Unique sink orientations for homogeneous linear inequalities and their alternative systems

Jourdain Lamperski∗
October 11, 2019

Abstract

It is not known, at the moment, whether there exists a strongly-polynomial algorithm for solving a system of linear inequalities. A noteworthy candidate is the simplex method, but despite a substantial amount of effort, it has not been shown that a version of the method runs in strongly-polynomial time. Motivated by this, Gärtner and Schurr [GS06] propose an alternative involving unique sink orientations (USOs) of the cube \( \{0,1\}^m \), which are orientations of the cube such that every oriented face contains a unique sink. This paper continues this line of work; given data \( A \in \mathbb{R}^{n \times m} \), we construct and study a random orientation of the cube \( \{0,1\}^m \) that admits the following property. It almost always is a USO with a unique sink that can be used to either compute a solution to the homogeneous system of linear inequalities \( A^\top x > 0 \) or compute a solution to the alternative system \( A\lambda = 0, \lambda \geq 0 \). We show that the USOs satisfy a certain local property that is not satisfied by all USOs that satisfy the Holt-Klee property, addressing a question of [Mor02, Jag06], which is motivated by the idea that such additional structure could be leveraged algorithmically, and is perhaps critical, to develop faster algorithms or a strongly-polynomial algorithm. We suspect that the tools and techniques developed in this work can be used to derive further properties. We also establish a connection to condition measure theory; we show that the unique sink can be used to compute primal and dual optimal solutions for second order cone problems, whose optimal values capture the extent of feasibility of the homogeneous or alternative system.

1 Introduction

It is not known, at the moment, whether there exists a strongly-polynomial algorithm for solving a system of linear inequalities. A noteworthy candidate is the simplex algorithm, but despite a substantial amount of effort, it has not been shown that a version of the method runs in strongly-polynomial time.

Gärtner and Schurr [GS06] propose an alternative to the simplex method involving unique sink orientations (USOs) of the cube \( \{0,1\}^m \), which are orientations of the cube such that

∗MIT Operations Research Center, jourdain@mit.edu.
every oriented face contains a unique sink, see Figure 1. (Stickney and Watson [SW78] introduced USOs as an abstraction of the \( P \)-matrix linear complementarity problem.) Gärtnner and Schurr construct a USO with the property that its unique sink can be used to either compute an optimal solution to a linear program, compute a certificate of infeasibility, or compute a certificate of unboundedness. And they show that if the unique sink of a USO can be computed in polynomial time (in the \textit{vertex evaluation model}), then linear programs can be solved in strongly-polynomial time. (In the vertex evaluation model, the only available operation takes as input a vertex in the orientation and outputs the direction of all of the arcs incident to that vertex.)

This motivates the problem to develop algorithms that, in the vertex evaluation model, compute the unique sink of a USO. There are a few such algorithms (the fastest algorithm, at the moment, is the Fibonacci Seesaw, see [SW01]), and it is known that the problem of computing the unique sink of a USO belongs to the complexity class PPAD [FGMS18]. PPAD contains several problems that are suspected to be hard, including the problem of finding a Nash equilibria [DGP06]. Accordingly, in this work, we are motivated to not pursue this direction of research, and instead, we consider a different line of attack, which we describe next.

The USOs constructed in [GS06] admit additional structure; they satisfy the Holt-Klee property: for any \( k \)-dimensional oriented face of the cube there exist \( k \) vertex-disjoint paths from its unique source to its unique sink. This raises the question: can this additional structure be leveraged algorithmically? It also raises the question: do these USOs satisfy other properties, and if so, can these properties be leveraged algorithmically? This latter question is raised in [Jag06], and it is motivated by the fact that such structure may be critical to develop a strongly-polynomial algorithm. (The same question is also considered in [Mor02] in the context of linear programs with a feasible region that is combinatorially equivalent to the cube.)

In this work, we address this question in the context of homogeneous linear inequalities. (The question was originally asked in the context of linear programming, but from a complexity stand point, linear programming and the problem of solving homogeneous linear inequalities are equivalent. And as stated before, the question was also asked in an even less general setting.) Given data \( A \in \mathbb{R}^{n \times m} \), we construct and study a random orientation of the cube \( \{0,1\}^m \) that almost surely contains a unique sink that can be used to either compute a solution to the homogenous system of strict linear inequalities

\[
(P) : \quad A^T x > 0
\]
when the set of solutions $K := \{ x \in \mathbb{R}^n : A^\top x > 0 \}$ is nonempty, or compute solution to the alternative system

$$(Alt) : \begin{align*}
A\lambda &= 0 \\
\lambda &\geq 0
\end{align*}$$

when such a solution exists. And of course, from a theorem of the alternative, $(Alt)$ is feasible when and only when $(P)$ is infeasible. Before stating our main results in Subsections 1.2-1.5, let us establish some preliminary notation in the following subsection.

### 1.1 Notation

Define $\mathbb{I} := \{1, ..., m\}$. For $I, K \subseteq \mathbb{I}$, define the set interval $[I, K] := \{ J \subseteq \mathbb{I} : I \subseteq J \subseteq K \}$. Let $a_1, \ldots, a_m$ denote the columns of $A$. For $I \subseteq \mathbb{I}$, let us collect the columns of $A$ indexed by $I$ into the matrix $A_I \in \mathbb{R}^{n \times |I|}$. Similarly, for $v \in \mathbb{R}^m$, let us collect the entries of $v$ indexed by $I$ into $v_I \in \mathbb{R}^{|I|}$. Let $e \in \mathbb{R}^m$ denote the vector of all ones. We will slightly abuse notation and write the $i$-th standard basis vector in $\mathbb{R}^k$ as $e_i$, where its dimension will be understood from context. Let $\| \cdot \|$ denote the euclidean norm. For a directed graph $G$, we will denote the vertex set as $V(G)$, the arc set as $A(G)$, and an arc from $u \in V(G)$ to $w \in V(G)$ as $(u, w) \in A(G)$. We will label the vertices of an orientation of the cube $\{0, 1\}^m$ with the subsets of $\mathbb{I}$.

### 1.2 Construction of random orientation

The construction of the orientation is straightforward to state, and we summarize it here. Like the construction given in [GS06], it uses a collection of least-squares solutions, and it requires randomization to avoid degeneracy. (The randomization does not change the set of solutions to the homogeneous system.)

Now let us construct the random orientation. Let $r \in \mathbb{R}^m$ such that $r > 0$, and for each $I \subseteq \mathbb{I}$, define

$$x(I, r) := \arg\min_x \| (A_I R_I)^\top x - e_I \|$$

s.t. $x \in \mathrm{span}(\{a_i\}_{i \in I})$,

where $R_I := \text{diag}(r_I)$ is the diagonal matrix with diagonal $r_I$.

**Remark 1.1.** In the objective, the matrix $R_I$ positively rescales the columns of $A_I$, and because the problem is constrained to $\mathrm{span}(\{a_i\}_{i \in I})$, it follows that $x(I, r)$ is the unique optimal solution to the least-squares problem.

Next we use the collection $\{x(I, r)\}_{I \subseteq \mathbb{I}}$ to construct a parameterized orientation $G(r)$ of the cube $\{0, 1\}^m$ as follows. For each $I \subseteq \mathbb{I}$ and $j \notin I$, define the orientation of the arc in $G(r)$ incident to $I$ and $I \cup j$ by

$$(I, I \cup j) \in A(G(r)) \iff r_j a_j^\top x(I, r) \leq 1.$$

Finally we introduce randomness. For $\epsilon \in \mathbb{R}$ such that $0 < \epsilon < 1$, define $r(\epsilon)$ to be a random $m$-dimensional vector with entries that are independent and uniformly distributed on the interval $[1 - \epsilon, 1 + \epsilon]$. We introduce randomness to obtain the following property:
Proposition 1.1. For $\epsilon \in \mathbb{R}$ such that $0 < \epsilon < 1$, it holds almost always that
\[ r_j(\epsilon) a_j^\top x(I, r(\epsilon)) \neq 1 \]
for all $I \subseteq [m]$ and $j \notin I$.

Proof. See Theorem 2.2 in [Jag06].

