On Generalized KKT Points of the Motzkin-Straus Program

Guglielmo Beretta⋆ (✉)[1,2][0000-0002-4757-1965], Alessandro Torcinovich⋆[1,3][0000-0001-8110-1791], and Marcello Pelillo[1,4,5]

[1] DAIS, Ca' Foscari University of Venice, Via Torino 155, 30170 Venice, Italy
guglielmo.beretta@unive.it (✉)
[2] DAUIN, Polytechnic University of Turin, Corso Castelfidardo 34/d, 10138 Turin, Italy
[3] D-INFK, ETH, Universitätstrasse 6, 8006 Zürich, Kanton Zürich, Switzerland
[4] College of Mathematical Medicine, Zhejiang Normal University, 688 Yingbin Road, Jinhua, Zhejiang Province, 321004 China
[5] European Centre for Living Technology, Ca' Bottacin, Dorsoduro 3911, Calle Crosera, 30123 Venezia, Italy

Abstract. In 1965, Motzkin and Straus established a profound connection between the clique number of a graph and the global maxima of a quadratic program defined on the standard simplex. Since then, a line of active and intensive research has been yielding heuristics and bounds concerning the maximum clique problem, thanks to the discoveries pertaining to local/global solutions of the Motzkin-Straus program. However, the Karush-Kuhn-Tucker (KKT) points thereof have received little to no attention in the literature. In this work, a parameterized version of the Motzkin-Straus program is discussed, and some results about its KKT points are obtained. What emerges is a connection between a generalized notion of KKT point and some regular structures contained in the graph.

Keywords: Clique · KKT points · Regular graphs · Replicator dynamics · Standard quadratic optimization.

1 Introduction

In 1965, Motzkin and Straus [12] conducted a study of the quadratic program

$$\underset{\boldsymbol{x} \in \Delta_n}{\text{maximize}} \quad f(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x}, \tag{1}$$

with $\boldsymbol{A}$ being the adjacency matrix of a graph G with n nodes, and Δ_n denoting the *standard simplex* in $\mathbb{R}^n$, i.e., the convex hull of the canonical basis of $\mathbb{R}^n$. They discovered that the value of Program (1) is $1-\omega(G)^{-1}$, where $\omega(G)$ stands for the clique number of G. In so doing, they exhibited a way to associate to

⋆ Equal contribution.

every maximum clique of G a global solution for Program (1), in such a way that distinct cliques correspond to distinct solutions.

However, as discussed in [14], the Motzkin-Straus program can exhibit *spurious* solutions, i.e., maximizers not associated with any maximum clique under this correspondence. Unfortunately, the presence of spurious solutions can be undesirable when using the Motzkin-Straus program to find maximum cliques, which is one of the remarkable applications of the results concerning the Motzkin-Straus program [13]. To address this issue, Bomze [3] introduced a modified version of Program (1), incorporating a convex regularization term into its objective function. This adjustment resulted in a program with local (global) solutions that are in one-to-one correspondence with maximal (maximum) cliques [3].

The Motzkin-Straus program and its variations have been extensively explored in the literature [2, 8, 13, 17, 18], inspiring various heuristics and bounds for the maximum clique problem, see e.g. [4, 6, 19, 23]. However, the majority of these studies focus on local/global maximizers of Program (1), with little or no emphasis on its Karush-Kuhn-Tucker (KKT) points [11] (with some exceptions, see e.g. [8]).

In contrast, our paper presents some results concerning the KKT points of a parametric version of the program introduced by Bomze in [3, 5]. Specifically, we outline a relation between a generalized notion of KKT point [3] and some structural information about the graph underlying the program. After extending some known results related to regular induced subgraphs, we leverage the notion of barycentric coordinates [10, 16] to obtain a convenient representation of points in the standard simplex. This representation helps investigate how the presence of (generalized) KKT points affects the combinatorial structure of the graph. We then apply the obtained results to a class of graphs that generalizes the notion of star graphs. Finally, we mention a connection with replicator dynamics, a model developed in evolutionary game theory [9] that found applications in diverse fields, including computer science.

An extended version of this work can be found in [1], which includes more details on the proofs, as well as additional properties regarding the generalized KKT points.

2 Notation

For $n \in \mathbb{N}$, we set $[n] = \{i \in \mathbb{N} \colon 1 \leq i \leq n\}$. We denote by $\mathbf{0}$ (resp. $\mathbf{1}$) a vector with every component equal to 0 (resp. 1). In this paper, $\operatorname{diag}(r_1, r_2, \ldots r_n)$ denotes the $n \times n$ diagonal matrix having on its main diagonal r_i as its i-th element. The *support* of a vector $\boldsymbol{x} \in \mathbb{R}^n$ is the set $\operatorname{supp}(\boldsymbol{x}) = \{i \in [n] \colon x_i \neq 0\}$. The *standard simplex* in $\mathbb{R}^n$ is the set $\Delta_n = \{\boldsymbol{x} \in \mathbb{R}^n \mid \mathbf{1}^\top \boldsymbol{x} = 1, \boldsymbol{x} \geq \mathbf{0}\}$. For a non-empty $S \subseteq [n]$, we define:

