Research Article

On Clustering Detection Based on a Quadratic Program in Hypergraphs

Qingsong Tang [ID]

College of Sciences, Northeastern University, Shenyang 110819, China

Correspondence should be addressed to Qingsong Tang; tangqs@mail.neu.edu.cn

A proper cluster is usually defined as maximally coherent groups from a set of objects using pairwise or more complicated similarities. In general hypergraphs, clustering problem refers to extraction of subhypergraphs with a higher internal density, for instance, maximal cliques in hypergraphs. The determination of clustering structure within hypergraphs is a significant problem in the area of data mining. Various works of detecting clusters on graphs and uniform hypergraphs have been published in the past decades. Recently, it has been shown that the maximum $\{1, 2\}$-clique size in $\{1, 2\}$-hypergraphs is related to the global maxima of a certain quadratic program based on the structure of the given nonuniform hypergraphs. In this paper, we first extend this result to relate strict local maxima of this program to certain maximal cliques including 2-cliques or $\{1, 2\}$-cliques. We also explore the connection between edge-weighted clusters and strictly local optimum solutions of a class of polynomials resulting from nonuniform $\{1, 2\}$-hypergraphs.

1. Introduction

Many important phenomena depend on the structures of graphs or hypergraphs, for example, the spread of disease in a society, image segmentation problems in image analysis, or feature extraction in networks. To understand hypergraph structure, we often start with the study of the subhypergraphs with denser relations inside and sparser connections to other subhypergraphs. Thus, detecting such hypergraph clusters of closely related objects remains one of the most interesting problems in the field of bioinformatics, social society, and data mining. Clustering is a process of partitioning a set of objects into meaningful subsets so that all objects in the same group are similar and objects in different groups are dissimilar. It is a method of data exploration and a way of looking for patterns or structure in the data that are of interest. The majority of approaches to clusters available in the literature assume that objects similarities are expressed as pairwise relations in networks in terms of 2-graphs. There is also study of pairwise clustering to edge-weighted and vertex-weighted graphs (see [1–6], respectively). For uniform hypergraphs, there are various works on

clustering with applications in different aspects, such as face clustering, perceptual grouping, and parametric motion segmentation, as well as image categorization using high order relations, since approximation of more complicated similarities in terms of pairwise interaction can lead to substantial loss of information (see [7–14]). In real-world cases, similarities in a group of objects may be more appropriate to be modeled in nonuniform weighted edges in general hypergraphs. As an illustration, think of a society of people with different income levels. It makes perfect sense to define similarity measures over one person and two persons that indicate how close they are. To be specific, two persons knowing each other would get pairwise weight 1 and weight 0 otherwise; this pairwise relationship can be modeled by the well-known adjacent matrix in this society; further, for a person labeled i with income bigger than certain amount, say m, we would assign this person weight 1 on a single edge $\{i\}$; for income less than this amount, we would assign this person weight 0 on the single edge $\{i\}$; this situation involving a subset in the society may be denoted by a vector $\vec{B}$. Naturally, the internal coherency of a cluster can be represented by a maximum optimization problem based on the

society as follows: $\max\left\{\overrightarrow{x}^T A \overrightarrow{x} + B\overrightarrow{x}\right\}$. This optimization is the same as the graph-Lagrangian formed from a nonuniform $\{1,2\}$-hypergraph that models the relationship in this society (the detailed definition is given in the next section). Therefore, it is interesting to detect different types of clusters, say $\{1,2\}$-cliques or 2-cliques. Clearly, this example can be generalized to any model fitting problem, where the deviation of a set of points from the model provides a measure of their dissimilarity. The problem of data clustering using more comprehensive dissimilarity (uniform or nonuniform) is usually referred to as hypergraph clustering, since we can represent any instance of this problem by means of a hypergraph, where vertices are the objects to be clustered and the (edge-weighted) hyperedges (uniform or nonuniform) encode different order similarities.

In 1965, Motzkin and Straus provided a solution to the maximum value of a class of homogeneous quadratic multilinear functions on n variable over the standard simplex of the n-dimension Euclidean space, where the homogeneous quadratic multilinear function is associated with the edge set of a graph with n vertices. Motzkin–Straus' result established a connection between the order of a maximum complete subgraph and the graph-Lagrangian of a graph. This result also provided a new proof of a theorem by Turán who pushed the development of the study of extremal problems in graph theory. In [15], Motzkin and Straus' result was extended to characterization of local maxima in simple graphs. For Motzkin and Straus' type result in nonuniform hypergraphs, recently, in [16], it has been shown that the global maxima of a certain quadratic program are related to the maximum $\{1,2\}$-clique size in $\{1,2\}$-hypergraphs.

In this paper, we extend uniform hypergraph clustering result to nonuniform hypergraphs and provide a solution to the maximum value of a class of nonhomogeneous multilinear functions in n variables over the standard simplex of the n-dimension Euclidean space. Specifically, we first extend this result to relate strict local maxima of this program to certain maximal cliques (either 2-cliques or $\{1,2\}$-cliques or both). We also explore the connection between edge-weighted clusters and strictly local optimum solutions of a class of polynomials in the given hypergraphs. Nonhomogeneous multilinear functions discussed in this paper are associated with nonuniform hypergraphs.

This paper is organized as follows. In Section 2, we give the brief introduction to main concepts, terminology, and related results. In Section 3, we list some useful lemmas. In Section 4, we present the characterization of certain maximal cliques (either 2-cliques or $\{1,2\}$-cliques or both) in terms of strictly local optimum solutions of a class of polynomials formed from unweighted $\{1,2\}$-graphs. In Section 5, we discuss the parametrization graph-Lagrangian and cliques in $\{1,2\}$-graphs. In Section 6, we extend the result in Section 4 to edge-weighted $\{1,2\}$-graphs in some way. Conclusions are given in Section 7.