In Section 3 we show that, under the setup of Proposition 1.1, it holds that $G(r)$ is a USO:

Theorem 1.1. Let $r \in \mathbb{R}^m$ such that $r > 0$, and suppose that for all $I \subseteq [m]$ and $j \notin I$, it holds that $r_j a_j^\top x(I, r) \neq 1$. Then $G(r)$ is a unique sink orientation.

The following corollary is immediate from Theorem 1.1 and Proposition 1.1.

Corollary 1.1. For $\epsilon \in \mathbb{R}$ such that $0 < \epsilon < 1$, it holds almost always that $G(r(\epsilon))$ is a unique sink orientation.

Remark 1.2. We will see in Section 2 that it is straightforward to compute $x(I, r)$ in strongly polynomial time with the Gram-Schmidt algorithm and Gaussian elimination. Accordingly, we can compute the orientations of the arcs in $G(r)$ in strongly-polynomial time. It follows that our construction, like the construction of [GS06], can be appropriately considered in the vertex evaluation model framework.

In Section 3 we also show that it is straightforward to compute a solution to either $(P)$ or $(Alt)$ using the unique sink:

Theorem 1.2. Let $r \in \mathbb{R}^m$ such that $r > 0$, and suppose that for all $I \subseteq [m]$ and $j \notin I$, it holds that $r_j a_j^\top x(I, r) \neq 1$. Let $S$ be the unique sink of $G(r)$.

1. If $(P)$ is feasible, then $x(S, r)$ is a solution to $(P)$.

2. Otherwise, if $(P)$ is infeasible, then $\lambda$ defined by
\[ \lambda_S = e_S - (A_S R_S)^\top x(S, r) \]
\[ \lambda_{[m]\setminus S} = 0 \]
is a solution to $(Alt)$.

1.3 A local property not satisfied by all Holt-Klee USOs

In this subsection we present our main result, a local property that is satisfied by the USOs constructed in Subsection 1.2, but that is not satisfied by all USOs that satisfy the Holt-Klee property.

First let us introduce a preliminary definition. For an orientation $G$ and each $I \subseteq [m]$ such that $|I| \leq m - 2$, define the directed graph $H_G(I)$ as follows. Let $[m] \setminus I$ be its vertex set, and let $(v, w)$ be an arc in $H_G(I)$ for $v, w \in [m] \setminus I$ if either
\[ (I, I \cup w), (I \cup \{v, w\}, I \cup v) \in \mathcal{A}(G) \text{ or } (I \cup w, I), (I \cup v, I \cup \{v, w\}) \in \mathcal{A}(G). \]

In Section 4 we show that under the usual setup of Proposition 1.1, in which case $G(r)$ is a USO, it turns out that $H_G(r)(I)$, where $I \subseteq [m]$ such that $|I| \leq m - 2$, is acyclic:
Theorem 1.3. Let \( r \in \mathbb{R}^m \) such that \( r > 0 \), and suppose that for all \( I \subseteq [m] \) and \( j \notin I \), it holds that \( r_j a_j^\top x(I, r) \neq 1 \). Also let \( I \subseteq [m] \) such that \( |I| \leq m - 2 \). Then \( H_{G(r)}(I) \) is acyclic.

It is straightforward to construct a USO of \( \{0, 1\}^3 \) that does not satisfy the property established in Theorem 1.3, see Figure 2. Observe that \( G_1 \) does not satisfy the additional property because \( H_{G_1}(\emptyset) \) is a cycle. However, \( G_1 \) is isomorphic to \( G_2 \), which satisfies the additional property.

![Figure 2: Two USOs \( G_1 \) and \( G_2 \) of \( \{0, 1\}^3 \) that satisfy the Holt-Klee property.](image)

### 1.4 A connection to condition measure theory

Here we discuss a connection to condition measure theory; we show that the unique sink of the USOs constructed in Subsection 1.2 can be used to compute primal and dual optimal solutions to certain second order cone problems, whose optimal values capture the extent of feasibility of \((P)\) or \((Alt)\).

Let \( r \in \mathbb{R}^m \) such that \( r > 0 \). Consider the following second order cone problem parameterized in terms of \( r \):

\[
\max \min_{\|x\| \leq 1} \max_{i \in [m]} r_i a_i^\top x = \max_{x, \eta} \eta \\
\text{s.t. } (AR)^\top x \geq \eta e, \|x\| \leq 1.
\]

When \((P)\) is feasible, \( r = e \), and each column of \( A \) has unit euclidean norm, the optimal value of (1), called the width of \( K \), is a condition measure that captures the extent of feasibility of \((P)\). The width is often a factor in complexity guarantees for algorithms that compute a solution to \((P)\), for example, the ellipsoid algorithm for conic problems [FV99] and the rescaled perceptron algorithm [DV08].

The dual of (1) is the parameterized second order cone problem (with parameter \( r \)):

\[
\min_{\lambda} \|AR\lambda\| \\
\text{s.t. } \lambda^\top e = 1, \lambda \geq 0.
\]
In Section 5 we show that when \((P)\) is feasible, and under the usual setup of Proposition 1.1, in which case \(G(r)\) is a USO, the unique sink of \(G(r)\) can be used to compute optimal solutions for (1) and (2):

**Theorem 1.4.** Let \(r \in \mathbb{R}^m\) such that \(r > 0\), and suppose that for all \(I \subseteq [m]\) and \(j \notin I\), it holds that \(r_j a_j^\top x(I, r) \neq 1\). Also suppose that \(S\) is the unique sink of \(G(r)\). If \((P)\) is feasible, then

\[
x = \frac{1}{\|x(S, r)\|}x(S, r)
\]

is optimal for (1) with parameter \(r\), and \(\lambda\) defined by

\[
\lambda_S = \frac{1}{e_S (R_S A_S^\top A_S R_S)^{-1} e_S} (R_S A_S^\top A_S R_S)^{-1} e_S
\]

\[
\lambda_{[m]\setminus S} = 0
\]

is optimal for (2) with parameter \(r\).

Now, let us consider the parametrized second order cone problem (with parameter \(r\)):

\[
\begin{aligned}
\min_{\lambda} & \quad \|\lambda\| \\
\text{s.t.} & \quad AR\lambda = 0 \\
& \quad e^\top \lambda = 1 \\
& \quad \lambda \geq 0.
\end{aligned}
\tag{3}
\]

When \((P)\) is infeasible, \(r = e\), and columns of \(A\) have unit euclidean norm, problem (3) is feasible and the optimal value can be interpreted as a condition measure that captures the extent of feasibility of \((Alt)\).

The dual of (3) is the parameterized second order cone problem (with parameter \(r\)):

\[
\begin{aligned}
\max_{x, \alpha, \gamma} & \quad \gamma \\
\text{s.t.} & \quad \|\alpha + \gamma e - (AR)^\top x\| \leq 1 \\
& \quad \alpha \geq 0.
\end{aligned}
\tag{4}
\]

In Section 5 we show that when \((P)\) is infeasible, and under the usual setup of Proposition 1.1, in which case \(G(r)\) is a USO, the unique sink of \(G(r)\) can be used to compute optimal solutions for (3) and (4):

**Theorem 1.5.** Let \(r \in \mathbb{R}^m\) such that \(r > 0\), and suppose that for all \(I \subseteq [m]\) and \(j \notin I\), it holds that \(r_j a_j^\top x(I, r) \neq 1\). Also suppose that \(S\) is the unique sink of \(G(r)\). If \((P)\) is infeasible, then \(\lambda\) defined by

\[
\lambda_S = \frac{1}{e_S (e_S - (A_S R_S)^\top x(S, r)) e_S - (A_S R_S)^\top x(S, r)}
\]

\[
\lambda_{[m]\setminus S} = 0
\]
is optimal for (3) with parameter \( r \), and \( (x, \alpha, \gamma) \) defined by

\[
x = \frac{1}{\|e_S - (A_S R_S)\top x(S,r)\|} x(S,r)
\]

\[
\alpha_{[m]\backslash S} = \frac{1}{\|e_S - (A_S R_S)\top x(S,r)\|} (A_{[m]\backslash S} \top x(S,r) - e_{[m]\backslash S})
\]

\[
\alpha_S = 0
\]

\[
\gamma = \frac{1}{\|e_S - (A_S R_S)\top x(S,r)\|}
\]

is optimal for (4) with parameter \( r \).

