EXTENDED FORMULATIONS FOR CONVEX HULLS OF SOME BILINEAR FUNCTIONS

AKSHAY GUPTE[1], THOMAS KALINOWSKI[2,3], FABIAN RIGTERINK[3], AND HAMISH WATERER[3]

Abstract. We consider the problem of characterizing the convex hull of the graph of a bilinear function f on the n-dimensional unit cube $[0,1]^n$. Extended formulations for this convex hull are obtained by taking subsets of the facets of the Boolean Quadric Polytope (BQP). Extending existing results, we propose a systematic study of properties of f that guarantee that certain classes of BQP facets are sufficient for an extended formulation. We use a modification of Zuckerberg's geometric method for proving convex hull characterizations [Geometric proofs for convex hull defining formulations, Operations Research Letters **44** (2016), 625–629] to prove some initial results in this direction. In particular, we provide small-sized extended formulations for bilinear functions whose corresponding graph is either a cycle with arbitrary edge weights or a clique or an almost clique with unit edge weights.

1. Introduction

An important technique in global optimization is the construction of convex envelopes for nonconvex functions, and there is a significant amount of literature on characterizing convex hulls of graphs of nonlinear functions, beginning with [Rik97; She97]; see also the book [LS13]. Rikun studies the question when this convex hull is a polyhedron and gives a complete characterization for functions on polyhedral domains. Even if the convex hull is a polyhedron there might be a very large number of facets, and this is reminiscent of a situation which is quite common in combinatorial optimization: for a natural mixed integer programming (MIP) formulation the convex hull of the feasible set can be described explicitly, but it is a polytope whose number of facets is exponential in the instance size. One approach that has been successful in this area is the use of *extended formulations* [CCZ10]. The basic idea is to introduce more variables in order to reduce the number of constraints. A reformulation of the convex hull with a polynomial number of constraints and polynomially many additional variables is called a *compact extended formulation*, and this is a key ingredient in so-called lift-and-project methods, and other related MIP formulation techniques [BCC93; SA90]. In this paper, we use a similar technique: instead of describing the convex hull of a graph of a bilinear function in the original variable space, we seek to describe it in a lifted space with as few inequalities as possible.

A bilinear function is a function $f:[0,1]^n\to\mathbb{R}$ of the form

$$f(\boldsymbol{x}) = \sum_{1\leqslant i<j\leqslant n} a_{ij}x_ix_j$$

with coefficients $a_{ij}\in\mathbb{R}$. The convex hull of the graph of f is the set

$$X(f) := \operatorname{conv}\{(\boldsymbol{x},z)\in[0,1]^n\times\mathbb{R}:\ z=f(\boldsymbol{x})\},$$

[1]School of Mathematical & Statistical Sciences, Clemson University, Clemson, USA
[2]School of Science and Technology, University of New England, Armidale, Australia
[3]School of Mathematical & Physical Sciences, University of Newcastle, Callaghan, Australia
E-mail addresses: agupte@clemson.edu, tkalinow@une.edu.au, fabian.rigterink@gmail.com, hamish.waterer@newcastle.edu.au.

Date: February 5, 2019.
2010 *Mathematics Subject Classification.* 90C57, 90C26, 52B12.
Key words and phrases. extended formulation, convex hull, bilinear, quadratic, boolean quadric polytope.
Research of AG was supported by the US Office of Naval Research grant N00014-16-1-2725.
Research of TK, FR, HW was supported by the ARC Linkage grant no. LP110200524, Hunter Valley Coal Chain Coordinator (hvccc.com.au) and Triple Point Technology (tpt.com).

which is a polytope since

$$X(f) = \operatorname{conv}\{(\boldsymbol{x}, z) \in \{0,1\}^n \times \mathbb{R} : \ z = f(\boldsymbol{x})\} \tag{1}$$

which was proved in [Rik97; She97]. These functions arise in many problem areas; see [DG15; Gup+13; Gup+17] and the references therein.

For bilinear functions, a natural setting for an extended formulation is to introduce additional variables y_{ij} representing the product $x_i x_j$ of two original variables for $a_{ij} \neq 0$. The classical McCormick inequalities [McC76] for relaxing each bilinear term are

$$y_{ij} \geqslant 0, \qquad y_{ij} \leqslant x_i, \qquad y_{ij} \leqslant x_j, \qquad x_i + x_j - y_{ij} \leqslant 1, \tag{2}$$

and they are exact at 0–1 points, that is, they imply $y_{ij} = x_i x_j$ when $x_i, x_j \in \{0,1\}$. The McCormick relaxation is the polytope

$$M := \left\{ (\boldsymbol{x}, \boldsymbol{y}, z) \in [0,1]^{n(n+1)/2} \times \mathbb{R} : \ z = \sum_{1 \leqslant i < j \leqslant n} a_{ij} y_{ij}, \ (2) \text{ for all } 1 \leqslant i < j \leqslant n \right\},$$

whose projection is typically a relaxation of $X(f)$. The cases where the projection of M is actually equal to $X(f)$ have been characterized in [MSF15] and independently in [Bol+17], and there are also some results in this regard for multilinear functions [LNL12]. In general, the McCormick relaxation can be quite weak [Bol+17], and the purpose of this paper is to investigate extended formulations for $X(f)$ obtained as strengthenings of the McCormick relaxation.

As is customary in the literature, let the functions $\operatorname{vex}[f] : [0,1]^n \to \mathbb{R}$ and $\operatorname{cav}[f] : [0,1]^n \to \mathbb{R}$, denoting the convex and concave envelopes, respectively, of f over $[0,1]^n$, be defined as

$$\operatorname{vex}[f](\boldsymbol{x}) = \min\{z : \ (\boldsymbol{x}, z) \in X(f)\}, \qquad \operatorname{cav}[f](\boldsymbol{x}) = \max\{z : \ (\boldsymbol{x}, z) \in X(f)\},$$

so that

$$X(f) = \{(\boldsymbol{x}, z) \in [0,1]^n \times \mathbb{R} : \ \operatorname{vex}[f](\boldsymbol{x}) \leqslant z \leqslant \operatorname{cav}[f](\boldsymbol{x})\}.$$

Introducing variables y_{ij} to represent the products $x_i x_j$, we are interested in describing $X(f)$ in terms of the x- and y-variables. To be more precise, we define a function $\pi[f] : \mathbb{R}^n \times \mathbb{R}^{n(n-1)/2} \to \mathbb{R}^{n+1}$ by

$$\pi[f](\boldsymbol{x}, \boldsymbol{y}) = \left(\boldsymbol{x}, \sum_{1 \leqslant i < j \leqslant n} a_{ij} y_{ij} \right),$$

and extend it to the power set of $\mathbb{R}^n \times \mathbb{R}^{n(n-1)/2}$ in the usual way:

$$\pi[f](P) = \{\pi[f](\boldsymbol{x}, \boldsymbol{y}) : \ (\boldsymbol{x}, \boldsymbol{y}) \in P\}$$

for every $P \subseteq \mathbb{R}^n \times \mathbb{R}^{n(n-1)/2}$. For a polytope P, let the functions $\operatorname{LB}_P[f] : [0,1]^n \to \mathbb{R}$ and $\operatorname{UB}_P[f] : [0,1]^n \to \mathbb{R}$ be defined as

$$\operatorname{LB}_P[f](\boldsymbol{x}) = \min \left\{ \sum_{1 \leqslant i < j \leqslant n} a_{ij} y_{ij} : \ (\boldsymbol{x}, \boldsymbol{y}) \in P \right\} = \min\{z : (\boldsymbol{x}, z) \in \pi[f](P)\},$$

$$\operatorname{UB}_P[f](\boldsymbol{x}) = \max \left\{ \sum_{1 \leqslant i < j \leqslant n} a_{ij} y_{ij} : \ (\boldsymbol{x}, \boldsymbol{y}) \in P \right\} = \max\{z : (\boldsymbol{x}, z) \in \pi[f](P)\},$$

respectively, so that

$$\pi[f](P) = \{(\boldsymbol{x}, z) \in [0,1]^n \times \mathbb{R} : \operatorname{LB}_P[f](\boldsymbol{x}) \leqslant z \leqslant \operatorname{UB}_P[f](\boldsymbol{x})\}. \tag{3}$$

Our aim is to find a polytope P such that $X(f) = \pi[f](P)$, so that this P is a compact extended formulation. Observe that

$$X(f) = \pi[f](P) \iff \operatorname{LB}_P[f](\boldsymbol{x}) = \operatorname{vex}[f](\boldsymbol{x}), \ \operatorname{UB}_P[f](\boldsymbol{x}) = \operatorname{cav}[f](\boldsymbol{x}), \quad \text{for all } \boldsymbol{x} \in [0,1]^n. \tag{4}$$

There are constructive methods for deriving extended formulations of $X(f)$ with exponentially many variables and facet-defining inequalities, such as using the extreme point characterization in (1) or the nontrivial approach of using the Sherali-Adams hierarchy [SA90] which can also be applied to more general nonlinear functions [BM14]. We restrict our attention to finding extended formulations in the quadratic space of $(\boldsymbol{x}, \boldsymbol{y})$ variables.

Padberg [Pad89] introduced the *Boolean Quadric Polytope* (BQP), which is the convex hull of the binary vectors satisfying the McCormick inequalities (2),

$$QP := \operatorname{conv}\left\{(\boldsymbol{x},\boldsymbol{y}) \in \{0,1\}^{n(n+1)/2} : \text{ (2) for all } 1 \leqslant i < j \leqslant n\right\}. \tag{5}$$

Since the McCormick inequalities are exact at 0–1 points, we have

$$QP = \operatorname{conv}\left\{(\boldsymbol{x},\boldsymbol{y}) \in \{0,1\}^{n(n+1)/2} : \; y_{ij} = x_i x_j \text{ for all } 1 \leqslant i < j \leqslant n\right\}.$$

It follows from (1) that QP is an extended formulation for $X(f)$:

$$X(f) = \pi[f](QP),$$

so that (4) implies $\operatorname{vex}[f](\boldsymbol{x}) = \mathrm{LB}_{QP}[f](\boldsymbol{x})$ and $\operatorname{cav}[f](\boldsymbol{x}) = \mathrm{UB}_{QP}[f](\boldsymbol{x})$. In fact, Burer and Letchford [BL09, Proposition 5] showed that

$$QP = \operatorname{conv}\left\{(\boldsymbol{x},\boldsymbol{y}) \in [0,1]^{n(n+1)/2} : \; y_{ij} = x_i x_j \text{ for all } 1 \leqslant i < j \leqslant n\right\}.$$

Padberg [Pad89] also extended the definition of BQP in the following sense:

$$QP(G) = \operatorname{conv}\left\{(\boldsymbol{x},\boldsymbol{y}) \in \{0,1\}^{n+m} : \text{ (2) for all } ij \in E\right\},$$

where $G = (V, E)$ is the edge weighted graph associated with the bilinear function f. This graph has the vertex set $V = [n] = \{1, \dots, n\}$, the edge set $E = \{\{i, j\} : a_{ij} \neq 0\}$, and the edge weights are given by a_{ij}. Note that this construction gives a one-to-one correspondence between bilinear functions and edge weighted graphs (without loops). Also, $QP = QP(K_n)$ where K_n is the complete graph, and $QP(G)$ is the projection of QP obtained by projecting out y_{ij}'s corresponding to $a_{ij} = 0$. Henceforth, we express

$$f(\boldsymbol{x}) = \sum_{ij \in E} a_{ij} x_i x_j,$$

where we use ij instead of $\{i, j\}$ to denote an edge when there is no danger of ambiguity. We call an edge $ij \in E$ *positive* if $a_{ij} > 0$ and *negative* if $a_{ij} < 0$. Abusing notation, we sometimes consider $\pi[f]$ as a function $\mathbb{R}^{n+m} \to \mathbb{R}$ where the value of m is clear from the context (the number of edges in the graph corresponding to the considered function f). This allows us to write $X(f) = \pi[f](QP(G))$.

The polytope QP, and in general $QP(G)$, has an exponential number of facets, not all of which are known and some of the known facets are NP-hard to separate [BM86; DL97; LS14; Pad89]. Furthermore, there are many graphs for which $QP(G)$ does not have a polynomial-sized extended formulation [AT15]. If we do not assume any structure on f and allow f to be arbitrary, then a complete characterization of $QP(G)$ seems necessary for convexifying f due to the following observation.

Remark 1. For every facet $\boldsymbol{\alpha}^T\boldsymbol{x} + \boldsymbol{\beta}^T\boldsymbol{y} \leqslant \alpha_0$ of $QP(G)$ there exists a bilinear function f such that $\boldsymbol{\alpha}^T\boldsymbol{x} + \boldsymbol{\beta}^T\boldsymbol{y} \leqslant \alpha_0$ is necessary to describe $X(f)$ in the sense that $\pi[f](P) \supsetneq X(f)$ for the polytope P obtained from QP by omitting $\boldsymbol{\alpha}^T\boldsymbol{x} + \boldsymbol{\beta}^T\boldsymbol{y} \leqslant \alpha_0$. To see this, just take $f(\boldsymbol{x}) = \sum_{ij \in E} \beta_{ij} x_i x_j$. Then $f(x) \leqslant \alpha_0 - \boldsymbol{\alpha}^T\boldsymbol{x}$ for every $\boldsymbol{x} \in [0,1]^n$, but there exists $(\boldsymbol{x}^*, \boldsymbol{y}^*) \in P$ with $\boldsymbol{\alpha}^T\boldsymbol{x}^* + \boldsymbol{\beta}^T\boldsymbol{y}^* > \alpha_0$, and therefore $\pi[f](\boldsymbol{x}^*, \boldsymbol{y}^*) \notin X(f)$.

