FULL LENGTH PAPER

Optimality conditions for maximizing a function over a polyhedron

William W. Hager · **James T. Hungerford**

Abstract   We present new first and second-order optimality conditions for maximizing a function over a polyhedron. These conditions are expressed in terms of the first and second-order directional derivatives along the edges of the polyhedron, and an edge description of the polyhedron. If the objective function is quadratic and edge-convex, and the constraint polyhedron includes a box constraint, then local optimality can be checked in polynomial time. The theory is applied to continuous formulations of the vertex and edge separator problems. It is seen that the necessary and sufficient optimality conditions for these problems are related to the existence of edges at specific locations in the graph.

W. W. Hager (✉) · J. T. Hungerford
Department of Mathematics, University of Florida,
PO Box 118105, Gainesville, FL 32611-8105, USA
e-mail: hager@ufl.edu
URL: http://www.math.ufl.edu/~hager

J. T. Hungerford
e-mail: freerad@ufl.edu
URL: http://www.math.ufl.edu/~freerad

1 Introduction

Consider the optimization problem

$$\max \quad f(\mathbf{x}) \quad \text{subject to } \mathbf{x} \in \mathcal{P}, \tag{1.1}$$

where $f : \mathcal{P} \to \mathbb{R}$ is continuously differentiable and $\mathcal{P} \subset \mathbb{R}^n$ is the polyhedron defined by

$$\mathcal{P} = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b} \right\} \tag{1.2}$$

for some matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ with $rank(\mathbf{A}) = n$ and some vector $\mathbf{b} \in \mathbb{R}^m$. Given any $\mathbf{x}^* \in \mathcal{P}$, the set of active constraints at $\mathbf{x}^*$ is defined by

$$\mathcal{A}(\mathbf{x}^*) = \left\{ i \in [1, m] : \mathbf{A}_i \mathbf{x}^* = b_i \right\},$$

where $\mathbf{A}_i$ denotes the ith row of $\mathbf{A}$. The cone $\mathcal{F}(\mathbf{x}^*)$ of first-order feasible directions at $\mathbf{x}^*$ is

$$\mathcal{F}(\mathbf{x}^*) = \{\mathbf{d} \in \mathbb{R}^n : \mathbf{A}_i \mathbf{d} \leq 0 \quad \text{for all } i \in \mathcal{A}(\mathbf{x}^*)\}. \tag{1.3}$$

The usual first-order necessary optimality condition for (1.1) may be stated in the following way: If $\mathbf{x}^*$ is a local maximizer of (1.1), then

$$\mathbf{x}^* \in \mathcal{P} \quad \text{and} \quad \nabla f(\mathbf{x}^*)\mathbf{d} \leq 0 \quad \text{for every } \mathbf{d} \in \mathcal{F}(\mathbf{x}^*). \tag{1.4}$$

This condition is based on the observation that for each $\mathbf{d} \in \mathcal{F}(\mathbf{x}^*)$, we have $\mathbf{x}^* + \alpha\mathbf{d} \in \mathcal{P}$ for $\alpha \geq 0$ sufficiently small. Hence, if $\mathbf{x}^*$ is a local maximizer, then

$$(f(\mathbf{x}^* + \alpha\mathbf{d}) - f(\mathbf{x}^*))/\alpha \leq 0$$

when $\alpha > 0$ is sufficiently small. Take the limit as α tends to zero from the right to obtain (1.4) [1, p. 94].

The first-order necessary condition (1.4) is equivalent to the well-known Karush–Kuhn–Tucker conditions: There exists $\boldsymbol{\lambda} \in \mathbb{R}^n$ such that

$$\boldsymbol{\lambda} \geq \mathbf{0}, \quad \mathbf{b} - \mathbf{A}\mathbf{x}^* \geq \mathbf{0}, \quad \boldsymbol{\lambda}^\mathsf{T}(\mathbf{b} - \mathbf{A}\mathbf{x}^*) = 0, \quad \text{and} \quad \nabla f(\mathbf{x}^*) - \boldsymbol{\lambda}^\mathsf{T}\mathbf{A} = \mathbf{0}. \tag{1.5}$$

References for the KKT conditions include [1,6,11,13].

If f is twice continuously differentiable, then the second-order necessary optimality condition states that any local maximizer $\mathbf{x}^*$ must satisfy (1.4) in addition to the following:

$$\mathbf{d}^\mathsf{T} \nabla^2 f(\mathbf{x}^*)\mathbf{d} \leq 0 \quad \text{for every} \quad \mathbf{d} \in \mathcal{C}(\mathbf{x}^*), \tag{1.6}$$

where $\mathcal{C}(\mathbf{x}^*)$ is the critical cone at $\mathbf{x}^*$ defined by

$$
\begin{aligned}
\mathcal{C}(\mathbf{x}^*) &= \{\mathbf{d} \in \mathcal{F}(\mathbf{x}^*): \nabla f(\mathbf{x}^*)\mathbf{d} = 0\} \\
&= \{\mathbf{d} \in \mathcal{F}(\mathbf{x}^*): \mathbf{A}_i\mathbf{d} = 0 \text{ for all } i \text{ for which } \lambda_i > 0\}. \quad (1.7)
\end{aligned}
$$

The second-order condition (1.6) is obtained from a Taylor expansion around $\mathbf{x} = \mathbf{x}^*$. That is, if $\mathbf{d} \in \mathcal{C}(\mathbf{x}^*)$, then $\nabla f(\mathbf{x}^*)\mathbf{d} = 0$, $\mathbf{x}^* + \alpha\mathbf{d} \in \mathcal{P}$ for $\alpha \geq 0$ sufficiently small, and we have

$$
\begin{aligned}
f(\mathbf{x}^* + \alpha\mathbf{d}) &= f(\mathbf{x}^*) + \alpha\nabla f(\mathbf{x}^*)\mathbf{d} + \frac{1}{2}\alpha^2\mathbf{d}^\mathsf{T}\nabla^2 f(\mathbf{x}_\alpha)\mathbf{d} \\
&= f(\mathbf{x}^*) + \frac{1}{2}\alpha^2\mathbf{d}^\mathsf{T}\nabla^2 f(\mathbf{x}_\alpha)\mathbf{d},
\end{aligned}
$$

where $\mathbf{x}_\alpha$ lies between $\mathbf{x}^*$ and $\mathbf{x}^* + \alpha\mathbf{d}$. It follows that

$$
\mathbf{d}^\mathsf{T}\nabla^2 f(\mathbf{x}_\alpha)\mathbf{d} = \frac{2(f(\mathbf{x}^* + \alpha\mathbf{d}) - f(\mathbf{x}^*))}{\alpha^2}. \quad (1.8)
$$

Since $\mathbf{x}^*$ is a local maximizer, $f(\mathbf{x}^* + \alpha\mathbf{d}) \leq f(\mathbf{x}^*)$ for $\alpha > 0$ sufficiently small. We take the limit as α tends to zero from the right in (1.8) to obtain (1.6).

Borwein [2] and Contesse [3] show that when f is quadratic, the conditions (1.4) and (1.6) are also sufficient for local optimality. For a more general function, a sufficient condition for $\mathbf{x}^*$ to be a local maximizer of (1.1) is that $\mathbf{x}^*$ satisfy (1.4) in addition to the following [13, Theorem. 12.6]:

$$
\mathbf{d}^\mathsf{T}\nabla^2 f(\mathbf{x}^*)\mathbf{d} < 0 \quad \text{for every } \mathbf{d} \in \mathcal{C}(\mathbf{x}^*), \quad \mathbf{d} \neq \mathbf{0}. \quad (1.9)
$$

Since the cones $\mathcal{F}(\mathbf{x}^*)$ and $\mathcal{C}(\mathbf{x}^*)$ are infinite sets, the conditions (1.4), (1.6), and (1.9) are in general difficult to verify. In fact, as shown in [12,16], checking local optimality for an indefinite quadratic programming problem can be NP-hard. In the current paper, we will derive new first-order necessary optimality conditions and new second-order sufficient conditions which reduce to examining finite subsets of the cones $\mathcal{F}(\mathbf{x}^*)$ and $\mathcal{C}(\mathbf{x}^*)$ corresponding to an edge description of the polyhedron $\mathcal{P}$. Necessary second-order conditions are derived for the case when the objective function is convex in the edge directions. In important special cases, such as polyhedra that include box constraints, we show that the size of the edge description is bounded by a polynomial in n. Consequently, local optimality for a quadratic, edge-convex objective function can be checked in polynomial time.

Edge-directions were introduced in Dantzig's simplex method [4] for linear programming in order to move from one basic feasible solution to another. As pointed out in [15], the use of edge-directions in optimization has been explored more recently in a number of works including [5,9,14,18]. Edge-directions have been studied in combinatorial optimization in the context of vertex enumeration problems [5,14], and in identifying conditions under which discrete optimization problems are equivalent to continuous optimization problems [9,18]. In [18] Tardella showed that if the objective

f in (1.1) is convex along all edge-directions, then problem (1.1) has a vertex maximizer. A similar result, in which $\mathcal{P}$ is replaced by a compact, convex set was obtained by Hwang and Rothblum [9].

In Sect. 2 we give edge reductions for the cone of first-order feasible directions and the critical cone. In Sect. 3 these reductions are exploited to obtain a new edge reduced version of the first-order necessary optimality condition, and the second-order sufficient optimality condition. Moreover, when the objective function is edge convex, we obtain a new second-order necessary optimality condition. In Sect. 4 we obtain an estimate on the number of edges of a polyhedron that includes a box constraint. This estimate together with the edge reduced necessary and sufficient optimality conditions imply polynomial complexity for a class of quadratic programs. In Sect. 5 we apply the new necessary and sufficient optimality conditions to a quadratic programming formulation of the vertex separator problem. We obtain easily checked conditions which can be stated in terms of properties of the graph. In Sect. 6 we carry out a similar analysis for the edge separator problem.