2. Definitions and Related Results

A hypergraph $H = (V, E)$ consists of a vertex set V and an edge set E, where every edge in E is a subset of V. The set $T(H) = \{|F|: F \in E\}$ is called the set of edge types of H. We also say that H is a $T(H)$-graph. For example, if $T(H) = \{1,2\}$, then we say that H is a $\{1,2\}$-graph. If all edges have the same cardinality r, then H is an r-uniform hypergraph. A 2-uniform graph is called a graph. A hypergraph is nonuniform if it has at least two edge types. For any $r \in T(H)$, the rth-level hypergraph H^r is the hypergraph consisting of all edges with r vertices of H and $E(H^r)$ denotes the edge set of H^r. We write H_n^T for a hypergraph H on n vertices with $T(H) = T$. Given a subset $U \subseteq V(H)$, the induced subgraph denoted by $H[U]$ is a hypergraph on U with the edge set $\{F \in E(H): F \subseteq U\}$. An edge $\{i_1, i_2, \ldots, i_r\}$ in a hypergraph is simply written as $i_1 i_2 \cdots i_r$ throughout the paper.

For a positive integer n, let $[n]$ denote the set $\{1, 2, \ldots, n\}$. For a finite set V and a positive integer i, let $\binom{V}{i}$ denote the family of all i-subsets of V. The complete hypergraph K_n^T is a hypergraph on n vertices with edge set $\bigcup_{i \in T} \binom{[n]}{i}$. For example, $K_n^{\{r\}}$ is the complete r-uniform hypergraph on n vertices. $K_n^{[r]}$ is the nonuniform hypergraph with all possible edges of cardinality at most r. The complete graph on n vertices $K_n^{\{2\}}$ is also called a clique. We also let $[k]^{\{r\}}$ represent the complete r-uniform hypergraph on vertex set $[k]$.

For a T-graph $H = (V, E)$, for $r \in T$, we denote the $(r-1)$-neighborhood of a vertex $i \in V$ by $E_i^r = \left\{ A \in \binom{V - \{i\}}{r-1} : A \cup \{i\} \in E \right\}$. Similarly, we will denote the $(r-2)$-neighborhood of a pair of vertices $i, j \in V$ by $E_{ij}^r = \left\{ B \in \binom{V - \{i,j\}}{r-2} : B \cup \{i,j\} \in E \right\}$. We denote the complement of E_i^r by $\overline{E_i^r} = \left\{ A \in \binom{V - \{i\}}{r-1} : A \cup \{i\} \in \binom{V}{r} \backslash E \right\}$. Denote $E_{i\backslash j}^r = E_i^r \cap \overline{E_j^r}$.

Definition 1. For an r-uniform graph H with the vertex set $[n]$, edge set $E(H)$, and a vector $\overrightarrow{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, we associate a homogeneous polynomial in n variables, denoted by $\lambda(H, \overrightarrow{x})$, as follows: $\lambda(H, \overrightarrow{x}) := \sum_{i_1 i_2 \cdots i_r \in E(H)} x_{i_1} x_{i_2} \ldots x_{i_r}$. Let $\Delta := \left\{ \overrightarrow{x} = (x_1, x_2, \ldots, x_n) : \sum_{i=1}^n x_i = 1, x_i \geq 0 \text{ for } i = 1, 2, \ldots, n \right\}$. The graph-Lagrangian of H, denoted by $\lambda(H)$, is the maximum of the above homogeneous multilinear polynomial of degree r over the standard simplex S. Precisely,

$$\lambda(H) := \max\{\lambda(H, \overrightarrow{x}): \overrightarrow{x} \in \Delta\}. \tag{1}$$

The value x_i is called the weight of the vertex i. A vector $\vec{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ is called feasible weighting for H if and only if $\vec{x} \in S$. A vector $\vec{y} \in S$ is called optimal weighting for H if and only if $\lambda(H, \vec{y}) = \lambda(H)$.

Remark 1. $\lambda(H)$ was called Lagrangian of H in literature [17–20]. The terminology "graph-Lagrangian" was suggested by Franco Giannessi.

The characteristic vector of a set $C \subseteq V$, denoted by $\vec{x}^C = (x_1^C, x_2^C, \ldots, x_n^C)$, is the vector in S defined as

$$x_i^C = \frac{1_{i \in C}}{|C|} \tag{2}$$

where $|C|$ denotes the cardinality of C and 1_P is the indicator function returning 1 if property P is satisfied and 0 otherwise.

In [21], Motzkin and Straus provided the following simple expression for the graph-Lagrangian of a 2-graph.

Theorem 1 (see Theorem 1 in [21]). *If H is a 2-graph with n vertices in which a largest clique has order t then $\lambda(H) = \lambda(K_t^{\{2\}}) = (1/2)(1 - (1/t))$. Furthermore, the characteristic vector of a maximum clique of H is optimal weighting for H.*

This result provided a solution to the optimization problem of this type of quadratic functions over the standard simplex of an Euclidean space.

In [16], Peng et al. generalized the concept of graph-Lagrangian to nonuniform hypergraphs as given below.

Definition 2. For a hypergraph H_n^T with $T(H) = T$ and a vector $\vec{x} = (x_1, \ldots, x_n) \in R^n$, define

$$\lambda'\left(H_n^T, \vec{x}\right) = \sum_{r \in T}\left(r! \sum_{i_1 i_2 \cdots i_r \in E(H^r)} x_{i_1} x_{i_2} \cdots x_{i_r}\right). \tag{3}$$

Let $\Delta = \{\vec{x} = (x_1, x_2, \ldots, x_n): \sum_{i=1}^n x_i = 1, x_i \geq 0 \text{ for } i = 1, 2, \ldots, n\}$. The Lagrangian of H_n^T, denoted by $\lambda'(H_n^T)$, is defined as

$$\lambda'\left(H_n^T\right) = \max\{\lambda'\left(H_n^T, \vec{x}\right): \vec{x} \in \Delta\}. \tag{4}$$

The value x_i is called the weight of vertex i. A vector $\vec{y} \in S$ is called optimal weighting for H if $\lambda'(H, \vec{y}) = \lambda'(H)$.