$$\begin{aligned}
\Delta_n(S) &= \{\boldsymbol{x} \in \Delta_n \colon \operatorname{supp}(\boldsymbol{x}) \subseteq S\}, \\
\operatorname{int}(\Delta_n(S)) &= \{\boldsymbol{x} \in \Delta_n \colon \operatorname{supp}(\boldsymbol{x}) = S\}, \\
\partial(\Delta_n(S)) &= \Delta_n(S) \setminus \operatorname{int}(\Delta_n(S)),
\end{aligned}$$

where $\Delta_n(S)$, $\mathrm{int}(\Delta_n(S))$ and $\partial(\Delta_n(S))$ are the *face* of Δ_n associated with S, its *(relative) interior* and its *(relative) boundary* respectively. Given $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k \in \mathbb{R}^n$, their *convex hull* is the set $\{\sum_{m=1}^k y_m \boldsymbol{x}_m \mid \boldsymbol{y} \in \Delta_k\}$, which we denote by $\mathrm{conv}(\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_k)$. In particular, we set $[\boldsymbol{x}_1, \boldsymbol{x}_2] = \mathrm{conv}(\boldsymbol{x}_1, \boldsymbol{x}_2)$ for the closed *segment* with endpoints $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$.

As for the graph-related notation, $G = (V, E)$ denotes in the sequel an unweighted undirected graph with no loops on the set V and with set of edges $E \subseteq \binom{V}{2}$. Given two distinct vertices $i, j \in V$, we call i a *neighbor* of j and write $i \sim j$ if $\{i, j\} \in E$. The *adjacency matrix* of G is the symmetric $n \times n$ matrix with coefficient ij equal to 1 if $i \sim j$ and equal to 0 otherwise. Given a non-empty $S \subseteq V$, we use $G[S]$ for the subgraph of G *induced by* S. A non-empty subset $C \subseteq V$ is called a *clique* if $i \sim j$ for every pair of distinct $i, j \in C$. The *degree* of a node in a graph is the number of neighbors that node has in the graph. A graph is said *regular* if every node in the graph has the same degree.

3 Parametric Motzkin-Straus Program

Consider a graph $G = (V, E)$ on a finite non-empty set V and set $n = |V|$. Without loss of generality, we assume $V = [n]$ to simplify the notation. Let $\boldsymbol{A}$ be the adjacency matrix of G and write $\boldsymbol{I}$ for the $n \times n$ identity matrix. Fix now a parameter $c \in \mathbb{R}$ and consider the quadratic program

$$\max_{\boldsymbol{x} \in \Delta_n} \quad f_c(\boldsymbol{x}) = \boldsymbol{x}^\top (\boldsymbol{A} + c\boldsymbol{I})\boldsymbol{x} \tag{2}$$

and the associated *Lagrangian*

$$\mathcal{L}(\boldsymbol{x}, \mu_0, \boldsymbol{\mu}) = f_c(\boldsymbol{x}) + \mu_0(\boldsymbol{1}^\top \boldsymbol{x} - 1) + \boldsymbol{\mu}^\top \boldsymbol{x},$$

where $(\boldsymbol{x}, \mu_0, \boldsymbol{\mu}) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n$. Program (2) is discussed in [3] for $c = \frac{1}{2}$ and in [5] in its more general formulation. Observe that Program (1) is precisely Program (2) in case $c = 0$.

Definition 1. *Given $\boldsymbol{x} \in \mathbb{R}^n$, we call $\boldsymbol{x}$ a* KKT point *for Program* (2)*, and write $\boldsymbol{x} \in \mathcal{KKT}(c)$, if $\boldsymbol{x} \in \Delta_n$ and some $(\mu_0, \boldsymbol{\mu}) \in \mathbb{R} \times \mathbb{R}^n$ exists such that*

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial \boldsymbol{x}}(\boldsymbol{x}, \mu_0, \boldsymbol{\mu}) = \boldsymbol{0}, & \text{(3a)} \\ \mu_i x_i = 0, \quad \textit{for every } i \in V, & \text{(3b)} \\ \boldsymbol{\mu} \geq \boldsymbol{0}. & \text{(3c)} \end{cases}$$

We call $\boldsymbol{x}$ a generalized[6] KKT *(gKKT)* point *for Program* (2)*, and we write $\boldsymbol{x} \in \mathrm{g}\mathcal{KKT}(c)$, if $\boldsymbol{x} \in \Delta_n$ and some $(\mu_0, \boldsymbol{\mu}) \in \mathbb{R} \times \mathbb{R}^n$ exists such that Eq.* (3a) *and Eq.* (3b) *hold.*

[6] Generalized KKT points appear also in [3] with a slightly different definition.

Clearly, $\mathcal{KKT}(c) \subseteq \mathrm{g}\mathcal{KKT}(c)$, and, given the meaning of the sign condition on Lagrange multipliers, every element of $\mathrm{g}\mathcal{KKT}(c)$ can be seen as a KKT point of a suitable modification of Program (2).

Proposition 1. *Let $\hat{\boldsymbol{x}} \in \Delta_n$ and set $S = \mathrm{supp}(\hat{\boldsymbol{x}})$. Then $\hat{\boldsymbol{x}} \in \mathrm{g}\mathcal{KKT}(c)$ if and only if $\hat{\boldsymbol{x}}$ is a KKT point for the program*

$$\underset{\boldsymbol{x} \in \mathrm{int}(\Delta_n(S))}{\text{maximize}} \quad f_c(\boldsymbol{x}).$$

Proof. By the first-order optimality conditions [11], only Lagrange multipliers corresponding to inequality constraints are required to be non-negative. □

To prove that Definition 1 is satisfied, it is helpful to use a simpler description of $\mathcal{KKT}(c)$ and $\mathrm{g}\mathcal{KKT}(c)$, as in [5,8].