Notation $\mathbf{0}$ and $\mathbf{1}$ denote vectors whose entries are all 0 and all 1 respectively, the dimensions should be clear from context. If $\mathbf{x} \in \mathbb{R}^n$, then the support of $\mathbf{x}$ is defined by $\text{supp}(\mathbf{x}) = \{i \in [1, n] : x_i \neq 0\}$. If $f : \mathbb{R}^n \to \mathbb{R}$, then $\nabla f(\mathbf{x})$ denotes the gradient of f at $\mathbf{x}$, a row vector, and $\nabla^2 f(\mathbf{x})$ is the Hessian. The positive span of a set $\mathcal{S}$, denoted $\text{span}_+(\mathcal{S})$, is the set of linear combinations of vectors in $\mathcal{S}$ with nonnegative coefficients. $|\mathcal{S}|$ denotes the number of elements in $\mathcal{S}$. If $\mathbf{A} \in \mathbb{R}^{m \times n}$ is a matrix and $\mathcal{I} \subset \{1, 2, \ldots, m\}$, then $\mathbf{A}_\mathcal{I}$ denotes the submatrix obtained by only including the rows in $\mathcal{I}$. The null space of $\mathbf{A}$ is denoted $\text{null}(\mathbf{A})$.

2 Edge reductions of $\mathcal{C}$ and $\mathcal{F}$

A *face* of the polyhedron $\mathcal{P}$ defined in (1.2) is a non-empty set of the form

$$\mathcal{H} = \{\mathbf{x} \in \mathcal{P} : \mathbf{A}_\mathcal{I}\mathbf{x} = \mathbf{b}_\mathcal{I}\}$$

for some $\mathcal{I} \subset \{1, 2, \ldots, m\}$. The dimension of the face $\mathcal{H}$ is one less than the maximum number of affinely independent points in $\mathcal{H}$ and is denoted $dim(\mathcal{H})$. If $dim(\mathcal{H}) = 0$, then $\mathcal{H}$ is a *vertex* of $\mathcal{P}$. If $dim(\mathcal{H}) = 1$, then $\mathcal{H}$ is an *edge* of $\mathcal{P}$. If $\mathcal{H}$ is an edge and $\mathcal{J}$ is the index set of constraints that are active at more than one point on the edge, then the set of solutions to $\mathbf{A}_\mathcal{J}\mathbf{x} = \mathbf{b}_\mathcal{J}$ is a line containing the edge, and $\text{null}(\mathbf{A}_\mathcal{J})$ is the collection of vectors with the same direction as that of the line. We refer to any nonzero vector $\mathbf{d} \in \text{null}(\mathbf{A}_\mathcal{J})$ as an *edge direction*. Note that if $\mathbf{d}$ is an edge direction for $\mathcal{P}$, then $-\mathbf{d}$ is an edge direction for $\mathcal{P}$. A set $\mathcal{D}$ is an *edge description* of $\mathcal{P}$ if for each edge of $\mathcal{P}$, there is a parallel direction in $\mathcal{D}$. If $-\mathbf{d} \in \mathcal{D}$ when $\mathbf{d} \in \mathcal{D}$, then we say that $\mathcal{D}$ is a *reflective edge description* of $\mathcal{P}$. Edge directions play a central role in our analysis. A polyhedron may have a huge number of vertices or edges, but a relatively small edge description. We will reformulate the first and second-order optimality conditions in terms of edge directions in $\mathcal{F}(\mathbf{x}^*)$ and $\mathcal{C}(\mathbf{x}^*)$.

Lemma 2.1 *Let $\mathcal{P}$ be the polyhedron defined in (1.2) where $\mathbf{A}$ has full column rank, and let $\mathcal{D}$ be a reflective edge description of $\mathcal{P}$. For any $\mathbf{x}^* \in \mathcal{P}$ and $\mathcal{I} \subset \mathcal{A}(\mathbf{x}^*)$, the cone defined by*

$$\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) = \{\mathbf{d} \in \mathcal{F}(\mathbf{x}^*) : \mathbf{A}_{\mathcal{I}}\mathbf{d} = \mathbf{0}\}$$

has the property that

$$\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) = span_+(\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) \cap \mathcal{D}). \tag{2.1}$$

Proof First, suppose that $\mathbf{x}^*$ is a vertex of $\mathcal{P}$, in which case $\mathcal{F}(\mathbf{x}^*)$ contains no lines. Since $\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) \subset \mathcal{F}(\mathbf{x}^*)$, $\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)$ also contains no lines. By a result of Klee [10, Corollary 2.3], there exist vectors $\mathbf{d}^1, \mathbf{d}^2, \ldots, \mathbf{d}^l \in \mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)$ such that for each i, $\mathbf{d}^i$ is an edge-direction of $\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)$ and

$$\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) = \mathrm{span}_+(\mathbf{d}^1, \mathbf{d}^2, \ldots, \mathbf{d}^l). \tag{2.2}$$

We will now show that each edge direction $\mathbf{d}^i$ for $\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)$ is also an edge direction of $\mathcal{P}$. Let $\mathcal{E}$ be an edge of $\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)$ associated with the edge direction $\mathbf{d}^i$. Since $\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)$ is a face of $\mathcal{F}(\mathbf{x}^*)$, $\mathcal{E}$ is also an edge of $\mathcal{F}(\mathbf{x}^*)$. Therefore, $\mathbf{x}^* + \mathcal{E}$ contains an edge of $\mathcal{P}$. Hence, $\mathbf{d}^i$ is also an edge direction of $\mathcal{P}$. Since $\mathcal{D}$ is reflective, it follows that each $\mathbf{d}^i$ is a positive multiple of an element $\boldsymbol{\delta}^i \in \mathcal{D}$. Since $\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)$ is a cone, $\boldsymbol{\delta}^i \in \mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)$. Hence, (2.2) with $\mathbf{d}^i$ replaced by $\boldsymbol{\delta}^i$ implies that (2.1) holds in the case that $\mathbf{x}^*$ is a vertex of $\mathcal{P}$.

Now suppose that $\mathbf{x}^*$ is not a vertex of $\mathcal{P}$. Since $\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)$ is a convex cone, we have

$$\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) = \mathrm{span}_+(\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)) \supset \mathrm{span}_+(\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) \cap \mathcal{D}).$$

We will show the reverse containment. Let $\mathbf{d} \in \mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)$. Consider the following face of $\mathcal{P}$:

$$F = \{\mathbf{x} \in \mathcal{P} : \mathbf{A}_{\mathcal{A}(\mathbf{x}^*)}\mathbf{x} = \mathbf{b}_{\mathcal{A}(\mathbf{x}^*)}\}.$$

Since $rank(\mathbf{A}) = n$, there exists a vertex $\mathbf{v}$ of the polyhedron F. This vertex is found by making a series of moves in the face; each move is in the null space associated with the active constraints while maintaining nonorthogonality to at least one of the rows of $\mathbf{A}$ associated with an inactive constraint. Each move stops at a point where a previously inactive constraint becomes active. Since $rank(\mathbf{A}) = n$, we eventually reach a vertex where n constraints are active. For this vertex we have $\mathcal{A}(\mathbf{v}) \supset \mathcal{A}(\mathbf{x}^*)$ and $rank(\mathbf{A}_{\mathcal{A}(\mathbf{v})}) = n$. Since $\mathbf{d} \in \mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) \subset \mathcal{F}(\mathbf{x}^*)$, there exists some $t > 0$ such that $\mathbf{v} + (\mathbf{x}^* - \mathbf{v}) + t\mathbf{d} \in \mathcal{P}$. This implies that $(\mathbf{x}^* - \mathbf{v}) + t\mathbf{d} \in \mathcal{F}(\mathbf{v})$. Furthermore, $\mathbf{A}_{\mathcal{I}}\mathbf{d} = \mathbf{0}$ since $\mathbf{d} \in \mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)$, and $\mathbf{A}_{\mathcal{I}}(\mathbf{x}^* - \mathbf{v}) = \mathbf{0}$ since $\mathcal{I} \subset \mathcal{A}(\mathbf{x}^*) \subset \mathcal{A}(\mathbf{v})$. Hence $\mathbf{A}_{\mathcal{I}}((\mathbf{x}^* - \mathbf{v}) + t\mathbf{d}) = \mathbf{0}$ and therefore $(\mathbf{x}^* - \mathbf{v}) + t\mathbf{d} \in \mathcal{K}_{\mathcal{I}}(\mathbf{v})$. Thus,

$$t\mathbf{d} \in (\mathbf{v} - \mathbf{x}^*) + \mathcal{K}_{\mathcal{I}}(\mathbf{v}). \tag{2.3}$$

We will show that the right side of (2.3) is contained in $\mathrm{span}_+(\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) \cap \mathcal{D})$.

Since $\mathcal{A}(\mathbf{v}) \supset \mathcal{A}(\mathbf{x}^*)$, we have $\mathcal{F}(\mathbf{v}) \subset \mathcal{F}(\mathbf{x}^*)$, and therefore $\mathcal{K}_{\mathcal{I}}(\mathbf{v}) \subset \mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)$. And since $\mathbf{v}$ is a vertex of F, it is also a vertex of $\mathcal{P}$. Earlier we showed that (2.1) holds at any vertex of $\mathcal{P}$. Therefore, replacing $\mathbf{x}^*$ by $\mathbf{v}$ in (2.1), we have

$$\mathcal{K}_{\mathcal{I}}(\mathbf{v}) = \text{span}_+(\mathcal{K}_{\mathcal{I}}(\mathbf{v}) \cap \mathcal{D}) \subset \text{span}_+(\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) \cap \mathcal{D}). \tag{2.4}$$

The last inclusion is since $\mathcal{A}(\mathbf{x}^*) \subset \mathcal{A}(\mathbf{v})$ which implies that $\mathcal{K}_{\mathcal{I}}(\mathbf{v}) \subset \mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)$.