Our approach. The polytope $QP(G)$ has a very rich combinatorial structure that is not known explicitly and is even hard to generate algorithmically. Also, since $QP(G)$ is an extension of $X(f)$ and two polytopes can project onto the same polytope, it is natural to expect that for certain bilinear functions f, or equivalently weighted graphs G, fully characterizing $QP(G)$ may be much more than what is actually necessary for convexifying f. These facts motivate us to search for graphs G for which we can identify polynomial-sized polytopes $P \supseteq QP(G)$ such that $\pi[f](P) = X(f)$. We would also like such a P to be minimal in the following sense. An extended formulation $P \supseteq QP(G)$ of $X(f)$ is said to be *minimal* if omitting any facet-defining inequality of P leads to a polytope $P' \supsetneq P$ with $\pi[f](P') \supsetneq X(f)$. In other words, we want to identify minimal classes of valid inequalities for QP which still ensure that the polytope defined by these inequalities satisfies $\pi[f](P) = X(f)$, the motivation being that $P \subseteq \mathbb{R}^{n(n+1)/2}$ might have significantly fewer facets than $X(f) \subseteq \mathbb{R}^{n+1}$.

Contributions and Outline. We identify two graph families for which a polynomial number of commonly known valid inequalities for $QP(G)$ are sufficient to convexify the corresponding f. These main results are stated in §2.2, following a short description of the valid inequalities considered by us. Another main contribution of this paper is to use a new technique for proving the tightness of our extended formulations. This technique is inspired by a recent work in the literature on geometric characterization of 0–1 polytopes, and is described in §3. We remove some of the technicalities of a result from the literature and apply our simplified result to state a description of the convex hulls of graphs of arbitrary nonlinear, and bilinear, functions over $\{0,1\}^n$. This is then used to prove our main results in §4. We hope that our successful use of this technique will inspire others to use it to prove more results on extended formulations for $X(f)$ or other combinatorial polytopes.

2. Background and Our Results

To state our main results, we first need to describe the families of valid inequalities and facets of BQP that we are interested in. Other classes of valid inequalities are also available in the literature, many of which are obtained by exploiting the linear bijection between BQP and the cut polytope [BM86; DL97; Sim90]. Separation questions related to some of these inequalities have also been addressed in [LS14].

2.1. Padberg's inequalities for BQP.

The inequalities derived by Padberg [Pad89] can be written down conveniently using the following notation, for $S \subseteq V$, $\hat{E} \subseteq E$:

$$E(S) = \{ij \in E : \ i, j \in S\}, \quad x(S) = \sum_{i \in S} x_i, \quad y(\hat{E}) = \sum_{ij \in \hat{E}} y_{ij}.$$

The inequalities that are relevant for our results are the following.

Triangle inequalities: for any $i, j, k \in V$,

$$x_i + x_j + x_k - y_{ij} - y_{ik} - y_{jk} \leqslant 1, \tag{6a}$$

$$-x_i + y_{ij} + y_{ik} - y_{jk} \leqslant 0, \tag{6b}$$

$$-x_j + y_{ij} - y_{ik} + y_{jk} \leqslant 0, \tag{6c}$$

$$-x_k - y_{ij} + y_{ik} + y_{jk} \leqslant 0, \tag{6d}$$

Clique inequalities: for any $S \subseteq V$ with $|S| \geqslant 3$ and integer α, $1 \leqslant \alpha \leqslant |S| - 2$,

$$\alpha x(S) - y(E(S)) \leqslant \frac{\alpha(\alpha+1)}{2} \tag{7}$$

Cycle inequalities: for every cycle $C \subseteq E$ and every subset $D \subseteq C$ with odd cardinality,

$$x(V_0) - x(V_1) + y(C \setminus D) - y(D) \leqslant (|D| - 1)/2, \tag{8}$$

where

$$V_0 = \{u \in V : \ e \cap \hat{e} = \{u\} \text{ for some } e, \hat{e} \in D\},$$
$$V_1 = \{u \in V : \ e \cap \hat{e} = \{u\} \text{ for some } e, \hat{e} \in C \setminus D\}.$$

Using only subsets of the known facets of QP, one can define different relaxations, and there exist a number of results on conditions on G which guarantee that $QP(G)$ is equal to certain relaxations. The LP relaxation of $QP(G)$ is the polytope in $[0,1]^{n+|E|}$ defined by the McCormick inequalities (2). Padberg showed this is equal to $QP(G)$ if and only if G is acyclic. The *cycle relaxation* of $QP(G)$, which is a strengthening of the McCormick relaxation by adding cycle inequalities (8) for each chordless cycle $C \subseteq E$ and each subset $D \subseteq C$ with $|D|$ odd, is equal to $QP(G)$ if and only if G is K_4-minor-free (series-parallel graphs) [Pad89; Sim90]. More characterizations for the cycle relaxation were obtained recently by [Mic18]. Another set of known results about Padberg's inequalities is that the triangle inequalities give the Chvátal-Gomory closure of the LP relaxation of $QP(K_n)$ [BCH92], and this was recently generalized to the odd cycle inequalities giving the Chvátal-Gomory closure of the LP relaxation of $QP(G)$ for arbitrary G [BGL18].

All of these literature results are about characterizing $QP(G)$, which, as we have explained in the introduction, might sometimes be more than what is required for convexifying f. One of our main results is that if G is a chordless cycle of length n, which is K_4-minor-free and so the corresponding f is convexified by all the 2^{n-1} cycle inequalities (8), then only *two* cycle inequalities are needed to get a polytope P with $\pi[f](P) = X(f)$. This extends to cactus graphs, that is, graphs in which every edge is contained in at most one cycle. The next section presents precise statements of our results. We clarify that none of these results implies anything about the description of $QP(G)$.

2.2. **Main results.** When $f(\boldsymbol{x}) = \sum_{1\leqslant i<j\leqslant n} x_ix_j$, so that the corresponding graph is the complete graph K_n with all edge weights equal to 1, then it is known that the McCormick inequalities $y_{ij} \leqslant \min\{x_i, x_j\}$ together with the clique inequalities (7) with $S = V$ suffice to describe the convex hull of the graph of f. In fact, this implies that the only lower bounds on the y-variables come from inequalities which can be written in the form $y(E) \geqslant \dots sx(V) - \binom{s+1}{2}$, so the convex envelope vex$[f]$ can be written in terms of the original variables $x_1, \dots, x_n$ and $z = y(E)$, and this is precisely the convex envelope characterization proved in [Rik97; She97]. We provide an alternative proof of this result in §3.2. As an extension, our first main result addresses the case of K_n^-, which is the graph obtained from K_n by deleting the edge between $n-1$ and n, with unit weights, for which the bilinear function is $f(\boldsymbol{x}) = \sum_{1\leqslant i<j\leqslant n-1} x_ix_j + \sum_{i=1}^{n-2} x_ix_n$.

Theorem 1 (Almost complete)**.** *If $G = K_n^-$ and all edge weights are equal to 1, then $X(f) = \pi[f](P)$ where $P \subseteq [0,1]^{n(n+1)/2}$ is described by $y(E) \geqslant 0$, the McCormick inequalities $y_{ij} \leqslant x_i$ and $y_{ij} \leqslant x_j$ for $1 \leqslant i < j \leqslant n$, together with the $3(n-2)$ inequalities*

$$2x_i + x_{n-1} + x_n - y_{i,n-1} - y_{in} \leqslant 2 \qquad i = 1, \dots, n-2 \tag{9}$$

$$s\left[x\left(V\setminus\{n-1,n\}\right) + \frac{x_{n-1}+x_n}{2}\right] - y\Big(E\left(V\setminus\{n-1,n\}\right)\Big) - \frac{1}{2}\sum_{i=1}^{n-2}\left(y_{i,n-1}+y_{in}\right) \leqslant \binom{s+1}{2}, \qquad s = 1,\dots,n-2, \tag{10}$$

$$sx(V) - y(E) - y_{n-1,n} \leqslant \binom{s+1}{2}, \qquad s = 1,\dots,n-2. \tag{11}$$

Moreover, this polytope P is a minimal extension of $X(f)$.

Constraints (9) are sums of McCormick inequalities $x_i+x_{n-1}-y_{i,n-1} \leqslant 1$ and $x_i+x_n-y_{in} \leqslant 1$. The other inequalities in the above theorem come from Padberg's clique inequalities: (10) are the averages of the clique inequalities for the two maximal cliques in G and (11) are the clique inequalities for K_n. It should be of no surprise that the non-McCormick inequalities required to describe $X(f)$ all have negative coefficients on the y variables, because only vex$[f](\boldsymbol{x})$ is unknown since two of the McCormick inequalities for every y_{ij} are known to describe cav$[f](\boldsymbol{x})$ when edge weights are non-negative ([LNL12; TRX12], see also Corollary 2 below). An interesting feature of this theorem is that the variable $y_{n-1,n}$ is used although it does not correspond to an edge in G. The point in Figure 1 illustrates that the clique inequalities for the cliques in the given graph are not sufficient in general. The point satisfies the clique inequalities for the seven 3-cliques, and for the two 4-cliques, as well as the cycle inequalities for all cycles in K_5^-, but $\pi[f](\boldsymbol{x},\boldsymbol{y}) = (1/2, 1/2, 1/2, 3/4, 1/4, 3/2) \notin X(f)$, because it violates the inequality

$$2(x_1+x_2+x_3+x_4) + x_5 - z \leqslant 3$$

which can be obtained by using $y_{45} \leqslant x_5$ to eliminate y_{45} from a clique inequality for the K_5.

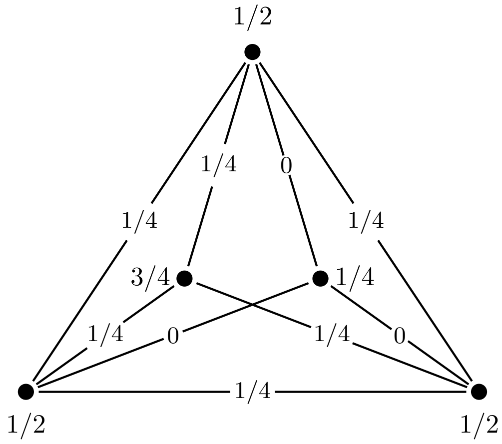

Figure 1. A point satisfying the clique and cycle inequalities involving only variables corresponding to edges of K_5^-.

From Theorem 1, we can also read off a description of $\mathrm{vex}[f](\boldsymbol{x})$ in terms of the original variables: $\mathrm{vex}[f](\boldsymbol{x}) = \max\{0,\, z_1,\, z_2\}$, where

$$z_1 = \max\left\{ sx(V) - \min\{x_{n-1}, x_n\} - \binom{s+1}{2} \,:\, 1 \leqslant s \leqslant n-2 \right\},$$

$$z_2 = \max\Bigg\{ s\left[x\left(V\setminus\{n-1,n\}\right) + \frac{x_{n-1}+x_n}{2}\right] - \binom{s+1}{2}$$
$$+ \frac{1}{2}\sum_{i=1}^{n-2}\max\{0,\, 2x_i + x_{n-1} + x_n - 2\} \,:\, 1 \leqslant s \leqslant n-2\Bigg\}.$$

The second main result is about chordless cycles. Let C_n denote an n-cycle with edges $\{i, i+1\}$ for $i \in [n-1]$, and the edge $\{1, n\}$. The vertex set and edge set are each in bijection with $[n]$ and hence indexed by $[n]$. For ease of notation, all indices have to be read modulo n in the obvious way. In particular, $n \equiv 0$ and $n+1 \equiv 1$. For $i = 1, \ldots, n$, edge $\{i, i+1\}$ is referred to as edge i and $a_{i,i+1}$ is written as a_i. The index set $[n]$ can be partitioned as $[n] = E^- \cup E^+$ based on the signs of the coefficients:

$$E^- = \{i \,:\, a_i < 0\}, \quad E^+ = \{i \;\, a_i > 0\}.$$

We define V^- (resp. V^+) to be the set of vertices $i \in [n]$ such that both of the edges incident with i correspond to negative (resp. positive) terms in $f(\boldsymbol{x})$. More formally,

$$V^- = \{i \in [n] :\ \{i-1, i\} \subseteq E^-\}, \quad V^+ = \{i \in [n] :\ \{i-1, i\} \subseteq E^+\}.$$

which, in general, may not partition $[n]$.

Theorem 2 (Cycles)**.** *For $G = C_n$, we have $\pi[f](P) = X(f)$ where $P \subseteq [0,1]^{2n}$ is the polytope described by the McCormick inequalities* (2) *and*

$$x(V^-) - x(V^+) + y(E^+) - y(E^-) \leqslant \left\lfloor \frac{|E^-|}{2} \right\rfloor, \tag{12}$$

$$x(V^+) - x(V^-) + y(E^-) - y(E^+) \leqslant \left\lfloor \frac{|E^+|}{2} \right\rfloor. \tag{13}$$

Moreover, inequality (12) *(resp.* (13)*) can be omitted if and only if $|E^-|$ (resp. $|E^+|$) is even, and then P is a minimal extension of $X(f)$.*

The polytope P in Theorem 2 is parametrized by edge weights $\boldsymbol{a}$ and should be read as $P_{\boldsymbol{a}}$, since (12) and (13) are constructed using the sign pattern on the edge weights. Obviously, since P has only two cycle inequalities, it is a weak relaxation of $QP(C_n)$, which we know is given by all the 2^{n-1} odd-cycle inequalities due to C_n being K_4-minor-free. When $|E^-|$ is odd, inequality (12) is Padberg's cycle inequality (8) corresponding to the odd subset $D = E^-$ in the graph C_n. This is because $D = E^- = E(C_n) \setminus E^+$ and every vertex in C_n having exactly two edges incident on it implies that $V_0 = V^-$ and $V_1 = V^+$. If $|E^-|$ is even, we will show in the proof of Theorem 2 that inequality (12) is a linear combination of McCormick inequalities. Analogous arguments hold for E^+ and (13). Note that inequalities (12) and (13) do not use any non-edge variables y_{ij}; in contrast, Theorem 1 presents a minimal extension for the chordal graph K_n^- using a non-edge variable $y_{n-1,n}$.