Proposition 2. *Let $\boldsymbol{x} \in \Delta_n$, set $\lambda = f_c(\boldsymbol{x})$ and for $\boldsymbol{M} = \boldsymbol{A} + c\boldsymbol{I}$, consider the statements:*

(S1) $(\boldsymbol{M}\boldsymbol{x})_i = \lambda$ *for every* $i \in \mathrm{supp}(\boldsymbol{x})$;
(S2) $(\boldsymbol{M}\boldsymbol{x})_i \leq \lambda$ *for every* $i \in V \setminus \mathrm{supp}(\boldsymbol{x})$.

Then:

(a) $(\boldsymbol{M}\boldsymbol{x})_i = (\boldsymbol{M}\boldsymbol{x})_j$ *for every* i, $j \in \mathrm{supp}(\boldsymbol{x})$ *if and only if* (S1) *holds;*
(b) $\boldsymbol{x} \in \mathcal{KKT}(c)$ *if and only if both* (S1) *and* (S2) *hold;*
(c) $\boldsymbol{x} \in \mathrm{g}\mathcal{KKT}(c)$ *if and only if* (S1) *holds.*

4 Characteristic Vectors and Convex Hulls

For a non-empty $S \subseteq V$, the *characteristic vector* representing S in Δ_n is the vector $\boldsymbol{x}^S \in \Delta_n$ defined by $x_i^S = 1/|S|$ for $i \in S$ and $x_i^S = 0$ for $i \in V \setminus S$. Characteristic vectors representing maximum (resp. maximal) cliques emerge as global (resp. local) solutions to Program (1), see e.g. [12, 14, 20]. Indeed, the correspondence mentioned in Sect. 1 maps every non-empty $S \subseteq V$ to its corresponding characteristic vector. We remark that $\boldsymbol{x}^S$ may be a KKT point even if S is not a clique, i.e, even if the graph $G[S]$ is not complete. What is known, however, is that if $\boldsymbol{x}^S$ is a gKKT then the graph $G[S]$ is necessarily regular, and this is independent of the value of c.

Proposition 3. *Let $\boldsymbol{x}$ be a characteristic vector. Then $\boldsymbol{x} \in \mathrm{g}\mathcal{KKT}(c)$ if and only if $G[\mathrm{supp}(\boldsymbol{x})]$ is regular.*

Proposition 3 is proved in [3] for the case $c = \frac{1}{2}$. The same proof, as mentioned in [5], holds also for the case $0 \leq c \leq 1$, and indeed it is valid for every $c \in \mathbb{R}$. We can say more in the following generalization of [5, Proposition 6].

Proposition 4. *Let $\boldsymbol{x} \in \Delta_n$ and suppose two distinct c_1, $c_2 \in \mathbb{R}$ exist such that $\boldsymbol{x} \in \mathrm{g}\mathcal{KKT}(c_j)$ for $j = 1, 2$. Then $\boldsymbol{x} = \boldsymbol{x}^S$ for some $S \subseteq V$ and $G[S]$ is regular.*

Proof. Set $S = \mathrm{supp}(\boldsymbol{x})$. By Proposition 2, there exist λ_1, $\lambda_2 \in \mathbb{R}$ such that for $j = 1, 2$, the equality $((\boldsymbol{A} + c_j \boldsymbol{I})\boldsymbol{x})_i = \lambda_j$ holds for every $i \in S$. Therefore, *every* non-zero component of $\boldsymbol{x}$ equals $(\lambda_1 - \lambda_2)/(c_1 - c_2)$. This entails $\boldsymbol{x} = \boldsymbol{x}^S$, hence $G[S]$ is regular by Proposition 3. □

Unlike characteristic vectors, which by Proposition 3 are in $\mathrm{g}\mathcal{KKT}(c)$ either for every value of c or for no value of c, the elements of Δ_n that are not characteristic vectors exhibit different behavior.

Proposition 5. *Let $\boldsymbol{x} \in \Delta_n$ and suppose $\boldsymbol{x}$ is* not *a characteristic vector. Then $\boldsymbol{x} \in \mathrm{g}\mathcal{KKT}(c)$ for at most one value of c.*

The proof of Proposition 5 is immediately obtained by considering the contrapositive statement of Proposition 4.

We now introduce a way to represent elements of Δ_n that is convenient for our subsequent discussion. To this end, we observe that for a family of non-empty and pairwise disjoint subsets of V, the corresponding characteristic vectors are linearly independent.