1.5 A characterization for USOs of the cube

In this subsection we describe a characterization for USOs of the cube that we needed to develop in order to prove Theorem 1.1. The characterization is similar to the characterization given in [SW01].

First we introduce a few preliminary definitions. For an orientation \( G \) of \( \{0,1\}^m \) and \( I, K \subseteq [m] \) such that \( I \subseteq K \), let \( G[I,K] \) denote the oriented face in \( G \) induced by the vertices \( [I,K] \), and for \( J \in [I,K] \), define the out-indices of \( J \) in \( G[I,K] \) as

\[
N_{G[I,K]}^+(J) := (K \setminus I) \cap \{i \in [m] : (I, I \cup i) \in A(G) \text{ or } (I, I \setminus i) \in A(G)\}
\]

and the in-indices of \( J \) in \( G[I,K] \) as

\[
N_{G[I,K]}^-(J) := (K \setminus I) \cap \{i \in [m] : (I \cup i, I) \in A(G) \text{ or } (I \setminus i, I) \in A(G)\}.
\]

Lastly, for \( J_1, J_2 \in [I,K] \), we will say \( J_1 \) and \( J_2 \) are antipodal in \([I,K]\) if \( J_1 = I \cup (K \setminus J_2) \).

In Subsection 2.2, we establish the following characterization for USOS of the cube:

**Theorem 1.6.** Let \( G \) be an orientation of the cube \( \{0,1\}^m \). Then, \( G \) is a unique sink orientation if and only if for all subsets \( I, K \subseteq [m] \) such that \( I \subseteq K \), and for all \( J_1, J_2 \in [I,K] \) such that \( J_1 \) and \( J_2 \) are antipodal in \([I,K]\), it holds that the collection of out-indices of \( J_1 \) in \( G[I,K] \) is not equal to the collection of out-indices of \( J_2 \) in \( G[I,K] \).

1.6 Organization

In Section 2 we establish some preliminaries; we introduce some additional definitions, we provide an alternative characterization for \( x(I,r) \), we establish the characterization for USOs described in Subsection 1.5, and we establish a few useful equations. In Sections 3-5 we establish the results described in Subsections 1.2-1.4, respectively.

2 Preliminaries

In this section we establish some preliminaries. Note that in each subsection we introduce some definitions that we will use for the remainder of the paper.
2.1 Alternative characterization of $x(I, r)$

Let $I \subseteq [m]$ and $r \in \mathbb{R}^m$ such that $r > 0$. Define $U_I$ to be a matrix with orthonormal rows that span the subspace $\text{span}(\{a_i\}_{i \in I})$. Note that $U_I U_I^T = I$ and $U_I^T U_I a = a$ for $a \in \text{span}(\{a_i\}_{i \in I})$.

**Remark 2.1.** The particular choice of $U_I$ is not important, but it will sometimes be useful to specify a choice in the proofs.

It will be convenient to define the following quantities:

- $A(I, r) := U_I A_I R_I$
- $c(I, r) := A(I, r) e_I$
- $M(I, r) := A(I, r) A(I, r)^T$

Note that these quantities are not invariant with respect to the choice of $U_I$.

It is straightforward to establish the following alternative characterization of $x(I, r)$ using these quantities:

$$x(I, r) = U_I^T M(I, r)^{-1} c(I, r).$$

**Remark 2.2.** Note that $x(I, r)$ is invariant with respect to the choice of $U_I$ because $x(I, r)$ is the unique optimal solution to an optimization problem. From this alternative characterization, we see how to compute $x(I, r)$ with the Gram-Schmidt algorithm and Gaussian elimination.

We establish some properties of $x(I, r)$:

**Proposition 2.1.** Let $I \subseteq [m]$ and $r \in \mathbb{R}^m$ such that $r > 0$. It holds that

$$A_I R_I^2 A_I^T x(I, r) = A_I R_I e_I. \quad (5)$$

And if $A_I$ is full column rank, then

$$x(I, r) = A_I R_I (R_I A_I^T A_I R_I)^{-1} e_I, \quad (6)$$

$$(A_I R_I)^T x(I, r) = e_I. \quad (7)$$

**Proof.** From the fact that $A_I = U_I^T U_I A_I$,

$$A_I R_I^2 A_I^T x(I) = U_I^T U_I A_I R_I^2 A_I^T x(I)$$

$$= U_I^T M(I, r) M(I, r)^{-1} c(I, r)$$

$$= U_I^T U_I A_I R_I e_I$$

$$= A_I R_I e_I,$$

where the last equality follow from $U_I^T U_I A_I = A_I$.

If $A_I$ is full column rank, then $A(I, r)$ is invertible, and we obtain

$$x(I, r) = U_I^T (R_I A_I^T U_I^T)^{-1} e_I$$

$$= U_I^T U_I A_I R_I (U_I A_I R_I)^{-1} (R_I A_I^T U_I^T)^{-1} e_I$$

$$= A_I R_I (R_I A_I^T U_I^T U_I A_I R_I)^{-1} e_I$$

$$= A_I R_I (R_I A_I^T A_I R_I)^{-1} e_I.$$
where the third and fourth equality follow from $U_I^t U_I A_I = A_I$.

If $A_I$ is full column rank, then by (6),

$$(A_I R_I)^\top x(I) = R_I A_I^\top x(I) = R_I A_I^\top A_I R_I (R_I A_I^\top)^{-1} e_I = e_I,$$

and thus the proof is complete.

It will also be convenient to define

$$\lambda(I, r) = \begin{cases} (R_I A_I^\top A_I R_I)^{-1} e_I & \text{rank}(A_I) = |I| \\ e_I - R_I A_I^\top x(I, r) & \text{rank}(A_I) < |I| \end{cases}$$

and to define the projection matrix

$$P(I, r) = \begin{cases} A_I R_I (R_I A_I^\top A_I R_I)^{-1} R_I A_I^\top & \text{rank}(A_I) = |I| \\ A(I, r)^\top M(I, r)^{-1} A(I, r) & \text{rank}(A_I) < |I|. \end{cases}$$

### 2.2 A characterization for USOs of the cube

In this subsection we prove Theorem 1.6. First, let us introduce some definitions and establish a lemma.

For an orientation $G$ of $\{0, 1\}^m$, we will say $a \in A(G)$ is aligned with $j \in [m]$ if there exists $I \subseteq [m]$ such that $a = (I \cup j, I)$ or $a = (I, I \cup j)$. And for each $I \subseteq [m]$, define $G_I$ to be the orientation obtained by changing the direction of all arcs in $G$ aligned with some $i \in I$.

**Lemma 2.1.** Let $I \subseteq [m]$ and $G$ be an orientation. If $G$ is a unique sink orientation, then $G_I$ is a unique sink orientation.

**Proof.** By induction, it is sufficient to show that $G_i$, for some $i \in I$, is a unique sink orientation. Let $I, J \subseteq [m]$ such that $I \subseteq J$. We will show that $G[I, J]$ contains a unique sink, implying $G_i$ is a unique sink orientation. If $i \in I$ or $i \notin J$, then $G[I, J] = G_i[I, J]$, and so, $G_i[I, J]$ contains a unique sink.

Accordingly, assume that $i \in J \setminus I$. Let $S_1$ denote the unique sink in $G[I, J]$. Suppose that $i \in S_1$. We claim that the unique sink in $G[I, J \setminus i] = G_i[I, J \setminus i]$, say $S_2$, is the unique sink in $G_i[I, J]$. From the definitions of $S_1$ and $S_2$,

$$N_{G[I, J]}^{-1}(S_2) = (J \setminus I) \setminus i.$$  \hfill (8)

From (8) and the definition of $G_i$,

$$N_{G_i[I, J]}^{-1}(S_2) = J \setminus I.$$  \hfill (9)

It immediately follows from (9) that $S_1$ is a sink in $G_i[I, J]$. Suppose there exists another sink $S_3 \neq S_2$ in $G_i[I, J]$. Note that $S_3 \neq S_1$ because by the definition of $G_i$, $S_1$ is not a sink in $G_i[I, J]$. From the definition of $G_i$, it follows that $S_3$ is either a sink in $G[I \cup i, J]$ or $G[I, J \setminus i]$, which yields a contradiction because $S_1$ and $S_2$ are the unique sinks of $G[I \cup i, J]$ and $G[I, J \setminus i]$, respectively.

