$ is *simple* [7] if $|A \cap V(\pi)| = 2$ for each edge A of $\boldsymbol{H}(\pi)$. Note that in a graph every path is simple.

Remark 1. Let $\pi = (a_0, A_1, a_1, \ldots, A_k, a_k)$ be a path between a and b in $\boldsymbol{H}$. Let $i(1) = \max\{h \leqslant k\colon\ a_0 \in A_h\}$. Then, $\pi_1 = (a_0, A_{i(1)}, a_{i(1)}, \ldots, A_k, a_k)$ is a path between a and b in $\boldsymbol{H}$. If $i(1) = k$ then π_1 is a simple path between a and b in $\boldsymbol{H}$. Otherwise, let $i(2) = \max\{i(1) \leqslant h \leqslant k\colon\ a_{i(1)} \in A_h\}$. Then, $\pi_2 = (a_0, A_{i(1)}, a_{i(1)}, A_{i(2)}, a_{i(2)}, \ldots, A_k, a_k)$ is a path between a and b in $\boldsymbol{H}$. If $i(2) = k$ then π_2 is a simple path between a and b in $\boldsymbol{H}$. And so on. Thus, we can construct a simple path between a and b in $\boldsymbol{H}$.

Remark 2. Let $\pi = (a_0, A_1, a_1, \ldots, A_k, a_k)$ be a simple path in $\boldsymbol{H}$. If $\boldsymbol{H}(\pi)$ contains a vertex c that is not in $V(\pi)$ and has degree 2 or more, then c is on the simple path $\pi' = (a_0, A_1, a_1, \ldots, A_{i'}, c, A_{i''}, a_{i''}, \ldots, A_k, a_k)$ where $i' = \min\{h \leqslant k\colon\ c \in A_h\}$ and $i'' = \max\{h \leqslant k\colon\ c \in A_h\}$.

A *simple circuit* [7] is a sequence $(a_0, A_1, a_1, \ldots, A_{k-1}, a_{k-1}, A_k, a_0)$, $k \geqslant 2$, where $(a_0, A_1, a_1, \ldots, A_{k-1}, a_{k-1})$ is a simple path and $A_k \cap \{a_0, a_1, \ldots, a_{k-1}\} = \{a_0, a_{k-1}\}$; the *length* of the simple circuit is the number k of its edges. A hypergraph $\boldsymbol{H}$ is *totally balanced* if $\boldsymbol{H}$ contains no simple circuit of length greater than 2.

A vertex of a hypergraph is a *nest vertex* [7] (corresponding to a *simple row* [1] of the vertex-edge incidence matrix of $\boldsymbol{H}$) if the edges containing it form a nested (that is, totally ordered with respect to set-inclusion) family of sets. A hypergraph is totally balanced if and only if every induced subhypergraph of $\boldsymbol{H}$ has a nest vertex [1,7]. Based on this characterization of totally balanced hypergraphs, Anstee and Farber [1] gave a recognition algorithm for totally balanced hypergraphs, which runs in $O(n^2m)$ time if the input hypergraph has n vertices and m edges and consists in recursively deleting nest vertices.

Let $\boldsymbol{H}$ be a connected hypergraph. The *sp-interval* between two vertices a and b of $\boldsymbol{H}$ is the set $I(a, b)$ which consists of every vertex on any simple path between a and b. A subset X of $V(\boldsymbol{H})$ is *sp-convex* if either X is empty or X contains $I(a, b)$ for every two vertices in X. The *sp-convex hull* of a subset X of $V(\boldsymbol{H})$ is the minimal superset of X that is *sp*-convex.

Let X be an *sp*-convex set of $\boldsymbol{H}$. A vertex v in X is an *extreme point* of X if the set $X\backslash\{v\}$ is *sp*-convex. The family of *sp*-convex sets of $\boldsymbol{H}$ is a *convex geometry* if every *sp*-convex set equals the *sp*-convex hull of the set of its extreme points. In [7] it was proven that this is the case if and only if $\boldsymbol{H}$ is totally balanced.