Definition 2. *Consider a family $\mathcal{F} = \{V_1, V_2, \ldots, V_k\}$ of pairwise disjoint non-empty subsets of V. Given $\boldsymbol{x} \in \mathrm{conv}(\boldsymbol{x}^{V_1}, \boldsymbol{x}^{V_2}, \ldots, \boldsymbol{x}^{V_k})$, the barycentric coordinates of $\boldsymbol{x}$ with respect to (the characteristic vectors representing the classes of) $\mathcal{F}$ is the unique*[7] *vector $\boldsymbol{y} = \mathrm{bary}_{\mathcal{F}}(\boldsymbol{x}) \in \Delta_k$ such that $\boldsymbol{x} = \sum_{\ell=1}^{k} y_\ell \boldsymbol{x}^{V_\ell}$.*

As an example, in case $\mathcal{F} = \{\{i\} \mid i \in V\}$, it is trivial to show that $\mathrm{bary}_{\mathcal{F}}(\boldsymbol{x}) = \boldsymbol{x}$ for every $\boldsymbol{x} \in \Delta_n$. Moreover, it is easy to see that $\mathrm{bary}_{\mathcal{F}}(\boldsymbol{x})$ lies in $\mathrm{int}(\Delta_k)$ if and only if $\mathrm{supp}(\boldsymbol{x}) = \cup_{\ell=1}^{k} V_\ell$. The barycentric-coordinates notation is conveniently combined with another concept, which qualifies a particular way to partition the support of an element of Δ_n.

Definition 3. *Let $\boldsymbol{x} \in \Delta_n$. We say that a partition $\mathcal{P}$ of* $\mathrm{supp}(\boldsymbol{x})$ separates distinct values *of $\boldsymbol{x}$ if for every i, $j \in \mathrm{supp}(\boldsymbol{x})$ the relation $x_i \neq x_j$ implies that the nodes i and j belong to distinct classes of $\mathcal{P}$.*

Choose now $\boldsymbol{x} \in \Delta_n$ and consider a partition $\mathcal{P} = \{V_1, V_2, \ldots, V_k\}$ of $\mathrm{supp}(\boldsymbol{x})$ separating distinct values of $\boldsymbol{x}$. In this setting, it is not hard to prove that $\boldsymbol{x} \in \mathrm{conv}(\boldsymbol{x}^{V_1}, \boldsymbol{x}^{V_2}, \ldots, \boldsymbol{x}^{V_k})$, hence $\mathrm{bary}_{\mathcal{P}}(\boldsymbol{x})$ is well defined.

Moreover, if $\boldsymbol{x}$ is also a gKKT point, then these barycentric coordinates satisfy some algebraic relations depending on the structure of G. Formulating this dependency involves the graph-theoretical notion of *density* [7]. For every non-empty S_1, $S_2 \subseteq V$, let $e_G(S_1, S_2)$ count the *ordered* pairs of adjacent nodes in the set $S_1 \times S_2$ and call *edge density* between S_1 and S_2 the ratio

$$d_G(S_1, S_2) = \frac{e_G(S_1, S_2)}{|S_1||S_2|}.$$

Suitable edge densities and barycentric coordinates can be combined to produce an alternative expression for the multiplication by $\boldsymbol{A}$.

[7] Strictly speaking, to have uniqueness, it is necessary to fix an enumeration of the classes in $\mathcal{F}$. However, we will omit writing it explicitly to simplify the notation.

Lemma 1. *Let $\mathcal{F} = \{V_1, V_2, \ldots, V_k\}$ be a family of pairwise disjoint non-empty subsets of V, let $\boldsymbol{x} \in \operatorname{conv}(\boldsymbol{x}^{V_1}, \boldsymbol{x}^{V_2}, \ldots, \boldsymbol{x}^{V_k})$ and set $\boldsymbol{y} = \operatorname{bary}_{\mathcal{F}}(\boldsymbol{x})$. Then for every node $i \in V$:*

$$(\boldsymbol{A}\boldsymbol{x})_i = \sum_{m=1}^{k} d_G(\{i\}, V_m) y_m.$$

Proof. Since $\boldsymbol{x} = \sum_\ell y_\ell \boldsymbol{x}^{V_\ell}$, then:

$$\begin{aligned}(\boldsymbol{A}\boldsymbol{x})_i &= \sum_{j=1}^{n} a_{ij} x_j = \sum_{m=1}^{k} \sum_{j \in V_m} a_{ij} x_j = \sum_{m=1}^{k} \sum_{j \in V_m} a_{ij} (y_m / |V_m|) \\ &= \sum_{m=1}^{k} e_G(\{i\}, V_m) y_m |V_m|^{-1} = \sum_{m=1}^{k} d_G(\{i\}, V_m) y_m.\end{aligned}$$

□

One last definition is needed to introduce the first main result of this section.

Definition 4. *For a finite family $\mathcal{F} = \{V_1, V_2, \ldots V_k\}$ of distinct non-empty subsets of V, the* edge density matrix *of $\mathcal{F}$ is the* symmetric *matrix $\boldsymbol{D} \in \mathbb{R}^{k \times k}$ with general coefficient $d_{\ell,m} = d_G(V_\ell, V_m)$.*

Theorem 1. *Let $\boldsymbol{x} \in \Delta_n$ and let $\mathcal{P} = \{V_1, V_2, \ldots V_k\}$ be a partition of $\operatorname{supp}(\boldsymbol{x})$ separating distinct values of $\boldsymbol{x}$. Call $\boldsymbol{D}$ the edge density matrix associated with $\mathcal{P}$ and set $\boldsymbol{\Lambda} = \operatorname{diag}(|V_1|, |V_2|, \ldots, |V_k|)$. If $\boldsymbol{x} \in \mathrm{g}\mathcal{KKT}(c)$,[8] then $\operatorname{bary}_{\mathcal{P}}(\boldsymbol{x})$ is a KKT point for the program:*

$$\underset{\boldsymbol{y} \in \operatorname{int}(\Delta_k)}{\text{maximize}} \quad \boldsymbol{y}^\top (\boldsymbol{D} + c\boldsymbol{\Lambda}^{-1}) \boldsymbol{y}. \tag{4}$$

