ORIGINAL PAPER

A generalization of the Motzkin–Straus theorem to hypergraphs

Samuel Rota Bulò · Marcello Pelillo

Abstract   In 1965, Motzkin and Straus established a remarkable connection between the global maxima of the Lagrangian of a graph G over the standard simplex and the clique number of G. In this paper, we provide a generalization of the Motzkin–Straus theorem to k-uniform hypergraphs (k-graphs). Specifically, given a k-graph G, we exhibit a family of (parameterized) homogeneous polynomials whose local (global) minimizers are shown to be in one-to-one correspondence with maximal (maximum) cliques of G.

1 Introduction

A *k-uniform hypergraph*, or simply a *k-graph*, is a pair $G = (V, E)$, where $V = \{1, \ldots, n\}$ is a finite set of *vertices* and $E \subseteq \binom{V}{k}$ is a set of k-subsets of V, each of which is called a *hyperedge*. 2-graphs are typically called *graphs*. The *complement* of a k-graph G is given by $\overline{G} = (V, \overline{E})$ where $\overline{E} = \binom{V}{k} \setminus E$. A subset of vertices $C \subseteq V$ is called a *hyperclique* if $\binom{C}{k} \subseteq E$. To improve readability, in the sequel we will drop the prefix "hyper" when referring to edges and cliques of a k-graph. A clique is said to be *maximal* if it is not contained in any other clique, while it is called *maximum* if it has maximum cardinality. The *clique number* of a k-graph G, denoted by $\omega(G)$, is defined as the cardinality of a maximum clique.

S. Rota Bulò (✉) · M. Pelillo
Dipartimento di Informatica, Università Ca' Foscari di Venezia,
via Torino 155, 30172 Venezia Mestre, Italy
e-mail: srotabul@dsi.unive.it

M. Pelillo
e-mail: pelillo@dsi.unive.it

Given a k-graph G with n vertices, the *Lagrangian* of G is the following homogeneous multilinear polynomial in n variables:

$$L_G(\mathbf{x}) = \sum_{e \in E} \prod_{i \in e} x_i \tag{1}$$

with a view to provide a new proof of Turan's theorem, in 1965 Motzkin and Straus [13] established a remarkable connection between the clique number of a graph G with n vertices and the maxima of its Lagrangian over the *standard simplex* of $\mathbb{R}^n$, which is defined as:

$$\Delta = \left\{ \mathbf{x} \in \mathbb{R}^n \ : \ \forall i = 1 \ldots n, \ x_i \geq 0 \text{ and } \sum_{i=1}^{n} x_i = 1, \right\}.$$

Theorem 1 (Motzkin–Straus) *Let G be a graph with clique number $\omega(G)$. If $\mathbf{x}^*$ is a maximizer of L_G over Δ, then*:

$$L_G(\mathbf{x}^*) = \frac{1}{2} \left[1 - \frac{1}{\omega(G)} \right].$$

The *characteristic vector* of a subset of vertices $S \subseteq V$ is the vector in Δ defined as:

$$\mathbf{x}_i^S = \frac{1_{i \in S}}{|S|},$$

where $|S|$ denotes the cardinality of S and 1_P is an indicator function returning 1 if property P is satisfied and 0 otherwise. Moreover the *support* of a vector $\mathbf{x} \in \Delta$, is the set of indices corresponding to nonnegative components of $\mathbf{x}$, i.e.

$$\sigma(\mathbf{x}) = \{i : x_i > 0, \ i = 1 \ldots n\}.$$

In the course of the proof of their result, Motzkin and Straus showed also that a subset of vertices C is a maximum clique of G if and only if its characteristic vector $\mathbf{x}^C$ is a global maximizer of L_G on Δ. However, not all maximizers of L_G over Δ are in the form of characteristic vectors and these *spurious* solutions were characterized in 1995 by Pelillo and Jagota [18]. They also provided an extension of the Motzkin–Straus result by showing a one-to-one connection between (non-spurious) *local* maximizers of L_G over Δ and *maximal* cliques of G. Later in 1997 [2], to overcome the problem of spurious solutions, Bomze introduced a regularized version of the Lagrangian defined as

$$L_G^\tau(\mathbf{x}) = L_G(\mathbf{x}) + \tau \sum_{i \in V} x_i^2$$

and proved a one-to-one correspondence between local (global) maximizers of L_G^τ over Δ and maximal (maximum) cliques of G provided that $0 < \tau < \frac{1}{2}$.