In [22], Peng and Yao gave a generalization of Motzkin–Straus result to {1, 2}-graphs.

Theorem 2 (see Theorem 1.4 in [22]). *If H is a {1, 2}-graph with n vertices and the order of its maximum complete {1, 2}-subgraph is t (where $t \geq 2$), then $\lambda'(H) = \lambda'(K_t^{\{2\}}) = 2 - (1/t)$. Furthermore, the characteristic vector of a maximum clique of H is optimal weighting for H.*

In [23, 24], Gu et al. and Tang et al. obtained more Motzkin–Straus results to some uniform hypergraphs.

3. Graph-Lagrangians and Cliques in {1, 2}-Graphs: Unweighted Case

There is a 1-to-1 connection between strictly local optimum and maximal cliques (2-cliques or {1, 2}-cliques) in {1, 2}-graphs.

Theorem 3. *A subset C of vertices in H is a maximum clique of a {1, 2}-hypergraph H if and only if its characteristic vector $\vec{x}^C$ is a global maximum of optimization problem (4).*

Proof. One direction is immediate from Theorem 2. For the other direction, suppose that $\vec{x}^C$ is a global maximum of optimization problem (4). From (4), $\lambda'(H, \vec{x}^C) = \lambda'(H, \vec{x}^U)$, where U is a maximum clique in H. Let $|C| = c$ and $|S| = s$. From Theorem 2, $\lambda'(H, \vec{x}^C) \leq 2 - (1/c)$ and $\lambda'(x^U) = 2 - (1/u)$. However, $\lambda'(H, \vec{x}^C) = \lambda'(H, \vec{x}^U)$ only if $2 - (1/u) = 2 - (1/c)$. This implies $c = s$. So C must be a clique from Theorem 2 and C is a maximum clique. $\qquad\square$

Proposition 1. *Let C be a subset of k vertices of a {1, 2}-hypergraph H. Then C is a maximal clique of H if and only if its characteristic vector $\vec{x}^C$ satisfies*

$$\frac{\partial \lambda'(H, \vec{x})}{\partial x_j} = \begin{cases} = \dfrac{3k - 2}{k}, & \text{if } j \in C, \\[2mm] \leq \dfrac{3k - 2}{k}, & \text{if } j \notin C. \end{cases} \tag{5}$$

Proof. Suppose that C is a maximal clique. From the definition of maximal clique

$$\frac{\partial \lambda'\left(H, \vec{x}^C\right)}{\partial x_j} = 1 + 2 \sum_{ij \in E^2(C)} x_i^C = 1 + 2\frac{k - 1}{k}$$
$$= \frac{3k - 2}{k}, \tag{6}$$

for all $j \in C$; and

$$\frac{\partial \lambda'\left(H, \vec{x}^C\right)}{\partial x_j} = 2 \sum_{ij \in E^2(C)} x_i x_j \leq \frac{3k - 2}{k} \tag{7}$$

for all $j \notin C$. Hence $\vec{x}^C$ satisfies (5).

For the other direction, if C is not a clique, then, for some vertex $j \in C$, there exists a vertex $i \in C$ satisfying $i \neq j$ and $ij \notin E^2(C)$ or $i \notin E^1(C)$. Hence,

$$\frac{\partial \lambda'\left(H, \vec{x}^C\right)}{\partial x_j} < 1 + 2 \sum_{ij \in E^2(C)} x_i^C = \frac{3k - 2}{k}. \tag{8}$$

This contradicts $\partial\lambda'(H,\overrightarrow{x}^C)/\partial x_j = (3k-2)/k$ for all $j \in C$. So C must be a clique. If C is not a maximal clique, then there must exist a clique C'^C. Let $j \in C' \smallsetminus C$. Then

$$\frac{\partial\lambda'\left(H,\overrightarrow{x}^C\right)}{\partial x_j} = 1 + 2 \sum_{ij \in E^2(C)} x_i^C = 3 > \frac{3k-2}{k}. \quad (9)$$

This contradicts $\partial\lambda'(H,\overrightarrow{x}^C)/\partial x_j \leq (3k-2)/k$ for $j \notin C$. $\quad\square$

Lemma 1 (KKT necessary condition, [25]). *If feasible weighting $\overrightarrow{x} = (x_1, x_2, \ldots, x_n)$ is a local solution of optimization problem (4), then there exists $\theta \in \mathbb{R}$ such that, for all $j \in [n]$,*

$$\frac{\partial\lambda'(H,\overrightarrow{x})}{\partial x_i} = \begin{cases} = \theta, & \text{if } j \in \sigma(\overrightarrow{x}), \\ \leq \theta, & \text{if } j \notin \sigma(\overrightarrow{x}). \end{cases} \quad (10)$$

The following corollary follows from Proposition 1 immediately.

Corollary 1. *If C is a maximal clique of a $\{1,2\}$-hypergraph H, then x^C satisfies the first-order KKT necessarily.*

Definition 3. For a $\{1,2\}$-hypergraph H, a maximal clique C is said to be strictly maximal if, for all $i \in (V\backslash C) \cap E^1$, the number of 2-edges crossing i and C is less than $|C| - 1$.

Note that, for C to be a maximal clique, it suffices that the number of edges crossing i and C be no more than $|C| - 1$ for all $i \in (V\backslash C) \cap E^1$. However, for C to be a strictly maximal clique, this number needs to be strictly less than $|C| - 1$.