Theorem 2 implies that if the bilinear function corresponds to an even cycle C_n having both $|E^+|$ and $|E^-|$ even, then the McCormick inequalities are sufficient to convexify the function. This implication is a special case of the following characterization by Boland et al. [Bol+17, Theorem 4]: for any bilinear function f, the McCormick relaxation projects onto $X(f)$ if and only if *every* cycle in the graph G of f has both $|E^+|$ and $|E^-|$ even. This naturally raises the question of what can be said about bilinear functions with odd cycles. In general, we should not expect to convexify a bilinear function using extended formulation for each cycle in the function, since the function may have large extension complexity, whereas any cycle has a small extended formulation given in Theorem 2. However, we show that for cactus graphs, which are graphs whose cycles are edge-disjoint, or equivalently, any two cycles have at most one common vertex, the bilinear function is convexified by considering each cycle individually.

Theorem 3. *If G is a cactus with k cycles, then for any edge weight vector $\boldsymbol{a}$, $X(f)$ is described by the McCormick inequalities and at most $2k$ cycle inequalities.*

This result is argued using the following consequence of the method described in Section 3: if f and g are two bilinear functions which share at most one variable, then $X(f+g)$ can be easily described in terms of $X(f)$ and $X(g)$ (see Corollary 3 for a precise statement).

We conclude this section by illustrating the reduction in the number of required inequalities for small examples. In Table 1 we compare the numbers of facets of $X(f) \subseteq \mathbb{R}^{n+1}$ to the numbers of inequalities describing a polytope $P \subseteq \mathbb{R}^{n(n+1)/2}$ with $\pi[f](P) = X(f)$. The facet numbers for $X(f)$ are determined using `polymake` [Ass+17], while the numbers in the columns for the polytopes P are determined as follows. For $G = K_n$ with unit coefficients, we can choose

TABLE 1. Numbers of facets for $X(f)$ and for our extended formulations.

	$G = K_n$		$G = K_n^-$		$G = C_n$	
n	$X(f)$	P	$X(f)$	P	$X(f)$	P
3	15	15	12	16	15	20
4	36	24	34	27	26	26
5	135	35	120	40	63	32
6	738	48	636	55	118	38
7	5,061	63	4,376	72	255	44
8	40,344	80	35,372	91	498	50

$P \subseteq \mathbb{R}^{n(n+1)/2}$ described by the following $n(n+2)$ inequalities:

- $2n$ variable bounds $0 \leqslant x_i \leqslant 1$ for $i \in [n]$,
- $n(n-1)$ McCormick upper bounds $y_{ij} \leqslant x_i$ and $y_{ij} \leqslant x_j$ for $ij \in E$,
- $n-1$ inequalities $y(E) \geqslant sx(V) - s(s+1)/2$ for $s \in [n-1]$, and
- one inequality $y(E) \geqslant 0$.

For $G = K_n^-$ with unit coefficients, the following $n^2 + 4n - 4$ inequalities are sufficient:

- $2n$ variable bounds $0 \leqslant x_i \leqslant 1$ for $i \in [n]$,
- $n(n-1)$ McCormick inequalities $y_{ij} \leqslant x_i$ and $y_{ij} \leqslant x_j$,
- $3n - 6$ inequalities (9), (10), (11),

- one inequality $y(E) \geqslant 0$.

For $G = C_n$, with arbitrary coefficients, the following $6n + 2$ inequalities are sufficient:

- $2n$ variable bounds $0 \leqslant x_i \leqslant 1$ for $i \in [n]$,
- $4n$ McCormick inequalities (2), and
- two inequalities (12) and (13).

3. A GEOMETRIC CHARACTERIZATION OF COMBINATORIAL POLYTOPES

3.1. **Zuckerberg's method.** Zuckerberg [Zuc04; Zuc16] developed a technique to prove convex hull characterizations for subsets of $\{0,1\}^n$. In this section, we simplify one of the technical results from this work and extend it to the convex hulls of graphs of arbitrary functions $\psi : \{0,1\}^n \to \mathbb{R}$.

We are interested in the convex hull of a set $\mathcal{F} \subseteq \{0,1\}^n$. Any such $\mathcal{F}$ can be represented as a finite combination of unions, intersections and complementations of the sets

$$A_i = \{\boldsymbol{\xi} \in \{0,1\}^n \,:\, \xi_i = 1\}, \quad i = 1, \dots, n,$$

and we fix such a representation $F(A_1, \dots, A_n)$. For instance, the set $X(f)$ for the function $f : [0,1]^2 \to [0,1]$ given by $f(x_1, x_2) = x_1 x_2$ is the convex hull of the set

$$\mathcal{F} = \{(0,0,0),\ (0,1,0),\ (1,0,0),\ (1,1,1)\},$$

which is given by the variable bounds $0 \leqslant x_1, x_2, x_3 \leqslant 1$ and the McCormick inequalities

$$x_3 \leqslant x_1, \qquad x_3 \leqslant x_2, \qquad x_3 \geqslant x_1 + x_2 - 1.$$

In this case $\mathcal{F}$ can be represented as

$$\mathcal{F} = F(A_1, A_2, A_3) = (A_1 \cap A_2 \cap A_3) \cup \left(\overline{A_1 \cup A_2 \cup A_3}\right) \cup \left(A_1 \cap \overline{A_2 \cup A_3}\right) \cup \left(A_2 \cap \overline{A_1 \cup A_3}\right), \quad (14)$$

where $\bar{\cdot}$ indicates the complement in $\{0,1\}^3$.

The main result in [Zuc16, Theorem 7] can be stated as follows. For every $\boldsymbol{x} \in [0,1]^n$, we have $\boldsymbol{x} \in \operatorname{conv}(\mathcal{F})$ if and only if there exist

- a set U with a collection $\mathcal{L}$ of subsets which contains the empty set and is closed under taking complements and finite unions, and
- a function $\mu : \mathcal{L} \to \mathbb{R}$ with $\mu(U) = 1$ and $\mu(L_1 \cup \cdots \cup L_k) = \mu(L_1) + \cdots + \mu(L_k)$ for any finite collection of pairwise disjoint elements $L_1, \dots, L_k \in \mathcal{L}$, and
- sets $X_1, \dots, X_n \in \mathcal{L}$ with $\mu(X_i) = x_i$ for all $i \in [n]$ and $\mu(F(X_1, \dots, X_n)) = 1$ (where the complement has to be taken in U instead of $\{0,1\}^n$).

To be precise, Zuckerberg states only one direction of this equivalence, but the other one is easy (see the proof of Theorem 4 below).

Example 1. For $\mathcal{F} = \{(0,0,0),\ (0,1,0),\ (1,0,0),\ (1,1,1)\}$ we take U to be the half-open interval $[0,1)$ with $\mathcal{L}$ being the collection of unions of finitely many half-open intervals and μ the Lebesgue measure (restricted to $\mathcal{L}$), that is,

$$\mathcal{L} = \left\{[a_1, b_1) \cup \cdots \cup [a_k, b_k) \;:\; 0 \leqslant a_1 < b_1 < a_2 < b_2 < \cdots < a_k < b_k \leqslant 1,\ k \in \mathbb{N}\right\}, \quad (15)$$

$$\mu(X) = (b_1 - a_1) + \cdots + (b_k - a_k) \qquad \text{for } X = [a_1, b_1) \cup \cdots \cup [a_k, b_k) \in \mathcal{L}. \quad (16)$$

For $(x_1, x_2, x_3) \in \operatorname{conv}(\mathcal{F})$ we set

$$X_1 = [0,\, x_1), \qquad X_2 = [x_1 - x_3,\, x_1 + x_2 - x_3), \qquad X_3 = [x_1 - x_3,\, x_1).$$

Then $\mu(X_i) = x_i$ for all $i \in \{1,2,3\}$, and from

$$\begin{aligned} F(X_1, X_2, X_3) &= (X_1 \cap X_2 \cap X_3) \cup \left(\overline{X_1 \cup X_2 \cup X_3}\right) \cup \left(X_1 \cap \overline{X_2 \cup X_3}\right) \cup \left(X_2 \cap \overline{X_1 \cup X_3}\right) \\ &= [x_1 - x_3, x_1) \cup [x_1 + x_2 - x_3, 1) \cup [0, x_1 - x_3) \cup [x_1, x_1 + x_2 - x_3) = [0,1) \end{aligned}$$

we get $\mu\left(F(X_1, X_2, X_3)\right) = 1$, as required. This provides a proof that the McCormick inequalities indeed characterize the convex hull of the set $\{(x_1,\, x_2,\, x_1 x_2) \,:\, x_1, x_2 \in [0,1]\}$. The construction is illustrated in Figure 2 for $\boldsymbol{x} = (0.5, 0.4, 0.1)$. The sets X_1, X_2 and X_3 do not

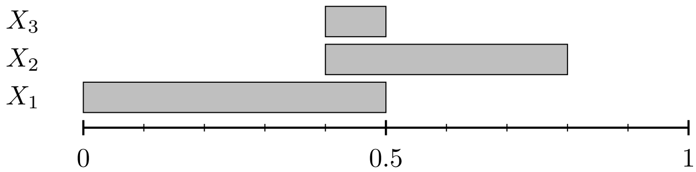

FIGURE 2. The sets X_1, X_2 and X_3 for $x_1 = 0.5$, $x_2 = 0.4$ and $x_3 = 0.1$ with $X_1 \cap X_2 \cap X_3 = [0.4, 0.5)$, $\overline{X_1 \cup X_2 \cup X_3} = [0.8, 1)$, $X_1 \cap \overline{X_2 \cup X_3} = [0, 0.4)$ and $X_2 \cap \overline{X_1 \cup X_3} = [0.5, 0.8)$.

only provide a certificate that $\boldsymbol{x} \in \operatorname{conv} \mathcal{F}$, but they also encode a representation of $\boldsymbol{x}$ as a convex combination of the elements of $\mathcal{F}$. To see this we associate with each $t \in [0,1)$ the vector $\boldsymbol{x}(t) = (x_1(t),\, x_2(t),\, x_3(t))$ with $x_i(t) = 1$ if $t \in X_i$ and $x_i(t) = 0$ if $t \notin X_i$. In our example

$$\boldsymbol{x}(t) = \begin{cases} (1,\,0,\,0) & \text{for } t \in X_1 \cap \overline{X_2 \cup X_3} = [0, 0.4),\\ (1,\,1,\,1) & \text{for } t \in X_1 \cap X_2 \cap X_3 = [0.4, 0.5),\\ (0,\,1,\,0) & \text{for } t \in X_2 \cap \overline{X_1 \cup X_3} = [0.5, 0.8),\\ (0,\,0,\,0) & \text{for } t \in \overline{X_1 \cup X_2 \cup X_3} = [0.8, 1), \end{cases}$$

which corresponds to the convex representation

$$\begin{pmatrix}0.5\\0.4\\0.1\end{pmatrix} = 0.4\begin{pmatrix}1\\0\\0\end{pmatrix} + 0.1\begin{pmatrix}1\\1\\1\end{pmatrix} + 0.3\begin{pmatrix}0\\1\\0\end{pmatrix} + 0.2\begin{pmatrix}0\\0\\0\end{pmatrix}. \tag{17}$$

Our main simplification of Zuckerberg's result is that the statement remains true if U and $\mathcal{L}$ are fixed as in Example 1. As a consequence, the condition $\mu(F(X_1,\dots,X_n)) = 1$ can be replaced by $F(X_1,\dots,X_n) = U$, and in fact, the set theoretic representation of $\mathcal{F}$ can be avoided completely using the standard identification of the elements of $\{0,1\}^n$ with subsets of $[n]$. More precisely, a vector $\boldsymbol{\xi} \in \{0,1\}^n$ is identified with the set $\{i \in [n] : \xi_i = 1\}$, so that in particular the elements of $\mathcal{F}$ are identified with subsets of $[n]$. The following theorem is a reformulation of [Zuc16, Theorem 7] which is more convenient for our purpose. We provide a complete proof in our setting because the proof is short and it would be a bit cumbersome to explain in detail how our variant can be obtained from the arguments in [Zuc16]. Following the proof of the theorem we explain in detail how Zuckerberg's original statement can be obtained as a consequence of our theorem (see Remark 2).