The argument for $i \notin S$ is similar to the argument above, and thus excluded. Thus, the proof is complete. \qed
We are now prepared to prove Theorem 1.6:

Proof of Theorem 1.6. Suppose that $G$ is a unique sink orientation. For the sake of contradiction, suppose there exist $I, J, K \subseteq [m]$ such that $I \subseteq J \subseteq K$ and

$$N_{G[I,K]}^+(J) = N_{G[I,K]}^+(I \cup (K \setminus J)).$$

For notational convenience, let $L = N_{G[I,K]}^+(J) = N_{G[I,K]}^+(I \cup (K \setminus J))$. Consider the orientation $G_L$, which by Lemma 2.1 is a unique sink orientation. From the definition of $G_L$, it holds that $J$ and $I \cup (K \setminus J)$ are sinks in $G_L[I,K]$. This contradicts the fact that $G_L[I,K]$ contains a unique sink.

Suppose for all subsets $I, J, K \subseteq [m]$ such that $I \subseteq J \subseteq K$,

$$N_{G[I,K]}^+(J) \neq N_{G[I,K]}^+(I \cup (K \setminus J)). \quad (10)$$

It is sufficient to show that for all $U, V \subseteq [m]$ such that $U \subseteq V$, it holds that $G[U,V]$ contains a unique sink. We prove this statement by induction on the cardinality of $V \setminus U$. The statement holds trivially for $|V \setminus U| = 1$ because $G[U,V]$ is either an arc from $U$ to $V$ or from $V$ to $U$. Suppose the statement holds for $k \geq 1$ and $|V \setminus U| = k + 1$. For the sake of contradiction, suppose $G[U,V]$ does not contain a unique sink. Choose any $v \in V \setminus U$. From the induction hypothesis, $G[U,V \setminus v]$ and $G[U \cup v, V]$ contain unique sinks, say $S_1$ and $S_2$, respectively. Suppose that neither $S_1$ nor $S_2$ is a unique sink in $G[U,V]$. Then, for $I = S_1 \cap S_2$, $J = S_1$, and $K = S_1 \cup S_2$,

$$N_{G[I,K]}^+(J) = N_{G[I,K]}^+(S_1) = v = N_{G[I,K]}^+(S_2) = N_{G[I,K]}^+(I \cup (K \setminus J)),$$

which contradicts (10). Hence, at least one of $S_1$ or $S_2$ is a sink in $G[U,V]$. Suppose both $S_1$ and $S_2$ are sinks in $G[U,V]$. Then, for $I = S_1 \cap S_2$, $J = S_1$, and $K = S_1 \cup S_2$,

$$N_{G[I,K]}^+(J) = N_{G[I,K]}^+(S_1) = \emptyset = N_{G[I,K]}^+(S_2) = N_{G[I,K]}^+(I \cup (K \setminus J)),$$

which contradicts (10). Thus, $G[U,V]$ contains a unique sink, completing the induction.

\[ \square \]

2.3 Preliminary results

In this subsection we establish three preliminary results. Let $I \subseteq [m]$ and $r \in \mathbb{R}^m$ such that $r > 0$.

The first result relates $x(I \setminus i, r)$, for $i \in I$, and $\lambda(I, r)$ when $A_I$ is full column rank. And the result follows from the formula that expresses an inverse in terms of Schur complements.

**Proposition 2.2.** Let $I \subseteq [m]$ and $r \in \mathbb{R}^m$ such that $r > 0$. Suppose that $A_I$ is full-column rank. Then for all $i \in I$ it holds that

$$\text{sgn}(1 - r_i a_i^T x(I \setminus i, r)) = \text{sgn}(\lambda_i(I, r)).$$

**Proof.** Let $i \in I$, let $X_I = [A_{I \setminus i} R_{I \setminus i} \; r_i a_i]$, and let $Q$ be the $|I| \times |I|$ permutation matrix such that $A_I R_I = X_I Q$. It follows from the definition of $\lambda(I, r)$ that

$$\lambda_i(I, r) = e_i^T (R_I A_I^T A_I R_I)^{-1} e_I = e_i^T Q^T (X_I^T X_I)^{-1} Q e_I = e_i^T (\tilde{A}_I^T \tilde{A}_I)^{-1} e_I, \quad (11)$$

10
and now let us provide an alternative expression for \((X_I^TX_I)^{-1}\). Note that
\[
X_I^TX_I = \begin{bmatrix} R_{I,i}A_{I,i}^T A_{I,i} R_{I,i} & R_{I,i}A_{I,i}^T r_i a_i \\ r_i a_i^T R_{I,i} A_{I,i} & r_i a_i^T r_i a_i \end{bmatrix}.
\] (12)

For notational convenience, let
\[
\theta = r_i a_i^T r_i a_i - r_i a_i^T P_{I\setminus i} r_i a_i \quad \text{and} \quad v = (R_{I,i} A_{I,i}^T A_{I,i} R_{I,i})^{-1} R_{I,i} A_{I,i}^T r_i a_i.
\]

Note that \(\theta > 0\) because \(P_{I,i}\) is a projection matrix and \(a_i\) is not in the span of the columns of \(A_{I,i}\). From (12) and the formula that expresses an inverse in terms of Schur complements, we have
\[
(X_I^TX_I)^{-1} = \begin{bmatrix} (R_{I,i} A_{I,i}^T A_{I,i} R_{I,i})^{-1} + \theta^{-1} vv^T & -\theta^{-1} v \\ -\theta^{-1} v^T & \theta^{-1} \end{bmatrix}.
\] (13)

Substituting (13) into (11) gives
\[
\lambda_i(I, r) = \theta^{-1}(1 - v^T e) = \theta^{-1}(1 - r_i a_i^T x(I \setminus i, r)),
\] (14)
and the desired result follows from (14) and the fact that \(\theta > 0\).

The second result relates \(x(I, r)\) and \(x(I \setminus i, r)\), for \(i \in I\), when \(\text{rank}(A) = \text{rank}(A_{I \setminus i})\). And the result follows from the Sherman-Morrison formula.

**Proposition 2.3.** Let \(I \subseteq [m]\) and \(r \in \mathbb{R}^m\) such that \(r > 0\). Suppose that \(\text{rank}(A) = \text{rank}(A_{I \setminus i})\). Then for all \(i \in I\) it holds that
\[
\text{sgn}(1 - r_i a_i^T x(I, r)) = \text{sgn}(1 - r_i a_i^T x(I \setminus i, r)).
\]

**Proof.** Because \(\text{rank}(A_{I \setminus i}) = \text{rank}(A)\), the columns of \(A_{I \setminus i}\) span the same subspace as the columns of \(A_I\), and so we may suppose that \(R_I = R_{I \setminus i}\). From the definition of \(x(I, r)\) and the Sherman-Morrison formula,
\[
r_i a_i^T x(I, r) = r_i a_i^T U_I^T M(I, r)^{-1} c(I, r)
= r_i a_i^T U_I^T M(I \setminus i, r)^{-1} c(I, r)
- (r_i a_i^T U_I^T M(I \setminus i, r)^{-1} U_I r_i a_i)(r_i a_i^T U_I^T M(I \setminus i, r)^{-1} c(I, r))
\bigg/ \left[ 1 + r_i a_i^T U_I^T M(I \setminus i, r)^{-1} U_I r_i a_i \right]
= r_i a_i^T x(I \setminus i, r) + r_i a_i^T R_I^T M(I \setminus i, r)^{-1} r_i a_i
\bigg/ \left[ 1 + r_i a_i^T U_I^T M(I \setminus i, r)^{-1} U_I r_i a_i \right].
\] (15)

It follows from (15) that
\[
a_i^T x(I, r) - 1 = \frac{r_i a_i^T x(I \setminus i, r) + r_i a_i^T U_I^T M(I \setminus i, r)^{-1} U_I r_i a_i}{1 + r_i a_i^T U_I^T M(I \setminus i, r)^{-1} U_I r_i a_i} - 1
= \frac{r_i a_i^T x(I \setminus i, r) - 1}{1 + r_i a_i^T U_I^T M(I \setminus i, r)^{-1} U_I r_i a_i},
\] (16)
and the desired result follows from (16).
The last result relates $x(W, r)$ and $x(V, r)$, where $V \subseteq W$, and it follows from algebraic manipulation.