Proof. Set $\boldsymbol{y} = \operatorname{bary}_{\mathcal{F}}(\boldsymbol{x})$. We may write $\boldsymbol{x} = \sum_\ell y_\ell \boldsymbol{x}^{V_\ell}$ by definition of $\boldsymbol{y}$. By Proposition 2 and Lemma 1, there exists $\lambda \in \mathbb{R}$ such that for every i in $\operatorname{supp}(\boldsymbol{x})$:

$$\lambda = ((\boldsymbol{A} + c\boldsymbol{I})\, \boldsymbol{x})_i = \sum_{m=1}^{k} d_G(\{i\}, V_m) y_m + c x_i.$$

For every $\ell \in [k]$, it follows that

$$\begin{aligned}\lambda &= \frac{1}{|V_\ell|} \sum_{i \in V_\ell} \left(\sum_{m=1}^{k} d_G(\{i\}, V_m) y_m + c x_i \right) \\ &= \frac{1}{|V_\ell|} \sum_{i \in V_\ell} \left(\sum_{m=1}^{k} \frac{e_G(\{i\}, V_m)}{|V_m|} y_m + c(y_\ell / |V_\ell|) \right) \\ &= \sum_{m=1}^{k} d_G(V_\ell, V_m) y_m + (c/|V_\ell|) y_\ell = \left(\left(\boldsymbol{D} + c\boldsymbol{\Lambda}^{-1} \right) \boldsymbol{y} \right)_\ell .\end{aligned}$$

Then $\boldsymbol{y}$ is a KKT point for Program (4). □

[8] The theorem is still true if we replace the condition $\boldsymbol{x} \in \mathrm{g}\mathcal{KKT}(c)$ with the more restrictive $\boldsymbol{x} \in \mathcal{KKT}(c)$.

We remark that Proposition 1 can be regarded as a trivial corollary of Theorem 1 when the partition considered is $\mathcal{P} = \{\{i\} : i \in \mathrm{supp}(\boldsymbol{x})\}$.

5 Highly Regular Families

The converse of Theorem 1 would be helpful in characterizing the elements of $\mathrm{g}\mathcal{KKT}(c)$. Unfortunately, the implication appearing in Theorem 1 cannot be replaced with a double implication. In fact, suppose G is the graph on $V = [4]$ with adjacency matrix:

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \tag{5}$$

and consider the vector $\boldsymbol{x} = \boldsymbol{x}^V = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$. For $V_1 = \{1, 2\}$ and $V_2 = \{3, 4\}$ the family $\mathcal{P} = \{V_1, V_2\}$ partitions $\mathrm{supp}(\boldsymbol{x})$, and in this case:

$$\boldsymbol{D} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \boldsymbol{\Lambda} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}. \tag{6}$$

Observe that $\mathrm{bary}_{\mathcal{P}}(\boldsymbol{x}) = \left(\frac{1}{2}, \frac{1}{2}\right)$ is a KKT point for Program (4) and also that $\boldsymbol{x} \notin \mathcal{KKT}(c)$, and this is true regardless of the value of c.

However, Theorem 1 admits a partial converse under stronger hypotheses, which can be expressed using the notion of highly regular families.

Definition 5. *Consider a finite family $\mathcal{F} = \{V_1, V_2, \dots V_k\}$ of pairwise disjoint non-empty subsets of V. We call $\mathcal{F}$* highly regular *for G if:*[9]

$$d_G(V_\ell, V_m) = d_G(\{i\}, V_m) \qquad \text{for every } \ell,\, m \in [k] \text{ and every } i \in V_\ell.$$

What follows is a criterion for assessing whether Definition 5 is satisfied.

Proposition 6. *A finite family $\mathcal{F} = \{V_1, V_2, \dots V_k\}$ of subsets of V is highly regular for G if and only if the following two conditions hold:*

(HR1) *for every $\ell \in [k]$ the set V_ℓ is non-empty and $G[V_\ell]$ is regular;*
(HR2) *for every distinct ℓ, $m \in [k]$ the sets V_ℓ, V_m are disjoint and each node in V_ℓ has the same amount of neighbors in V_m.*

We are now in the position to prove Theorem 2.

Theorem 2. *Let $\boldsymbol{x} \in \Delta_n$ and let $\mathcal{P} = \{V_1, V_2, \dots V_k\}$ be a partition of* $\mathrm{supp}(\boldsymbol{x})$ *separating distinct values of $\boldsymbol{x}$. Call $\boldsymbol{D}$ the edge density matrix associated with $\mathcal{P}$ and set $\boldsymbol{\Lambda} = \mathrm{diag}(|V_1|, |V_2|, \dots, |V_k|)$. Assume $\mathcal{P}$ is highly regular for G. Then $\boldsymbol{x} \in \mathrm{g}\mathcal{KKT}(c)$ if and only if $\mathrm{bary}_{\mathcal{P}}(\boldsymbol{x})$ is a KKT point for Program (4).*

[9] Equivalently, for every non-empty $X \subseteq V_\ell$ and every non-empty $Y \subseteq V_m$ we have $d_G(X, V_m) = d_G(V_\ell, V_m) = d_G(V_\ell, Y)$.