Theorem 2 (Bomze) *Let G be a graph and $0 < \tau < \frac{1}{2}$. A vector $\mathbf{x} \in \Delta$ is a local (global) maximizer of L_G^τ over Δ if and only if it is the characteristic vector of a maximal (maximum) clique of G.*

The Motzkin–Straus result and its extension were successfully employed in spectral graph theory to provide upper and lower bounds to the clique number of graphs [3, 22,23] and in optimization to provide heuristics for the maximum clique problem [2,4,8,15,17,19,20]. Furthermore, the Motzkin–Straus theorem has been generalized to vertex-weighted graphs [8] and edge-weighted graphs [16].

In this paper, we are interested in studying a generalization of the Motzkin–Straus Theorem to hypergraphs. As it turns out, L_G cannot be directly used to extend the Motzkin–Straus theorem to k-graphs. Frankl and Rödl in 1984 [7] proved that by taking a maximizer $\mathbf{x}^*$ of L_G with support as small as possible, the subhypergraph induced by S is a 2-cover, i.e. a hypergraph such that every pair of vertices is contained in some hyperedge. Since 2-covers in graphs are basically cliques, we could expect a possible generalization of the Motzkin–Straus theorem where the clique number is replaced by the size l of the maximum 2-cover in the hypergraph. However $\mathbf{x}^*$ is not necessarily in the form of a characteristic vector, and it is not in general possible to express $L_G(\mathbf{x}^*)$ as a function of l. Nevertheless, this result was used by Mubay [14] to achieve a bound for $L_G(\mathbf{x}^*)$ in terms of l on k-graphs and he used it to provide an hypergraph extension of the Turán's theorem. A further attempt to generalize the Motzkin–Straus theorem to hypergraphs is due to Sós and Straus [21]. However, their solution applies to *conformal k-graphs*, i.e. complete-subgraph graphs of ordinary graphs. This restricts the applicability of this theorem to a class of hypergraphs isomorphic to a subclass of 2-graphs having cliques of cardinality $\geq k$.

Our contribution with this paper is to provide a generalization of the Motzkin–Straus as well as Bomze's Theorems to k-graphs. Specifically, we present a continuous characterization of maximal cliques in k-graphs in terms of minimizers of a particular (parametrized) homogeneous polynomial over the standard simplex.

This shift from the combinatorial to the continuous domain may reveal unexpected properties of the original problem as well as it could serve as the basis for the development of continuous-based heuristics for the maximum clique problem on hypergraphs.

2 A continuous characterization of maximal cliques in k-graphs

Given a k-graph G, consider the following non-linear program.

$$\text{minimize} \quad h_{\overline{G}}(\mathbf{x}) = L_{\overline{G}}(\mathbf{x}) + \tau \sum_{i=1}^{n} x_i^k$$

$$\text{subject to} \quad \mathbf{x} \in \Delta, \tag{2}$$

where $\tau \in \mathbb{R}$ and $L_{\overline{G}}$ is the Lagrangian of the complement of G. In order to simplify the notation we write h instead of $h_{\overline{G}}$ where the context is non ambiguous.

A *local solution* of problem (2) is a vector $\mathbf{x} \in \Delta$ for which there exists a local neighborhood $N_\mathbf{x}$ such that $h(\mathbf{y}) \geq h(\mathbf{x})$ for all $\mathbf{y} \in N_\mathbf{x}$, while a *global solution* is

a vector $\mathbf{x} \in \Delta$ such that $h(\mathbf{y}) \geq h(\mathbf{x})$, for every $\mathbf{y} \in \Delta$. We say that $\mathbf{x}$ is a *strict* global/local solution if the inequalities are strict when $\mathbf{y} \neq \mathbf{x}$. We define the *plateau* of (2) generated by $\mathbf{x}$ as the smallest closed connected set U containing $\mathbf{x}$ such that $f(\mathbf{y}) = f(\mathbf{x})$ for all $\mathbf{y} \in U$. Note that the plateau generated by a strict local solution $\mathbf{x}$ is the singleton $\{\mathbf{x}\}$.