**Lemma 2.2.** Let $V, W \subseteq [m]$ such that $V \subseteq W$, and let $r \in \mathbb{R}^m$ such that $r > 0$. Then

$$x(W, r) = x(V, r) + U_W^T M(W, r)^{-1} U_W A_W V R_W V (e_W - R_W V A_W^T x(V)).$$

**Proof.** Let $\hat{x}(V, r) = U_W x(V, r)$ and $\hat{x}(W, r) = U_W x(W, r)$. Note that $x(V, r)$ is in the span of the columns of $A_W$. Because $U_W^T \hat{x}(W, r) = U_W^T U_W x(W, r) = x(W, r)$, it holds that

$$M(W, r) \hat{x}(W, r) = U_W A_W R_W^2 A_W^T x(W, r) = U_W A_W R_W e_W = U_W A_W R_W V e_W + U_W A_W V R_W V e_W V = U_W A_W R_W^2 A_V^T x(V, r) + U_W A_W V R_W V e_W V = U_W A_W R_W^2 A_V^T U_W \hat{x}(V, r) + U_W A_W R_W V e_W V = M(W, r) \hat{x}(V, r) + U_W A_W V R_W V (e_W V - R_W V A_W^T U_W \hat{x}(V, r)) = M(W, r) \hat{x}(V, r) + U_W A_W V R_W V (e_W V - R_W V A_W^T x(V, r)),$$

(17)

where the second equality follows from $A_W R_W^2 A_W^T x(W) = A_W R_W e_W$, the fourth equality follows from $A_V R_W^2 A_V^T x(V) = A_V R_V e_V$, the fifth equality follows from $x(V, r) = U_W^T U_W x(V, r) = U_W^T \hat{x}(V, r)$, and the last equality $U_W^T \hat{x}(V, r) = U_W^T U_W x(V, r) = x(V, r)$.

Left-multiplying (17) by $U_W^T M(W, r)^{-1}$, and then using $\hat{x}(V, r) = U_W x(V, r)$ and $\hat{x}(W, r) = U_W x(W, r)$, gives the desired result. 

$\square$

### 3 Construction of unique sink orientation

First we prove Theorem 1.1, and this will require three lemmas:

**Lemma 3.1.** Let $U \subseteq [m]$ and $I, J, K \subseteq U$ such that $I$, $J$, and $K$ form a partition of $U$. Also let $r \in \mathbb{R}^m$ such that $r > 0$. If $A_{I \cup J}$ and $A_{I \cup K}$ are full column rank, then

$$\sum_{j \in J} \lambda_j (I \cup J, r) (1 - r_j a_j^T x(I \cup K, r)) + \sum_{k \in K} \lambda_k (I \cup K, r) (1 - r_k a_k^T x(I \cup J, r)) \geq 0,$$
Proof. From Proposition 2.1, it follows that \( R_I^T A_J^T x(I \cup K, r) = e_I \), and so
\[
\sum_{j \in J} \lambda_j(I \cup J, r)(1 - r_j a_j^T x(I \cup K, r)) \\
= \sum_{j \in I \cup J} \lambda_j(I \cup J, r)(1 - r_j a_j^T x(I \cup K, r)) \\
= \lambda(I \cup J, r)^T (e_{I \cup J} - R_{I \cup J} A_{I \cup J}^T x(I \cup K, r)) \\
= e_{I \cup J}(R_{I \cup J} A_{I \cup J}^T A_{I \cup J} R_{I \cup J})^{-1} e_{I \cup J} - e_{I \cup J}(R_{I \cup J} A_{I \cup J}^T A_{I \cup J} R_{I \cup J})^{-1} R_{I \cup J} A_{I \cup J}^T x(I \cup K, r) \\
= e_{I \cup J}(R_{I \cup J} A_{I \cup J}^T A_{I \cup J} R_{I \cup J})^{-1} e_{I \cup J} - x(I \cup J, r)^T x(I \cup K, r) \\
\geq e_{I \cup J}(R_{I \cup J} A_{I \cup J}^T A_{I \cup J} R_{I \cup J})^{-1} e_{I \cup J} - [(x(I \cup J, r)^T x(I \cup J, r))(x(I \cup K, r)^T x(I \cup K, r))]^{\frac{1}{2}} \\
= e_{I \cup J}(R_{I \cup J} A_{I \cup J}^T A_{I \cup J} R_{I \cup J})^{-1} e_{I \cup J} \\
- [(e_{I \cup J}(R_{I \cup J} A_{I \cup J}^T A_{I \cup J} R_{I \cup J})^{-1} e_{I \cup J}))^{\frac{1}{2}} \\
- [(e_{I \cup J}(R_{I \cup J} A_{I \cup J}^T A_{I \cup J} R_{I \cup J})^{-1} e_{I \cup J}))^{\frac{1}{2}}]]^{\frac{1}{2}}. \tag{18}
\]
where the third equality follows from the definition of \( \lambda(I \cup J) \), the inequality follows from the Cauchy-Schwarz inequality, and the last equality follows from substituting
\[
x(I \cup J, r) = A_{I \cup J} R_{I \cup J}(R_{I \cup J} A_{I \cup J}^T A_{I \cup J} R_{I \cup J})^{-1} e_{I \cup J}
\]
and
\[
x(I \cup K, r) = A_{I \cup K} R_{I \cup K}(R_{I \cup K} A_{I \cup K}^T A_{I \cup K} R_{I \cup K})^{-1} e_{I \cup K}.
\]
A similar argument to above yields
\[
\sum_{k \in K} \lambda_k(I \cup K, r)(1 - r_k a_k^T x(I \cup J)) \\
\geq e_{I \cup K}(R_{I \cup K} A_{I \cup K}^T A_{I \cup K} R_{I \cup K})^{-1} e_{I \cup K} \\
- [(e_{I \cup K}(R_{I \cup K} A_{I \cup K}^T A_{I \cup K} R_{I \cup K})^{-1} e_{I \cup K}))^{\frac{1}{2}} \\
- [(e_{I \cup K}(R_{I \cup K} A_{I \cup K}^T A_{I \cup K} R_{I \cup K})^{-1} e_{I \cup K}))^{\frac{1}{2}}]]^{\frac{1}{2}}. \tag{19}
\]
Adding (18) and (19) gives
\[
\sum_{j \in J} \lambda_j(I \cup J, r)(1 - a_j^T x(I \cup K, r)) + \sum_{k \in K} \lambda_k(I \cup K, r)(1 - a_k^T x(I \cup J, r)) \\
\geq e_{I \cup J}(R_{I \cup J} A_{I \cup J}^T A_{I \cup J} R_{I \cup J})^{-1} e_{I \cup J} + e_{I \cup K}(R_{I \cup K} A_{I \cup K}^T A_{I \cup K} R_{I \cup K})^{-1} e_{I \cup K} \\
- 2[(e_{I \cup K}(R_{I \cup K} A_{I \cup K}^T A_{I \cup K} R_{I \cup K})^{-1} e_{I \cup K}))^{\frac{1}{2}} \\
- (e_{I \cup J}(R_{I \cup J} A_{I \cup J}^T A_{I \cup J} R_{I \cup J})^{-1} e_{I \cup J}))^{\frac{1}{2}}]]^{\frac{1}{2}} \\
\geq 0,
\]
where the last inequality follows from arithmetic-geometric mean inequality. The proof is complete. \(\square\)

Lemma 3.2. Let \( U \subseteq [m] \) and \( I, J, K \subseteq U \) such that \( I, J, \) and \( K \) form a partition of \( U \). Also let \( r \in \mathbb{R}^m \) such that \( r > 0 \). Then
\[
(R_{I \cup J} A_{I \cup K}^T x(I \cup J, r) - e_{I \cup K})^T (R_{I \cup K} A_{I \cup K}^T x(I \cup K, r) - e_{I \cup K}) \geq 0.
\]
Proof. Define $g \in \mathbb{R}^m$ by
\[ g_K = e_K - R_K A_K^T x(I \cup J, r) \]
and define $h \in \mathbb{R}^m$ by
\[ h_J = e_J - R_J A_J^T x(I \cup K, r) \]
and $h_{[m]\setminus J} = 0$.