Proof. Assume $\boldsymbol{y} = \mathrm{bary}_{\mathcal{P}}(\boldsymbol{x})$ is a KKT point for Program (4) and assume $\mathcal{P}$ is highly regular. Then $\mathrm{supp}(\boldsymbol{y}) = [k]$, and there exists $\lambda \in \mathbb{R}$ such that for every $\ell \in [k]$:

$$\big((\boldsymbol{D} + c\boldsymbol{\Lambda}^{-1})\boldsymbol{y}\big)_\ell = \lambda.$$

Pick any i in the support of $\boldsymbol{x}$. The node i is in V_ℓ for some $\ell \in [k]$ and $d_G(\{i\}, V_m) = d_G(V_\ell, V_m)$ since $\mathcal{P}$ is highly regular. By Lemma 1:

$$((\boldsymbol{A} + c\boldsymbol{I})\boldsymbol{x})_i = \sum_{m=1}^{k} d_G(\{i\}, V_m)y_m + cx_i = \sum_{m=1}^{k} d_G(V_\ell, V_m)y_m + (c/|V_\ell|)y_\ell = \lambda.$$

Then $\boldsymbol{x} \in \mathrm{g}\mathcal{KKT}(c)$ by Proposition 2. The other implication follows immediately by Theorem 1. □

6 Applications for Bipartitions

We provide here an application of the previous results to a highly regular family of the form $\mathcal{P} = \{V_1, V_2\}$. Let $[d_{ij}]_{i,j=1}^2$ be the density matrix associated with $\mathcal{P}$ and set $\alpha_1(\mathcal{P}) = |V_2|(d_{12} - d_{22})$ and $\alpha_2(\mathcal{P}) = |V_1|(d_{21} - d_{11})$. We observe that, since $\mathcal{P}$ is highly regular, $G[V_1 \cup V_2]$ is a regular graph if and only if $\alpha_1(\mathcal{P}) = \alpha_2(\mathcal{P})$. For a proof of this technical result, see [1]. We can now state Corollary 1, which treats the case $G[V_1 \cup V_2]$ regular.

Corollary 1. *Let $\mathcal{P} = \{V_1, V_2\}$ be highly regular for G and assume $G[V_1 \cup V_2]$ is a regular graph. Then $c^* = \alpha_1(\mathcal{P}) = \alpha_2(\mathcal{P})$ is such that:*

- *If $c = c^*$, then $\mathrm{g}\mathcal{KKT}(c) \cap [\boldsymbol{x}^{V_1}, \boldsymbol{x}^{V_2}] = [\boldsymbol{x}^{V_1}, \boldsymbol{x}^{V_2}]$;*
- *If $c \neq c^*$, then $\mathrm{g}\mathcal{KKT}(c) \cap [\boldsymbol{x}^{V_1}, \boldsymbol{x}^{V_2}] = \{\boldsymbol{x}^{V_1}, \boldsymbol{x}^{V_2}, \boldsymbol{x}^{V_1 \cup V_2}\}$.*

The proof of Corollary 1 is similar to that of Corollary 2 here proved, which deals with the case $G[V_1 \cup V_2]$ not regular.

Corollary 2. *Let $\mathcal{P} = \{V_1, V_2\}$ be highly regular for G and assume $G[V_1 \cup V_2]$ is* not *a regular graph. Then the interval $I = \mathrm{conv}(\alpha_1(\mathcal{P}), \alpha_2(\mathcal{P}))$ is such that:*

- *If $c \in I$, then $\mathrm{g}\mathcal{KKT}(c) \cap [\boldsymbol{x}^{V_1}, \boldsymbol{x}^{V_2}] = \{\boldsymbol{x}^{V_1}, \boldsymbol{x}^{V_2}\}$;*
- *If $c \notin I$, then $\mathrm{g}\mathcal{KKT}(c) \cap [\boldsymbol{x}^{V_1}, \boldsymbol{x}^{V_2}] = \{\boldsymbol{x}^{V_1}, \boldsymbol{x}^{V_2}, \boldsymbol{x}_c\}$, where*

$$\boldsymbol{x}_c = \sum_{i=1}^{2} \frac{(c - \alpha_i(\mathcal{P}))|V_i|}{(c - \alpha_1(\mathcal{P}))|V_1| + (c - \alpha_2(\mathcal{P}))|V_2|} \boldsymbol{x}^{V_i}.$$

Proof. For every $c \in \mathbb{R}$, both $\boldsymbol{x}^{V_1}$ and $\boldsymbol{x}^{V_2}$ are in $\mathrm{g}\mathcal{KKT}(c)$ by Proposition 3. Consider now some $\boldsymbol{x} \neq \boldsymbol{x}^{V_1}, \boldsymbol{x}^{V_2}$. By Theorem 2 and thanks to some elementary algebraic manipulations, it is not hard to see that $\boldsymbol{x} \in \mathrm{g}\mathcal{KKT}(c) \cap [\boldsymbol{x}^{V_1}, \boldsymbol{x}^{V_2}]$ if and only if $\boldsymbol{x} = y_1\boldsymbol{x}^{V_1} + y_2\boldsymbol{x}^{V_2}$, where $(y_1, y_2)^\top \in \mathrm{int}(\Delta_2)$ and

$$(c - \alpha_2(\mathcal{P}))y_1/|V_1| = (c - \alpha_1(\mathcal{P}))y_2/|V_2|. \tag{7}$$

For positive y_1, y_2 a solution to Eq. (7) is possible only in case $c \notin I$, and if that occurs, then it is easy to check that $\boldsymbol{x} = \boldsymbol{x}_c$. □

Corollary 2 can be applied to *star graphs*, a class of complete bipartite graphs in which one node, called center, is an endpoint of every edge in the graph [7]. Indeed, we are about to show that it also holds for a more general class of graphs.