The Karush–Kuhn–Tucker (KKT) necessary conditions [12] for a vector $\mathbf{x}$ to be a local solution of (2) are as follows. There should exists $\lambda \in \mathbb{R}$ such that for all $j \in V$,

$$\partial_j h(\mathbf{x}) \begin{cases} = \lambda & \text{if } j \in \sigma(\mathbf{x}), \\ \geq \lambda & \text{if } j \notin \sigma(\mathbf{x}). \end{cases} \tag{3}$$

Here, $\partial_j h(\mathbf{x})$ denotes the partial derivative of h with respect to x_j, i.e.

$$\partial_j h(\mathbf{x}) = \sum_{e \in \overline{E}} 1_{j \in e} \prod_{i \in e \setminus \{j\}} x_i + \tau k x_j^{k-1},$$

and similarity $\partial_{j\ell} h(\mathbf{x})$ will denote the partial derivative with respect to x_j and x_ℓ, i.e.

$$\partial_{j\ell} h(\mathbf{x}) = 1_{j \neq \ell} \sum_{e \in \overline{E}} 1_{j,\ell \in e} \prod_{i \in e \setminus \{j,\ell\}} x_i + 1_{j=\ell} \tau k(k-1) x_j^{k-2}.$$

A sufficient condition for $\mathbf{x}$ to be a local solution of program (2) is to be a KKT point and to have the Hessian matrix of h in $\mathbf{x}$ positive definite on the subspace $M(\mathbf{x})$ defined as

$$M(\mathbf{x}) = \left\{ \boldsymbol{\varepsilon} \in \mathbb{R}^n : \sum_{i=1}^n \varepsilon_i = 0, \text{ and } \varepsilon_j = 0 \text{ for all } j \text{ such that } \partial_j h(\mathbf{x}) > \lambda \right\},$$

where the Hessian matrix of h in $\mathbf{x}$ is defined as

$$H(\mathbf{x}) = \left[\partial_{j\ell} h(\mathbf{x}) \right]_{j,\ell \in V}.$$

In other words, if $\mathbf{x}$ is a KKT point and for all $\boldsymbol{\varepsilon} \in M(\mathbf{x}) \setminus \{\mathbf{0}\}$, $\boldsymbol{\varepsilon}' H(\mathbf{x}) \boldsymbol{\varepsilon} > 0$, then $\mathbf{x}$ is a (strict) local solution of (2).

In order to grasp the intuition behind the choice of Program (2) let us investigate some elementary properties of the minimizers of the first term and the second one if considered separately. If we take any vector $\mathbf{x}$ in the simplex whose support is a clique of G, then trivially $L_{\overline{G}}(\mathbf{x})$ attains its global minimum at 0. Vice versa, for any clique C of G, every vector in the simplex with support C is a global minimizer of $L_{\overline{G}}$ over Δ. Hence, the role of the first term is to force the minimizers to have a clique as support. As for the second term, trivially, the minimizer of $\sum_{i=1}^n x_i^k$ over Δ is the simplex barycenter, i.e. the characteristic vector of V. Therefore, ideally, the role of the second term is to enforce the minimizers to have a maximal support and the form of characteristic vector. By linearly combining the two terms and by adequate choices of

τ, our aim is to achieve a continuous characterization of maximal (maximum) cliques in k-graphs which consists in a one-to-one correspondence between the set of maximal (maximum) cliques of a k-graph G, and the local (global) solutions of (2), generalizing in this way both the Motzkin–Straus and Bomze's Theorems to k-graphs.