Theorem 4. *Let $\mathcal{F} \subseteq \{0,1\}^n$, $\boldsymbol{x} \in [0,1]^n$, and let $\mathcal{L}$ and μ be defined by (15) and (16). Then $\boldsymbol{x} \in \operatorname{conv}(\mathcal{F})$ if and only if there are sets $X_1,\dots,X_n \in \mathcal{L}$ such that $\mu(X_i) = x_i$ for all $i \in [n]$, and $\{i \in [n] : t \in X_i\} \in \mathcal{F}$ for every $t \in [0,1)$.*

Proof. Let us fix an ordering $\boldsymbol{\xi}^1,\dots,\boldsymbol{\xi}^{|\mathcal{F}|}$ of $\mathcal{F}$ (In Example 1, the ordering would be the one in which the elements of $\mathcal{F}$ appear in the convex combination (17)), and suppose $\boldsymbol{x} \in \operatorname{conv}(\mathcal{F})$, say $\boldsymbol{x} = \lambda_1\boldsymbol{\xi}^1 + \dots + \lambda_{|\mathcal{F}|}\boldsymbol{\xi}^{|\mathcal{F}|}$ with $\lambda_1 + \dots + \lambda_{|\mathcal{F}|} = 1$ and $\lambda_k \geqslant 0$ for all $k \in [|\mathcal{F}|]$. We define a partition $U = I_1 \cup \dots \cup I_{|\mathcal{F}|}$ by setting $I_1 = [0, \lambda_1)$ and $I_k = [\lambda_1 + \dots + \lambda_{k-1},\, \lambda_1 + \dots + \lambda_k)$ for $k \in \{2,\dots,|\mathcal{F}|\}$. For

$$X_i = \bigcup_{k\,:\,\xi_i^k = 1} I_k,$$

we have, for every $i \in [n]$,

$$\mu(X_i) = \sum_{k\,:\,\xi_i^k=1} \mu(I_k) = \sum_{k\,:\,\xi_i^k=1} \lambda_k = \sum_{k=1}^{|\mathcal{F}|} \lambda_k \xi_i^k = x_i,$$

and, for every $t \in [0,1)$, there is a unique index k with $t \in I_k$, and then

$$\{i \in [n] : t \in X_i\} = \{i \in [n] : \xi_i^k = 1\} = \boldsymbol{\xi}^k \in \mathcal{F},$$

as required. Conversely, if X_i are sets with the described properties, we can set

$$\lambda_k = \mu\left(\left\{t \ : \ \{i \in [n] \ : \ t \in X_i\} = \boldsymbol{\xi}^k\right\}\right)$$

for $k = 1, \dots, |\mathcal{F}|$ to obtain the required convex representation $\boldsymbol{x} = \lambda_1\boldsymbol{\xi}^1 + \cdots + \lambda_{|\mathcal{F}|}\boldsymbol{\xi}^{|\mathcal{F}|}$. To see this note that by assumption, for every $t \in [0,1)$, there is a unique $k \in [|\mathcal{F}|]$ with $\boldsymbol{\xi}^k = \{i \in [n] \, : \, \xi_i^k = 1\}$, and then for every $j \in [n]$, $t \in X_j$ if and only if $\xi_j^k = 1$. In other words,

$$X_j = \bigcup_{k\,:\,\xi_j^k = 1} \left\{t \ : \ \{i \in [n] \ : \ t \in X_i\} = \boldsymbol{\xi}^k\right\},$$

and using this we can verify that $\lambda_1\boldsymbol{\xi}^1 + \cdots + \lambda_{|\mathcal{F}|}\boldsymbol{\xi}^{|\mathcal{F}|} = \boldsymbol{x}$: for every $j \in [n]$,

$$\sum_{k=1}^n \lambda_k \xi_j^k = \sum_{k=1}^n \mu\left(\left\{t \ : \ \{i \in [n] \, : \, t \in X_i\} = \boldsymbol{\xi}^k\right\}\right)\xi_j^k$$
$$= \sum_{k\,:\,\xi_j^k=1} \mu\left(\left\{t \ : \ \{i \in [n] \, : \, t \in X_i\} = \boldsymbol{\xi}^k\right\}\right) = \mu\left(\bigcup_{k\,:\,\xi_j^k=1} \left\{t \ : \ \{i \in [n] \ : \ t \in X_i\} = \boldsymbol{\xi}^k\right\}\right)$$
$$= \mu(X_j) = x_j. \qquad \square$$

Remark 2. Theorem 4 implies Zuckerberg's criterion since for any subsets $X_1, \dots, X_n \subseteq [0,1)$,

$$F(X_1, \dots, X_n) = \{t \in [0,1) \, : \, \{i \in [n] \, : \, t \in X_i\} \in \mathcal{F}\}.$$

This can be seen by induction on the structure of the formula F. The base case is $\mathcal{F} = A_j$ for some $j \in [n]$. Then

$$F(X_1, \dots, X_n) = X_j = \{t \in [0,1) \, : \, j \in \{i \in [n] \, : \, t \in X_i\}\}$$
$$= \{t \in [0,1) \, : \, \{i \in [n] \, : \, t \in X_i\} \in A_j\},$$

as required. For the induction step we have either

(1) $\mathcal{F} = F(A_1, \dots, A_n) = \overline{F_1(A_1, \dots, A_n)}$, or
(2) $\mathcal{F} = F(A_1, \dots, A_n) = F_1(A_1, \dots, A_n) \cup F_2(A_1, \dots, A_n)$, or
(3) $\mathcal{F} = F(A_1, \dots, A_n) = F_1(A_1, \dots, A_n) \cap F_2(A_1, \dots, A_n)$,

and in each case we can verify the statement.

Case 1: $\mathcal{F} = \overline{F_1(A_1, \dots, A_n)}$. By induction,

$$F_1(X_1, \dots, X_n) = \{t \in [0,1) \, : \, \{i \in [n] \, : \, t \in X_i\} \in F_1(A_1, \dots, A_n)\},$$

and then

$$t \in F(X_1, \dots, X_n) \iff t \notin F_1(X_1, \dots, X_n) \iff \{i \in [n] \, : \, t \in X_i\} \notin F_1(A_1, \dots, A_n)$$
$$\iff \{i \in [n] \, : \, t \in X_i\} \in \overline{F_1(A_1, \dots, A_n)} = \mathcal{F}.$$

Case 2: $\mathcal{F} = F_1(A_1, \dots, A_n) \cup F_2(A_1, \dots, A_n)$. By induction,

$$F_k(X_1, \dots, X_n) = \{t \in [0,1) \, : \, \{i \in [n] \, : \, t \in X_i\} \in F_k(A_1, \dots, A_n)\},$$

for $k \in \{1,2\}$, and then

$$t \in F(X_1, \dots, X_n) \iff t \in F_1(X_1, \dots, X_n) \cup F_2(X_1, \dots, X_n)$$
$$\iff \{i \in [n] \, : \, t \in X_i\} \in F_1(A_1, \dots, A_n) \cup F_2(A_1, \dots, A_n) = \mathcal{F}.$$

Case 3: $\mathcal{F} = F_1(A_1, \dots, A_n) \cap F_2(A_1, \dots, A_n)$. By induction,

$$F_k(X_1, \dots, X_n) = \{t \in [0,1) \, : \, \{i \in [n] \, : \, t \in X_i\} \in F_k(A_1, \dots, A_n)\},$$

for $k \in \{1,2\}$, and then

$$t \in F(X_1, \dots, X_n) \iff t \in F_1(X_1, \dots, X_n) \cap F_2(X_1, \dots, X_n)$$
$$\iff \{i \in [n] \, : \, t \in X_i\} \in F_1(A_1, \dots, A_n) \cap F_2(A_1, \dots, A_n) = \mathcal{F}.$$

Next we extend the statement of Theorem 4 so that it applies to the convex hull $X(\psi)$ of the graph of an arbitrary function $\psi : \{0,1\}^n \to \mathbb{R}$. For sets $X_1, \dots, X_n \in \mathcal{L}$ we partition U into 2^n subsets $R_{\boldsymbol{\xi}}(X_1,\dots,X_n)$, $\boldsymbol{\xi} \in \{0,1\}^n$, defined by

$$R_{\boldsymbol{\xi}}(X_1,\dots,X_n) = \{t \in [0,1) \,:\, \{i \in [n] \;:\; t \in X_i\} = \boldsymbol{\xi}\}.$$

Let us define two functions $\psi_-, \psi_+ : [0,1]^n \to \mathbb{R}$ as

$$\psi_-(\boldsymbol{x}) = \min\left\{\sum_{\boldsymbol{\xi}\in\{0,1\}^n} \mu(R_\xi(X_1,\dots,X_n))\psi(\boldsymbol{\xi}) \;:\; X_i \in \mathcal{L},\ \mu(X_i) = x_i \text{ for all } i \in [n]\right\},$$

$$\psi_+(\boldsymbol{x}) = \max\left\{\sum_{\boldsymbol{\xi}\in\{0,1\}^n} \mu(R_\xi(X_1,\dots,X_n))\psi(\boldsymbol{\xi}) \;:\; X_i \in \mathcal{L},\ \mu(X_i) = x_i \text{ for all } i \in [n]\right\}.$$

Theorem 5. *For every function $\psi : \{0,1\}^n \to \mathbb{R}$,*

$$X(\psi) = \left\{(\boldsymbol{x}, z) \in [0,1]^n \times \mathbb{R} :\ \psi_-(\boldsymbol{x}) \leqslant z \leqslant \psi_+(\boldsymbol{x})\right\}.$$

Proof. First suppose $(\boldsymbol{x}, z) \in X(\psi)$, say

$$(\boldsymbol{x},\, z) = \sum_{k=1}^{2^n} \lambda_k \left(\boldsymbol{\xi}^k,\, \psi\left(\boldsymbol{\xi}^k\right)\right),$$

where $\boldsymbol{\xi}^1, \dots, \boldsymbol{\xi}^{2^n}$ is a fixed ordering of $\{0,1\}^n$. The sets $X_i \in \mathcal{L}$ with $\mu(X_i) = x_i$ are defined exactly as in the proof of Theorem 4: For the partition $U = I_1 \cup \dots \cup I_{2^n}$ with $I_1 = [0, \lambda_1)$ and $I_k = [\lambda_1 + \dots + \lambda_{k-1},\, \lambda_1 + \dots + \lambda_k)$ for $k \in \{2, \dots, 2^n\}$, we set

$$X_i = \bigcup_{k \,:\, \xi_i^k = 1} I_k.$$

For every $k \in [2^n]$, $R_{\boldsymbol{\xi}^k}(X_1,\dots,X_n) = I_k$, hence $\mu(R_{\boldsymbol{\xi}^k}(X_1,\dots,X_n)) = \lambda_k$. With

$$z = \sum_{k=1}^{2^n} \lambda_k \psi\left(\boldsymbol{\xi}^k\right) = \sum_{k=1}^{2^n} \mu\left(R_{\boldsymbol{\xi}^k}(X_1,\dots,X_n)\right) \psi\left(\boldsymbol{\xi}^k\right)$$

it follows that $\psi_-(\boldsymbol{x}) \leqslant z \leqslant \psi_+(\boldsymbol{x})$.

For the converse, suppose $\psi_-(\boldsymbol{x}) \leqslant z \leqslant \psi_+(\boldsymbol{x})$, and let $(X_1,\dots,X_n)$ and $(X_1',\dots,X_n')$ be optimizers for the problems defining $\psi_-(\boldsymbol{x})$ and $\psi_+(\boldsymbol{x})$, respectively. We write $z = t\psi_-(\boldsymbol{x}) + (1-t)\psi_+(\boldsymbol{x})$ for some $t \in [0,1]$, and set

$$\lambda(\boldsymbol{\xi}) = t\mu(R_{\boldsymbol{\xi}}(X_1,\dots,X_n)) + (1-t)\mu(R_{\boldsymbol{\xi}}(X_1',\dots,X_n'))$$

for all $\boldsymbol{\xi} \in \{0,1\}^n$. This gives the required convex representation

$$(\boldsymbol{x}, z) = \sum_{\boldsymbol{\xi}\in\{0,1\}^n} \lambda(\boldsymbol{\xi})(\boldsymbol{\xi}, \psi(\boldsymbol{\xi})).$$

□

For the particular case that the function ψ has the form $\psi(\boldsymbol{x}) = f(\boldsymbol{x}) = \sum_{ij\in E} a_{ij}x_ix_j$, Theorem 5 yields the following interpretation for $\mathrm{cav}[f](\boldsymbol{x})$ and $\mathrm{vex}[f](\boldsymbol{x})$.