Next we show that

$$(\mathbf{v} - \mathbf{x}^*) \in \text{span}_+(\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) \cap \mathcal{D}). \tag{2.5}$$

Since $\mathbf{v}$ is a vertex of F, the cone $\mathcal{F}_F(\mathbf{v})$ of first-order feasible directions for F at $\mathbf{v}$ contains no lines. Since $\mathbf{x}^* - \mathbf{v} \in \mathcal{F}_F(\mathbf{v})$, it again follows from the result of Klee [10, Corollary 2.3] that $\mathbf{x}^* - \mathbf{v}$ is a positive linear combination of edge directions $\mathcal{D}' \subset \mathcal{D}$ for F. Since $\mathcal{D}$ is reflective, we may choose $\mathcal{D}'$ to be reflective. Hence, we have

$$\mathbf{v} - \mathbf{x}^* = -(\mathbf{x}^* - \mathbf{v}) \in \text{span}_+(\mathcal{D}'). \tag{2.6}$$

Since any edge-direction of F lies in $\text{null}(\mathbf{A}_{\mathcal{A}(\mathbf{x}^*)}) \subset \mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)$, we have $\mathcal{D}' = \mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) \cap \mathcal{D}'$. Therefore, by (2.6)

$$(\mathbf{v} - \mathbf{x}^*) \in \text{span}_+(\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) \cap \mathcal{D}') \subset \text{span}_+(\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) \cap \mathcal{D}), \tag{2.7}$$

which establishes (2.5).

Combining (2.3) with (2.4) and (2.7), we have $t\mathbf{d} \in \text{span}_+(\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) \cap \mathcal{D})$. Since $t > 0$, this implies

$$\mathbf{d} \in \text{span}_+(\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*) \cap \mathcal{D}).$$

Since $\mathbf{d}$ was an arbitrary vector in $\mathcal{K}_{\mathcal{I}}(\mathbf{x}^*)$, the proof is complete.   □

Corollary 2.2 *Let $\mathcal{P}$ be the polyhedron defined in (1.2) where $\mathbf{A}$ has full column rank, and let $\mathcal{D}$ be a reflective edge description of $\mathcal{P}$. For any $\mathbf{x}^* \in \mathcal{P}$, we have*

$$\mathcal{F}(\mathbf{x}^*) = \text{span}_+(\mathcal{F}(\mathbf{x}^*) \cap \mathcal{D}). \tag{2.8}$$

Furthermore, if $\mathbf{x}^*$ satisfies the first-order optimality conditions (1.5) for problem (1.1), then

$$\mathcal{C}(\mathbf{x}^*) = \text{span}_+(\mathcal{C}(\mathbf{x}^*) \cap \mathcal{D}). \tag{2.9}$$

Proof The identity (2.8) follows immediately from Lemma 2.1 by taking $\mathcal{I} = \emptyset$. If $\mathbf{x}^*$ satisfies (1.5), then define $\mathcal{I} = \{i : \lambda_i > 0\}$. By complementary slackness, $\mathcal{I} \subset \mathcal{A}(\mathbf{x}^*)$. And by the Definition (1.7) for the critical cone, we have

$$\mathcal{C}(\mathbf{x}^*) = \mathcal{K}_{\mathcal{I}}(\mathbf{x}^*).$$

Hence, (2.9) also follows from Lemma 2.1.   □

The identities (2.8) and (2.9) are related to a *conformal circuit decomposition* [9,17]. Consider the polyhedron $\mathcal{Q}$ defined by

$$\mathcal{Q} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{Bx} = \mathbf{b} \text{ and } \mathbf{x} \geq 0\},$$

where $\mathbf{B} \in \mathbb{R}^{m \times n}$ and $\mathbf{b} \in \mathbb{R}^m$. For any $\mathbf{x} \in \mathcal{Q}$ we have

$$\mathcal{F}(\mathbf{x}) = \{\mathbf{d} \in \mathbb{R}^n : \mathbf{Bd} = \mathbf{0} \text{ and } d_j \geq 0 \text{ whenever } x_j = 0\}.$$

A *circuit* of the matrix $\mathbf{B}$ is a non-zero vector $\mathbf{d} \in \text{null}(\mathbf{B})$ such that $||\mathbf{d}||_\infty = 1$, and $supp(\mathbf{d})$ is inclusion-minimal. That is, if $\mathbf{d}' \in \text{null}(\mathbf{B})$ and $\text{supp}(\mathbf{d}') \subset \text{supp}(\mathbf{d})$, then $\mathbf{d}' = \alpha\mathbf{d}$ for some scalar α. As pointed out in [15], it follows from [17, ex. 10.14, p. 506] that any non-zero vector $\mathbf{d} \in \text{null}(\mathbf{B})$ has a *conformal circuit decomposition*; that is, there exist scalars $\alpha_i > 0$ and circuits $\mathbf{d}^i$ of $\mathbf{B}$ such that

$$\mathbf{d} = \sum_i \alpha_i \mathbf{d}^i, \tag{2.10}$$

where for each i we have $d_j^i d_j > 0$ for all $j \in \text{supp}(\mathbf{d}^i)$. Hence if $\mathbf{d} \in \mathcal{F}(\mathbf{x})$, then $\mathbf{d}^i \in \mathcal{F}(\mathbf{x})$ for each i. If every circuit of $\mathbf{B}$ is parallel to an edge of $\mathcal{Q}$, then (2.8) follows from the decomposition (2.10). A similar argument proves (2.9). However, it is not true in general that every circuit of $\mathbf{B}$ is parallel to an edge of $\mathcal{Q}$ (see [9]), and conversely, some edges may not be circuits. Our proof of (2.8) and (2.9) is valid for any polyhedron $\mathcal{P}$ of the form (1.2), as long as the matrix $\mathbf{A}$ has full column rank. When the columns of $\mathbf{A}$ are linearly dependent, Lemma 2.1 may not hold; for example, consider the case where $\mathcal{P}$ is a half-space.

The assumption that $\mathbf{A}$ has full column rank in Lemma 2.1 is used to ensure the existence of an extreme point (a vertex) of $\mathcal{P}$. However, Corollary 2.2 is false if the polyhedron $\mathcal{P}$ is replaced by a general convex set containing an extreme point. For example, consider the unit sphere

$$\mathcal{X} = \left\{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\right\}.$$

Every point on the boundary of $\mathcal{X}$ is an extreme point. However, $\mathcal{X}$ contains no extreme directions. Therefore $\mathcal{D} = \emptyset$. So (2.8) would imply that $\mathcal{F}(\mathbf{x}^*) = \emptyset$ for every $\mathbf{x}^* \in \mathcal{X}$. This is clearly false; for instance if $\mathbf{x}^* = (-1, 0)$, then $\mathcal{F}(\mathbf{x}^*) = \{\mathbf{d} \in \mathbb{R}^2 : d_1 \geq 0\}$. Thus, Lemma 2.1 and Corollary 2.2 are properties of polyhedra, not general convex sets.

3 Edge reductions of optimality conditions

In this section we focus on edge reductions of the first and second-order optimality conditions. We start with the first-order optimality conditions:

Proposition 3.1 *If $f : \mathcal{P} \to \mathbb{R}$ is continuously differentiable, $\mathbf{A} \in \mathbb{R}^{m \times n}$ has full column rank, and $\mathcal{D}$ is a reflective edge description of $\mathcal{P}$, then $\mathbf{x}^*$ satisfies the first-order optimality conditions (1.4) if and only if*

$$\mathbf{x}^* \in \mathcal{P} \quad and \quad \nabla f(\mathbf{x}^*)\mathbf{d} \leq 0 \text{ for every } \mathbf{d} \in \mathcal{F}(\mathbf{x}^*) \cap \mathcal{D}. \tag{3.1}$$

Proof Obviously, (1.4) implies (3.1). Conversely, suppose $\mathbf{x}^*$ satisfies (3.1) and let $\mathbf{d} \in \mathcal{F}(\mathbf{x}^*)$. By Corollary 2.2, we have

$$\mathbf{d} = \sum_{i=1}^{k} \alpha_i \mathbf{d}^i,$$

for some vectors $\mathbf{d}^i \in \mathcal{F}(\mathbf{x}^*) \cap \mathcal{D}$ and some scalars $\alpha_i > 0$. Thus (3.1) implies

$$\nabla f(\mathbf{x}^*)\mathbf{d} = \sum_{i=1}^{k} \alpha_i \nabla f(\mathbf{x}^*)\mathbf{d}^i \leq 0,$$

which yields (1.4). □

Since the conditions (1.4), (1.5), and (3.1) are all equivalent, we will refer to these conditions collectively as the "first-order optimality conditions" for (1.1). Next, we consider second-order optimality conditions.

Proposition 3.2 *Suppose $f : \mathcal{P} \to \mathbb{R}$ is twice continuously differentiable, $\mathbf{A}$ has full column rank, $\mathcal{D}$ is a reflective edge description of $\mathcal{P}$, and $\mathbf{x}^*$ satisfies the first-order optimality conditions.*

1. $\mathbf{x}^*$ is a local maximizer of (1.1) if

$$(\mathbf{d}^1)^T \nabla^2 f(\mathbf{x}^*)\mathbf{d}^2 < 0 \text{ for every } \mathbf{d}^1, \mathbf{d}^2 \in \mathcal{C}(\mathbf{x}^*) \cap \mathcal{D}. \tag{3.2}$$

If f is quadratic, then the strict inequality in (3.2) can be replaced by an inequality.
2. If $\mathbf{x}^*$ is a local maximizer of (1.1) and

$$\mathbf{d}^T \nabla^2 f(\mathbf{x}^*)\mathbf{d} \geq 0 \quad for \text{ every } \mathbf{d} \in \mathcal{D}, \tag{3.3}$$

then

$$(\mathbf{d}^1)^T \nabla^2 f(\mathbf{x}^*)\mathbf{d}^2 \leq 0 \quad for \text{ every } \mathbf{d}^1, \mathbf{d}^2 \in \mathcal{C}(\mathbf{x}^*) \cap \mathcal{D}. \tag{3.4}$$

Proof Suppose that $\mathbf{x}^*$ satisfies the first-order optimality conditions. To prove part 1, we need only show that (3.2) implies the usual second-order sufficient condition (1.9). By Corollary 2.2, for any $\mathbf{d} \in \mathcal{C}(\mathbf{x}^*)$ we have

$$\mathbf{d} = \sum_{i=1}^{k} \alpha_i \mathbf{d}^i,$$

for some vectors $\mathbf{d}^i \in \mathcal{C}(\mathbf{x}^*) \cap \mathcal{D}$ and some scalars $\alpha_i > 0$. Therefore

$$\mathbf{d}^\mathsf{T} \nabla^2 f(\mathbf{x}) \mathbf{d} = \left(\sum_{i=1}^{k} \alpha_i \mathbf{d}^i \right)^\mathsf{T} \nabla^2 f(\mathbf{x}) \left(\sum_{i=1}^{k} \alpha_i \mathbf{d}^i \right) = \sum_{i,j=1}^{k} \alpha_i \alpha_j \mathbf{d}^i \nabla^2 f(\mathbf{x}) \mathbf{d}^j < 0,$$

if (3.2) holds. Consequently, (1.9) holds and $\mathbf{x}^*$ is a local maximizer.