Lemma 1 *Let G be a k-graph and let $\mathbf{x}$ be a local (global) solution of (2) with $\tau > 0$. If $C = \sigma(\mathbf{x})$ is a clique of G then it is a maximal (maximum) clique and $\mathbf{x}$ is the characteristic vector of C.*

Proof Since $\mathbf{x}$ is a local solution of (2), it satisfies the KKT conditions (3). Therefore for all $j \in C$ we have that $\lambda = \tau k x_j^{k-1} > 0$ and it follows that $\mathbf{x}$ is the characteristic vector of C. Moreover if there exists a larger clique D that contains C, then there is a vertex $j \in D \setminus C$ such that $\partial_j h(\mathbf{x}) = 0 < \lambda$. This contradicts conditions (3). Hence, C is a maximal clique of G.

Finally, $h(\mathbf{x}) = \tau |\sigma(\mathbf{x})|^{1-k}$ attains its global minimum only if $\mathbf{x}$ is the characteristic vector of a maximum clique. □

Lemma 2 *Let G be a k-graph and $\mathbf{x}$ a local (global) solution of (2). If both the following conditions hold*

1. $0 < \tau \le \frac{1}{k(k-1)}$,
2. $\sigma(\mathbf{x})$ *is minimum among all the solutions in the plateau of (2) generated by $\mathbf{x}$, if $k = 2$ and $\tau = \frac{1}{2}$,*

then $\mathbf{x}$ is the characteristic vector of a maximal (maximum) clique of G.

Proof We claim that the support of $\mathbf{x}$ is a clique of G. Otherwise, suppose that an edge $\widetilde{e} \subseteq \sigma(\mathbf{x})$ is missing. Let $j, \ell \in \widetilde{e}$ such that $x_j \le x_\ell \le \min_{i \in \widetilde{e} \setminus \{j, \ell\}} x_i$ and take $\mathbf{y} = \mathbf{x} + \varepsilon(\mathbf{e}^j - \mathbf{e}^\ell) \in \Delta$, where $\mathbf{e}^j$ denotes a zero vector except for the jth element set to 1 and where $0 < \varepsilon \le x_\ell$.

We study the sign of $h(\mathbf{y}) - h(\mathbf{x})$ in a neighborhood of $\mathbf{x}$ as $\varepsilon \to 0$ by means of the Taylor expansion of h truncated at the second-order term, where the first-order term cancels out since $\mathbf{x}$ satisfies (3), as it is a local solution of (2), and thereby $\partial_j h(\mathbf{x}) = \partial_\ell h(\mathbf{x})$:

$$
\begin{aligned}
h(\mathbf{y}) - h(\mathbf{x}) &= \frac{\varepsilon^2}{2} \left[\partial_{jj} h(\mathbf{x}) + \partial_{\ell\ell} h(\mathbf{x}) - 2\partial_{j\ell} h(\mathbf{x}) \right] + \mathcal{O}(\varepsilon^3) \\
&= \frac{\varepsilon^2}{2} \left[\tau k(k-1) \left(x_j^{k-2} + x_\ell^{k-2} \right) - 2 \sum_{e \in \overline{E}} 1_{j,\ell \in e} \prod_{i \in e \setminus \{j, \ell\}} x_i \right] + \mathcal{O}(\varepsilon^3).
\end{aligned}
$$

Let $\mu = 2 \sum_{e \in \overline{E}} 1_{j,\ell \in e} \prod_{i \in e \setminus \{j, \ell\}} x_i - (x_j^{k-2} + x_\ell^{k-2})$. Clearly $\mu \ge 0$, because at least $\widetilde{e}$ is in $\overline{E}$, and x_j and x_ℓ are the smallest components with indices in $\widetilde{e}$.

Then we can write

$$
\begin{aligned}
h(\mathbf{y}) - h(\mathbf{x}) &= \frac{\varepsilon^2}{2} \left[\tau k(k-1)(x_j^{k-2} + x_\ell^{k-2}) - \mu - (x_j^{k-2} + x_\ell^{k-2}) \right] + \mathcal{O}(\varepsilon^3) \\
&= \frac{\varepsilon^2}{2} \left\{ \left(x_j^{k-2} + x_\ell^{k-2} \right) [\tau k(k-1) - 1] - \mu \right\} + \mathcal{O}(\varepsilon^3). \quad (4)
\end{aligned}
$$

Note from (4) that the second-order term is nonpositive and becomes zero only if $\tau = \frac{1}{k(k-1)}$ and $\mu = 0$. We proceed now by distinguishing 3 cases, each of which yields a contradiction, thereby proving that $\sigma(\mathbf{x})$ is a clique of G. This in conjunction with Lemma 1 concludes the proof.