Corollary 1. *For the bilinear function $f(\boldsymbol{x})$, we have*

$$\mathrm{vex}[f](\boldsymbol{x}) = \min\left\{\sum_{ij\in E} a_{ij}\mu(X_i \cap X_j) \;:\; X_i \in \mathcal{L},\ \mu(X_i) = x_i \textit{ for all } i \in [n]\right\},$$

$$\mathrm{cav}[f](\boldsymbol{x}) = \max\left\{\sum_{ij\in E} a_{ij}\mu(X_i \cap X_j) \;:\; X_i \in \mathcal{L},\ \mu(X_i) = x_i \textit{ for all } i \in [n]\right\}.$$

In particular, for a polytope $P \subseteq \mathbb{R}^{n(n+1)/2}$ with $P \subseteq X(f)$, we have $\pi[f](P) = X(f)$ if and only if for every $\boldsymbol{x} \in [0,1]^n$, there exist $X_1, \dots, X_n \in \mathcal{L}$ and $X_1', \dots, X_n' \in \mathcal{L}$ with $\mu(X_i) = \mu(X_i') = x_i$ for all $i \in [n]$, and

$$\sum_{ij \in E} a_{ij}\mu(X_i \cap X_j) = \mathrm{LB}_P[f](\boldsymbol{x}), \qquad \sum_{ij \in E} a_{ij}\mu(X_i' \cap X_j') = \mathrm{UB}_P[f](\boldsymbol{x}).$$

Proof. We observe that

$$\begin{aligned}
\sum_{\boldsymbol{\xi} \in \{0,1\}^n} \mu\left(R_{\boldsymbol{\xi}}(X_1,\dots,X_n)\right) f(\boldsymbol{\xi}) &= \sum_{\boldsymbol{\xi} \in \{0,1\}^n} \mu\left(R_{\boldsymbol{\xi}}(X_1,\dots,X_n)\right) \sum_{ij \in E} a_{ij}\xi_i\xi_j \\
&= \sum_{ij \in E} a_{ij} \sum_{\boldsymbol{\xi} \in \{0,1\}^n} \mu\left(R_{\xi}(X_1,\dots,X_n)\right) \xi_i\xi_j \\
&= \sum_{ij \in E} a_{ij} \sum_{\boldsymbol{\xi} \in \{0,1\}^n \,:\, \xi_i = \xi_j = 1} \mu\left(\{t \in [0,1) \;:\; \{k \in [n] \;:\; t \in X_k\} = \boldsymbol{\xi}\}\right) \\
&= \sum_{ij \in E} a_{ij}\mu\left(\bigcup_{\boldsymbol{\xi} \in \{0,1\}^n \,:\, \xi_i = \xi_j = 1} \{t \in [0,1) \;:\; \{k \in [n] \;:\; t \in X_k\} = \boldsymbol{\xi}\}\right) \\
&= \sum_{ij \in E} a_{ij}\mu\left(X_i \cap X_j\right).
\end{aligned}$$

□

Example 2. We can use Corollary 1 to prove again that $X(f)$ for $f(x_1,x_2) = x_1x_2$ is given by the McCormick inequalities. From

$$\mu(X_1 \cap X_2) \leqslant \min\{\mu(X_1), \mu(X_2)\} = \min\{x_1, x_2\},$$

it follows that $\mathrm{cav}[f](x) \leqslant \min\{x_1,x_2\}$, and with $X_1 = [0,x_1)$, $X_2 = [0,x_2)$ we see that this bound can be achieved for all $x_1, x_2 \in [0,1]$, and therefore the concave envelope is given by $x_3 \leqslant x_1$ and $x_3 \leqslant x_2$. Similarly, from

$$\mu(X_1 \cap X_2) \geqslant \max\{0, \mu(X_1) + \mu(X_2) - 1\} = \max\{0, x_1 + x_2 - 1\},$$

it follows that $\mathrm{vex}[f](x) \geqslant \min\{0, x_1 + x_2 - 1\}$, and with $X_1 = [0,x_1)$, $X_2 = [1-x_2, 1)$ we see that this bound can be achieved for all $x_1, x_2 \in [0,1]$, and therefore the convex envelope is given by $x_3 \geqslant 0$ and $x_3 \geqslant x_1 + x_2 - 1$.

This example illustrates the use of Corollary 1 to characterize $X(f)$ in the simplest possible case, the function $f(x_1,x_2) = x_1x_2$. It contains the main idea of the proofs of our main results presented in the next section. The difference is, that for more complicated functions the choice of the sets X_i is not obvious, and can depend on the particular vector $\boldsymbol{x}$ for which we want to find $\mathrm{vex}[f](\boldsymbol{x})$.

Example 3. For $n = 5$ consider the function

$$f(\boldsymbol{x}) = \sum_{1 \leqslant i < j \leqslant 5} x_ix_j,$$

corresponding to the graph shown on the left in Figure 3, and the point $\boldsymbol{x} = (0.6,\ 0.3,\ 0.3,\ 0.9,\ 0.4)$. Corollary 1 can be used to verify that $\mathrm{vex}[f](\boldsymbol{x}) \leqslant 2$. For this purpose, consider the sets

$$X_1 = [0, 0.6), \qquad X_2 = [0.6, 0.9), \qquad X_3 = [0, 0.2) \cup [0.9, 1),$$
$$X_4 = [0, 0.1) \cup [0.2, 1), \qquad X_5 = [0.1, 0.5),$$

illustrated on the right in Figure 3. There are six vectors $\boldsymbol{\xi}$ with non-empty $R_{\boldsymbol{\xi}}(X_1, \dots, X_5)$:

$$\begin{aligned}
R_{(1,0,1,1,0)}(X_1,\dots,X_5) &= [0, 0.1), \\
R_{(1,0,1,0,1)}(X_1,\dots,X_5) &= [0.1, 0.2), \\
R_{(1,0,0,1,1)}(X_1,\dots,X_5) &= [0.2, 0.5), \\
R_{(1,0,0,1,0)}(X_1,\dots,X_5) &= [0.5, 0.6), \\
R_{(0,1,0,1,0)}(X_1,\dots,X_5) &= [0.6, 0.9), \\
R_{(0,0,1,1,0)}(X_1,\dots,X_5) &= [0.9, 0.1).
\end{aligned}$$

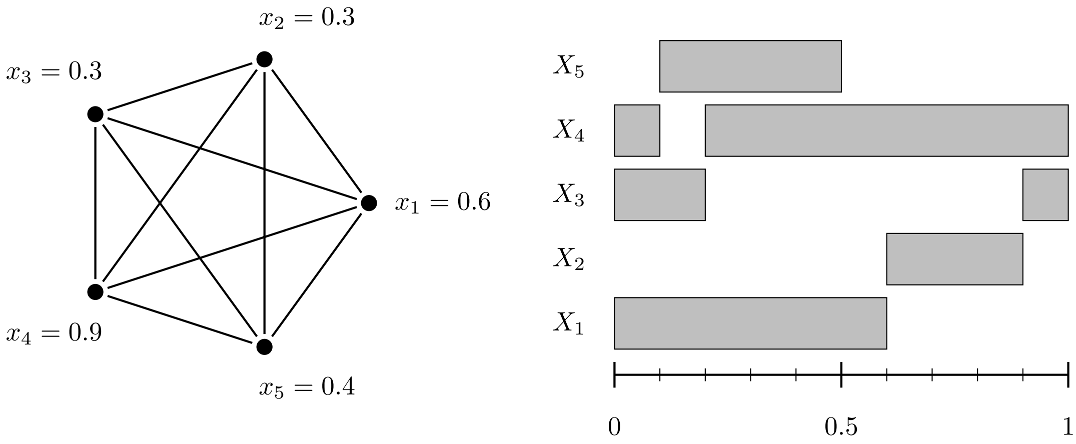

FIGURE 3. Illustration of the sets X_i in Example 3.

So these sets $X_1, \dots, X_5$ correspond to the representation

$$\begin{pmatrix}0.6\\0.3\\0.3\\0.9\\0.4\end{pmatrix} = 0.1\begin{pmatrix}1\\0\\1\\1\\0\end{pmatrix} + 0.1\begin{pmatrix}1\\0\\1\\0\\1\end{pmatrix} + 0.3\begin{pmatrix}1\\0\\0\\1\\1\end{pmatrix} + 0.1\begin{pmatrix}1\\0\\0\\1\\0\end{pmatrix} + 0.3\begin{pmatrix}0\\1\\0\\1\\0\end{pmatrix} + 0.1\begin{pmatrix}0\\0\\1\\1\\0\end{pmatrix},$$

and extending this to the function value coordinate, we get

$$\begin{aligned}\operatorname{vex}[f](\boldsymbol{x}) \leqslant 0.1f(1,0,1,1,0) + 0.1f(1,0,1,0,1) + 0.3f(1,0,0,1,1) + 0.1f(1,0,0,1,0)\\ + 0.3f(0,1,0,1,0) + 0.1f(0,0,1,1,0) = 0.3 + 0.3 + 0.9 + 0.1 + 0.3 + 0.1 = 2.\end{aligned}$$

The following result, which was proved in [LNL12] (see also [TRX12]), is an immediate consequence of Corollary 1.

Corollary 2. *If all edges are positive, then* $\operatorname{cav}[f](\boldsymbol{x}) = \sum_{ij\in E} a_{ij}\min\{x_i, x_j\}$.

Proof. With $X_i = [0, x_i)$ for all $i \in [n]$, we get

$$\operatorname{cav}[f](\boldsymbol{x}) \geqslant \sum_{ij\in E} a_{ij}\mu(X_i \cap X_j) = \sum_{ij\in E} a_{ij}\mu([0, \min\{x_i, x_j\}) = \sum_{ij\in E} a_{ij}\min\{x_i, x_j\}.$$

On the other hand, for any feasible choice of the sets X_i, we have, for all $i, j \in [n]$,

$$\mu(X_i \cap X_j) \leqslant \min\{\mu(X_i),\, \mu(X_j)\} = \min\{x_i, x_j\}.$$

With the assumption that $a_{ij} > 0$ for all $ij \in E$, this implies

$$\operatorname{cav}[f](\boldsymbol{x}) \leqslant \sum_{ij\in E} a_{ij}\min\{x_i, x_j\}.$$ □

As another consequence, we can combine convex hull characterizations of graphs of two bilinear functions if they share at most one variable.

Corollary 3. *Let* $f\colon [0,1]^k \to \mathbb{R}$ *and* $g\colon [0,1]^{n-k+1} \to \mathbb{R}$ *be two bilinear functions given as*

$$f(\boldsymbol{x}) = \sum_{1\leqslant i<j\leqslant k} a_{ij}x_ix_j, \qquad g(\boldsymbol{x}) = \sum_{k\leqslant i<j\leqslant n} a_{ij}x_ix_j,$$

so that f depends only on variables $x_1, \dots, x_k$, and g depends only on variables $x_k, \dots, x_n$. Let $P, Q \subseteq [0,1]^{n(n+1)/2}$ be polytopes with $\pi[f](P) = X(f)$ and $\pi[g](Q) = X(g)$, such that P is described by inequalities involving only the variables $x_1, \dots, x_k$ and y_{ij} with $1 \leqslant i < j \leqslant k$, and Q is described by inequalities involving only the variables $x_k, \dots, x_n$ and y_{ij} with $k \leqslant i < j \leqslant k$. Then $\pi[f+g](P \cap Q) = X(f+g)$.

Proof. Fix $\boldsymbol{x} \in [0,1]^n$. By assumption and Corollary 1, there are sets $X_1, \dots, X_n \in \mathcal{L}$ with $\mu(X_i) = x_i$ for all $i \in [k]$, and sets $X'_k, \dots, X'_n \in \mathcal{L}$ with $\mu(X'_i) = x_i$ for all $i \in [k,n]$, such that

$$\sum_{1 \leqslant i<j \leqslant k} a_{ij}\mu(X_i \cap X_j) = \mathrm{LB}_P[f](\boldsymbol{x}), \qquad \sum_{k \leqslant i<j \leqslant n} a_{ij}\mu(X'_i \cap X'_j) = \mathrm{LB}_Q[g](\boldsymbol{x}).$$

Applying a measure preserving bijection $[0,1) \to [0,1)$ that maps $\mathcal{L}$ to $\mathcal{L}$ and X'_k to X_k, we can assume that $X_k = X'_k$, and then the sets $X_1, \dots, X_{k-1}, X_k = X'_k, X'_{k+1}, \dots, X'_n$ provide a certificate for $\mathrm{vex}[f+g](\boldsymbol{x}) = \mathrm{LB}_{P \cap Q}[f+g](\boldsymbol{x})$:

$$\sum_{1 \leqslant i<j \leqslant k} a_{ij}\mu(X_i \cap X_j) + \sum_{k \leqslant i<j \leqslant n} a_{ij}\mu(X'_i \cap X'_j) = \mathrm{LB}_P[f](\boldsymbol{x}) + \mathrm{LB}_Q[g](\boldsymbol{x}) = \mathrm{LB}_{P \cap Q}[f+g](\boldsymbol{x}).$$

The same argument works for $\mathrm{cav}[f+g](\boldsymbol{x}) = \mathrm{UB}_{P \cap Q}[f+g](\boldsymbol{x})$. □

3.2. **Alternative proof for cliques.** In order to illustrate the utility of the geometric characterization of 0–1 polytopes in a simpler setting than what we have for our main results, we start with an alternative proof for the following result that was proved in [Rik97; She97].

Theorem 6. *If $G = K_n$, and all edge weights are equal to 1, then $X(f) = \pi[f](P)$ where $P \subseteq [0,1]^{n(n+1)/2}$ is the polytope described by the inequalities $y_{ij} \leqslant x_i$ and $y_{ij} \leqslant x_j$ for all $ij \in E$, together with*

$$sx(V) - y(E) \leqslant s(s+1)/2, \;\; s = 1, \dots, n-1. \tag{18}$$

To be precise, the result from [Rik97; She97] is just one half of Theorem 6: in our notation it says that

$$\mathrm{vex}[f](\boldsymbol{x}) = \max\left\{0, \, \max\left\{sx(V) - s(s+1)/2 \, : \, s = 1, 2, \dots, n-1\right\}\right\}.$$

The correspondence between this statement in the original space $\mathbb{R}^{n+1}$ and our version in the extended space $\mathbb{R}^{n(n+1)/2}$ comes from the fact that the only constraints enforcing lower bounds on the y-variables are (18), which put lower bounds on $y(E)$ which is precisely the term corresponding to $f(\boldsymbol{x})$ when the products $x_i x_j$ are replaced by the variables y_{ij}. The other half of Theorem 6 is an immediate consequence of Corollary 2: In order to describe $\mathrm{cav}[f](\boldsymbol{x})$ it is sufficient to require $y_{ij} \leqslant \min\{x_i, x_j\}$ for all $ij \in E$. In view of Corollaries 1 and 2, Theorem 6 is a consequence of the following lemma.

Lemma 1. *For every $\boldsymbol{x} \in [0,1]^n$,*

$$\mathrm{LB}_P[f](\boldsymbol{x}) = \mathrm{vex}[f](\boldsymbol{x}) = s(x_1 + \dots + x_n) - \binom{s+1}{2},$$

where $s = \lfloor x_1 + \dots + x_n \rfloor$.