If f is quadratic and the strict inequality (3.2) is replaced by an inequality, then the same argument yields (1.6), which is sufficient for local optimality when f is quadratic [2,3]. This completes the proof of part 1.

Next, we consider part 2. The second-order necessary condition (1.6) states that

$$\mathbf{d}^\mathsf{T} \nabla^2 f(\mathbf{x}^*) \mathbf{d} \leq 0 \quad \text{for every } \mathbf{d} \in \mathcal{C}(\mathbf{x}^*). \tag{3.5}$$

If (3.3) holds, then

$$\mathbf{d}^\mathsf{T} \nabla^2 f(\mathbf{x}^*) \mathbf{d} = 0 \quad \text{for every } \mathbf{d} \in \mathcal{C}(\mathbf{x}^*) \cap \mathcal{D}. \tag{3.6}$$

Since $\mathcal{C}(\mathbf{x}^*)$ is a convex cone, it follows that for any $\mathbf{d}^1, \mathbf{d}^2 \in \mathcal{C}(\mathbf{x}^*) \cap \mathcal{D}$, we have $\mathbf{d}^1 + \mathbf{d}^2 \in \mathcal{C}(\mathbf{x}^*)$. Therefore, (3.5) and (3.6) imply

$$0 \geq (\mathbf{d}^1 + \mathbf{d}^2)^\mathsf{T} \nabla^2 f(\mathbf{x}^*) (\mathbf{d}^1 + \mathbf{d}^2) = 2(\mathbf{d}^1)^\mathsf{T} \nabla^2 f(\mathbf{x}^*) \mathbf{d}^2,$$

which yields (3.4). □

The following corollary is an immediate consequence of Proposition 3.2.

Corollary 3.3 *Let $\mathcal{P}$ be the polyhedron defined in (1.2) where $\mathbf{A}$ has full column rank, and let $\mathcal{D}$ be a reflective edge description of $\mathcal{P}$. If f is quadratic and (3.3) holds, then $\mathbf{x}^*$ is a local maximizer of (1.1) if and only if (3.1) (the first-order condition) and (3.4) (the second-order condition) hold.*

Remark We now observe that if f is quadratic, (3.3) holds, and $\mathbf{x}^*$ satisfies the first-order optimality condition (3.1), then when any of the conditions (3.4) are violated, there is a easily computed ascent direction. In particular, if $\mathbf{Q} = \nabla^2 f$ then by a Taylor expansion, we have

$$f(\mathbf{x}^* + \alpha \mathbf{d}) = f(\mathbf{x}^*) + \frac{1}{2} \alpha^2 \mathbf{d}^\mathsf{T} \mathbf{Q} \mathbf{d} \tag{3.7}$$

for any $\mathbf{d} \in \mathcal{C}(\mathbf{x}^*)$ and $\alpha > 0$, since $\nabla f(\mathbf{x}^*)\mathbf{d} = 0$ by (1.7). If $\mathbf{d}^\mathsf{T}\mathbf{Q}\mathbf{d} > 0$ for some $\mathbf{d} \in \mathcal{C}(\mathbf{x}^*) \cap \mathcal{D}$, then $\mathbf{d}$ is an ascent direction. Otherwise, $\mathbf{d}^\mathsf{T}\mathbf{Q}\mathbf{d} \leq 0$ for every $\mathbf{d} \in \mathcal{C}(\mathbf{x}^*) \cap \mathcal{D}$. So by (3.3), $\mathbf{d}^\mathsf{T}\mathbf{Q}\mathbf{d} = 0$ for every $\mathbf{d} \in \mathcal{C}(\mathbf{x}^*) \cap \mathcal{D}$. If $(\mathbf{d}^1)^\mathsf{T}\mathbf{Q}\mathbf{d}^2 > 0$ for some $\mathbf{d}^1$ and $\mathbf{d}^2 \in \mathcal{C}(\mathbf{x}^*) \cap \mathcal{D}$, then $(\mathbf{d}^i)^\mathsf{T}\mathbf{Q}\mathbf{d}^i = 0$ for $i = 1, 2$. Since $\mathbf{d}^1 + \mathbf{d}^2 \in \mathcal{C}(\mathbf{x}^*)$, it follows from (3.7) that

$$f\left(\mathbf{x}^* + \alpha\left(\mathbf{d}^1 + \mathbf{d}^2\right)\right) = f\left(\mathbf{x}^*\right) + \alpha^2\left(\mathbf{d}^1\right)^\mathsf{T}\mathbf{Q}\mathbf{d}^2 > f\left(\mathbf{x}^*\right)$$

for any $\alpha > 0$. This shows that $\mathbf{d} = \mathbf{d}^1 + \mathbf{d}^2 \in \mathcal{C}(\mathbf{x}^*)$ is an ascent direction.

4 A complexity result

We define a *minimal edge description* of $\mathcal{P}$ to be an edge description which does not properly contain any other edge description. Under the assumptions of Corollary 3.3, the computational complexity of checking whether a given point is a local maximizer is proportional to the size of a minimal edge description squared. Suppose the number of columns of $\mathbf{A}$ is held fixed, but the number of rows of $\mathbf{A}$ is allowed to vary. A trivial upper bound on the size of a minimal edge description is $\binom{m}{n-1}$. Since this is a polynomial in m when n is fixed, it follows that local optimality can be checked in polynomial in m time.

If m is fixed and n is allowed to vary, then we cannot apply Corollary 3.3 since the columns of $\mathbf{A}$ are linearly dependent when $n > m$. On the other hand, we can apply Corollary 3.3 when the polyhedron includes a box constraint:

$$\mathcal{P}_B = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A}\mathbf{x} \leq \mathbf{b} \text{ and } \boldsymbol{\ell} \leq \mathbf{x} \leq \mathbf{u}\}. \tag{4.1}$$

Here $\mathbf{A} \in \mathbb{R}^{m \times n}$ is any matrix, $\boldsymbol{\ell} \in \mathbb{R}^n$, and $\mathbf{u} \in \mathbb{R}^n$. Due to the box constraint, the constraint matrix associated with the polyhedron has full column rank. Hence, we consider the optimization problem

$$\max f(\mathbf{x}) \text{ subject to } \mathbf{x} \in \mathcal{P}_B. \tag{4.2}$$

We now show that for the box constrained polyhedron where m is fixed and n is allowed to increase, the size of a minimal edge description is bounded by a polynomial in n.

Corollary 4.1 *If m is held fixed, $f : \mathbb{R}^n \to \mathbb{R}$ is quadratic, and for some edge description $\mathcal{D}$ of $\mathcal{P}_B$, we have*

$$\mathbf{d}^T \nabla^2 f \mathbf{d} \geq 0 \quad \text{for every } \mathbf{d} \in \mathcal{D},$$

then local optimality of any feasible point in problem (4.2) can be checked in time that is bounded by a polynomial in n.

Proof To prove this result, we obtain a bound on the size of a minimal edge description for $\mathcal{P}_B$. Any edge of $\mathcal{P}_B$ is contained in the solution set of a linear system of the form

$\mathbf{M}\mathbf{x} = \mathbf{c}$ where $\mathbf{M} \in \mathbb{R}^{(n-1) \times n}$ has rank $n - 1$. An edge direction is any nonzero vector in null$(\mathbf{M})$; $\mathbf{M}$ will be called the *edge matrix*. Note that multiplying a row of $\mathbf{M}$ and the corresponding component of $\mathbf{c}$ by -1 has no effect on either null$(\mathbf{M})$ or the solution set of $\mathbf{M}\mathbf{x} = \mathbf{c}$. When building an edge matrix corresponding to some edge of $\mathcal{P}_B$, up to m of the rows of $\mathbf{M}$ can be taken from rows of $\mathbf{A}$ while the remaining rows are taken from the identity matrix; these latter rows correspond to the constraints $x_i \geq \ell_i$ or $x_i \leq u_i$ that are active on the edge. For an edge along which either $x_i = \ell_i$ or $x_i = u_i$, the corresponding row of the edge matrix will be the ith row of the identity matrix. Hence, when constructing the edge matrix, it does not matter whether x_i is at the upper bound or at the lower bound, all that matters is whether the ith box constraint is active. For $n > m$, an upper bound on the number of edge directions is given by the expression

$$\sum_{i=1}^{m} \binom{n}{n-1-i} \binom{m}{i}. \tag{4.3}$$

Here i represents the number of rows of $\mathbf{A}$ to be inserted into the edge matrix and $n - 1 - i$ is the number of rows of the identity matrix to be inserted into the edge matrix. There are $\binom{m}{i}$ different collections of rows of $\mathbf{A}$ that could be placed in the edge matrix. And for each selection of the rows in $\mathbf{A}$, another $n - 1 - i$ rows are selected from the identity matrix. There are $\binom{n}{n-1-i}$ different collections of rows from the identity matrix. Since m is fixed, the expression (4.3) is a polynomial in n. $\qquad \square$

Note that we could have either $\ell_i = -\infty$ or $u_i = +\infty$, but not both, and Corollary 4.1 still holds. The reason is that the columns of the constraint matrix remain linearly independent when either one of the constraints $\ell_i \leq x_i$ or $x_i \leq u_i$ is dropped (but not both).