Lemma 2. *Consider a $\{1,2\}$-hypergraph H which contains two cliques C and D of equal cardinality $|C| = |D| = k$. Let $m = |C\backslash D| = |D\backslash C| = m \leq k$. Then, for every $\alpha_1, \alpha_2 \geq 0$ satisfying $\alpha_1 + \alpha_2 = 1$, we have the following:*

(a) If H has exactly $m(m-1)$ edges crossing $C\backslash D$ and $D\backslash C$, then $\lambda'(H, \alpha_1\overrightarrow{x}^C + \alpha_1\overrightarrow{x}^D) = \lambda'(H, \overrightarrow{x}^C)$

(b) If H has fewer than $m(m-1)$ edges crossing $C\backslash D$ and $D\backslash C$, then $\lambda'(H, \alpha_1\overrightarrow{x}^C + \alpha_1\overrightarrow{x}^D) < \lambda'(H, \overrightarrow{x}^C)$

The proof of this lemma is similar to the proof of Theorem 6 in [15]. So we omit the details here.

Lemma 3. *Let $\overrightarrow{x}$ be a strict local maximum of optimization problem (4); then, $\forall i, j \in \sigma(\overrightarrow{x})$, there exists an edge $e \in E(H)$ such that $\{i, j\} \subseteq e$.*

Proof. Suppose, for a contradiction, that there exist i and j in $\sigma(\overrightarrow{x})$ such that $\{i, j\} \not\subseteq e$ for any $e \in E(H)$. We define new weighting $\overrightarrow{y}$ for H as follows. Let $\delta < x_j$ be an arbitrarily small positive constant. Let $y_l = x_l$ for $l \neq i, j$, $y_i = x_i + \delta$, and $y_j = x_j - \delta$; then $\overrightarrow{y}$ is clearly legal weighting for H, and

$$\lambda'(H,\overrightarrow{y}) - \lambda'(H,\overrightarrow{x}) = x_j\left(\frac{\partial\lambda'(H,\overrightarrow{x})}{\partial x_i} - \frac{\partial\lambda'(H,\overrightarrow{x})}{\partial x_j}\right)$$

$$- x_j^2\frac{\partial^2\lambda'(H,\overrightarrow{x})}{\partial x_i \partial x_j} = 0. \quad (11)$$

This contradicts $\overrightarrow{x}$ being a strict local maximum of optimization problem (4). Hence Lemma 3 holds.

Now we are ready to prove the main result of this section. $\quad\square$

Theorem 4. *Let C be a subset of k vertices of a $\{1,2\}$-hypergraph H.*

(a) If $E^1 \cap C \neq \varphi$, then C is a strict maximal $\{1,2\}$-clique of H if and only if $\overrightarrow{x}^C$ is a strict local maximum of optimization problem (4)

(b) If $E^1 \cap C = \varphi$, then C is a strict maximal $\{2\}$-clique of H if and only if $\overrightarrow{x}^C$ is a strict local maximum of optimization problem (4)

Proof. (a) Suppose that $\overrightarrow{x}^C$ is a strict local maximum of optimization problem (4); then the KKT conditions (10) hold for some θ. We will show that $\theta = ((3k-2)/k)$, where $k = |C|$. Then, by Proposition 1, C is a maximal clique. Suppose that $\theta \neq ((3k-2)/k)$ for a contradiction. For every two vertices i, j in C, $ij \in E^2$ by Lemma 3. We will show that all the vertices in C are contained in E^1. Then C is a clique and $\theta \neq ((3k-2)/k)$. Assume that there exist some vertices in C not contained in E^1. Since $E^1 \cap C \neq \varphi$, there must exist a vertex $i \in C$ contained in E^1. Assume that $j \in C$ but it is not contained in E^1; then

$$\frac{\partial\lambda'(H,\overrightarrow{x})}{\partial x_i} = 1 + 2 \cdot \frac{k-1}{k},$$

$$\frac{\partial\lambda'(H,\overrightarrow{x})}{\partial x_j} = 2 \cdot \frac{k-1}{k}. \quad (12)$$

This contradicts $(\partial\lambda'(H,\overrightarrow{x})/\partial x_i) = (\partial\lambda'(H,\overrightarrow{x})/\partial x_j)$ by Lemma 1 Hence, all the vertices in C are contained in E^1.

To see C as a strictly maximal clique, suppose to the contrary that $j \in E^1\backslash C$ adjacent to exactly $k-1$ vertex in C, and let i denote the only vertex in C not adjacent to j. Then set $D = j \cup (C\backslash\{i\})$ as a clique of the same cardinality as C. Because $m = |C\backslash D| = |D\backslash C| = 1$, there are no edges crossing CD and $D\backslash C$, since i, j are nonadjacent. From Lemma 2, for all $\alpha \in [0, 1]$, we have $\lambda'(H, \overrightarrow{x}^C) = \lambda'(H, \alpha_1\overrightarrow{x}^C + \alpha_1\overrightarrow{x}^D)$ which contradicts the hypothesis that $\overrightarrow{x}^C$ is a strict maximum of optimization problem (4). This proves the first part of the theorem.

For the other part, suppose that C is a strictly maximal clique. To prove that x^C is a strict local maximum of optimization problem (4), we apply the second-order sufficiency conditions for constrained optimization. First, from

Corollary 1, x^C satisfies the KKT conditions. Note that, in this case, the Lagrange multipliers μ_i's are given by

$$\mu_i = \frac{3k-2}{k} - 1_{i \in E^1} - \sum_{ij \in E^2} x_j. \quad (13)$$

It remains to be shown that the Hessian of the Lagrangian associated with the optimization in (4) is negative definite on the subspace

$$M = \left\{ \overrightarrow{y} \in R^n, \sum_{i=1}^n y_i = 0 \text{ and } y_i = 0, \text{for all } i \in I \right\} \quad (14)$$

where $I = \{i: x_i^c = 0 \text{ and } \mu_i > 0\}$. Since C is a strict maximal clique, we have

$$\sum_{ij \in E^2} x_j < \frac{3k-2}{k}, \quad (15)$$

for all $i \notin C$. Now, let $\overrightarrow{y} \in M \setminus \{0\}$; then

$$\lambda'(H, \overrightarrow{y}) = \sum_{i \in E^1} y_j + \sum_{ij \in E^2} y_i y_j = \sum_{i \in C} y_i \left(\sum_{j \in C} y_j - y_i \right) \quad (16)$$
$$= -\sum_{i \in C} y_i^2 < 0.$$