From Lemma 4 with $V = I \cup J$ and $W = U$, it holds that
\[
R_U A_U^T x(U, r) = R_U A_U^T x(I \cup J, r) + A(U, r)^T M(U, r)^{-1} A(U, r) g_U.
\]
Equating (20) and (21), and rearranging, gives
\[
R_U A_U^T x(U, r) = R_U A_U^T x(I \cup K, r) + A(U, r)^T M(U, r)^{-1} A(U, r) h_U.
\]
Similarly, from Lemma 4 with $V = I \cup K$ and $W = U$, it holds that
\[
R_U A_U^T x(I \cup J, r) = R_U A_U^T x(I \cup K, r) + A(U, r)^T M(U, r)^{-1} A(U, r) h_U - g_U.
\]
Left-multiplying the left-hand side of (22) by $(h_U - g_U)^T$ gives
\[
(h_U - g_U)^T (R_U A_U^T x(I \cup J, r) - R_U A_U^T x(I \cup K, r))
= (e_J - R_J A_J^T x(I \cup K, r))^T R_J A_J^T x(I \cup J, r)
\]
and left-multiplying the right-hand side of (22) by $(h_U - g_U)^T$ gives
\[
(h_U - g_U)^T A(U, r)^T M(U, r)^{-1} A(U, r) (h_U - g_U)
\leq (h_U - g_U)^T (h_U - g_U)
= (e_J - R_J A_J^T x(I \cup K, r))^T (e_J - R_J A_J^T x(I \cup K, r))
+ (e_K - R_K A_K^T x(I \cup J, r))^T (e_K - R_K A_K^T x(I \cup J, r)),
\]
and the desired result follows from (25).
Lemma 3.3 below establishes the main result we will need for the construction. Consider the following interpretation: for \( I, J, K \subseteq [m] \) such that \( J \in [I, K] \), the lemma states that there exists \( \ell \in K \setminus I \) such that \( x(J \setminus \ell, \hat{r}) \) and \( x(I \cup (K \setminus J)) \setminus \ell, \hat{r} \) admit at least some similar structure in the sense that either (26) or (27) holds.

**Lemma 3.3.** Let \( r \in \mathbb{R}^m \) such that \( r > 0 \), and suppose that for all \( I \subseteq [m] \) and \( j \notin I \), it holds that \( r_j a_j^\top x(I, r) \neq 1 \). Then for all \( I, J, K \subseteq [m] \) such that \( J \in [I, K] \), there exists \( \ell \in K \setminus I \) such that one of the following holds:

\[
\begin{align*}
&\sum_{\ell \in K \setminus J} (1 - r_{\ell} a_{\ell}^\top x(J, r)) \lambda_{\ell}(I \cup (K \setminus J), r) + \sum_{\ell \in J} (1 - r_{\ell} a_{\ell}^\top x(I \cup (K \setminus I)), r) \lambda_{\ell}(J, r) < 0, \tag{29} \\
&\text{where} \ \sum_{\ell \in K \setminus J} (1 - r_{\ell} a_{\ell}^\top x(J, r)) \lambda_{\ell}(I \cup (K \setminus J), r) < 0.
\end{align*}
\]

which will contradict Lemma 3.1. Let \( \ell \in J \). From (28),

\[
\text{sgn}(1 - r_{\ell} a_{\ell}^\top x(I \cup (K \setminus J)), r) = \text{sgn}(r_{\ell} a_{\ell}^\top x(J \setminus \ell, r) - 1),
\]

and from Proposition 2.2,

\[
\text{sgn}(r_{\ell} a_{\ell}^\top x(J \setminus \ell, r) - 1) = \text{sgn}(-\lambda_{\ell}(J, r)).
\]

From (30) and (31),

\[
\sum_{\ell \in J} (1 - r_{\ell} a_{\ell}^\top x(I \cup (K \setminus I)), r) \lambda_{\ell}(J, r) < 0.
\]

A similar argument establishes

\[
\sum_{\ell \in K \setminus J} (1 - r_{\ell} a_{\ell}^\top x(J, r)) \lambda_{\ell}(I \cup (K \setminus J), r) < 0.
\]

Adding (32) and (33) gives (29), a contradiction.

**Case 1.** Suppose that \( A_J \) and \( A_{I \cup (K \setminus J)} \) are full column rank. We show that

\[
\sum_{\ell \in K \setminus J} (1 - r_{\ell} a_{\ell}^\top x(J, r)) \lambda_{\ell}(I \cup (K \setminus J), r) + \sum_{\ell \in J} (1 - r_{\ell} a_{\ell}^\top x(I \cup (K \setminus I)), r) \lambda_{\ell}(J, r) < 0,
\]

Without loss of generality, suppose rank\((A_J) < |J|\). We show that

\[
(R_{K \setminus I} A_{K \setminus I}^\top x(J, r) - e_{K \setminus I})^\top (R_{K \setminus I} A_{K \setminus I}^\top x(I \cup (K \setminus J), r) - e_{K \setminus I}) < 0,
\]

and (34) will contradict Lemma 3.2. From (28) and Proposition 2.3,

\[
(R_{J \setminus I} A_{J \setminus I}^\top x(J, r) - e_{J \setminus I})^\top (R_{J \setminus I} A_{J \setminus I}^\top x(I \cup (K \setminus J), r) - e_{J \setminus I}) < 0.
\]
Suppose that \( \text{rank}(A_{I \cup (K \setminus J)}) < |I \cup (K \setminus J)| \). Then a similar argument yields
\[
(R_{K \setminus J}A_{K \setminus J}^T x(J, r) - e_{K \setminus J})^T (R_{K \setminus J}A_{K \setminus J}^T x(I \cup (K \setminus J), r) - e_{K \setminus J}) < 0.
\]
(36)

Adding (35) and (36) gives (34), a contradiction. Suppose that \( \text{rank}(I \cup (K \setminus J)) = |I \cup (K \setminus J)| \). Then
\[
(R_{K \setminus I}A_{K \setminus I}^T x(J, r) - e_{K \setminus I})^T (R_{K \setminus I}A_{K \setminus I}^T x(I \cup (K \setminus J), r) - e_{K \setminus I}) < 0,
\]
(37)
where the first equality follows from \( A_{J \setminus J}^T x(I \cup (K \setminus J)) = e_{J \setminus J} \), and the inequality follows from (35). The result follows because (37) yields (34).

We are now prepared to prove Theorem 1.1:

**Proof of Theorem 1.1.** For notational convenience, let \( G = G(r) \). Let \( I, J, K \subseteq [m] \) such that \( I \subseteq J \subseteq K \). To establish \( G \) is a unique sink orientation, it is sufficient to show, by Theorem 1.6, that
\[
N_{G_{I \setminus I}, K}^+(J) \neq N_{G_{I \setminus I}, K}^+(I \cup (K \setminus J)).
\]
(38)

Lemma 3.3 gives that there exists \( \ell \in K \setminus I \) such that either (26) or (27) holds. Suppose that (26) holds. And, without loss of generality, suppose \( \ell \in J \). Note that it follows that \( (I \cup (K \setminus J)) \setminus \ell = I \cup (K \setminus J) \). From the construction of \( G \) and (26),
\[
(I, J \setminus \ell) \in \mathcal{A}(G),
\]
\[
(I \cup (K \setminus J) \cup \ell, I \cup (K \setminus J)) \in \mathcal{A}(G).
\]
(39)

From (39), it holds that \( \ell \in N_{G_{I \setminus I}, K}^+(J) \) and \( \ell \notin N_{G_{I \setminus I}, K}^+(I \cup (K \setminus J)) \). And it follows that (38) holds. The argument for the case in which (27) holds is similar. Thus the proof is complete.