In [18] Tardella gives a condition that ensures edge convexity for a function defined over a polyhedron $\mathcal{P}_B$ that includes a box constraint. If $f : \mathcal{X} \to \mathbb{R}$ where $\mathcal{X} \subset \mathbb{R}^n$ is a convex set, then f is $k-convex$ over $\mathcal{X}$ if $f(\alpha \mathbf{x} + (1-\alpha)\mathbf{y}) \leq \alpha f(\mathbf{x}) + (1-\alpha) f(\mathbf{y})$ for every $\alpha \in [0, 1]$ and for every $\mathbf{x}, \mathbf{y} \in \mathcal{X}$ such that $x_i = y_i$ for at least $n - k$ indices $i \in [1, n]$. With this terminology, Tardella's result is as follows:

Proposition 4.2 *Suppose $f : \mathcal{P} \to \mathbb{R}$ is twice continuously differentiable, $rank(\mathbf{A}) = k - 1$ for some k, and $\mathcal{D}$ is an edge description of $\mathcal{P}_B$. If f is $k - convex$ over $\mathcal{P}_B$, then*

$$\mathbf{d}^T \nabla^2 f(\mathbf{x}^*) \mathbf{d} \geq 0 \quad \text{for every } \mathbf{d} \in \mathcal{D}.$$

We combine Corollary 3.3, 4.1, and Proposition 4.2 to obtain the following result:

Corollary 4.3 *If $rank(\mathbf{A}) = k - 1$ for some k, the objective function f is quadratic, f is $k - convex$ over $\mathcal{P}_B$, and $\mathcal{D}$ is a reflective edge description of $\mathcal{P}_B$, then a feasible point $\mathbf{x}^*$ of problem (4.2) is locally optimal if and only if (3.1) and (3.4) hold. Hence, under these assumptions, local optimality in (4.2) can be checked in time bounded by a polynomial in n.*

In the next two sections, we apply Corollary 3.3 to two box constrained quadratic programs arising in discrete optimization, the vertex and edge separator problems. We show that these programs satisfy the edge-convexity condition (3.3), from which it follows that a vertex maximizer exists [18] and local optimality can be checked in polynomial time (Corollary 4.1).

5 Vertex separator problem

Let G be an undirected graph on vertex set $\mathcal{V} = \{1, 2, \ldots, n\}$ with nonnegative vertex weights $c_1, c_2, \ldots, c_n$, not all zero. Here we use the word *vertex* to refer to a node in a graph and *edge* to refer to a connection between two vertices in a graph. While these words were also defined in the context of polyhedra, the intended meaning should always be clear from context. Let $\mathbf{A}$ be the adjacency matrix for G; that is, $a_{ij} = 1$ if there is an edge between vertices i and j and $a_{ij} = 0$ otherwise. Given positive integers $\ell_a \leq u_a$ and $\ell_b \leq u_b$, the *vertex separator problem* is to partition the vertices of G into three disjoint sets $\mathcal{A}$, $\mathcal{B}$, and $\mathcal{S}$ such that the following conditions are satisfied:

1. There are no edges between the sets $\mathcal{A}$ and $\mathcal{B}$.
2. The size constraints $\ell_a \leq |\mathcal{A}| \leq u_a$ and $\ell_b \leq |\mathcal{B}| \leq u_b$ are satisfied.
3. The sum of the weights of vertices in $\mathcal{S}$ is minimal.

In [7] Hager and Hungerford develop an equivalence between the vertex separator problem and the following quadratic program:

$$\max_{\mathbf{x}, \mathbf{y} \in \mathbb{R}^n} \quad \mathbf{c}^\mathsf{T}(\mathbf{x} + \mathbf{y}) - \gamma \mathbf{x}^\mathsf{T}(\mathbf{A} + \mathbf{I})\mathbf{y} \tag{5.1}$$

subject to $\mathbf{0} \leq \mathbf{x} \leq \mathbf{1}$, $\mathbf{0} \leq \mathbf{y} \leq \mathbf{1}$, $\ell_a \leq \mathbf{1}^\mathsf{T}\mathbf{x} \leq u_a$, and $\ell_b \leq \mathbf{1}^\mathsf{T}\mathbf{y} \leq u_b$.

Here γ is any scalar greater than or equal to $\max\{c_i : i = 1, 2, \ldots, n\}$. Given any solution to (5.1), there exists an easily computed binary solution $(\mathbf{x}, \mathbf{y})$ such that an optimal partition is

$$\mathcal{A} = \{i : x_i = 1\}, \quad \mathcal{B} = \{i : y_i = 1\}, \quad \text{and} \quad \mathcal{S} = \{i : x_i = y_i = 0\}. \tag{5.2}$$

Let f be the objective function in (5.1):

$$f(\mathbf{x}, \mathbf{y}) = \mathbf{c}^\mathsf{T}(\mathbf{x} + \mathbf{y}) - \gamma \mathbf{x}^\mathsf{T}(\mathbf{A} + \mathbf{I})\mathbf{y}.$$

We will apply the first and second-order optimality conditions of Sect. 3 to the vertex separator problem. Since (5.1) involves 4 lower bounds and 4 upper bounds, the standard statement of the KKT conditions (1.5) involves 8 multipliers, 8 inequality constraints, and 8 complementary slackness conditions. A more compact way of expressing these 16 conditions is as follows: If $(\mathbf{x}, \mathbf{y})$ is a local maximizer of (5.1), then there exist multipliers $\boldsymbol{\mu}^a$ and $\boldsymbol{\mu}^b \in \mathbb{R}^n$ and λ^a and $\lambda^b \in \mathbb{R}$ such that

$$\nabla f(\mathbf{x}, \mathbf{y}) + \boldsymbol{\mu}^a + \boldsymbol{\mu}^b + \lambda^a \mathbf{1} + \lambda^b \mathbf{1} = \mathbf{0}, \tag{5.3}$$

where

$$\boldsymbol{\mu}^a \in \mathcal{M}(\mathbf{x}), \quad \boldsymbol{\mu}^b \in \mathcal{M}(\mathbf{y}), \quad \lambda^a \in \mathcal{L}(\mathbf{x}, \ell_a, u_a), \quad \lambda^b \in \mathcal{L}(\mathbf{y}, \ell_b, u_b)$$

with

$$\mathcal{M}(\mathbf{z}) = \{\boldsymbol{\mu} \in \mathbb{R}^n : \mu_i z_i \geq \max\{\mu_i, 0\} \text{ for all } 1 \leq i \leq n\},$$

and

$$\mathcal{L}(\mathbf{z}, \ell, u) = \left\{ \lambda \in \mathbb{R} : \lambda \mathbf{1}^\mathsf{T} \mathbf{z} \geq \max\{\lambda u, \lambda \ell\} \right\}.$$

According to [7], the following set is a reflective edge description of the polyhedron associated with (5.1):

$$\mathcal{D} = \bigcup_{\substack{i,j=1 \\ i \neq j}}^n \{\pm[\mathbf{e}_i, \mathbf{0}], \ \pm[\mathbf{0}, \mathbf{e}_i], \ [\mathbf{0}, \mathbf{e}_i - \mathbf{e}_j], \ [\mathbf{e}_i - \mathbf{e}_j, \mathbf{0}]\}.$$

Since

$$\nabla^2 f = \begin{pmatrix} \mathbf{0} & -\gamma(\mathbf{A} + \mathbf{I}) \\ -\gamma(\mathbf{A} + \mathbf{I}) & \mathbf{0} \end{pmatrix},$$

it can be checked that $\mathbf{d}^\mathsf{T}(\nabla^2 f)\mathbf{d} \geq 0$ for every $\mathbf{d} \in \mathcal{D}$. Therefore, by Corollary 3.3, a feasible point $(\mathbf{x}, \mathbf{y})$ of (5.1) is a local maximizer if and only if the first-order optimality conditions (5.3) hold and

$$(\mathbf{d}^1)^\mathsf{T}(\nabla^2 f)\mathbf{d}^2 \leq 0 \quad \text{for every } \mathbf{d}^1, \ \mathbf{d}^2 \in \mathcal{C}(\mathbf{x}, \mathbf{y}) \cap \mathcal{D}. \tag{5.4}$$

Table 1 gives all the different possible values for $(\mathbf{d}^1)^\mathsf{T}(\nabla^2 f)\mathbf{d}^2$, where $\mathbf{d}^1$ and $\mathbf{d}^2$ are edge directions, in terms of $\mathbf{H} = \mathbf{A} + \mathbf{I}$. Since $\mathcal{D}$ is described in terms of 6 different vectors, there are 36 products $(\mathbf{d}^1)^\mathsf{T}(\nabla^2 f)\mathbf{d}^2$ corresponding to the 6×6 different pairs among the vectors describing $\mathcal{D}$. However, 15 of these products are known by symmetry. The remaining 21 products are shown in Table 1. The blank entries correspond to entries known from symmetry.

Suppose $(\mathbf{x}, \mathbf{y})$ satisfies the first-order optimality conditions (5.3). The condition (5.4) states that $(\mathbf{x}, \mathbf{y})$ is a local maximizer if and only if the entries of Table 1 are non-positive whenever the vectors in the corresponding row and column lie in $\mathcal{C}(\mathbf{x}, \mathbf{y})$. For example, if $(\mathbf{x}, \mathbf{y})$ is a local maximizer, then whenever $(\mathbf{e}_i, \mathbf{0})$ and $(\mathbf{0}, -\mathbf{e}_k)$ lie in the cone $\mathcal{C}(\mathbf{x}, \mathbf{y})$ for some i and k, we must have $h_{ik} \leq 0$ (since $\gamma > 0$). This implies that $a_{ik} = 0$. Moreover, since $h_{ii} = 1$, it follows that both $(\mathbf{e}_i, \mathbf{0})$ and $(\mathbf{0}, -\mathbf{e}_i)$ cannot be contained in $\mathcal{C}(\mathbf{x}, \mathbf{y})$ at a local maximizer.