This completes the proof. (b) This is similar to that in (a). We omit the details here. □

4. The Parametrization Lagrangian and Cliques in $\{1, 2\}$-Graphs

Definition 4. For a hypergraph H_n^T with H_n^T and a vector $\overrightarrow{x} = (x_1, \dots, x_n) \in R^n$, define

$$\lambda'\left(H_n^T, \overrightarrow{x}\right) = \sum_{r \in T} \left(r! \sum_{i_1 i_2 \cdots i_r \in E(H^r)} x_{i_1} x_{i_2} \dots x_{i_r} \right). \quad (17)$$

Let $PS = \{\overrightarrow{x} = (x_1, x_2, \dots, x_n): \sum_{i=1}^n x_i = 1, \sum_{i=1}^n x_i^2 = 1/s, x_i \geq 0 \text{ for } i = 1, 2, \dots, n\}$, where s is a real number between $[1, n]$. The parametrization Lagrangian of H_n^T, denoted by $\lambda_s'(H_n^T)$, is defined as

$$\lambda_s'\left(H_n^T\right) := \max\left\{\lambda'\left(H_n^T, \overrightarrow{x}\right): \overrightarrow{x} \in PS\right\}. \quad (18)$$

The value x_i is called the weight of vertex i. A vector $\overrightarrow{y} \in PS$ is called optimal weighting for H if $\lambda_s'(H, \overrightarrow{y}) = \lambda_s'(H)$.
Let

$$X(s, C) := \{\overrightarrow{x} | \overrightarrow{x} \in PS \text{ and } x_u = 0 \forall u \notin C\}. \quad (19)$$

The following is easy to see.

Lemma 4. *The set $X(s, C)$ is nonempty if and only if $1 \leq s \leq |C|$. For $s = |C|$, $X(s, C)$ consists simply of the character vector of C.*

Theorem 5. *Let H be a $\{1, 2\}$-hypergraph with clique number ω. Then*

(a) $\lambda_s'(H) = 2 - (1/s)$ for $1 \leq s \leq \omega$

(b) $\lambda_s'(H) \leq 2 - (1/\omega) < 2 - (1/s)$ for $\omega \leq s \leq n$

(c) $\lambda_s'(H) = 2 - (1/s)$ if and only if $s \leq \omega$

(d) $\lambda'(H) = \max_{1 \leq s \leq n} \lambda_s'(H) = 2 - (1/\omega)$

(e) For $s \leq \omega$, the set of global optimal solutions of (18) is given by

$$\cup \{(Xs, C) | C \text{ is a clique}, s \leq | \quad (20)$$

(f) The set of global optimal solutions of $\lambda_\omega'(H)$ is $(1_C/\omega)$, where C is an (optimal) $\{1, 2\}$-clique of order ω. Hence, there is a one-to-one correspondence between the global optimal solutions of $\lambda_\omega'(H)$ and the optimal cliques in H.

Proof. Proof of (a). Let $s \leq \omega$. Let $\overrightarrow{x}$ be any feasible solution of $\lambda_s'(H)$. Then

$$\lambda_s'(H) = \sum_{i \in E^1} x_i + 2 \sum_{ij \in E^2} x_i x_j$$

$$\leq 1 + (x_1 + x_2 + \cdots + x_n)^2 - \sum_{i=1}^n x_i^2 - 2 \sum_{ij \in \overline{E}^2} x_i x_j$$

$$\leq 2 - \frac{1}{s}. \quad (21)$$

On the other side, let C be a clique of size s in H for some $S \geq s$ (note that s is not assumed to be integral). Since $\omega \geq s$, $X(s, C)$ is not empty. Now, for an arbitrary $x \in X(s, C)$,

$$\lambda_s'(H) = \sum_{i \in C} x_i + 2 \sum_{ij \in C} x_i x_j$$

$$1 + \left(\sum_{i \in C} x_i \right)^2 - \sum_{i \in C} x_i^2 \quad (22)$$

$$2 - \frac{1}{s}.$$

Hence, $\lambda_s'(H) = 2 - (1/s)$ for $1 \leq s \leq \omega$.

Proof of (b). Let $s > \omega$. Since any $\overrightarrow{x}$ that is feasible for $\lambda_s'(H)$ is also feasible for $\lambda'(H)$, and since $\lambda'(H) = 2 - 1/\omega$, we have $\lambda_s'(H) < \lambda'(H) = 2 - 1/\omega$.

Proof of (c). Since $2 - (1/s)$ is an increasing function of s, (a) and (b) together imply (c).

Proof of (d). Note that the feasible region of $\lambda'(H)$ is the feasible region of $\lambda_s'(H)$ for s in the range $[1, n]$. Combining with (a) and (b) implies (d).

Proof of (e). By equation (22), every $x \in X(s, C)$, where C is a clique of size at least s, satisfies

$$\lambda_s'(H) = 2 - \frac{1}{s}. \quad (23)$$

On the other side, for an arbitrary x of $L(s)$, we have

$$\lambda'_s(H) = 2 - \frac{1}{s} - \sum_{i \in \overline{E}^1} x_i - 2 \sum_{ij \in \overline{E}^2} x_i x_j. \quad (24)$$

Hence, if $\lambda'_s(H) = 2 - 1/s$, then $\sum_{i \in \overline{E}^1} x_i + \sum_{ij \in \overline{E}^2} x_i x_j = 0$. This happens if and only if

$$\begin{aligned} x_i &= 0, \text{whenever } i \in \overline{E}^1, \\ x_i x_j &= 0, \text{whenever } ij \in \overline{E}^2. \end{aligned} \quad (25)$$

So the support $\sigma(\overrightarrow{x})$ of $\overrightarrow{x}$ forms a clique in H. Let C be this clique. Clearly, $\overrightarrow{x} \in X(x, C)$. Lemma 4 implies that $s \le |C|$.