Case 1: $0 < \tau < \frac{1}{k(k-1)}$ or $\mu > 0$.

In this case, $h(\mathbf{y}) - h(\mathbf{x})$ is strictly negative for sufficiently small values of ε, contradicting the local minimality of $\mathbf{x}$.

Case 2: $\tau = \frac{1}{k(k-1)}$, $k = 2$ and $\mu = 0$.

For $k = 2$ we have trivially that $h(\mathbf{y}) - h(\mathbf{x}) = 0$ for any admissible ε. Hence, $\mathbf{y}$ is in the plateau of (2) generated by $\mathbf{x}$. By taking $\varepsilon = x_\ell$, we contradict the minimality of the support size of $\mathbf{x}$ stated in condition 2, because $\sigma(\mathbf{y}) = \sigma(\mathbf{x}) - 1$.

Case 3: $\tau = \frac{1}{k(k-1)}$, $k > 2$ and $\mu = 0$.

Note that if $\mu = 0$, then $\widetilde{e}$ is the only edge in $\overline{E}$ with vertices in $\sigma(\mathbf{x})$ that contains both j and ℓ. Moreover x_i is constant for all $i \in \widetilde{e}$. It follows that we could have arbitrarily chosen j, ℓ in $\widetilde{e}$ for the construction of $\mathbf{y}$. Hence, for every pair of vertices in $\widetilde{e}$ there exists only one edge in $\overline{E}$ with vertices in $\sigma(\mathbf{x})$ containing them, namely $\widetilde{e}$.

Let $m \in \arg\min_{i \in \widetilde{e} \setminus \{j, \ell\}} x_i$ and take $\mathbf{z} = \mathbf{x} + \varepsilon \left[(\mathbf{e}^j + \mathbf{e}^\ell)/2 - \mathbf{e}^m \right] \in \Delta$ where $0 < \varepsilon \le x_m$. We study the sign of $h(\mathbf{z}) - h(\mathbf{x})$ in a neighborhood of $\mathbf{x}$ as $\varepsilon \to 0$ by means of the Taylor expansion of h truncated at the third-order term. Here again, the first-order term cancels out since $\mathbf{x}$ satisfies (3) and therefore we obtain

$$
\begin{aligned}
&h(\mathbf{z}) - h(\mathbf{x}) \\
&= \frac{\varepsilon^2}{2} \left[\frac{\partial_{jj} h(\mathbf{x}) + \partial_{\ell\ell} h(\mathbf{x})}{4} + \partial_{mm} h(\mathbf{x}) - \partial_{jm} h(\mathbf{x}) - \partial_{\ell m} h(\mathbf{x}) + \frac{\partial_{j\ell} h(\mathbf{x})}{2} \right] \\
&\quad + \frac{\varepsilon^3}{6} \left[\frac{\partial_{jjj} h(\mathbf{x}) + \partial_{\ell\ell\ell} h(\mathbf{x})}{8} - \partial_{mmm} h(\mathbf{x}) - \frac{3}{2} \partial_{j\ell m} h(\mathbf{x}) \right] + \mathcal{O}(\varepsilon^4),
\end{aligned} \tag{5}
$$

where $\partial_{uvw} h$ denotes the partial derivative of h with respect to x_u, x_v and x_w, i.e.

$$
\partial_{uvw} h(\mathbf{x}) = 1_{u \ne v} 1_{u \ne w} 1_{v \ne w} \sum_{e \in \overline{E}} 1_{u,v,w \in e} \prod_{i \in e \setminus \{u,v,w\}} x_i + 1_{u=v=w} (k-2) x_u^{k-3}.
$$