Proof. Since $\mathrm{vex}[f](\boldsymbol{x}) \geqslant \mathrm{LB}_P[f](\boldsymbol{x})$ it is sufficient to show that

$$\mathrm{vex}[f](\boldsymbol{x}) \leqslant s(x_1 + \dots + x_n) - \binom{s+1}{2}, \tag{19}$$

$$\mathrm{LB}_P[f](\boldsymbol{x}) \geqslant s(x_1 + \dots + x_n) - \binom{s+1}{2}. \tag{20}$$

In order to show (19) using Corollary 1 we concatenate intervals of lengths $x_1, \dots, x_n$ and obtain sets $X_i \subseteq [0,1)$ by interpreting the result modulo $\mathbb{Z}$. More formally, the sets X_i are defined as follows: for $i = 1$, put $X_1 = [0, x_1)$. Now let $i \geqslant 2$, suppose X_{i-1} has been defined already, set $b = \sup X_{i-1}$, and put

$$X_i = \begin{cases} [b, b + x_i) & \text{if } b + x_i \leqslant 1, \\ [b, 1) \cup [0, x_i - (1-b)) & \text{if } b + x_i > 1. \end{cases}$$

This construction is illustrated in Figure 4 where $s = \lfloor 2.5 \rfloor = 2$. Now $[0,1) = A \cup B$, where

$$A = \{t \in [0,1) \; : \; |\{i \in [n] \; : \; t \in X_i\}| = s\}, \quad B = \{t \in [0,1) \; : \; |\{i \in [n] \; : \; t \in X_i\}| = s+1\},$$

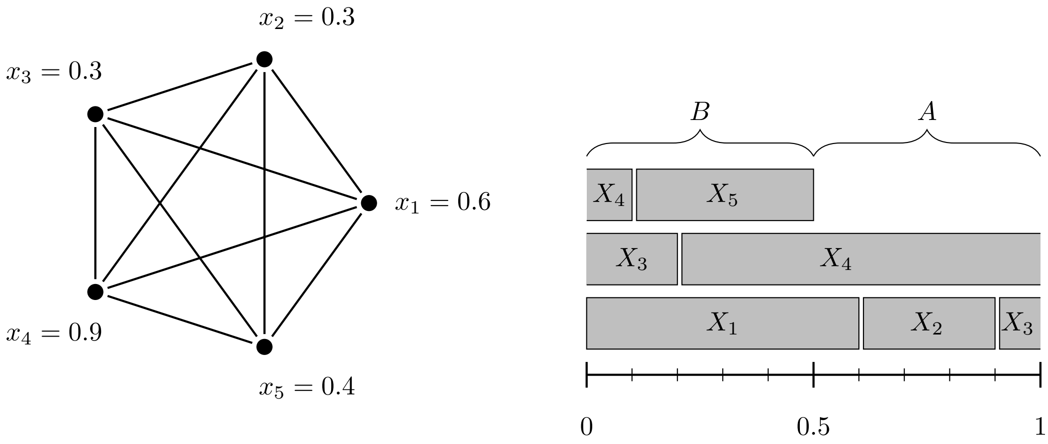

FIGURE 4. Illustration of the construction of the sets X_i in the proof of Lemma 1.

and $\mu(B) = (x_1 + \cdots + x_n) - s$, $\mu(A) = s + 1 - (x_1 + \cdots + x_n)$. Therefore,

$$\begin{aligned}\sum_{1\leqslant i<j\leqslant n} \mu(X_i \cap X_j) &= \binom{s}{2}\mu(A) + \binom{s+1}{2}\mu(B) \\ &= \binom{s}{2}\left[s+1-(x_1+\cdots+x_n)\right] + \binom{s+1}{2}\left[(x_1+\cdots+x_n)-s\right] \\ &= (x_1+\cdots+x_n)\left[\binom{s+1}{2} - \binom{s}{2}\right] + (s+1)\binom{s}{2} - s\binom{s+1}{2} \\ &= s(x_1+\cdots+x_n) - \binom{s+1}{2},\end{aligned}$$

and this implies (19). If $s < n-1$ then (20) is obvious: for $s = 0$ the right-hand side is zero, and for $1 \leqslant s \leqslant n-1$,

$$\sum_{1\leqslant i<j\leqslant n} y_{ij} \geqslant s(x_1+\cdots+x_n) - \binom{s+1}{2}$$

is one of the clique inequalities. Finally, if $s = n$ then

$$\sum_{1\leqslant i<j\leqslant n} y_{ij} = \binom{n}{2} = n^2 - \binom{n+1}{2} = n(x_1+\cdots+x_n) - \binom{n+1}{2}. \qquad \square$$

4. PROOFS OF MAIN RESULTS

In Sections 4.1 and 4.2 we will use Corollary 1 to prove Theorems 1 and 2.

4.1. **Proof of Theorem 1.** We will establish this result by proving $\mathrm{vex}[f](\boldsymbol{x}) \leqslant \mathrm{LB}_P[f](\boldsymbol{x})$. Without loss of generality we assume $x_n \leqslant x_{n-1}$ and $x_1 \geqslant x_2 \geqslant \cdots \geqslant x_{n-2}$, and we proceed as follows. We start with $X_n = [0, x_n)$ and $X_{n-1} = [0, x_{n-1})$ and construct the sets $X_1, \dots, X_{n-2}$ as described in Algorithm 1.

Example 4. For $n = 6$ two different outcomes of Algorithm 1 are illustrated in Figure 5: For $\boldsymbol{x} = (0.9, 0.6, 0.2, 0.1, 0.6, 0.4)$ the algorithm terminates with $b = 1$, while for $\boldsymbol{x} = (0.9, 0.8, 0.2, 0.1, 0.6, 0.4)$, we have still $b < 1$ in the end.

Setting $y_{ij} = \mu(X_i \cap X_j)$ for all $ij \in E$, Corollary 1 implies $\mathrm{vex}[f](\boldsymbol{x}) \leqslant \sum_{ij\in E} y_{ij}$, and it is sufficient to show that $\boldsymbol{y}$ is an optimal solution for the LP defining $\mathrm{LB}_P[f](\boldsymbol{x})$:

$$\text{Minimize } \sum_{ij\in E} y_{ij} \text{ subject to the McCormick inequalities (2) and (10)–(11).}$$

Algorithm 1 Construction of the sets X_i in the proof of Theorem 1

$(a,b) \leftarrow (x_n, x_{n-1})$
for $k = 1, 2, \dots, n-2$ **do**
  if $x_k \leqslant 1-b$ **then**
    $X_k \leftarrow [b, b+x_k)$
    $b \leftarrow b + x_k$
  else if $x_k \leqslant 1-a$ **then**
    $X_k \leftarrow [b,1) \cup [a, a+x_k+b-1)$
    $(a,b) \leftarrow (a+x_k+b-1, 1)$
  else
    $X_k \leftarrow [a,1) \cup [0, x_k+a-1)$
    $a \leftarrow x_k + a - 1$

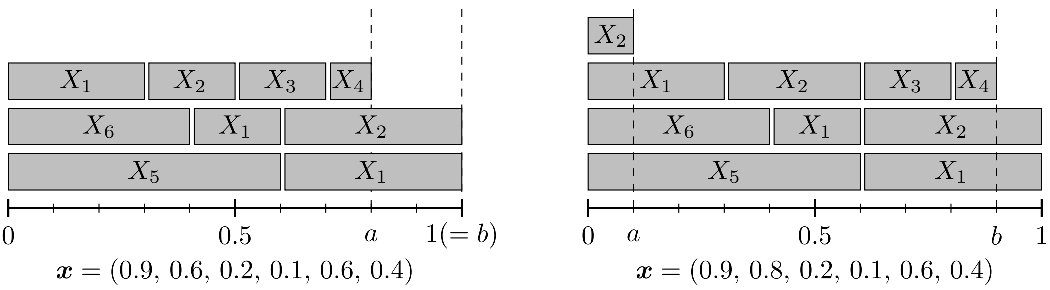

FIGURE 5. The sets X_i constructed by Algorithm 1 for two vectors $\boldsymbol{x}$.

Our argument will be based on expressing $\sum_{ij\in E} y_{ij}$ in terms of the variables x_i and then arguing the inequalities listed in Theorem 1 imply that this expression is a lower bound for $\mathrm{LB}_P[f](\boldsymbol{x})$. In other words, we need to verify $\sum_{ij\in E} y_{ij} \leqslant \sum_{ij\in E} y'_{ij}$ for every $\boldsymbol{y}'$ with $(\boldsymbol{x},\boldsymbol{y}') \in P$. The following observations turn out to be useful.

- During the runtime of the algorithm the parameter b is never decreasing.
- After every step of Algorithm 1 there is an integer s such that

$$|\{i \in [n] \,:\, t \in X_i\}| = \begin{cases} s+2 & \text{for } t < a, \\ s+1 & \text{for } a \leqslant t < b, \\ s & \text{for } b \leqslant t < 1 \end{cases}$$

  if $b < 1$, and

$$|\{i \in [n] \,:\, t \in X_i\}| = \begin{cases} s+1 & \text{for } t < a, \\ s & \text{for } a \leqslant t < 1, \end{cases}$$

  if $b = 1$.
- If $b = 1$ at termination of the algorithm, and s is the integer with $|\{i \in [n] \,:\, t \in X_i\}| = s$ for $t \in [a,1)$ then $a = x_1 + x_2 + \dots + x_n - s$.
- If $b < 1$ at termination of the algorithm, and s is the integer with $|\{i \in [n] \,:\, t \in X_i\}| = s$ for $t \in [b,1)$ then $a = x_1 + x_2 + \dots + x_s + x_n - s$ and $b = x_{s+1} + x_{s+2} + \dots + x_{n-1}$.

Case 1: $b = 1$. Then $1 \leqslant s \leqslant n-1$. As in the proof of Lemma 1,

$$\sum_{ij\in E} y_{ij} = \sum_{1\leqslant i<j\leqslant n} y_{ij} - y_{n-1,n} = s(x_1+\dots+x_n) - x_n - \binom{s+1}{2}.$$

For $s \leqslant n-2$, we use (11):

$$s(x_1+\dots+x_n) - x_n - \binom{s+1}{2} \leqslant sx(V) - y'_{n,n-1} - \binom{s+1}{2} \overset{(11)}{\leqslant} y'_{ij}(E).$$

For $s = n-1$, we combine (9) and (10):

$$y'(E) = \left[y'(E(V\setminus\{n-1,n\})) + \frac{1}{2}\sum_{i=1}^{n-2}\left(y'_{i,n-1}+y'_{in}\right)\right] + \frac{1}{2}\sum_{i=1}^{n-2}\left(y'_{i,n-1}+y'_{in}\right)$$
$$\overset{(9),(10)}{\geqslant} (n-2)\left[x(V\setminus\{n-1,n\}) + \frac{x_{n-1}+x_n}{2}\right] - \binom{n-1}{2}$$
$$+ x\left(V\setminus\{n-1,n\}\right) + \frac{(n-2)(x_{n-1}+x_n)}{2} - (n-2)$$
$$= (n-1)x(V) - (x_{n-1}+x_n) - \binom{n}{2} + 1 \geqslant (n-1)x(V) - x_n - \binom{n}{2},$$

as required.

Case 2: $b < 1$. Then $0 \leqslant s \leqslant n-2$, and

$$\sum_{ij\in E} y_{ij} = a\left[\binom{s+2}{2} - 1\right] + (x_n - a)\left[\binom{s+1}{2} - 1\right] + (b - x_n)\binom{s+1}{2} + (1-b)\binom{s}{2}$$
$$= a(s+1) - x_n + bs + \binom{s}{2}$$
$$= (x_1 + \cdots + x_s + x_n - s)(s+1) - x_n + (x_{s+1} + \cdots + x_{n-1})s + \binom{s}{2}$$
$$= sx(V) + (x_1 + \cdots + x_s) - s - \binom{s+1}{2}$$
$$= s\left[x\left(V\setminus\{n-1,n\}\right) + \frac{x_{n-1}+x_n}{2}\right] + \frac{1}{2}\sum_{i=1}^{s}(2x_i + x_{n-1} + x_n - 2) - \binom{s+1}{2}.$$

Next we verify that this is a lower bound for $\mathrm{LB}_P[f](\boldsymbol{x})$. We start with the inequality

$$\sum_{ij\in E} y'_{ij} = \sum_{ij\in E(V\setminus\{n-1,n\})} y'_{ij} + \sum_{i=1}^{n-2}\left(y'_{i,n-1}+y'_{in}\right)$$
$$\geqslant \left[\sum_{ij\in E(V\setminus\{n-1,n\})} y'_{ij} + \frac{1}{2}\sum_{i=1}^{n-2}\left(y'_{i,n-1}+y'_{in}\right)\right] + \frac{1}{2}\sum_{i=1}^{s}\left(y'_{i,n-1}+y'_{in}\right).$$

We use (9) and (10) to bound the second and the first part, respectively:

$$\sum_{ij\in E(V\setminus\{n-1,n\})} y'_{ij} + \frac{1}{2}\sum_{i=1}^{n-2}\left(y'_{i,n-1}+y'_{in}\right) \overset{(10)}{\geqslant} s\left[x\left(V\setminus\{n-1,n\}\right) + \frac{x_{n-1}+x_n}{2}\right] - \binom{s+1}{2},$$
$$\frac{1}{2}\sum_{i=1}^{s}\left(y'_{i,n-1}+y'_{in}\right) \overset{(9)}{\geqslant} \frac{1}{2}\sum_{i=1}^{s}\left(2x_i + x_{n-1} + x_n - 2\right),$$

and as a consequence $y'(E) \geqslant y(E)$, as required.