3. Background

A brute-force method for constructing the *sp*-convex hull of a vertex set $X \subseteq V(\boldsymbol{H})$ begins by setting $Y := X$; then, till we can no longer enlarge Y, we repeatedly add to Y the set $I(a, b)$ for every two vertices a and b in Y. Unfortunately, this procedure is not efficient because, for a given value of Y it is NP-hard to compute $I(a, b)$ for two given vertices a and b in Y. To see it, let $G(\boldsymbol{H})$ be the bipartite graph with bipartition $(V(\boldsymbol{H}), \boldsymbol{H})$ where there is an arc (a, A), $a \in V(\boldsymbol{H})$ and $A \in \boldsymbol{H}$, if and only if $a \in A$. For convenience, we call the elements of $V(\boldsymbol{H})$ and $\boldsymbol{H}$ the *vertex-nodes* and *edge-nodes* of $G(\boldsymbol{H})$, respectively. Note that a path in $\boldsymbol{H}$ is simple if and only if it is a chordless path in $G(\boldsymbol{H})$, that is, no edge-node on the path is adjacent to three vertex-nodes on the path. As proven in [9], given three vertices a, b and c of a bipartite graph it is NP-complete to decide whether or not c is on a chordless path between a and b. In other words, it is NP-complete to decide whether or not c belongs to $I(a, b)$.

In the special case that $\boldsymbol{H}$ is a totally balanced hypergraph (in which case the family of *sp*-convex sets of $\boldsymbol{H}$ is a convex geometry), the following result easily entails the problem of computing *sp*-convex hulls is polynomial.

Proposition 1. *(See Corollary 5.8 in [7].) Let $\boldsymbol{H}$ be a totally balanced and connected hypergraph. A subset X of $V(\boldsymbol{H})$, is sp-convex if and only if there is an ordering $a_1, a_2, \ldots, a_m$ of the vertices in $V(\boldsymbol{H})\backslash X$ such that, for all $i = 1, \ldots, m$, a_i is a nest vertex of the subhypergraph of $\boldsymbol{H}$ induced by $X \cup \{a_i, a_{i+1}, \ldots, a_m\}$.*

Corollary 1. *Let $\boldsymbol{H}$ be a totally balanced and connected hypergraph with n vertices and m edges, and let X be a subset of $V(\boldsymbol{H})$. The sp-convex hull of X can be constructed in $O(n^2m)$ time.*

Proof. By Proposition 1, the *sp*-convex hull of X can be obtained by repeatedly deleting the nest vertices of $\boldsymbol{H}$ that do not belong to X. Therefore, the *sp*-convex hull problem reduces to a selective deletion of nest vertices of $\boldsymbol{H}$, which can be done in $O(n^2m)$ time using the above-mentioned Anstee–Farber algorithm. □

4. Computing *sp*-convex hulls

In this section we shall state a characterization of *sp*-convex sets which leads to a polynomial algorithm for finding the *sp*-convex hull of a given vertex set in an arbitrary hypergraph. To achieve this, we need the following definition.

Let X be a subset of $V(H)$. Two edges A and B of $\boldsymbol{H}$ are *connected outside X* (*X-connected*, for short), written $A \equiv_X B$, if

$A = B$ or

$(A \cap B)\backslash X \neq \varnothing$ or there exists an edge C of $\boldsymbol{H}$ such that

$$(A \cap C)\backslash X \neq \varnothing \quad \text{and} \quad C \equiv_X B.$$

The edge relation $\equiv_X$ is an equivalence relation; the classes of the resultant partition of $\boldsymbol{H}$ will be referred to as the *X-connected components* of $\boldsymbol{H}$, and $\boldsymbol{H}$ is *X-connected* if it has exactly one X-connected component. For an X-connected component $\boldsymbol{C}$ of $\boldsymbol{H}$, we call the set $X \cap V(\boldsymbol{C})$ the *boundary* of $\boldsymbol{C}$. In what follows, given two distinct vertices a and b in $X \cap V(\boldsymbol{C})$, by $\boldsymbol{C}_{a,b}$ we denote the hypergraph obtained from $\boldsymbol{C}$ by deleting the vertices in $X\backslash\{a,b\}$ and the edges that contain both a and b. Note that $\boldsymbol{C}_{a,b}$ need not contain a (or b) (see the example below).

Theorem 1. *A vertex set X is sp-convex if and only if either $|X| \leqslant 1$ or, for every nontrivial X-connected component $\boldsymbol{C}$ of $\boldsymbol{H}$ with $|X \cap V(\boldsymbol{C})| > 1$ and for every two distinct vertices a and b in the boundary of $\boldsymbol{C}$, there exists no path between a and b in $\boldsymbol{C}_{a,b}$.*