Table 1 Values of $(\mathbf{d}^1)^{\mathsf{T}}(\nabla^2 f)\mathbf{d}^2$ for $\mathbf{d}^1, \mathbf{d}^2 \in \mathcal{D}$

	$(\mathbf{e}_k, \mathbf{0})$	$(\mathbf{0}, \mathbf{e}_k)$	$(-\mathbf{e}_k, \mathbf{0})$	$(\mathbf{0}, -\mathbf{e}_k)$
$(\mathbf{e}_i, \mathbf{0})$	0	$-\gamma h_{ik}$	0	γh_{ik}
$(\mathbf{0}, \mathbf{e}_i)$		0	γh_{ik}	0
$(-\mathbf{e}_i, \mathbf{0})$			0	$-\gamma h_{ik}$
$(\mathbf{0}, -\mathbf{e}_i)$				0
	$(\mathbf{e}_k - \mathbf{e}_l, \mathbf{0})$			$(\mathbf{0}, \mathbf{e}_k - \mathbf{e}_l)$
$(\mathbf{e}_i, \mathbf{0})$	0			$-\gamma(h_{ki} - h_{li})$
$(\mathbf{0}, \mathbf{e}_i)$	$-\gamma(h_{ki} - h_{li})$			0
$(-\mathbf{e}_i, \mathbf{0})$	0			$-\gamma(h_{li} - h_{ki})$
$(\mathbf{0}, -\mathbf{e}_i)$	$-\gamma(h_{li} - h_{ki})$			0
$(\mathbf{e}_i - \mathbf{e}_j, \mathbf{0})$	0			$-\gamma(h_{ik} - h_{il} - h_{jk} + h_{jl})$
$(\mathbf{0}, \mathbf{e}_i - \mathbf{e}_j)$				0

In order to verify either the first or second-order optimality conditions, we need to determine when a vector in the edge description $\mathcal{D}$ lies in $\mathcal{F}(\mathbf{x}, \mathbf{y})$ or $\mathcal{C}(\mathbf{x}, \mathbf{y})$. From the Definitions (1.3) and (1.7), we have

$$
\mathcal{F}(\mathbf{x}, \mathbf{y}) = \left\{ (\mathbf{d}^1, \mathbf{d}^2) \in \mathbb{R}^{2n} : \begin{array}{l} \mathbf{1}^{\mathsf{T}}\mathbf{d}^1 \le 0 \text{ if } \mathbf{1}^{\mathsf{T}}\mathbf{x} = u^a, \ d_i^1 \le 0 \ \forall \ i \text{ with } x_i = 1, \\ \mathbf{1}^{\mathsf{T}}\mathbf{d}^1 \ge 0 \text{ if } \mathbf{1}^{\mathsf{T}}\mathbf{x} = \ell^a, \ d_i^1 \ge 0 \ \forall \ i \text{ with } x_i = 0, \\ \mathbf{1}^{\mathsf{T}}\mathbf{d}^2 \le 0 \text{ if } \mathbf{1}^{\mathsf{T}}\mathbf{y} = u^b, \ d_i^2 \le 0 \ \forall \ i \text{ with } y_i = 1, \\ \mathbf{1}^{\mathsf{T}}\mathbf{d}^2 \ge 0 \text{ if } \mathbf{1}^{\mathsf{T}}\mathbf{y} = \ell^b, \ d_i^2 \ge 0 \ \forall \ i \text{ with } y_i = 0 \end{array} \right\},
$$

$$
\mathcal{C}(\mathbf{x}, \mathbf{y}) = \left\{ \mathbf{d} \in \mathcal{F}(\mathbf{x}, \mathbf{y}) : \begin{array}{l} \mathbf{1}^{\mathsf{T}}\mathbf{d}^1 = 0 \text{ if } \lambda^a \ne 0, \ d_i = 0 \ \forall \ i \text{ with } \mu_i^a \ne 0 \\ \mathbf{1}^{\mathsf{T}}\mathbf{d}^2 = 0 \text{ if } \lambda^b \ne 0, \ d_i = 0 \ \forall \ i \text{ with } \mu_i^b \ne 0 \end{array} \right\}.
$$

Now define the following sets:

$$
\mathcal{A}_0 = \left\{ i : x_i > 0, \ \mu_i^a = 0 \right\} \quad \text{and} \quad \bar{\mathcal{A}}_0 = \left\{ i : x_i < 1, \ \mu_i^a = 0 \right\},
$$
$$
\mathcal{B}_0 = \left\{ i : y_i > 0, \ \mu_i^b = 0 \right\} \quad \text{and} \quad \bar{\mathcal{B}}_0 = \left\{ i : y_i < 1, \ \mu_i^b = 0 \right\}.
$$

Table 2 shows when each element of $\mathcal{D}$ also lies in $\mathcal{C}(\mathbf{x}, \mathbf{y})$. Combining the information in Tables 1 and 2, we will establish the following theorem.

Theorem 5.1 *If* $(\mathbf{x}, \mathbf{y})$ *is feasible in (5.1), then* $(\mathbf{x}, \mathbf{y})$ *is a local maximizer of (5.1) if and only if* (V1)–(V5) *hold:*

(V1) *The first-order conditions (5.3) hold.*
(V2) *Suppose* $\lambda^a = 0$.
    a. If $\mathbf{1}^{\mathsf{T}}\mathbf{x} < u_a$, $i \in \bar{\mathcal{A}}_0$, $j \in \bar{\mathcal{B}}_0$, *and* $k \in \mathcal{B}_0$, *then* $h_{ij} \ge h_{ik}$.
    b. If $\mathbf{1}^{\mathsf{T}}\mathbf{x} > \ell_a$, $i \in \mathcal{A}_0$, $j \in \mathcal{B}_0$, *and* $k \in \bar{\mathcal{B}}_0$, *then* $h_{ij} \ge h_{ik}$.

Table 2 A description of $\mathcal{D} \cap \mathcal{C}(\mathbf{x}, \mathbf{y})$

$(\mathbf{e}_i, \mathbf{0}) \in \mathcal{C}(\mathbf{x}, \mathbf{y})$	$\Leftrightarrow$	$\mathbf{1}^{\mathsf{T}}\mathbf{x} < u_a$ $x_i < 1$ $\mu_i^a = \lambda^a = 0$	$\Leftrightarrow$	$\mathbf{1}^{\mathsf{T}}\mathbf{x} < u_a$ $\lambda^a = 0$ $i \in \bar{\mathcal{A}}_0$
$(\mathbf{0}, \mathbf{e}_i) \in \mathcal{C}(\mathbf{x}, \mathbf{y})$	$\Leftrightarrow$	$\mathbf{1}^{\mathsf{T}}\mathbf{y} < u_b$ $y_i < 1$ $\mu_i^b = \lambda^b = 0$	$\Leftrightarrow$	$\mathbf{1}^{\mathsf{T}}\mathbf{y} < u_b$ $\lambda^b = 0$ $i \in \bar{\mathcal{B}}_0$
$(-\mathbf{e}_i, \mathbf{0}) \in \mathcal{C}(\mathbf{x}, \mathbf{y})$	$\Leftrightarrow$	$\mathbf{1}^{\mathsf{T}}\mathbf{x} > \ell_a$ $x_i > 0$ $\mu_i^a = \lambda^a = 0$	$\Leftrightarrow$	$\mathbf{1}^{\mathsf{T}}\mathbf{x} > \ell_a$ $\lambda^a = 0$ $i \in \mathcal{A}_0$
$(\mathbf{0}, -\mathbf{e}_i) \in \mathcal{C}(\mathbf{x}, \mathbf{y})$	$\Leftrightarrow$	$\mathbf{1}^{\mathsf{T}}\mathbf{y} > \ell_b$ $y_i > 0$ $\mu_i^b = \lambda^b = 0$	$\Leftrightarrow$	$\mathbf{1}^{\mathsf{T}}\mathbf{y} > \ell_b$ $\lambda^b = 0$ $i \in \mathcal{B}_0$
$(\mathbf{e}_i - \mathbf{e}_j, \mathbf{0}) \in \mathcal{C}(\mathbf{x}, \mathbf{y})$	$\Leftrightarrow$	$x_i < 1 \text{ and } x_j > 0$ $\mu_i^a = \mu_j^a = 0$	$\Leftrightarrow$	$i \in \bar{\mathcal{A}}_0 \text{ and } j \in \mathcal{A}_0$
$(\mathbf{0}, \mathbf{e}_i - \mathbf{e}_j) \in \mathcal{C}(\mathbf{x}, \mathbf{y})$	$\Leftrightarrow$	$y_i < 1 \text{ and } y_j > 0$ $\mu_i^b = \mu_j^b = 0$	$\Leftrightarrow$	$i \in \bar{\mathcal{B}}_0 \text{ and } j \in \mathcal{B}_0$

(V3) *Suppose $\lambda^b = 0$.*
    a. *If $\mathbf{1}^{\mathsf{T}}\mathbf{y} < u_b$, $i \in \bar{\mathcal{B}}_0$, $j \in \bar{\mathcal{A}}_0$, and $k \in \mathcal{A}_0$, then $h_{ij} \geq h_{ik}$.*
    b. *If $\mathbf{1}^{\mathsf{T}}\mathbf{y} > \ell_b$, $i \in \mathcal{B}_0$, $j \in \mathcal{A}_0$, and $k \in \bar{\mathcal{A}}_0$, then $h_{ij} \geq h_{ik}$.*
(V4) *Suppose $\lambda^a = \lambda^b = 0$.*
    a. *If $\mathbf{1}^{\mathsf{T}}\mathbf{x} > \ell_a$ and $\mathbf{1}^{\mathsf{T}}\mathbf{y} < u_b$, then $\mathcal{A}_0 \cap \bar{\mathcal{B}}_0 = \emptyset$ and $h_{ij} = 0$ whenever $i \in \mathcal{A}_0$*
      and $j \in \bar{\mathcal{B}}_0$.
    b. *If $\mathbf{1}^{\mathsf{T}}\mathbf{x} < u_a$ and $\mathbf{1}^{\mathsf{T}}\mathbf{y} > \ell_b$, then $\bar{\mathcal{A}}_0 \cap \mathcal{B}_0 = \emptyset$ and $h_{ij} = 0$ whenever $i \in \bar{\mathcal{A}}_0$*
      and $j \in \mathcal{B}_0$.
(V5) *If $i \in \bar{\mathcal{A}}_0$, $j \in \mathcal{A}_0$, $k \in \bar{\mathcal{B}}_0$, and $l \in \mathcal{B}_0$, then $h_{ik} + h_{jl} \geq h_{il} + h_{jk}$.*

Proof First, suppose that the conditions (V1)–(V5) are satisfied. We wish to show that (5.4) holds, which implies that $(\mathbf{x}, \mathbf{y})$ is a local maximizer. Consider any entry in Table 1 that is potentially positive, such as γh_{ik}. Referring to the row and column edge directions in the table, we only need to consider this possibility if (a) both $(\mathbf{0}, \mathbf{e}_i)$ and $(-\mathbf{e}_k, \mathbf{0}) \in \mathcal{C}(\mathbf{x}, \mathbf{y})$ or (b) both $(\mathbf{e}_i, \mathbf{0})$ and $(\mathbf{0}, -\mathbf{e}_k) \in \mathcal{C}(\mathbf{x}, \mathbf{y})$. If (b) holds, then by Table 2, we have $\mathbf{1}^{\mathsf{T}}\mathbf{x} < u_a$, $\lambda^a = 0$, $i \in \bar{\mathcal{A}}_0$, $\mathbf{1}^{\mathsf{T}}\mathbf{y} > \ell_b$, $\lambda^b = 0$, and $k \in \mathcal{B}_0$. By (V4b), $i \neq k$ and $h_{ik} = 0$. If (a) holds, then by Table 2, we have $\mathbf{1}^{\mathsf{T}}\mathbf{x} > \ell_a$, $\lambda^a = 0$, $k \in \mathcal{A}_0$, $\mathbf{1}^{\mathsf{T}}\mathbf{y} < u_b$, $\lambda^b = 0$, and $i \in \bar{\mathcal{B}}_0$. By (V4a), $i \neq k$ and $h_{ik} = 0$. Thus the proof boils down to the following: (i) locate each potentially positive entry in Table 1, (ii) extract the associated edge directions from the corresponding row and column, (iii) use Table 2 to deduce associated conditions on the constraints and

the multipliers when these edge directions lie in $\mathcal{C}(\mathbf{x}, \mathbf{y})$, and (iv) use the conclusions in (V2)–(V5) to deduce that the entry in Table 1 must be nonpositive.