Definition 6. *A graph $G' = (V', E')$ is a* generalized star *with* core H' *if:*

(GS1) *H' is a proper non-empty subset of V';*
(GS2) *Every node in H' is adjacent to every node in $V' \setminus H'$;*
(GS3) *H' is a clique, but V' is not a clique;*
(GS4) *The induced subgraph $G[V' \setminus H']$ is regular.*

Trivially, a star graph with center $\boldsymbol{z}$ is a generalized star with core $\{\boldsymbol{z}\}$.

Corollary 3. *Let H, P be disjoint subsets of V such that $G[H \cup P]$ is a generalized star with core H. There exists an integer $b > 1$ such that, if $c \notin [1, b]$, then $\mathrm{g}\mathcal{KKT}(c)$ contains a vector with support $H \cup P$.*

Proof. Set $h = |H|, p = |P|$ and assume $G[P]$ is a d-regular graph. By hypothesis, the family $\{P, H\}$ is highly regular for G, with edge density matrix:

$$\boldsymbol{D} = \begin{pmatrix} d/p & 1 \\ 1 & 1 - 1/h \end{pmatrix},$$

and $b = p - d$ satisfies $1 < b < p$ since $G[P]$ is not complete. For $c \in \mathbb{R} \setminus [1, b]$, set $y_1 = (c-1)p/[(c-1)p + (c-b)h]$ and $y_2 = (c-b)h/[(c-1)p + (c-b)h]$, and observe that $\boldsymbol{x}_c = y_1 \boldsymbol{x}^P + y_2 \boldsymbol{x}^H$ is an element of $\mathrm{g}\mathcal{KKT}(c)$ by Corollary 2. $\square$

7 Link with the Replicator Dynamics

The results we have discussed can be reinterpreted in the context of *replicator dynamics*, a term denoting a class of dynamical systems studied in evolutionary game theory. Introduced to describe the coexistence of interacting and self-replicating species [9, 21, 22], replicator dynamics has found applications also in economics and social sciences [22], as well as in computer science, see e.g., [3, 18]. To derive one of its possible formulations, consider a matrix $\boldsymbol{M} \in \mathbb{R}^{n \times n}$ and the associated ordinary differential equation:

$$\dot{x}_i = x_i[(\boldsymbol{Mx})_i - \boldsymbol{x}^\top \boldsymbol{Mx}], \quad i = 1, \dots, n. \tag{8}$$

It can be proved [9] that Δ_n is invariant under the flow defined by Eq. (8). The continuous-time *replicator dynamics*[10] with payoff-matrix $\boldsymbol{M}$ is the dynamics defined on Δ_n by Eq. (8). We recall that a point $\boldsymbol{z} \in \Delta_n$ is *stationary* with respect to (8) if the constant function $\boldsymbol{x}(t) = \boldsymbol{z}$ is a solution of Eq. (8). Bomze [3] provides a characterization[11] of stationary points of Eq. (8). The following theorem is a rewording of that characterization specialized to the case $\boldsymbol{M} = c\boldsymbol{I}$.

[10] This is not the only possible replicator dynamics with $\boldsymbol{M}$ as payoff-matrix [9, 15].

[11] The reason why Bomze considers generalized KKT points in [3] is precisely due to the content of Theorem 3.

Theorem 3. *A point $\boldsymbol{x} \in \Delta_n$ is stationary for the replicator dynamics with payoff-matrix $\boldsymbol{A} + c\boldsymbol{I}$ if and only if $\boldsymbol{x} \in \mathrm{g}\mathcal{KKT}(c)$.*

Thanks to Theorem 3, every result contained in this work concerning $\mathrm{g}\mathcal{KKT}(c)$ can be restated as a result on stationary points of the replicator dynamics with payoff-matrix $\boldsymbol{A} + c\boldsymbol{I}$.

8 Conclusion

In this article, we have discussed some properties of Karush-Kuhn-Tucker (KKT) points associated with a parametric version of the Motzkin-Straus program. We have extended some known results regarding characteristic vectors and KKT points. We have then studied the generalized KKT points by representing them as convex combinations of characteristic vectors, thanks to a suitable use of barycentric coordinates. This has led to Theorem 1 and Theorem 2, which clarify how passing to barycentric coordinates can help in describing the generalized KKT points. To this end, we introduced some convenient ways to describe how a family of sets of nodes splits both the support of a generalized KKT point (partition separating distinct values) and the edges of the graph (highly regular family). Finally, the case of a highly regular bipartition has enabled the identification of some generalized KKT points associated with generalized star graphs.