Let \( S \) be the unique sink of \( G(r) \); the following properties are immediate from the construction of \( G(r) \):
\[
(A_{[m] \setminus S} R_{[m] \setminus S})^T x(S, r) > e_{[m] \setminus S},
\]
(40)
\[
r_a^T x(S \setminus i, r) < 1 \text{ for all } i \in S.
\]
(41)

Next we prove Theorem 1.2; the result follows from the two lemmas below.

**Lemma 3.4.** Let \( r \in \mathbb{R}^m \) such that \( r > 0 \), and suppose that for all \( I \subseteq [m] \) and \( j \notin I \), it holds that \( r_j a_j^T x(I, r) \neq 1 \). Also suppose that \( (P) \) is feasible. Then \( \text{rank}(A_S) = |S| \), and \( x(S, r) \) is a solution to \( (P) \).

**Proof.** Let \( S \) denote the unique sink of \( G(r) \). Suppose that \( \text{rank}(A_S) < |S| \). We will show that \( \lambda \) defined as

\[
\lambda_S = \lambda(S, r)
\]
\[
\lambda_{[m] \setminus S} = 0
\]

16
3.4 establishes that \( \lambda \) that \((40)\) follows from \((40)\) that \( x \)

Proof. Let \( U \)

consider \( U \)

hypothesis of the theorem, it holds for every arc \((C)\). Note that from the definition of \( H \)

Suppose for the sake of contradiction that \( \lambda \)

Proof of Theorem 1.3.

In this section we prove Theorem 1.3:

\[ A R \lambda = A_S R S \lambda (S, r) \]
\[ = U^\top_S U_S A_S R S \lambda (S, r) \]
\[ = U^\top_S U_S A_S R S (e_S - R_S A_S^\top x(S, r)) \]
\[ = U^\top_S (U_S A_S R_S e_s - M(S, r) M(S, r)^{-1} U_S A_S R_S e_s), \]
\[ = 0. \]

where the second equality follows from the fact that \( U^\top_S U_S A_S = A_S \). Further observe that \( \lambda_S = e_S - (A_S R_S)^\top x(S, r) > 0 \) from \((41)\) and Proposition 2.3. Hence \( \lambda \geq 0 \), and thus, \( \lambda(S, r) \)

is feasible for \((Alt)\), a contradiction.

Because \( \text{rank}(A_S) = |S| \), it follows from Proposition 2.1 that \( R_S A_S^\top x(S, r) = e_S \). Accordingly, from \((40)\), we have that \( x(S, r) \) is a solution to \((P)\).

\[ \text{Lemma 3.5.} \text{ Let} \ r \in \mathbb{R}^m \text{ such that} \ r > 0, \text{ and suppose that for all} \ I \subseteq [m] \text{ and} \ j \notin I, \text{ it holds that} \ r_j a_j^\top x(I, r) \neq 1. \text{ Also suppose that} \ (P) \text{ is infeasible. Then} \ \text{rank}(A_S) < |S|, \text{ and} \ \lambda \text{ defined by} \]
\[ \lambda_S = \lambda(S, r), \]
\[ \lambda_{[m] \setminus S} = 0. \]

is a solution to \((Alt)\).

Proof. Let \( S \) denote the unique sink of \( G \). Suppose \( \text{rank}(A_S) = |S| \). Then, \( A_S^\top x(S) = e_S \), and it follows from \((40)\) that \( x(S, r) \) is a solution to system \((P)\). This contradicts the hypothesis that \((P)\) is infeasible. Because \( \text{rank}(A_S) < |S| \), the argument given in the proof of Theorem 3.4 establishes that \( \lambda \) is feasible for \((Alt)\).

The theorem in the introduction is immediate.

4 A local property not satisfied by all Holt-Klee USOs

In this section we prove Theorem 1.3:

Proof of Theorem 1.3. Suppose for the sake of contradiction that \( H_{G(r)}(I) \) contains a directed cycle \( C \). Note that from the definition of \( H_{G(r)}(I) \), the definition of \( G(r) \), and the hypothesis of the theorem, it holds for every arc \((j, k) \in \mathcal{A}(C) \)

\[ \text{sgn}(1 - r_k a_k^\top x(I, r)) \neq \text{sgn}(1 - r_k a_k^\top x(I \cup j)). \]

(42)

Fix any choice for \( U_I \), and for each vertex \( j \in V(C) = [m] \setminus I \), it will be convenient to consider \( U_{I \cup j} \) defined as follows:

1. If \( a_j \in \text{span} \{ a_i \}_{i \in I} \), define \( U_{I \cup j} = U_I \).

2. If \( a_j \notin \text{span} \{ a_i \}_{i \in I} \), take \( u_j \in \text{span} \{ a_i \}_{i \in I} \cap S^{n-1} \) such that \( u_j \) together with the rows of \( U_I \) form an orthonormal collection that spans the subspace \( \text{span} \{ a_i \}_{i \in I \cup j} \), and define \( U_{I \cup j} \) by collecting \( u_j \) and the rows of \( U_I \) into the rows of \( U_{I \cup j} \).
For each arc \((j, k) \in A(C)\), it will be convenient to define
\[
p(j, k) := r_k a_k^T U_{I \cup j}^T M(I \cup j)^{-1} U_{I \cup j} r_j a_j.
\]

Lemma and the definition of \(p(j, k)\) imply
\[
r_k a_k^T x(I \cup j, r) = r_k a_k^T x(I, r) + p(j, k) (1 - r_j a_j^T x(I, r)).
\]
(43)

For \((j, k) \in A(C)\) such that \(r_k a_k^T x(I, r) < 1\), it follows from (43) that
\[
p(j, k) (1 - r_j a_j^T x(I, r)) = r_k a_k^T x(I \cup j, r) - r_k a_k^T x(I, r) > 1 - r_k a_k^T x(I, r) > 0,
\]
(44)

where the first inequality follows from (42) and the theorem hypothesis. It follows from a similar argument that for \((j, k) \in A(C)\) such that \(r_k a_k^T x(I, r) > 1\),
\[
p(j, k) (r_j a_j^T x(I, r) - 1) \geq r_k a_k^T x(I, r) - 1 > 0.
\]
(45)

From (44), (45), and rearranging,
\[
\prod_{(j, k) \in A(C)} p(j, k) \prod_{j \in V(C): r_j a_j^T x(I, r) < 1} (1 - r_j a_j^T x(I, r)) \prod_{j \in V(C): r_j a_j^T x(I, r) > 1} (r_j a_j^T x(I, r) - 1) > 0,
\]

and hence
\[
\prod_{(j, k) \in A(C)} p(j, k) > 1.
\]
(46)

Now we consider three cases:

**Case 1.** Suppose that \(a_j \in \text{span}\{a_i\}_{i \in I}\) for all \(j \in V(C)\). Note that this implies \(U_{I \cup j} = U_I\). It follows from the Sherman-Morrison formula and \(U_{I \cup j} = U_I\) that
\[
p(j, k) = r_k a_k^T U_{I \cup j}^T M(I \cup j)^{-1} U_{I \cup j} r_j a_j
= r_k a_k^T U_I^T M(I, r)^{-1} U_I r_j a_j - \frac{(r_k a_k^T U_I^T M(I, r)^{-1} U_I r_j a_j)(r_j a_j^T U_I^T M(I, r)^{-1} U_I r_j a_j)}{1 + r_j a_j^T U_I^T M(I, r)^{-1} U_I r_j a_j}
= \frac{r_k a_k^T U_I^T M(I, r)^{-1} U_I r_j a_j}{1 + r_j a_j^T U_I^T M(I, r)^{-1} U_I r_j a_j}.
\]
(47)
From (46) and (47), it holds that
\[
\prod_{j \in V(C)} (1 + r_j a_j^T U_j M(I, r)^{-1} U_j r_j a_j) < \prod_{(j,k) \in A(C)} r_k a_k^T U_j M(I, r)^{-1/2} M(I, r)^{-1/2} U_j r_j a_j
\]
\[= \prod_{(j,k) \in A(C)} r_k a_k^T U_j M(I, r)^{-1/2} (r_j a_j^T U_j M(I, r)^{-1/2} U_j r_j a_j)^{1/2}
\]
\[= \prod_{j \in V(C)} r_j a_j^T U_j M(I, r)^{-1} U_j r_j a_j,
\] (48)
where the first equality follows because \(M(I, r)\) is positive definite, the second inequality follows from the Cauchy-Schwarz inequality, and the last equality follows from the fact that \(C\) is a cycle. Inequality (48) yields a contradiction, completing Case 1.