Conversely, suppose that $(\mathbf{x}, \mathbf{y})$ is a local maximizer, or equivalently, suppose that (V1) and (5.4) hold. We must show that (V2)–(V5) are satisfied. We start with (V2a) and suppose that $\lambda^a = 0$, $\mathbf{1}^\mathsf{T}\mathbf{x} < u_a$, $i \in \bar{\mathcal{A}}_0$, $j \in \bar{\mathcal{B}}_0$, and $k \in \mathcal{B}_0$. By the first and last rows of Table 2, we have $(\mathbf{e}_i, \mathbf{0})$ and $(\mathbf{0}, \mathbf{e}_j - \mathbf{e}_k) \in \mathcal{C}(\mathbf{x}, \mathbf{y})$. By (5.4) and Table 1, $h_{ji} - h_{ki} \geq 0$. By symmetry, this is equivalent to $h_{ij} \geq h_{ik}$. This establishes (V2a). Thus the proof of the converse proceeds as follows: When any of the hypotheses in (V2)–(V5) are fulfilled, we use Table 2 to determine two edge directions $\mathbf{d}^1$ and $\mathbf{d}^2$ that must lie in $\mathcal{C}(\mathbf{x}, \mathbf{y})$. We use Table 1 to obtain the product $(\mathbf{d}^1)^\mathsf{T}(\nabla^2 f)\mathbf{d}^2$, and we use (5.4) to obtain a relation between elements of $\mathbf{H}$ which corresponds to the conclusion appearing in (V2)–(V5). □

6 Edge separator problem

A related problem in discrete optimization is the *edge separator problem*: Given an undirected, unweighted graph G and positive integers $\ell \leq u$, partition the vertices of G into disjoint sets $\mathcal{A}$ and $\mathcal{B}$ such that $\ell \leq |\mathcal{A}| \leq u$, and the number of edges between vertices in different sets is minimized. Let $\mathbf{A}$ be the adjacency matrix for the graph G. In [8] Hager and Krylyuk showed that the edge separator problem can be formulated as the following continuous quadratic program:

$$\max \quad -(\mathbf{1} - \mathbf{x})^\mathsf{T}(\mathbf{A} + \mathbf{I})\mathbf{x} \tag{6.1}$$
$$\text{subject to} \quad \mathbf{0} \leq \mathbf{x} \leq \mathbf{1} \quad \text{and} \quad \ell \leq \mathbf{1}^\mathsf{T}\mathbf{x} \leq u.$$

In particular, given any solution to (6.1), there exists an easily computed binary solution $\mathbf{x}$ and the optimal edge separator is given by

$$\mathcal{A} = \{i : x_i = 1\} \quad \text{and} \quad \mathcal{B} = \{i : x_i = 0\}.$$

Observe that the vertex separator problem (5.1) with the additional constraints $\mathbf{x} + \mathbf{y} = \mathbf{1}$, $\ell_b = 0$, and $u_b = n$, becomes the edge separator problem (6.1). Consequently, the derivation of optimality conditions for the edge separator problem is similar to that of the vertex separator problem.

Let g denote the objective function in (6.1):

$$g(\mathbf{x}) = -(\mathbf{1} - \mathbf{x})^\mathsf{T}(\mathbf{A} + \mathbf{I})\mathbf{x}$$

The first-order optimality conditions for (6.1) can be stated in the following way: If $\mathbf{x}$ is a local maximizer of (6.1), then there exist multipliers

$$\boldsymbol{\mu} \in \mathcal{M}(\mathbf{x}) \quad \text{and} \quad \lambda \in \mathcal{L}(\mathbf{x}, \ell, u)$$

such that

$$\nabla g(\mathbf{x}) + \boldsymbol{\mu} + \lambda \mathbf{1} = \mathbf{0}. \tag{6.2}$$

Similar to the edge description for the vertex separator problem derived in [7], a reflective edge description for the edge separator problem is given by

Table 3 Values of $(\mathbf{d}^1)^\mathsf{T}(\nabla^2 g)\mathbf{d}^2$ for $\mathbf{d}^1, \mathbf{d}^2 \in \mathcal{D}$

	$\mathbf{e}_k$	$-\mathbf{e}_k$	$\mathbf{e}_k - \mathbf{e}_l$
$\mathbf{e}_i$	h_{ik}	$-h_{ik}$	$h_{ik} - h_{il}$
$-\mathbf{e}_i$		h_{ik}	$h_{il} - h_{ik}$
$\mathbf{e}_i - \mathbf{e}_j$			$h_{ik} - h_{jk} - h_{il} + h_{jl}$

$$\mathcal{D} = \bigcup_{\substack{i,j=1 \\ i \neq j}}^{n} \{\pm\mathbf{e}_i, \ \mathbf{e}_i - \mathbf{e}_j\}.$$

Since $\nabla^2 g = (\mathbf{A} + \mathbf{I})$, we have

$$\mathbf{e}_i^\mathsf{T}(\nabla^2 g)\mathbf{e}_i = 1 \quad \text{and} \quad (\mathbf{e}_i - \mathbf{e}_j)^\mathsf{T}(\nabla^2 g)(\mathbf{e}_i - \mathbf{e}_j) = 2 - 2a_{ij} \quad \text{for } i \neq j. \quad (6.3)$$

The second term in (6.3) is non-negative since $a_{ij} \leq 1$. Hence,

$$\mathbf{d}^\mathsf{T}(\nabla^2 g)\mathbf{d} \geq 0 \quad \text{for all } \mathbf{d} \in \mathcal{D},$$

which shows that the hypotheses of Corollary 3.3 are satisfied. Consequently, a feasible point $\mathbf{x}$ of (6.1) is a local maximizer if and only if the first-order optimality conditions (6.2) hold and

$$(\mathbf{d}^1)^\mathsf{T}(\nabla^2 g)\mathbf{d}^2 \leq 0 \quad \text{for every } \mathbf{d}^1, \mathbf{d}^2 \in \mathcal{C}(\mathbf{x}) \cap \mathcal{D}. \quad (6.4)$$

Table 3 gives all the different possible values for $(\mathbf{d}^1)^\mathsf{T}(\nabla^2 g)\mathbf{d}^2$ when $\mathbf{d}^1$ and $\mathbf{d}^2$ are edge directions. Since $\mathcal{D}$ is described in terms of 3 distinct types of vectors, we obtain a 3 by 3 matrix of products; the entries left blank are known from symmetry.

Next, we construct the sets $\mathcal{F}(\mathbf{x})$ and $\mathcal{C}(\mathbf{x})$ using the definitions (1.3) and (1.7):

$$\mathcal{F}(\mathbf{x}) = \left\{ \mathbf{d} \in \mathbb{R}^n : \begin{array}{ll} \mathbf{1}^\mathsf{T}\mathbf{d} \leq 0 \text{ if } \mathbf{1}^\mathsf{T}\mathbf{x} = u, & d_i \leq 0 \ \forall \ i \text{ with } x_i = 1, \\ \mathbf{1}^\mathsf{T}\mathbf{d} \geq 0 \text{ if } \mathbf{1}^\mathsf{T}\mathbf{x} = l, & d_i \geq 0 \ \forall \ i \text{ with } x_i = 0 \end{array} \right\},$$

$$\mathcal{C}(\mathbf{x}) = \left\{ \mathbf{d} \in \mathcal{F}(\mathbf{x}) : \mathbf{1}^\mathsf{T}\mathbf{d} = 0 \text{ if } \lambda \neq 0 \quad \text{and} \quad d_i = 0 \ \forall \ i \text{ with } \mu_i \neq 0 \right\}.$$

Table 4 shows when each element of $\mathcal{D}$ also lies in $\mathcal{C}(\mathbf{x}, \mathbf{y})$. The table makes use of the following sets:

$$\mathcal{A}_0 = \{i : x_i > 0, \quad \mu_i = 0\} \quad \text{and} \quad \bar{\mathcal{A}}_0 = \{i : x_i < 1, \quad \mu_i = 0\}.$$

Theorem 6.1 *If $\mathbf{x}$ is feasible in (6.1), then $\mathbf{x}$ is a local maximizer of (6.1) if and only if (E1)–(E3) hold:*

Table 4 A description of $\mathcal{D} \cap \mathcal{C}(\mathbf{x})$

		$\mathbf{1}^\mathsf{T}\mathbf{x} < u$		$\mathbf{1}^\mathsf{T}\mathbf{x} < u$
$\mathbf{e}_i \in \mathcal{C}(\mathbf{x})$	$\Leftrightarrow$	$x_i < 1$	$\Leftrightarrow$	$\lambda = 0$
		$\mu_i = \lambda = 0$		$i \in \bar{\mathcal{A}}_0$
		$\mathbf{1}^\mathsf{T}\mathbf{x} > \ell$		$\mathbf{1}^\mathsf{T}\mathbf{x} > \ell$
$-\mathbf{e}_i \in \mathcal{C}(\mathbf{x})$	$\Leftrightarrow$	$x_i > 0$	$\Leftrightarrow$	$\lambda = 0$
		$\mu_i = \lambda = 0$		$i \in \mathcal{A}_0$
$\mathbf{e}_i - \mathbf{e}_j \in \mathcal{C}(\mathbf{x})$	$\Leftrightarrow$	$x_i < 1 \text{ and } x_j > 0$	$\Leftrightarrow$	$i \in \bar{\mathcal{A}}_0 \text{ and } j \in \mathcal{A}_0$
		$\mu_i = \mu_j = 0$		

(E1) *The first-order conditions* (6.2) *hold.*
(E2) *If* $i \in \bar{\mathcal{A}}_0$ *and* $j \in \mathcal{A}_0$, *then* $h_{ij} = 1$.
(E3) *Suppose* $\lambda = 0$.
     a. If $\mathbf{1}^\mathsf{T}\mathbf{x} < u$, *then* $\bar{\mathcal{A}}_0 = \emptyset$.
     b. If $\mathbf{1}^\mathsf{T}\mathbf{x} > \ell$, *then* $\mathcal{A}_0 = \emptyset$.