Proof. (only if) Assume that X is *sp*-convex. Let $\boldsymbol{C}$ be any nontrivial X-connected component of $\boldsymbol{H}$ with $|X \cap V(\boldsymbol{C})| > 1$, and let a and b be two distinct vertices in the boundary of $\boldsymbol{C}$. If a or b is not a vertex of $\boldsymbol{C}_{a,b}$ then trivially there exists no path between a and b in $\boldsymbol{C}_{a,b}$. Assume that both a and b are vertices of $\boldsymbol{C}_{a,b}$. By construction of $\boldsymbol{C}_{a,b}$, a and b are not adjacent in $\boldsymbol{C}_{a,b}$. Moreover, if a and b were connected in $\boldsymbol{C}_{a,b}$, then by Remark 1 there would exist a simple path $\pi_{a,b} = (a_0, B_1, a_1, \ldots, B_k, a_k)$, $k \geqslant 2$, between a and b in $\boldsymbol{C}_{a,b}$. Therefore, there would exist a simple path $\pi = (a_0, A_1, a_1, \ldots, A_k, a_k)$ between a and b in $\boldsymbol{H}$ where A_h is an edge of $\boldsymbol{C}$ being the disjoint union of B_h with some subset of $X\backslash\{a,b\}$, for all h. But then one would have $V(\pi)\backslash X \neq \varnothing$ which contradicts the hypothesis that X is *sp*-convex.

(if) Assume that, for every nontrivial X-connected component $\boldsymbol{C}$ of $\boldsymbol{H}$ with $|X \cap V(\boldsymbol{C})| > 1$ and for every two distinct vertices a and b in the boundary of $\boldsymbol{C}$, there exists no path between a and b in $\boldsymbol{C}_{a,b}$. Suppose by contradiction that X is not *sp*-convex. Then, there would exist a simple path π between two vertices a and b in X in $\boldsymbol{H}$ such that $V(\pi)\backslash X \neq \varnothing$. Let c be a vertex on π that does not belong to X, let u be the last vertex on π that is in X and precedes c in π and let v be the first vertex on π that is in X and follows c. Then u, v and c are vertices of some nontrivial X-connected component $\boldsymbol{C}$ of $\boldsymbol{H}$; furthermore, u and v belong to the boundary of $\boldsymbol{C}$ and are connected in $\boldsymbol{C}_{u,v}$, which contradicts the hypothesis. □

Example. Let $\boldsymbol{H} = \{A_1, A_2, A_3, A_4, A_5\}$ where $A_1 = \{1,2\}$, $A_2 = \{1,2,3\}$, $A_3 = \{3,4\}$, $A_4 = \{3,4,5\}$. The hypergraph $\boldsymbol{H}$ is shown in Fig. 1.

Let $X = \{1,3,4\}$. The X-components of $\boldsymbol{H}$ are shown in Fig. 2 and $\boldsymbol{C}$ is the only the X-component of $\boldsymbol{H}$ that is not a trivial hypergraph. The boundary of $\boldsymbol{C}$ is {1, 3}. The hypergraph $\boldsymbol{C}_{13}$ is shown in Fig. 3.

Since 3 is not a vertex of $\boldsymbol{C}_{13}$, there exists no path joining 1 and 3 in $\boldsymbol{C}_{13}$. By Theorem 1 the set X is *sp*-convex, which is confirmed by the fact that the only simple paths joining two vertices in X are: $(1, A_2, 3), (1, A_2, 3, A_3, 4)$, $(1, A_2, 3, A_4, 4), (3, A_3, 4), (3, A_4, 4)$.

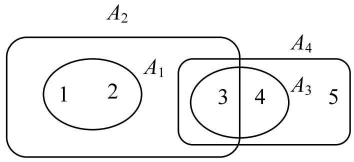

Fig. 1.

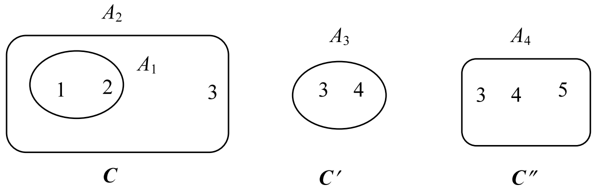

Fig. 2.

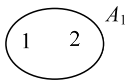

Fig. 3.

Using Theorem 1 we easily obtain a polynomial algorithm for computing the *sp*-convex hull of a given vertex set X. However, we can speed up the construction of the *sp*-convex hull of X using Remark 2. Suppose that $\boldsymbol{C}$ is a nontrivial X-connected component of $\boldsymbol{H}$ and $\pi = (a_0, A_1, a_1, A_2, \ldots, A_k, a_k)$ is a simple path between two distinct vertices a and b in the boundary of $\boldsymbol{C}$ and assume that $\boldsymbol{C}_{a,b}(\pi)$ contains a vertex c of degree 2 or more which is not in X. From Remark 2 we know that another simple path between a and b in $\boldsymbol{C}_{a,b}$ is given by $\pi' = (a_0, A_1, a_1, \ldots, A_{i'}, c, A_{i''}, a_{i''}, \ldots, A_k, a_k)$ where $i' = \min\{h \leqslant k\colon\ c \in A_h\}$ and $h'' = \max\{h \leqslant k\colon\ c \in A_h\}$. Thus, we obtain the following algorithm.