5. Maximum Vertex-Weighted Cliques in $\{1, 2\}$-Graphs

Given a nonnegative weight vector $\overrightarrow{w}_v$, for any subset C of the vertex set, $w_v(C)$ denotes the sum of the weights of vertex in C. The vertex-weighted clique number $\omega(\overrightarrow{w}_v, H)$ is the maximum of $w_v(C)$ over all cliques C of G. Note that $\omega(e, H)$ is the usual clique number $\omega(H)$ of the hypergraph. Given a positive weight vector $\overrightarrow{w}_v$, define a set of matrices as follows:

$$\mathcal{M}(\overrightarrow{w}_v, H) = \left\{ B | B_{ii} = \frac{1}{w_{v_i}} \forall i, B_{ij} + B_{ji} \ge \frac{1}{w_{v_i}} + \frac{1}{w_{v_j}} \forall ij \in \overline{E}^2, B_{ij} = 0 \forall ij \in E^2 \right\}. \quad (26)$$

For a given a matrix $B \in \mathcal{M}(\overrightarrow{w}_v, H)$, consider the following optimization problem:

$$L(H, \overrightarrow{w}_v) = \min \left\{ -\sum_{i \in E^1} \frac{x_i^2}{w_{v_i}} + 2\overrightarrow{x}^T B \overrightarrow{x} \mid \overrightarrow{x} \in S \right\}. \quad (27)$$

Theorem 6. *Let H be a $\{1, 2\}$-graph. Then $\omega(\overrightarrow{w}_v, H) = (1/L(H, \overrightarrow{w}_v))$ for any positive weight vector $\overrightarrow{w}$, and $B \in \mathcal{M}(\overrightarrow{w}, H)$.*

In the Proof of Theorem 6, we will impose an additional condition on a solution $\overrightarrow{x} = (x_1, x_2, \ldots, x_n)$ to a global optimum $\overrightarrow{x}$ to problem (27): $(*) |\{i : x_i > 0\}|$ is minimal; that is, if $\overrightarrow{y}$ is a feasible solution for H satisfying $|i : y_i > 0| < |i : x_i > 0|$, then $L(H, \overrightarrow{y}) < L(H)$. We need the following lemmas.

Lemma 5. *Let $\overrightarrow{x}$ be a global optimum of optimization problem (27) with minimum support; then there exists an edge $e \in E(H)$ such that $\{i, j\} \subseteq e \forall i, j \in \sigma(\overrightarrow{x})$.*

Proof. Let $\overrightarrow{x}$ be a global optimum of optimization problem (27) with minimum support. Let $f(\overrightarrow{x}) = -\sum_{i \in E^1} x_i^2 / w_{v_i} + 2\overrightarrow{x}^T B \overrightarrow{x}$. Suppose, for a contradiction, that there exist i and j in $\sigma(\overrightarrow{x})$ such that $\{i, j\} \not\subseteq e$ for any $e \in E(H)$. We define a new feasible solution $\overrightarrow{y}$ to (27) as follows. Let $y_l = x_l$ for $l \ne i, j$, $y_i = x_i + x_j$, and $y_j = x_j - x_j = 0$; then $\overrightarrow{y}$ is clearly a feasible solution (27) with smaller support compared to $\overrightarrow{x}$. By KKT necessary condition $(\partial f(\overrightarrow{x}) / \partial x_i) = (\partial f(\overrightarrow{x}) / \partial x_j)$, and

Proof of (f). The result follows from (e) and Lemma 4. □

$$\begin{aligned} f(\overrightarrow{y}) - f(\overrightarrow{x}) &= x_j \left(\frac{\partial f(\overrightarrow{x})}{\partial x_i} - \frac{\partial f(\overrightarrow{x})}{\partial x_j} \right) \\ &\quad + 2x_j^2 \left(B_{ii} + B_{jj} - 1_{i \in E^1} \frac{1}{w_{v_i}} - 1_{j \in E^1} \frac{1}{w_{v_j}} - B_{ij} - B_{ji} \right) \\ &= 2x_j^2 \left(B_{ii} + B_{jj} - 1_{i \in E^1} \frac{1}{w_{v_i}} - 1_{j \in E^1} \frac{1}{w_{v_j}} - B_{ij} - B_{ji} \right) \\ &\le 0, \end{aligned} \quad (28)$$

since $B_{ij} + B_{ji} \ge (1/w_{v_i}) + (1/w_{v_j}) \ge B_{ii} + B_{jj} - 1_{i \in E^1} (1/w_{v_i}) - 1_{j \in E^1} (1/w_{v_i})$. This contradicts $\overrightarrow{x}$ being a global optimum of optimization problem (27) with minimum support. □

Claim 1. Either $i \in E^1$ for all $i \in \sigma(\overrightarrow{x})$ or $i \notin E^1$ for all $i \in \sigma(\overrightarrow{x})$.

Proof. Suppose that $i \in E^1$ but $j \notin E^1$ for a contradiction. By the KKT condition, $(\partial f(\overrightarrow{x}) / \partial x_i) = (\partial f(\overrightarrow{x}) / \partial x_j)$. By Lemma 5, $\forall i, j \in \sigma(\overrightarrow{x}), ij \in E^2$; therefore $0 = (2x_j / w_{v_i})$. This is a contradiction.

Now we are ready to prove Theorem 6. □

Proof. of Theorem 6. Let $\overrightarrow{x}$ be a global optimum of optimization problem (27) with minimum support. By Lemma 5 and Claim 1, $\sigma(\overrightarrow{x})$ induces $\{1, 2\}$-clique or a 2-clique of H.