Proof First, suppose that the conditions (E1)–(E3) are satisfied. We wish to show that (6.4) holds, which implies that $\mathbf{x}$ is a local maximizer. By Table 4, $\mathbf{e}_i \in \mathcal{C}(\mathbf{x})$ if and only if $\mathbf{1}^\mathsf{T}\mathbf{x} < u$, $\lambda = 0$, and $i \in \bar{\mathcal{A}}_0$. But by (E3a), $\bar{\mathcal{A}}_0 = \emptyset$ if $\lambda = 0$ and $\mathbf{1}^\mathsf{T}\mathbf{x} < u$. Hence, $\mathcal{D} \cap \mathcal{C}(\mathbf{x})$ contains no elements of the form $\mathbf{e}_i$. Similarly, $\mathcal{D} \cap \mathcal{C}(\mathbf{x})$ contains no elements of the form $-\mathbf{e}_i$, and the only elements in $\mathcal{D} \cap \mathcal{C}(\mathbf{x})$ have the form $\mathbf{e}_i - \mathbf{e}_j$. If $\mathbf{e}_i - \mathbf{e}_j \in \mathcal{C}(\mathbf{x})$, then by Table 4, $i \in \bar{\mathcal{A}}_0$ and $j \in \mathcal{A}_0$. By (E2) we conclude that $h_{ij} = 1$. If $i \neq j$, this implies that $a_{ij} = 1$ and consequently,

$$(\mathbf{e}_i - \mathbf{e}_j)^\mathsf{T}(\nabla^2 g)(\mathbf{e}_i - \mathbf{e}_j) = 2 - 2a_{ij} = 0.$$

If $(\mathbf{e}_i - \mathbf{e}_j)$ and $(\mathbf{e}_k - \mathbf{e}_l)$ are distinct elements of $\mathcal{C}(\mathbf{x})$, then by Table 4, $i, k \in \bar{\mathcal{A}}_0$ and $j, l \in \mathcal{A}_0$. By (E2) and the symmetry of $\mathbf{H}$, it follows that $h_{jk} = 1 = h_{il}$. Since $h_{ik} \leq 1$ and $h_{jl} \leq 1$, we conclude that

$$h_{ik} - h_{jk} - h_{il} + h_{jl} \leq 0.$$

This completes the proof of (6.4), and $\mathbf{x}$ is a local maximizer.

Conversely, suppose that $\mathbf{x}$ is a local maximizer, or equivalently, suppose that (E1) and (6.4) hold. We show that (E2) and (E3) hold. First consider (E3a) and suppose that $\lambda = 0$ and $\mathbf{1}^\mathsf{T}\mathbf{x} < u$. If $i \in \bar{\mathcal{A}}_0$, then by Table 4, $\mathbf{e}_i \in \mathcal{C}(\mathbf{x})$. By Table 3 and (6.4), $h_{ii} \leq 0$, which is impossible since $h_{ii} = 1$. Hence, $\bar{\mathcal{A}}_0 = \emptyset$. (E3b) is established in a similar fashion.

Now consider (E2). If $i \in \bar{\mathcal{A}}_0$ and $j \in \mathcal{A}_0$, then by Table 4, $\mathbf{e}_i - \mathbf{e}_j \in \mathcal{C}(\mathbf{x})$. By (6.4), it follows that when $i \neq j$,

$$(\mathbf{e}_i - \mathbf{e}_j)^\mathsf{T}(\nabla^2 g)(\mathbf{e}_i - \mathbf{e}_j) = 2 - 2a_{ij} \leq 0.$$

Hence, $a_{ij} = 1 = h_{ij}$. If $i = j$, then $h_{ii} = 1$ since the diagonal of $\mathbf{H}$ is the diagonal of $\mathbf{I}$. This completes the proof. $\square$

The conditions in Theorem 6.1 were also obtained by Hager and Krylyuk in [8], however, the analysis there was specifically tailored to the edge separator problem, while here these conditions are deduced from the general theory for polyhedral maximization of an edge-convex quadratic.

7 Conclusions

In this paper, we have developed new optimality conditions for the maximization of a function over a polyhedron $\mathcal{P}$. We have shown in Corollary 2.2 that when the constraint matrix describing the polyhedron has full column rank, then the cone $\mathcal{F}(\mathbf{x}^*)$ of first-order feasible directions is a conical combination of edge directions of the polyhedron:

$$\mathcal{F}(\mathbf{x}^*) = \text{span}_+(\mathcal{F}(\mathbf{x}^*) \cap \mathcal{D}); \tag{7.1}$$

here $\mathcal{D}$ is a reflective edge description of $\mathcal{P}$. Similarly, if $\mathbf{x}^*$ satisfies the first-order optimality conditions, then

$$\mathcal{C}(\mathbf{x}^*) = \text{span}_+(\mathcal{C}(\mathbf{x}^*) \cap \mathcal{D}), \tag{7.2}$$

where $\mathcal{C}(\mathbf{x}^*)$ is the critical cone. A consequence of (7.1), established in Proposition 3.1, is that the usual first-order optimality condition only needs to be checked for a finite set $\mathcal{F}(\mathbf{x}^*) \cap \mathcal{D}$. A consequence of (7.2), established in Proposition 3.2, is a new second-order sufficient optimality condition, and a new second-order necessary optimality condition for problems where the objective function is convex along the edge-directions of $\mathcal{P}$.

In [12,16] it is shown that checking local optimality for an indefinite quadratic programming problem can be NP-hard. On the other hand, Corollary 4.1 shows that when $\mathcal{P}$ also includes box constraints and the objective function is edge-convex, local optimality can be checked in polynomial time since the size of the edge description is bounded by a polynomial in the problem dimension.

The necessary and sufficient optimality conditions are applied to quadratic programming formulations of the vertex and edge separator problems. Our theory applies to these problems since the objective function is convex along the edges of the constraint polyhedron. The analysis of local optimality reduces to testing conditions (V1)–(V5) for the vertex separator problem and conditions (E1)–(E3) for the edge separator problem. For either problem, local optimality is easily checked. In general, when the objective function in a polyhedral constrained optimization problem is convex, our optimality conditions are applicable since the objective function is trivially edge convex. Hence, our theory could be used to test for local optimality of any feasible point in the algorithm of [19].

References

1. Bertsekas, D.P.: Nonlinear Programming. Athena Scientific, Belmont (1999)
2. Borwein, J.: Necessary and sufficient conditions for quadratic minimality. Numer. Funct. Anal. Optim. **5**, 127–140 (1982)
3. Contesse, L.: Une caractérisation complète des minima locaux en programmation quadratique. Numer. Math. **34**, 315–332 (1980)
4. Dantzig, G.: Maximization of a linear function of variables subject to linear inequalities. In: Koopman, T.C. (eds), Activity Analysis of Production and Allocation, Cowles Commission Monograph, vol. 13, Chap. XXI. Wiley, New York (1951)
5. Gritzmann, P., Sturmfels, B.: Minkowski addition of polytopes: computational complexity and applications to Gröbner bases. SIAM J. Disc. Math. **6**, 246–269 (1993)
6. Griva, I., Nash, S.G., Sofer, A.: Linear and Nonlinear Optimization. SIAM, Philadelphia (2009)
7. Hager, W.W., Hungerford, J.T.: A Continuous Quadratic Programming Formulation of the Vertex Separator Problem. Technical Report. Department of Mathematics, University of Florida. http://www.math.ufl.edu/~hager/papers/GP/vertex.pdf (2012)
8. Hager, W.W., Krylyuk, Y.: Graph partitioning and continuous quadratic programming. SIAM J. Disc. Math. **12**, 500–523 (1999)
9. Hwang, F., Rothblum, U.: Directional-quasi-convexity, asymmetric Schur-convexity and optimality of consecutive partitions. Math. Oper. Res. **21**, 540–554 (1996)
10. Klee, V.: Some characterizations of convex polyhedra. Acta Math. **102**, 79–107 (1959)
11. Luenberger, D.G., Ye, Y.: Linear and Nonlinear Programming. Springer, Berlin (2008)
12. Murty, K.G., Kabadi, S.N.: Some NP-complete problems in quadratic and linear programming. Math. Program. **39**, 117–129 (1987)
13. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer, New York (1999)
14. Onn, S., Rothblum, U.: Convex combinatorial optimization. Discret. Comput. Geom. **32**, 549–566 (2004)
15. Onn, S., Rothblum, U., Tangir, Y.: Edge-directions of standard polyhedra with applications to network flows. J. Glob. Optim. **33**, 109–122 (2005)
16. Pardalos, P.M., Schnitger, G.: Checking local optimality in constrained quadratic programming can be NP-hard. Oper. Res. Lett. **7**, 33–35 (1988)
17. Rockafellar, R.T.: Network Flows and Monotropic Optimization. Wiley, New York (1984)
18. Tardella, F.: On the equivalence between some discrete and continuous optimization problems. Ann. Oper. Res. **25**, 291–300 (1990)
19. Tuy, H.: Concave programming under linear constraints. Soviet Math. **5**, 1439–1440 (1964)