By the observation made at the beginning of this case and by setting $\xi = x_j$, it follows that $\forall u, v \in \widetilde{e}.\partial_{uv} h(\mathbf{x}) = \xi^{k-2}$, and $\forall u \in \widetilde{e}.\partial_{uuu} h(\mathbf{x}) = (k-2)\xi^{k-3}$ and finally $\partial_{j\ell m} h(\mathbf{x}) = \xi^{k-3}$. Hence, the sign of $h(\mathbf{z}) - h(\mathbf{x})$ for sufficiently small values of ε is given by the sign of $-\frac{\varepsilon^3}{8} k \xi^{k-3}$ which is clearly negative and this contradicts the local minimality of $\mathbf{x}$.   □

An interesting observation suggested by Lemma 2 is that all minimizers of (2) are strict provided that condition 1 holds, excepting the case $k = 2$ and $\tau = \frac{1}{2}$. In fact, the existence of non strict minimizers would imply, through the lemma, the existence of non isolated characteristic vectors, which is impossible. Hence, the only case where nonstrict solutions may appear in our formulation corresponds to the maximization of the graph Lagrangian over Δ, that we have already seen to be affected by the possible

presence of spurious solutions. Indeed, Bomze's regularization coincides with choices of τ strictly less than $\frac{1}{2}$, which in fact guarantees to have only strict solutions.

The following theorem provides the generalization of the Motzkin–Straus Theorem (1) to k-graphs alluded to at the end of the previous section. Basically, we show that the global minimum of $h_{\overline{G}}$ over Δ can be expressed as a function of the clique number of G. This establishes a link between the combinatorial problem of finding the clique number of a k-graph, and the global optimization of a homogeneous polynomial over the simplex in the continuous domain.

Theorem 3 *Let G be a k-graph with clique number $\omega(G)$. Then h attains its minimum over Δ at $\tau \, \omega(G)^{1-k}$ provided that $0 < \tau \le \frac{1}{k(k-1)}$.*

Proof Let $\mathbf{x}$ be a global solution of (2) with support as small as possible. Then by Lemma 2 we have that $\mathbf{x}$ is the characteristic vector of a maximum clique of G. It follows that $h(\mathbf{x}) = \tau |\sigma(\mathbf{x})|^{1-k} = \tau \omega(G)^{1-k}$. □

Note that this result is equivalent to the original Motzkin–Straus Theorem (1) for graphs, if we take $k = 2$ and $\tau = \frac{1}{2}$. In fact, in this case we obtain

$$L_G(\mathbf{x}) = \sum_{\{i,j\}\in E} x_i x_j = \frac{1}{2} - \sum_{\{i,j\}\in \overline{E}} x_i x_j - \frac{1}{2}\sum_{i=1}^{n} x_i^2 = \frac{1}{2} - h(\mathbf{x})$$

and it follows that

$$\max_{\mathbf{x}\in\Delta} L_G(\mathbf{x}) = \frac{1}{2} - \min_{\mathbf{x}\in\Delta} h(\mathbf{x}) = \frac{1}{2} - \frac{1}{2\omega(G)} = \frac{1}{2}\left[1 - \frac{1}{\omega(G)}\right].$$

As example of interesting application of Theorem 3, we propose an upper bound to the maximum number of edges that a k-graph without p-cliques can contain. This kind of problems arise in the extremal graph theory field, and are usually called Turán problems. In particular, this problem is still open. Unfortunately, the bound that we obtain with the following result is not the best known (see [5,6]), however this example shows that the shift from the combinatorial to the continuous domain, may reveal interesting theoretical properties.

Corollary 1 *A k-graph $G = (V, E)$ of order n, which contains no p-clique with $p \ge k$, has*

$$|E| \le \binom{n}{k} - \frac{n}{k(k-1)}\left\{\left[\frac{n}{(p-1)}\right]^{k-1} - 1\right\}.$$

Proof If G contains no p-clique then for all $\mathbf{x} \in \Delta$ we have by Theorem 3 that $h(\mathbf{x}) \ge \frac{(p-1)^{1-k}}{k(k-1)}$. In particular if we take the barycenter of the simplex as $\mathbf{x}$, by trivial calculations we obtain the result. □

The next result is the converse of Lemma 2 and is instrumental to prove Theorem 4, namely our generalization of Bomze's Theorem to k-graphs.