We conclude by mentioning two possible research directions in which the study here presented could proceed. On the one hand, it could be worth looking for a weaker version of Theorem 2 that does not involve highly regular families. This could result in a more accurate description of KKT points and generalized KKT points. On the other hand, a promising path concerns the connection with the replicator dynamics. Theorem 1 and Theorem 2, in the context of replicator dynamics, describe a correspondence between stationary points associated with two distinct replicator dynamics, occurring on simplices of different dimensionality. It is worth investigating whether passing to barycentric coordinates and simulating the dynamics occurring on the smaller simplex could help in finding stationary points for the other dynamics. Besides that, we do not exclude that the replicator dynamics, other than a tool for detecting cliques, could offer a more powerful tool for graph analysis, and help in identifying regular structures within a given graph.

Acknowledgments. Guglielmo Beretta's scholarship is funded jointly by the Ca' Foscari University of Venice and by the Polytechnic University of Turin. The authors would like to thank Morteza H. Chehreghani for his insightful advice and Sebastiano Smaniotto for his help in the initial stages of this work.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Beretta, G., Torcinovich, A., Pelillo, M.: A Note on the KKT Points for the Motzkin-Straus Program. Preprint at https://arxiv.org/abs/2305.08519 (2023)

2. Bomze, I.M., Pelillo, M., Stix, V.: Approximating the maximum weight clique using replicator dynamics. IEEE Trans. Neural Netw. **11**(6), 1228–1241 (2000). https://doi.org/10.1109/72.883403
3. Bomze, I.M.: Evolution towards the maximum clique. J. Glob. Optim. **10**(2), 143–164 (1997). https://doi.org/10.1023/A:1008230200610
4. Bomze, I.M., Budinich, M., Pardalos, P.M., Pelillo, M.: The Maximum Clique Problem. In: Du, D.Z., Pardalos, P.M. (eds.) Handbook of Combinatorial Optimization, pp. 1–74. Springer, Boston, MA, USA (1999)
5. Bomze, I.M., Budinich, M., Pelillo, M., Rossi, C.: Annealed replication: a new heuristic for the maximum clique problem. Discret. Appl. Math. **121**(1-3), 27–49 (2002). https://doi.org/10.1016/S0166-218X(01)00233-5
6. Budinich, M.: Exact bounds on the order of the maximum clique of a graph. Discrete Appl. Math. **127**(3), 535–543 (2003). https://doi.org/10.1016/S0166-218X(02)00386-4
7. Diestel, R.: Graph theory. Springer, Heidelberg, Germany (2005)
8. Gibbons, L.E., Hearn, D.W., Pardalos, P.M., Ramana, M.V.: Continuous characterizations of the maximum clique problem. Math. Oper. Res. **22**(3), 754–768 (1997). https://doi.org/10.1287/moor.22.3.754
9. Hofbauer, J., Sigmund, K.: Evolutionary Games and Population Dynamics. Cambridge University Press, Cambridge, U.K. (1998)
10. Hormann, K., Sukumar, N.: Generalized barycentric coordinates in computer graphics and computational mechanics. CRC Press, Boca Raton, FL, USA (2017)
11. Luenberger, D.G., Ye, Y.: Linear and Nonlinear Programming. Springer, Cham (2016)
12. Motzkin, T.S., Straus, E.G.: Maxima for graphs and a new proof of a theorem of Turán. Can. J. Math. **17**, 533–540 (1965). https://doi.org/10.4153/CJM-1965-053-6
13. Pardalos, P.M., Phillips, A.T.: A global optimization approach for solving the maximum clique problem. Int. J. Comput. Math. **33**, 209–216 (1990). https://doi.org/10.1080/00207169008803851
14. Pelillo, M., Jagota, A.: Feasible and infeasible maxima in a quadratic program for maximum clique. J. Artif. Neural Netw. **2**(4), 411–420 (1995)
15. Pelillo, M., Torsello, A.: Payoff-monotonic game dynamics and the maximum clique problem. Neural Comput. **18**(5), 1215–1258 (2006). https://doi.org/10.1162/neco.2006.18.5.1215
16. Quarteroni, A.: Numerical Models for Differential Problems. Springer, Cham, 3rd edn. (2017)
17. Rota Bulò, S., Pelillo, M.: A generalization of the Motzkin–Straus theorem to hypergraphs. Optim. Lett. **3**(2), 287–295 (2009). https://doi.org/10.1007/s11590-008-0108-3
18. Rota Bulò, S., Pelillo, M.: Dominant-set clustering: A review. Eur. J. Oper. Res. **262**(1), 1–13 (2017). https://doi.org/10.1016/j.ejor.2017.03.056
19. Stozhkov, V., Pastukhov, G., Boginski, V., Pasiliao, E.L.: New analytical lower bounds on the clique number of a graph. Optim. Methods Softw. **32**(2), 336–368 (2017). https://doi.org/10.1080/10556788.2016.1172578
20. Tang, Q., Zhang, X., Zhao, C., Zhao, P.: On the maxima of Motzkin-Straus programs and cliques of graphs. J. Glob. Optim. **84**(4), 989–1003 (2022). https://doi.org/10.1007/s10898-022-01187-3
21. Taylor, P.D., Jonker, L.B.: Evolutionarily stable strategies and game dynamics. Math. Biosci. **40**, 145–156 (1978). https://doi.org/10.1016/0025-5564(78)90077-9

22. Weibull, J.: Evolutionary game theory. MIT Press, Cambridge, MA, USA (1995)
23. Wilf, H.S.: Spectral bounds for the clique and independence numbers of graphs. J. Comb. Theory, Series B **40**(1), 113–117 (1986). https://doi.org/10.1016/0095-8956(86)90069-9