If $\sigma(\overrightarrow{x})$ induces a $\{1, 2\}$-clique, then

$$f(\overrightarrow{x}) = -\sum_{i \in E^1} \frac{x_i^2}{w_{v_i}} + 2\overrightarrow{x}^T B \overrightarrow{x} = \sum_{i \in \sigma(\overrightarrow{x})} \frac{x_i^2}{w_{v_i}}, \qquad (29)$$

and its minimum over the simplex is at

$$x_i = \begin{cases} & \text{if } i \in \sigma(\overrightarrow{x}), \\ 0, & \text{if } i \notin \sigma(\overrightarrow{x}). \end{cases} \qquad (30)$$

for $i = 1, 2, \ldots, n$. So the optimal value of $f(\overrightarrow{x})$ is $1/w_v(\sigma(\overrightarrow{x}))$. For the solution to be global optimal, $w_v(\sigma(\overrightarrow{x}))$ must be the maximum $\{1, 2\}$-clique in H.

If $\sigma(\overrightarrow{x})$ induces a 2-clique, then

$$f(\overrightarrow{x}) = -\sum_{i \in E^1} \frac{x_i^2}{w_{v_i}} + 2\overrightarrow{x}^T B \overrightarrow{x} = 2 \sum_{i \in \sigma(\overrightarrow{x})} \frac{x_i^2}{w_{v_i}}, \qquad (31)$$

and the left is similar to the case where $\sigma(\overrightarrow{x})$ induces $\{1, 2\}$-clique. □

6. Dominant Set for $\{1, 2\}$-Graphs

Let $G = (V, E)$ be an edge-weighted graph with edge weight $w_{ij} > 0$ for $ij \in E$. The weighted adjacency matrix $A = (a_{ij})_n$ is defined as $a_{ij} = w_{ij}$ if $ij \in E$ and $a_{ij} = 0$ otherwise. The average weighted degree of i with regard to S is defined as

$$\text{awdeg}_S(i) = \frac{1}{|S|} \sum_{j \in S} a_{ij}. \qquad (32)$$

If $j \notin S$, define

$$\varphi_S(i, j) = a_{ij} - \text{awdeg}_S(i). \qquad (33)$$

The weight of i with regard to S is defined as

$$W_S(i) = \begin{cases} 1, & \text{if } |S| = 1, \\ \sum_{j \in S\{i\}} \varphi_{S\{i\}}(j, i) w_{S\{i\}}(j), & \text{otherwise}, \end{cases} \qquad (34)$$

$$W(S) = \sum_{i \in S} w_S(i).$$

Definition 5 (see [4]). A nonempty subset of vertices $S \subseteq V$ such that $W(T) > 0$ for any nonempty $T \subseteq S$ is said to be dominant if

(1) $W_S(i) > 0$ for all $i \in S$

(2) $W_{S*\{i\}}(i) < 0$ for all $i \notin S$

Set the weighted characteristic vector $\overrightarrow{x}^S \in \triangle$ as follows:

$$x_i^S = \begin{cases} \dfrac{w_S(i)}{W(S)}, & \text{if } i \in S, \\ 0, & \text{otherwise}. \end{cases} \qquad (35)$$

Pavan and Pelillo connect the dominant set to the following quadratic program:

$$\text{maximize} \, \overrightarrow{x}^T A \overrightarrow{x} \qquad (36)$$
$$\text{subject to } \in \triangle,$$

and establish a correspondence between the global (local) maxima of (36) and the dominant sets of a graph.

Theorem 7 (see [4]). *If S is a dominant subset of vertices, then its weighted characteristic vector $\overrightarrow{x}^S$ is a strict local solution of program (36). Conversely, if $\overrightarrow{x}^*$ is a strict local solution of program (36), then its support $\sigma = \sigma(\overrightarrow{x}^*)$ is a dominant set, provided that $w_{\sigma \cup \{i\}}(i) \neq 0$ for all $i \notin \sigma$.*

Here we consider the quadratic program related to edge-weighted $\{1, 2\}$-graph. Let H be a $\{1, 2\}$-graph on vertex set $[n]$ with edge sets $E^1 \cup E^2$. Let $\overrightarrow{u} \geq 0$ be the edge weight vector of H^1 and let $A \geq 0$ be the edge weight matrix of H^2.

Let $p_i = u_i$ if $\{i\} \in E^1$ and $p_i = 0$ if $\{i\} \notin E^1$; let $q_{ij} = a_{ij} - p_i$ if $ij \in E^2$ and $q_{ij} = -p_i$ if $ij \notin E^2$; that is, $a_{ij} = q_{ij} + p_i$. Consider the following quadratic program:

$$\text{maximize} f(\overrightarrow{x}) = \overrightarrow{p}^T \overrightarrow{x} + \overrightarrow{x}^T Q \overrightarrow{x} \qquad (37)$$
$$\text{subject to } \in \triangle,$$

where $\overrightarrow{p} = (p_1, \ldots, p_n)^T$ and $Q = (q_{ij})_n$. A vector $\overrightarrow{x} \in \triangle$ satisfies the Karush-Kuhn-Tucker (KKT) conditions for problem (37), that is, the first-order necessary conditions for local optimality, if there exist $n + 1$ real constants (Lagrange multipliers) $\mu_1, \ldots, \mu_n$ and λ with $\mu_i \geq 0$ for all $i = 1, \ldots, n$, such that

$$p_i + (Q\overrightarrow{x})_i + \theta + \mu_i = 0, \qquad (38)$$

for all $i = 1, \ldots, n$, $\sum_{i=1}^n x_i = 1$, and $\sum_{i=1}^n x_i \mu_i = 0$. Note that $\sum_{i=1}^n x_i = 1$ and $a_{ij} = Q_{ij} + p_i$. Equality (38) is equivalent to

$$(A\overrightarrow{x})_i + \theta + \mu_i = 0, \qquad (39)$$

for all $i = 1, \ldots, n$, $\sum_{i=1}^n x_i = 1$, and $\sum_{i=1}^n x_i \mu_i = 0$. So, if we define the dominant set of H as the dominant set of H^2, then, by Theorem 7, we have the following.