Lemma 3 *Let G be a k-graph and* $\mathbf{x}^C$ *the characteristic vector of a maximal (maximum) clique C of G. Then* $\mathbf{x}^C$ *is a strict local (global) solution of* (2) *provided that* $0 < \tau < \frac{1}{k}$.

Proof For simplicity let $\mathbf{x} = \mathbf{x}^C$. We will show that $\mathbf{x}$ is a strict local solution of (2) by proving that it satisfies the sufficient conditions introduced at the beginning of this section. First we prove that $\mathbf{x}$ satisfies (3) and then we show that $H(\mathbf{x})$ is positive definite on the subspace $M(\mathbf{x})$.

For all $j \in \sigma(\mathbf{x})$ we have $\partial_j h(\mathbf{x}) = \tau k |C|^{1-k} = \lambda$, while for all $\ell \notin \sigma(\mathbf{x})$ we have $\partial_\ell h(\mathbf{x}) \geq |C|^{1-k} > \lambda$, since $\sigma(\mathbf{x})$ is a maximal clique and therefore at least one edge joining ℓ and $k-1$ vertices in C is missing. Hence, $\mathbf{x}$ is a KKT point.

Moreover all eigenvalues of $H(\mathbf{x})|_{\sigma(\mathbf{x})}$, i.e. the Hessian in $\mathbf{x}$ restricted to the support of $\mathbf{x}$, are positive. In fact, $H(\mathbf{x})|_{\sigma(\mathbf{x})}$ is a diagonal matrix with positive diagonal entries

$$H(\mathbf{x})|_{\sigma(\mathbf{x})} = \tau k(k-1)|C|^{2-k} I,$$

where I is the identity matrix. This implies that $H(\mathbf{x})$ is positive definite on the subspace $M(\mathbf{x})$.

Finally, $h(\mathbf{x}^C) = \tau |C|^{1-k}$ attains its global minimum where C is as large as possible, i.e. a maximum clique. $\square$

Note that $\mathbf{x}^C$ is a KKT point even for $\tau = \frac{1}{k}$. However in this case, $H(\mathbf{x})|_{\sigma(\mathbf{x})} = \alpha I$ does not imply that $H(\mathbf{x})$ is positive-definite on $M(\mathbf{x})$, in general, if $\partial_\ell h(\mathbf{x}) = \lambda$ may happen for some $\ell \in \sigma(\mathbf{x})$.

Theorem 4 *Let G be a k-graph and* $0 < \tau \leq \frac{1}{k(k-1)}$ *(with strict inequality for* $k = 2$*). A vector* $\mathbf{x} \in \Delta$ *is a local (global) solution of* (2) *if and only if it is the characteristic vector of a maximal (maximum) clique of G.*

Proof It follows from Lemmas 2 and 3. $\square$

Note that if we take $k = 2$ and $0 < \tau < \frac{1}{2}$ then local (global) minimizers of h correspond to local (global) maximizers of $L_G^{\frac{1}{2}-\tau}$. In fact

$$h(\mathbf{x}) = \sum_{\{i,j\}\in\overline{E}} x_i x_j + \tau \sum_{i=1}^n x_i^2 = \frac{1}{2} - \sum_{\{i,j\}\in E} x_i x_j + \left(\tau - \frac{1}{2}\right)\sum_{i=1}^n x_i^2$$

$$= \frac{1}{2} - \left[\sum_{\{i,j\}\in E} x_i x_j + \left(\frac{1}{2} - \tau\right)\sum_{i=1}^n x_i^2\right] = \frac{1}{2} - L_G^{\frac{1}{2}-\tau}(\mathbf{x}).$$

Since $0 < \frac{1}{2} - \tau < \frac{1}{2}$, this is an equivalent formulation of Bomze's Theorem on graphs in terms of a minimization problem.