SPCH algorithm
Input: a connected hypergraph $\boldsymbol{H}$ and a subset X of $V(\boldsymbol{H})$.
Output: the *sp*-convex hull of X in the variable Y.
begin
$Y := \varnothing$;
$Z := X$;
while $Y \neq Z$ **do**
  begin
  $Y := Z$;
  for every nontrivial Y-connected component $\boldsymbol{C}$ of $\boldsymbol{H}$ **do**
    for every two distinct vertices a and b in the boundary of $\boldsymbol{C}$ that are connected in $\boldsymbol{C}_{a,b}$ **do**
      begin
      find a simple path π between a and b in $\boldsymbol{C}_{a,b}$;
      add to Z the vertices of $\boldsymbol{C}_{a,b}(\pi)$ with degree 2 or more
      end
  end
end

We will evaluate the complexity of the SPCH algorithm in terms of the number n of vertices of $\boldsymbol{H}$, of the number m of edges of $\boldsymbol{H}$ and of the size $s = \sum_{A \in \boldsymbol{H}} |A|$ of $\boldsymbol{H}$.

We make use of the bipartite graph $G(\boldsymbol{H})$ to represent $\boldsymbol{H}$. Thus, $G(\boldsymbol{H})$ is connected and has $m+n$ nodes and s arcs.

For a given value of Y, we mark the vertex-nodes of $G(\boldsymbol{H})$ that belong to Y. Then, we can construct the Y-connected components of $\boldsymbol{H}$ with their boundaries in $O(s)$ time and their number is $O(m)$. For a given Y-connected component $\boldsymbol{C}$ of $\boldsymbol{H}$ there exist $O(n^2)$ pair of vertices in the boundary of $\boldsymbol{C}$. Let $\{a, b\}$ be a pair of vertices in the boundary of $\boldsymbol{C}$. In the bipartite graph $G(\boldsymbol{C})$ we unmark a and b and we mark the edge-nodes adjacent to both a and b. Thus, we can construct $G(\boldsymbol{C}_{a,b})$ by ignoring the marked nodes of $G(\boldsymbol{C})$ and, if a and b are connected in $G(\boldsymbol{C}_{a,b})$, in $O(s)$ time we can construct a shortest path π between a and b in $G(\boldsymbol{C}_{a,b})$ and find the set Y' of vertex-nodes of $G(\boldsymbol{C}_{a,b}(\pi))$ with degree 2 or more. Note that π is also a chordless path in $G(\boldsymbol{C}_{a,b})$ and, hence, a simple path between a and b in $\boldsymbol{C}_{a,b}$. Finally, we can add Y' to Z in $O(n)$ time. Therefore, since $n < s$, processing a given value of Y requires $O(n^2ms)$ time. Since Y can assume $O(n)$ distinct values the complexity of the SPCH algorithm is $O(n^3ms)$.

References

[1] R.P. Anstee, M. Farber, Characterization of totally balanced matrices, J. Algorithms 5 (1984) 215–230.

[2] J. Cáceres, A. Márquez, M.L. Puertas, Steiner distance and convexity in graphs, European J. Combin. 29 (2008) 726–736.

[3] M. Changat, J. Mathew, On triangle path convexity, Discrete Math. 206 (1999) 91–95.

[4] M. Changat, S. Klavzar, H.H. Mulder, The all paths transit function on a graph, Czechoslovak Math. J. 126 (2001) 439–448.

[5] P. Duchet, Convex sets in graphs II: Minimal path convexity, J. Combin. Theory Ser. B 44 (1988) 307–316.

[6] P. Duchet, Hypergraphs, in: R.L. Graham, M. Grötschel, L. Lovász (Eds.), Handbook of Combinatorics, vol. I, North-Holland, 1995, pp. 381–432.

[7] M. Farber, R.E. Jamison, Convexity in graphs and hypergraphs, SIAM J. Algebraic and Discrete Methods 7 (1986) 433–444.

[8] F.M. Malvestuto, Canonical and monophonic convexities in hypergraphs, Discrete Math. 309 (2009) 4287–4298.

[9] M. Mezzini, On the complexity of finding chordless paths in bipartite graphs and some interval operators in graphs and hypergraphs, Theoret. Comput. Sci. 411 (2010) 1212–1220.

[10] M.H. Nelsen, O.R. Oellermann, Steiner tree and convex geometries, SIAM J. Discrete Math. 23 (2009) 680–693.