Theorem 8. *If S is a dominant subset of vertices, then its weighted characteristic vector $\overrightarrow{x}^S$ is a strict local solution of program (37). Conversely, if $\overrightarrow{x}^*$ is a strict local solution of program (37), then its support $\sigma = \sigma(\overrightarrow{x}^*)$ is a dominant set, provided that $w_{\sigma \cup \{i\}}(i) \neq 0$ for all $i \notin \sigma$.*

7. Conclusion

In this paper, we study the connection between the local maxima of a class of quadratic program and certain maximal cliques including 2-cliques or $\{1, 2\}$-cliques of $\{1, 2\}$-hypergraphs. We also explore the connection between edge-weighted clusters and strictly local optimum solutions of a class of polynomials resulting from nonuniform $\{1, 2\}$-hypergraphs. In the future, we will try to extend these results to general hypergraphs.

Data Availability

No data were used to support this study.

Conflicts of Interest

The author declares no conflicts of interest.

Acknowledgments

References

[1] I. M. Bomze, "Evolution towards the maximum clique," *Journal of Global Optimization*, vol. 10, no. 2, pp. 143–164, 1997.

[2] M. Budinich, "Exact bounds on the order of the maximum clique of a graph," *Discrete Applied Mathematics*, vol. 127, no. 3, pp. 535–543, 2003.

[3] S. Busygin, "A new trust region technique for the maximum weight clique problem," *Discrete Applied Mathematics*, vol. 154, no. 15, pp. 2080–2096, 2006.

[4] M. Pavan and M. Pelillo, "Dominant sets and pairwise clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 1, pp. 167–172, 2007.

[5] J. Hou and M. Pelillo, "A simple feature combination method based on dominant sets," *Pattern Recognition*, vol. 46, no. 11, pp. 3129–3139, 2013.

[6] L. E. Gibbons, D. W. Hearn, P. M. Pardalos, and M. V. Ramana, "Continuous characterizations of the maximum clique problem," *Mathematics of Operations Research*, vol. 22, no. 3, pp. 754–768, 1997.

[7] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognition Letters*, vol. 31, no. 8, pp. 651–666, 2010.

[8] S. Agarwal, J. Lim, L. Zelnik-Manor, P. Perona, D. Kriegman, and S. Belongie, "Beyond pairwise clustering," in *Proceedings of the IEEE Conference Computer Vision and Pattern Recognition*, vol. 2, pp. 838–845, San Diego, CA, USA, June 2005.

[9] V. M. Govindu, "Tensor decomposition for geometric grouping and segmentation," in *Proceedings of the IEEE Conference Computer Vision and Pattern Recognition*, pp. 1150–1157, San Diego, CA, USA, June 2005.

[10] A. Shashua, R. Zass, and T. Hazan, "Multi-way clustering using super-symmetric non-negative tensor factorization," *Computer Vision-ECCV 2006*, vol. 3954, pp. 595–608, 2006.

[11] Y. Huang, Q. Liu, F. Lv, Y. Gong, and D. N. Metaxas, "Unsupervised image categorization by hypergraph partition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 6, pp. 1266–1273, 2011.

[12] S. Rota Bulo and M. Pelillo, "A game-theoretic approach to hypergraph clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 6, pp. 1312–1327, 2013.

[13] H. Liu, L. J. Latecki, and S. Yan, "Dense subgraph partition of positive hypergraphs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 3, pp. 541–554, 2015.

[14] H. Liu, L. J. Latecki, and S. Yan, "Robust clustering as ensembles of affinity relations," in *Proceedings of the 24th Annual Conference on Neural Information Processing Systems 2010*, Vancouver, CanadaAdvances in Neural Information Processing Systems 23, Vancouver, Canada, December 2010.

[15] M. Pelillo and A. Jagota, "Feasible and infeasible maxima in a quadratic program for maximum clique," *Journal of Artificial Neural Networks*, vol. 2, pp. 411–420, 1995.

[16] Y. J. Peng, H. Peng, Q. S. Tang, and C. Zhao, "An extension of Motzkin-Straus theorem to non-uniform hypergraphs and its applications," *Discrete Applied Mathematics*, vol. 200, pp. 170–175, 2015.

[17] P. Frankl and Z. Füredi, "Extremal problems whose solutions are the blowups of the small witt-designs," *Journal of Combinatorial Theory-Series A*, vol. 52, no. 1, pp. 129–147, 1989.

[18] Y. Peng and C. Zhao, "A Motzkin-Straus type result for 3-uniform hypergraphs," *Graphs and Combinatorics*, vol. 29, no. 3, pp. 681–694, 2013.

[19] P. Frankl and V. Rödl, "Hypergraphs do not jump," *Combinatorica*, vol. 4, no. 2-3, pp. 149–159, 1984.

[20] J. M. Talbot, "Lagrangians of hypergraphs," *Combinatorics, Probability & Computing*, vol. 11, no. 2, pp. 199–216, 2002.

[21] T. S. Motzkin and E. G. Straus, "Maxima for graphs and a new proof of a theorem of Turán," *Canadian Journal of Mathematics*, vol. 17, pp. 533–540, 1965.

[22] Y. J. Peng and Y. P. Yao, "On Motzkin-Straus type of results and Frankl-Fúedi conjecture for hypergraphs," 2013, https://arxiv.org/abs/1312.3034.

[23] R. Gu, X. Li, Y. Peng, and Y. Shi, "Some Motzkin-Straus type results for non-uniform hypergraphs," *Journal of Combinatorial Optimization*, vol. 31, no. 1, pp. 223–238, 2016.

[24] Q. Tang, Y. Peng, X. Zhang, and C. Zhao, "On Motzkin-Straus type results for non-uniform hypergraphs," *Journal of Combinatorial Optimization*, vol. 34, no. 2, pp. 504–521, 2016.

[25] D. G. Luenberger and Y. Ye, *Linear and Nonlinear Programming*, Springer Science Business Media, LLC, Berlin, Germany, 3rd edition, 2008.