Before proceeding to the second and third cases, let us establish a preliminary result. For \(j \in V(C)\) such that \(a_j \notin \text{span}\{\{a_i\}_{i \in I}\}\) it holds that
\[M(I \cup j, r) \frac{1}{r_j a_j^T u_j} e_j = A(I \cup j, r) \frac{1}{r_j a_j^T u_j} A(I \cup j, r)^T e_j
\]
\[= A(I \cup j, r) \frac{1}{r_j a_j^T u_j} R_{I \cup j}^T A_{I \cup j}^T u_j^T
\]
\[= A(I \cup j, r) e_j = U_{I \cup j} r_j a_j,
\] (49)
where the second equality follows from the fact that \(r_i a_i^T u_j = 0\) for all \(i \in I\) as \(u_j \in \text{span}\{\{a_i\}_{i \in I}\}\). It follows from (49) that \(M(I \cup j)^{-1} U_{I \cup j} r_j a_j = \frac{1}{r_j a_j^T u_j} e_j\), and hence for \((j,k) \in A(C)\),
\[p(j, k) = \frac{1}{r_j a_j^T u_j} r_k a_k^T U_{I \cup j}^T e_j = \frac{r_k a_k^T u_j}{r_j a_j^T u_j}.
\] (50)

**Case 2.** Suppose that there exist \(v, w \in V(C)\) such that \(a_v \in \text{span}\{\{a_i\}_{i \in I}\}\) and \(a_w \notin \text{span}\{\{a_i\}_{i \in I}\}\). Then because \(C\) is a cycle, there exists an arc \((j,k) \in A(C)\) such that \(a_j \notin \text{span}\{\{a_i\}_{i \in I}\}\) and \(a_k \in \text{span}\{\{a_i\}_{i \in I}\}\). It follows from (50) that \(p(v,w) = 0\), which contradicts (46), completing Case 2.

**Case 3.** Suppose that \(a_j \notin \text{span}\{\{a_i\}_{i \in I}\}\) for all \(j \in V(C)\). Then from (46) and (50),
\[1 < \prod_{(j,k) \in A(C)} p(j,k) = \prod_{(j,k) \in A(C)} \frac{r_k a_k^T u_j}{r_j a_j^T u_j} = \prod_{(j,k) \in A(C)} \frac{r_k a_k^T u_j}{r_k a_k^T u_k} = \prod_{(j,k) \in A(C)} \left| \frac{u_j^T a_k}{u_k^T a_k} \right| \leq 1,
\] (51)
where the second equality follows from the fact that \(C\) is a cycle, the third equality follows from the fact that the product is greater than one (from the first inequality), and the last inequality follows from \(|u_j^T a_k| \leq |u_k^T a_k|\) for all \(u \in \text{span}\{\{a_i\}_{i \in I}\}^\perp \cap S^{n-1}\) and \(k \in V(C)\). Inequality (51) yields a contradiction, completing the proof. \(\square\)
5 A connection to condition measure theory

First we prove Theorem 1.4:

Proof of Theorem 1.4. Clearly $x$ is feasible for (1). Accordingly, it is sufficient to show $\lambda$ is feasible for (2) and the objective value of (2) evaluated at $\lambda$ equals the objective value of (1) evaluated at $x$. From (41) and Proposition 2.2, it follows that $\lambda_S > 0$, and hence, $\lambda \geq 0$. Also, $e^\top \lambda = e_S^\top \lambda_S = 1$. Thus, $\lambda$ is feasible for (2).

The objective value of (2) evaluated at $\lambda$ is

$$
\|AR\lambda\| = \frac{1}{e_S^\top \lambda(S, r)} \|A_S R_S \lambda(S, r)\| = \frac{1}{(e_S^\top (R_S A_S^\top A_S R_S)^{-1} e_S)^{1/2}},
$$

(52)

where the first equality follows from the definition of $\lambda$, and the second equality follows from substituting $\lambda(S, r) = (R_S A_S^\top A_S R_S)^{-1} e_S$.

The objective value of (1) evaluated at $x$ is

$$
\min_{i \in [m]} a_i^\top \left( \frac{1}{\|x(S, r)\|} x(S, r) \right) = \frac{1}{\|x(S, r)\|} \min_{i \in S} a_i^\top x(S, r)
$$

$$
= \frac{1}{\|x(S, r)\|} = \frac{1}{(e_S^\top (R_S A_S^\top A_S R_S)^{-1} e_S)^{1/2}};
$$

(53)

where the first equality follows from $R_S A_S^\top x(S, r) = e_S$ and (40), the second equality follows from $R_S A_S^\top x(S, r) = e_S$, and the last equality follows from substituting $x(S, r) = A_S R_S (R_S A_S^\top A_S R_S)^{-1} e_S$.

From (52) and (53), the objective value of (2) evaluated at $\lambda$ equals the objective value of (1) evaluated at $x$. □

Next we prove Theorem 1.5:

Proof of Theorem 1.5. It is sufficient to show that $\lambda$ is feasible for (3), $(x, \alpha, \gamma)$ is feasible for (4), and the objective of (3) evaluated at $\lambda$ is equal to the objective of (4) evaluated at $(x, \alpha, \gamma)$.

Observe that $AR\lambda = \frac{1}{e_S^\top \lambda(S, r)} A_S R_S \lambda(S, r) = 0$. Also observe that $\lambda(S, r) = e_S - R_S A_S^\top x(S, r) > 0$, where the strict inequality follows from (41) and Proposition 2.3. Hence, $\lambda(S, r) \geq 0$. Finally, $e^\top \lambda = e_S^\top \lambda_S = 1$. Thus, $\lambda$ is feasible for (3).

Observe that

$$
(\alpha + \gamma e - (AR)^\top x)_{[m]\setminus S} = \alpha_{[m]\setminus S} + \gamma e_{[m]\setminus S} - R_{[m]\setminus S} A_{[m]\setminus S}^\top x = 0,
$$

(54)

where the last equality follows from the definitions of $x$, $\alpha_{[m]\setminus S}$, and $\gamma$. Also, observe that

$$
(\alpha + \gamma e - RA^\top x)_S = \alpha_S + \gamma e_S - R_S A_S^\top x = \frac{1}{\|\lambda(S, r)\|} (e_S - R_S A_S^\top x(S, r))
$$

$$
= \frac{1}{\|\lambda(S, r)\|} \lambda(S, r),
$$

(55)
where the third equality follows from the definitions of $x$, $\alpha$, and $\gamma$. From (54) and (55), it follows that $\|\alpha + \gamma e - RA^\top x\| = 1$. Also $\alpha_{[m] \setminus S} = \frac{1}{\lambda_{[S \setminus r]}(S,R)} \left( R_{[m] \setminus S} A^\top_{[m] \setminus S} x(S,r) - e_{[m] \setminus S} \right) > 0$, where the inequality follows from (40). Thus, $(x, \alpha, \gamma)$ is feasible for (4).

The objective of (3) evaluated at $\lambda$ is

$$\|\lambda\| = \frac{\|\lambda(S,r)\|}{\epsilon^\top S\lambda(S,r)} = \frac{(e^\top_S(I - P(S,r))^2e_S)^{1/2}}{(e^\top_S(I - P(S,r))e_S)^{1/2}} = \frac{1}{(e^\top_S(I - P_S)e_S)^{1/2}},$$

where the last equality follows from the fact that $I - P(S,r)$ is a projection matrix. Also, the objective of (4) evaluated at $(x, \alpha, \gamma)$ is

$$\gamma = \frac{1}{\|\lambda(S,r)\|} = \frac{1}{(e^\top_S(I - P(S,r))e_S)^{1/2}} = \frac{1}{e^\top_S(I - P(S,r))e_S},$$

where again the last equality follows from the fact that $I - P(S,r)$ is a projection matrix. Thus, the objectives are equal. \[ \square \]

References