Theorem 4 may be directly applied for devising continuous-based heuristics for the maximum clique problem on k-graphs, which is a relatively unexplored, though interesting, hard combinatorial problem. In particular, one may generalize the maximum

clique heuristic proposed by Pelillo [17], by employing the Baum–Eagon inequality [1] for the minimization of $h_{\overline{G}}(\mathbf{x})$ over Δ. Alternatively, the optimization literature offers several techniques for the global minimization of polynomials over the simplex [9–11], which can serve well our purpose.

References

1. Baum, L.E., Eagon, J.A.: An inequality with applications to statistical estimation for probabilistic functions of Markov processes and to a model for ecology. Bull. Am. Math. Soc. **73**, 360–363 (1967)
2. Bomze, I.M.: Evolution towards the maximum clique. J. Global Optimiz. **10**(2), 143–164 (1997)
3. Budinich, M.: Exact bounds on the order of the maximum clique of a graph. Discrete Appl. Math. **127**, 535–543 (2003)
4. Busygin, S.: A new trust region technique for the maximum weight clique problem. Discrete Appl. Math. **304**(4), 2080–2096 (2006)
5. Caen, D.D.: Extension of a theorem of Moon and Moser on complete hypergraphs. Ars Combin. **16**, 5–10 (1983)
6. Caen, D.D.: Linear constraints related to Turán's problem. Congressus Numerantium **39**, 291–303 (1983)
7. Frankl, P., Rödl, V.: Hypergraphs do not jump. J. Combinatorica **4**, 149–159 (1984)
8. Gibbons, L.E., Hearn, D.W., Pardalos, P.M.: A continuous based heuristic for the maximum clique problem. In: Cliques, Coloring and Satisfiability: 2nd DIMACS Impl. Chall., vol. 26, pp. 103–124 (1996)
9. Klerk, E.D.: The complexity of optimizing over a simplex, hypercube or sphere: a short survey. Central Eur. J. Oper. Res. **16**, 111–125 (2008)
10. Klerk, E.D., Laurent, M., Parrilo, P.A.: A PTAS for the minimization of polynomials of fixed degree over the simplex. Theor. Comp. Sci. **361**(2), 210–225 (2006)
11. Lasserre, J.B.: Global optimization with polynomials and the problem of moments. SIAM J. Optim. **11**(3), 796–817 (2001)
12. Luenberger, D.G.: Linear and Nonlinear Programming. Addison Wesley, Reading (1984)
13. Motzkin, T.S., Straus, E.G.: Maxima for graphs and a new proof of a theorem of Turán. Canad. J. Math. **17**, 533–540 (1965)
14. Mubay, D.: A hypergraph extension of Turán's theorem. J. Combin. Theory Ser. B **96**, 122–134 (2006)
15. Pardalos, P.M., Phillips, A.T.: A global optimization approach for solving the maximum clique problem. Int. J. Comput. Math. **33**, 209–216 (1990)
16. Pavan, M., Pelillo, M.: Dominant sets and pairwise clustering. IEEE Trans. Pattern Anal. Mach. Intell. **29**(1), 167–172 (2007)
17. Pelillo, M.: Relaxation labeling networks for the maximum clique problem. J. Artif. Neural Netw. **2**, 313–328 (1995)
18. Pelillo, M., Jagota, A.: Feasible and infeasible maxima in a quadratic program for maximum clique. J. Artif. Neural Netw. **2**, 411–420 (1995)
19. Pelillo, M., Torsello, A: Payoff-monotonic game dynamics and the maximum clique problem. Neural Comput. **18**(5), 1215–1258 (2006)
20. Rota Bulò, S., Torsello, A., Pelillo, M.: A continuous-based approach for partial clique enumeration. Graph-Based Representations Patt. Recogn. **4538**, 61–70 (2007)
21. Sós, V., Straus, E.G.: Extremal of functions on graphs with applications to graphs and hypergraphs. J. Combin. Theory Ser. B **63**, 189–207 (1982)
22. Wilf, H.S.: The eigenvalues of a graph and its chromatic number. J. Lond. Math. Soc. **42**, 330–332 (1967)
23. Wilf, H.S.: Spectral bounds for the clique and independence numbers of graphs. J. Combin. Theory Ser. B **40**, 113–117 (1986)