The second part of the theorem (that P is a minimal extension of $X(f)$) is proved by identifying, for each inequality listed in Theorem 1, a point $(\boldsymbol{x},\boldsymbol{y})$ which is contained in the polytope P' obtained from P by omitting the inequality, such that $\pi[f](\boldsymbol{x},\boldsymbol{y}) \notin X(f)$. This is described in detail in Appendix A.

4.2. **Proof of Theorem 2.** Throughout all indices are in $[n]$ and have to be read modulo n in the obvious way. In particular, $n \equiv 0$ and $n+1 \equiv 1$. The n-cycle corresponds to the function

$$f(\boldsymbol{x}) = \sum_{i=1}^{n} a_i x_i x_{i+1},$$

where $\boldsymbol{a}$ is arbitrary with $a_i \neq 0$ for all $i \in [n]$. Let $P \subseteq [0,1]^{2n}$ be the polytope described by the McCormick inequalities together with (12) and (13). We claim that $\pi[f](P) = X(f)$, and that (12) (resp. (13)) can be omitted if $|E^-|$ (resp. $|E^+|$) is even.

We need to show that for every $\boldsymbol{x} \in [0,1]^n$ we have $\operatorname{vex}[f](\boldsymbol{x}) = \mathrm{LB}_P[f](\boldsymbol{x})$ and $\operatorname{cav}[f](\boldsymbol{x}) = \mathrm{UB}_P[f](\boldsymbol{x})$. We present the argument for $\operatorname{cav}[f](\boldsymbol{x}) = \mathrm{UB}_P[f](\boldsymbol{x})$ in detail, as $\operatorname{vex}[f](\boldsymbol{x}) = \mathrm{LB}_P[f](\boldsymbol{x})$ can be proved similarly. Fix $\boldsymbol{x} \in [0,1]^n$ and put

$$\mu_i = \min\{x_i, x_{i+1}\},$$
$$\eta_i = \max\{0, x_i + x_{i+1} - 1\},$$
$$A = x(V^+) - x(V^-) + \left\lfloor \frac{|E^-|}{2} \right\rfloor.$$

Then

$$\mathrm{UB}_P[f](\boldsymbol{x}) = \max\left\{\sum_{i=1}^n a_i y_i \ :\ \eta_i \leqslant y_i \leqslant \mu_i,\ y(E^+) - y(E^-) \leqslant A\right\} \tag{21}$$

$$= \min\left\{\sum_{i=1}^n (\mu_i z_i - \eta_i w_i) + A\alpha \ :\ z_i - w_i + \alpha \geqslant a_i \text{ for } i \in E^+,\right.$$
$$\left. z_i - w_i - \alpha \geqslant a_i \text{ for } i \in E^-,\ z_i, w_i, \alpha \geqslant 0\right\}. \tag{22}$$

W.l.o.g. we assume $|a_n| \leqslant |a_i|$ for all $i \in [n]$, and we define sets $X_i \subseteq [0,1)$ as follows. Let $X_1 = [0, x_1)$ and

$$X_2 = \begin{cases} [0, x_2) & \text{if } a_1 > 0, \\ [1 - x_2, 1) & \text{if } a_1 < 0. \end{cases}$$

For $i \geqslant 3$ we set $E_i^- = \{j \ :\ i \leqslant j \leqslant n,\ a_j < 0\}$ and define X_i depending on the parity of $|E_i^-|$ and the sign of a_{i-1}. Intuitively, we can think of filling a bucket of capacity x_i from the reservoirs $R_1 = X_{i-1} \setminus X_1$, $R_2 = X_{i-1} \cap X_1$, $R_3 = [0,1) \setminus (X_1 \cup X_{i-1})$ and $R_4 = X_1 \setminus X_{i-1}$, and there are two objectives:

(1) If $a_{i-1} > 0$ we want to maximize $\mu(X_{i-1} \cap X_i)$, so the reservoirs R_1 and R_2 are used before R_3 and R_4. If $a_{i-1} < 0$ then we want to minimize $\mu(X_{i-1} \cap X_i)$, so R_3 and R_4 are used before R_1 and R_2.

(2) If $|E_i^-|$ is odd we want to minimize $\mu(X_1 \cap X_i)$, so R_1 is used before R_2 and R_3 before R_4. If $|E_i^-|$ is even we want to maximize $\mu(X_1 \cap X_i)$, so R_2 is used before R_1 and R_4 before R_3.

Algorithm 2 `Bucket`(Y_1, Y_2, Y_3, Y_4, x)

Input: A partition $[0,1) = Y_1 \cup Y_2 \cup Y_3 \cup Y_4$ (with $Y_k \in \mathcal{L}$) and capacity $x \in [0,1]$

Initialize $X \leftarrow \emptyset$ and $k \leftarrow 1$
while $\mu(X) < x$ **do**
  Let $Z \in \mathcal{L}$ be a subset of Y_k with $\mu(Z) = \min\{\mu(Y_k),\ x - \mu(X)\}$
  $X \leftarrow X \cup Z$
  $k \leftarrow k + 1$

Output: $X \in \mathcal{L}$ with $\mu(X) = x$

More formally,

$$X_i = \begin{cases} \texttt{Bucket}(R_1, R_2, R_3, R_4, x_i) & \text{if } a_{i-1} > 0 \text{ and } |E_i^-| \text{ odd}, \\ \texttt{Bucket}(R_2, R_1, R_4, R_3, x_i) & \text{if } a_{i-1} > 0 \text{ and } |E_i^-| \text{ even}, \\ \texttt{Bucket}(R_3, R_4, R_1, R_2, x_i) & \text{if } a_{i-1} < 0 \text{ and } |E_i^-| \text{ odd}, \\ \texttt{Bucket}(R_4, R_3, R_2, R_1, x_i) & \text{if } a_{i-1} < 0 \text{ and } |E_i^-| \text{ even}. \end{cases}$$

where the function `Bucket` is described in Algorithm 2. Note that this corresponds to a solution with

$$y_i = \mu(X_i \cap X_{i+1}) = \begin{cases} \mu_i & \text{for } i \in E^+ \setminus \{n\}, \\ \eta_i & \text{for } i \in E^- \setminus \{n\}. \end{cases}$$

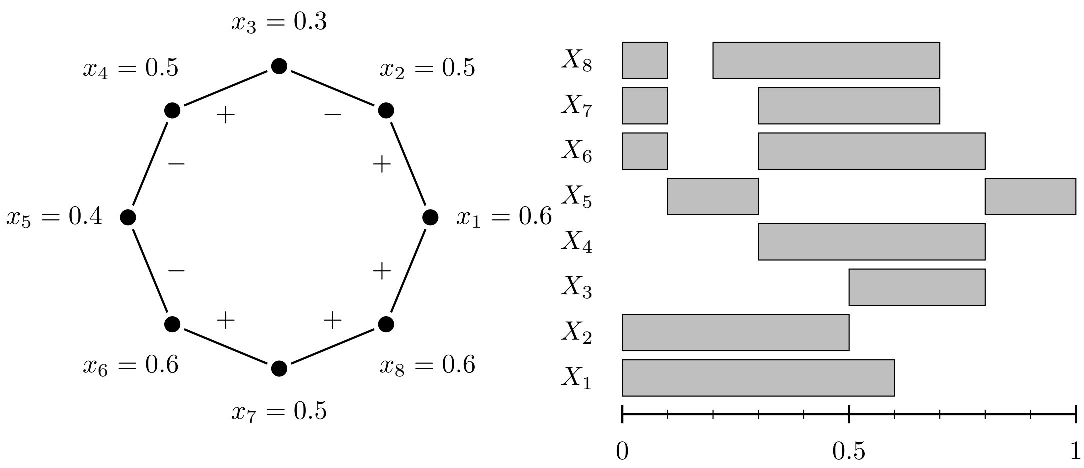

FIGURE 6. Constructing the sets X_i, where the edge labels indicate the sign of the coefficient a_i.

Example 5. The construction is illustrated in Figure 6, and with $|a_i| = 1$ for all $i \in [8]$ the corresponding objective value is 2.3 and we conclude $\text{cav}[f](\boldsymbol{x}) = \text{UB}_P[f](\boldsymbol{x}) = 2.3$ because (12) becomes

$$x(V^-) - x(V^+) + y(E^+) - y(E^-) = 0.4 - 1.7 + f(\boldsymbol{x}) \leqslant \left\lfloor \frac{|E^-|}{2} \right\rfloor = 1.$$

Next we define the *defects*

$$\delta_i = \begin{cases} \mu(X_1 \cap X_i) - \max\{0, x_1 + x_i - 1\} & \text{if } |E_i^-| \text{ is odd,} \\ \min\{x_i, x_1\} - \mu(X_1 \cap X_i) & \text{if } |E_i^-| \text{ is even.} \end{cases}$$

We have $\delta_i \geqslant 0$ for all $i \in [n]$ and

$$\text{cav}[f](\boldsymbol{x}) \geqslant \sum_{i=1}^{n} a_i \mu(X_i \cap X_{i+1}) = \sum_{i \in E^+} a_i \mu_i + \sum_{i \in E^-} a_i \eta_i - |a_n| \delta_n,$$

so in order to complete the proof of the claim $\text{cav}[f](\boldsymbol{x}) = \text{UB}_P[f](\boldsymbol{x})$ it is sufficient to prove

$$\text{UB}_P[f](\boldsymbol{x}) \leqslant \sum_{i \in E^+} a_i \mu_i + \sum_{i \in E^-} a_i \eta_i - |a_n| \delta_n. \tag{23}$$

If $\delta_n = 0$ then this follows immediately from the McCormick inequalities which imply

$$a_i y_i \leqslant \begin{cases} a_i \mu_i & \text{for } i \in E^+, \\ a_i \eta_i & \text{for } i \in E^-. \end{cases}$$

For $\delta_n > 0$ the claim is a consequence of the following lemma.

Lemma 2. *If* $\delta_n > 0$ *then* $\delta_n = \sum_{i \in E^+} \mu_i - \sum_{i \in E^-} \eta_i - A.$

Using Lemma 2, we can prove inequality (23) by LP duality. If $a_n > 0$ we define a solution for the dual problem (22) by $\alpha = a_n$ and (using the assumption that $|a_n| \leqslant |a_i|$ for all $i \in [n]$)

$$z_i = \begin{cases} a_i - a_n & \text{for } i \in E^+, \\ 0 & \text{for } i \in E^-, \end{cases} \qquad w_i = \begin{cases} 0 & \text{for } i \in E^+, \\ -a_i - a_n & \text{for } i \in E^-. \end{cases}$$

This is a feasible solution for (22) with objective value

$$\sum_{i\in E^+} \mu_i(a_i - a_n) + \sum_{i\in E^-} \eta_i(a_i + a_n) + Aa_n = \sum_{i\in E^+} a_i\mu_i + \sum_{i\in E^-} a_i\eta_i + a_n\left(A - \sum_{i\in E^+}\mu_i + \sum_{i\in E^-}\eta_i\right)$$
$$= \sum_{i\in E^+} a_i\mu_i + \sum_{i\in E^-} a_i\eta_i - a_n\delta_n.$$

Similarly, for $a_n < 0$ we define a solution for the dual problem (22) by $\alpha = -a_n$ and

$$z_i = \begin{cases} a_i + a_n & \text{for } i \in E^+, \\ 0 & \text{for } i \in E^-, \end{cases} \qquad w_i = \begin{cases} 0 & \text{for } i \in E^+, \\ a_n - a_i & \text{for } i \in E^-. \end{cases}$$

with objective value

$$\sum_{i\in E^+} \mu_i(a_i + a_n) + \sum_{i\in E^-} \eta_i(a_i - a_n) - Aa_n = \sum_{i\in E^+} a_i\mu_i + \sum_{i\in E^-} a_i\eta_i + a_n\left(\sum_{i\in E^+}\mu_i - \sum_{i\in E^-}\eta_i - A\right)$$
$$= \sum_{i\in E^+} a_i\mu_i + \sum_{i\in E^-} a_i\eta_i + a_n\delta_n.$$

Before proving Lemma 2 we show that the sequence $(\delta_i)_{i=2,\dots,n}$ is decreasing, and therefore in the proof of Lemma 2 we may assume $\delta_i > 0$ for all $i \in [n]$.

Lemma 3. *For all $i \in \{3, \dots, n\}$ we have $\delta_i \leqslant \delta_{i-1}$.*

Proof. We check the following implications.

(1) $|E_{i-1}|$ odd, $a_{i-1} > 0 \implies \delta_i = \delta_{i-1} - \begin{cases} \max\{0,\, x_{i-1} - x_i,\, (x_1 + x_i - 1)\} & \text{if } x_1 + x_{i-1} \leqslant 1, \\ \max\{0,\, 1 - x_1 - x_i,\, x_i - x_{i-1}\} & \text{if } x_1 + x_{i-1} \geqslant 1. \end{cases}$

(2) $|E_{i-1}|$ odd, $a_{i-1} < 0 \implies \delta_i = \delta_{i-1} - \begin{cases} \max\{0,\, x_1 - x_i,\, (x_{i-1} + x_i - 1)\} & \text{if } x_1 + x_{i-1} \leqslant 1, \\ \max\{0,\, 1 - x_{i-1} - x_i,\, x_i - x_1\} & \text{if } x_1 + x_{i-1} \geqslant 1. \end{cases}$

(3) $|E_{i-1}|$ even, $a_{i-1} > 0 \implies \delta_i = \delta_{i-1} - \begin{cases} \max\{0,\, x_{i-1} - x_i,\, (x_i - 1)\} & \text{if } x_{i-1} \leqslant x_1, \\ \max\{0,\, x_1 - x_i,\, x_i - x_{i-1}\} & \text{if } x_{i-1} \geqslant x_1. \end{cases}$

(4) $|E_{i-1}|$ even, $a_{i-1} < 0 \implies \delta_i = \delta_{i-1} - \begin{cases} \max\{0,\, 1 - x_1 - x_i,\, x_{i-1} + x_i - 1\} & \text{if } x_{i-1} \leqslant x_1, \\ \max\{0,\, 1 - x_{i-1} - x_i,\, x_1 + x_i - 1\} & \text{if } x_{i-1} \geqslant x_1. \end{cases}$

∎

Proof of Lemma 2. If $|E^-|$ is even then $\delta_2 = 0$, hence $\delta_i = 0$ for all $i \in [2, n]$ by Lemma 3, and there is nothing to do. For odd $|E^-|$ we proceed by induction on n. Note that the construction of the sets X_i, the definition of the numbers δ_i, and the statement of the lemma depend only on sequences $(x_1, \dots, x_n)$ and $(a_1, \dots, a_n)$ (actually only on the signs of the a_i). Therefore we can use $n = 2$ as the base case. Then $|E^-| = 1$, $V^+ = V^- = \emptyset$, $A = 0$, and we have the following two cases.

Case 1: if $a_2 < 0 < a_1$, then $\mu(X_1 \cap X_2) = \min\{x_1, x_2\}$ and

$$\delta_2 = \mu(X_1 \cap X_2) - \max\{0, x_1 + x_2 - 1\} = \mu_1 - \eta_2.$$

Case 2: if $a_1 < 0 < a_2$, then $\mu(X_1 \cap X_2) = \max\{0, x_1 + x_2 - 1\}$ and

$$\delta_2 = \min\{x_1, x_2\} - \mu(X_1 \cap X_2) = \mu_2 - \eta_1.$$

Now let $n \geqslant 3$ and set

$$\gamma = \sum_{i\in E^+} \mu_i - \sum_{i\in E^-} \eta_i - A = \sum_{i\in E^+} \mu_i - \sum_{i\in E^-} \eta_i - x(V^+) + x(V^-) - \left\lfloor \frac{|E^-|}{2} \right\rfloor,$$

so that our aim becomes to show $\delta_n = \gamma$. Applying the induction hypothesis to the sets $X_1, \dots, X_{n-1}$ that are obtained by applying the construction for the sequences $(x_1, \dots, x_{n-1})$ and $(a_1, \dots, a_{n-2}, \operatorname{sign}(a_{n-1}a_n))$ (the number of negative terms is still odd), we get

$$\delta_{n-1} = \begin{cases} \gamma - \mu_{n-1} - \mu_n + x_n + \min\{x_{n-1}, x_1\} & \text{if } n-1, n \in E^+, \\ \gamma + \eta_{n-1} + \eta_n - x_n - x_{n-1} - x_1 + \min\{x_{n-1}, x_1\} + 1 & \text{if } n-1, n \in E^-, \\ \gamma - \mu_{n-1} + \eta_n + x_{n-1} - \max\{0, x_{n-1} + x_1 - 1\} & \text{if } n-1 \in E^+, n \in E^-, \\ \gamma + \eta_{n-1} - \mu_n + x_1 - \max\{0, x_{n-1} + x_1 - 1\} & \text{if } n-1 \in E^-, n \in E^+. \end{cases}$$

Now we discuss the four cases separately. The detailed case analysis can be found in Appendix B.

Case 1: $n-1, n \in E^+$. We may assume $\mu(X_1 \cap X_{n-1}) < x_n < x_1 + x_{n-1} - \mu(X_1 \cap X_{n-1})$ since otherwise $\delta_n = 0$.
Case 1.1: $x_{n-1} \leqslant x_1$. In this case $\delta_{n-1} = \mu(X_{n-1} \setminus X_1)$.
Case 1.2: $x_{n-1} > x_1$. In this case $\delta_{n-1} = \mu(X_1 \setminus X_{n-1})$.

Case 2: $n-1, n \in E^-$. We may assume $1 - \mu(X_1 \cup X_{n-1}) < x_n < 1 - \mu(X_1 \cap X_{n-1})$ since otherwise $\delta_n = 0$.
Case 2.1: $x_{n-1} \leqslant x_1$. In this case $\delta_{n-1} = \mu(X_{n-1} \setminus X_1) = \mu(X_1 \cup X_{n-1}) - x_1$.
Case 2.2: $x_{n-1} > x_1$. In this case $\delta_{n-1} = \mu(X_1 \setminus X_{n-1}) = \mu(X_1 \cup X_{n-1}) - x_{n-1}$.

Case 3: $n-1 \in E^+$, $n \in E^-$. We may assume $\mu(X_{n-1} \setminus X_1) < x_n < 1 - \mu(X_1 \setminus X_{n-1})$ since otherwise $\delta_n = 0$.
Case 3.1: $x_1 + x_{n-1} \leqslant 1$. In this case $\delta_{n-1} = \mu(X_1 \cap X_{n-1})$.
Case 3.2: $x_1 + x_{n-1} > 1$. In this case $\delta_{n-1} = \mu([0,1) \setminus (X_1 \cup X_{n-1}))$.

Case 4: $n-1 \in E^-$, $n \in E^+$. We may assume $\mu(X_{n-1} \setminus X_1) < x_n < 1 - \mu(X_{n-1} \setminus X_1)$ since otherwise $\delta_n = 0$.
Case 4.1: $x_1 + x_{n-1} \leqslant 1$. In this case $\delta_{n-1} = \mu(X_1 \cap X_{n-1})$.
Case 4.2: $x_1 + x_{n-1} > 1$. In this case $\delta_{n-1} = \mu([0,1) \setminus (X_1 \cup X_{n-1}))$. □

Finally, to complete our proof of Theorem 2, we need to verify necessity of the McCormick inequalities (2) and the inequalities (12) and (13). For each of these inequalities, we exhibit a point $(\boldsymbol{x}, \boldsymbol{y}) \in P'$, where P' is the polytope obtained by dropping an inequality from P, such that $\pi[f](\boldsymbol{x}, \boldsymbol{y}) \notin X(f)$.

The inequality $y_i \leqslant x_i$. Let $x_{i+1} = 1$ and $x_j = 0$ for all $j \in [n] \setminus \{i+1\}$, so that $(\boldsymbol{x}, z) \in X(f) \iff z = 0$. Setting $y_i = 1$ and $y_j = 0$ for all $j \in [n] \setminus \{i\}$, we obtain a point $(\boldsymbol{x}, \boldsymbol{y})$ with $\pi[f](\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, a_i) \notin X(f)$. If $E^- \neq \{i+1\}$ and $E^+ \neq \{i+1\}$, then $(\boldsymbol{x}, \boldsymbol{y})$ satisfies (12) and (13), so $(\boldsymbol{x}, \boldsymbol{y}) \in P'$, as required. But if $E^- = \{i+1\}$ (resp. $E^+ = \{i+1\}$) then $(\boldsymbol{x}, \boldsymbol{y})$ is cut off by (12) (resp. (13)). In these cases we use the point given by $x_{i+1} = x_{i+2} = 1$, $x_j = 0$ for all $j \in [n] \setminus \{i+1, i+2\}$, $y_i = y_{i+1} = 1$ and $y_j = 0$ for all $j \in [n] \setminus \{i, i+1\}$ instead.

The inequality $y_i \leqslant x_{i+1}$ can be treated similarly.

The inequality $y_i \geqslant x_i + x_{i+1} - 1$. Let $x_i = x_{i+1} = 1$ and $x_j = 0$ for all $j \in [n] \setminus \{i, i+1\}$, so that $(\boldsymbol{x}, z) \in X(f) \iff z = a_i$. Setting $\boldsymbol{y} = \boldsymbol{0}$, we obtain a point $(\boldsymbol{x}, \boldsymbol{y})$ with $\pi[f](\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, 0) \notin X(f)$. If $E^+ \neq \{i-1, i, i+1\}$ and $E^- \neq \{i-1, i, i+1\}$ then $(\boldsymbol{x}, \boldsymbol{y}) \in P'$ and we are done. Otherwise, we use the point given by $x_i = x_{i+1} = x_{i+2} = 1$, $x_j = 0$ for all $j \in [n] \setminus \{i, i+1, i+2\}$, $y_{i+1} = 1$ and $y_j = 0$ for all $j \in [n] \setminus \{i+1\}$.

The inequality (12). This is Padberg's cycle inequality corresponding to the subset $D = E^-$, and hence is known to be implied by the McCormick inequalities when $|E^-|$ is even. Now let $|E^-|$ be odd, say $|E^-| = 2k+1$ for some non-negative integer k. Set $x_i = 1/2$ for all $i \in [n]$.

Then $(\boldsymbol{x},\boldsymbol{y})\in P$ implies

$$y(E^+)-y(E^-)\overset{(12)}{\leqslant} x(V^+)-x(V^-)+k=\frac{|V^+|-|V^-|}{2}+k=\frac{n-2(2k+1)}{2}+k$$
$$=\frac{n-2k-2}{2}=\frac{|E^+|-1}{2},$$

and consequently, for every $(\boldsymbol{x},z)\in X(f)$,

$$z=\sum_{i=1}^n a_iy_i=\sum_{i\in E^+}a_iy_i+\sum_{i\in E^-}a_iy_i<\frac{1}{2}\sum_{i\in E^+}a_i.$$

Setting $y_i=1/2$ for all $i\in E^+$ and $y_i=0$ for all $i\in E^-$, we obtain a point $(\boldsymbol{x},\boldsymbol{y})\in P'$ with $\pi[f](\boldsymbol{x},\boldsymbol{y})=\left(\boldsymbol{x},\,\frac{1}{2}\sum_{i\in E^+}a_i\right)\notin X(f)$.

The inequality (13). The arguments here are similar to those for (12).

4.3. **Proof of Theorem 3.** Let $G=(V,E)$ be a cactus graph with k cycles and arbitrary edge weights. We want to show that for the corresponding function f, $X(f)=\pi[f](P)$ where P is described by the McCormick inequalities and at most $2k$ cycle inequalities. We prove this by induction on k, the number of cycles. For $k=0$, G is a tree, and the McCormick inequalities are sufficient. For $k=1$, we proceed by induction on the number of edges that are not contained in the cycle. If there are no such edges then G is a cycle, and the claim follows from Theorem 2. Otherwise, there are two graphs $G_1=(V_1,E_1)$ and $G_2=(V_2,E_2)$ with $V_1\cup V_2=V$, $E_1\cup E_2=E$, $|V_1\cap V_2|=1$, $E_2\neq\emptyset$, and such that the cycle of G is contained in G_1. Let f_1 and f_2 be the bilinear functions corresponding to the graphs G_1 and G_2, respectively. Since G_2 is cycle-free, $X(f_2)$ is described by the McCormick inequalities, and by induction, $X(f_1)$ is described by the McCormick inequalities and at most two cycle inequalities. Now the result follows from Corollary 3. For $k\geqslant 2$, there are two graphs $G_1=(V_1,E_1)$ and $G_2=(V_2,E_2)$ with $V_1\cup V_2=V$, $E_1\cup E_2=E$, $|V_1\cap V_2|=1$, such that $k=k_1+k_2$ where k_i $(i\in\{1,2\})$ is the number of cycles in graph G_i. Again, let f_1 and f_2 be the bilinear functions corresponding to the graphs G_1 and G_2, respectively. By induction, $X(f_1)$ is described by the McCormick inequalities and at most $2k_1$ cycle inequalities, and $X(f_2)$ is described by the McCormick inequalities and at most $2k_2$ cycle inequalities. The result follows from Corollary 3.

5. Conclusion and open problems

We have used an extension of Zuckerberg's geometric method for characterizing convex hulls of subsets of the discrete n-cube to find extended formulations for the convex hulls of graphs of bilinear functions corresponding to almost complete graphs with unit weights and cactus graphs with arbitrary weights. We think that this approach can be used in more general situations, but additional insights are needed to avoid the tedious case discussions as in the proof of Theorem 2.

A natural next test case for the method is the class of wheels. A wheel W_{n-1} is the graph with vertex set $V=\{1,\dots,n\}$ for $n\geqslant 5$, and edge set

$$E=\{\{1,2\},\{2,3\},\dots,\{n-2,n-1\}\}\cup\{\{i,n\}\,:\,1\leqslant i\leqslant n-1\}.$$

That is, W_{n-1} is the cycle C_{n-1} with spokes from the center vertex n. Since W_{n-1} has a K_4-minor, we know that $QP(W_n)$ needs more than the cycle inequalities for its description. One extra facet-defining inequality for $QP(W_{n-1})$ is

$$\lfloor (n-1)/2\rfloor\,x_n+\sum_{i=1}^n x_i-y(E)\leqslant\lfloor (n-1)/2\rfloor\,. \tag{24}$$

which can be argued by first principles. Another inequality comes from the cut polytope of a graph. Barahona and Mahjoub [BM86] introduced $\mathsf{CUT}(G)$ as the convex hull of incidence vectors of the cuts in G, and showed that for every odd bicycle wheel in G, there is a corresponding facet-defining inequality for $\mathsf{CUT}(G)$. An *odd bicycle wheel* is the graph $W_{n-1}+\{v\}$ for odd $n-1$, where $G+\{v\}$ is the graph obtained from G by joining every vertex of G to a new vertex v. When $G=W_{n-1}$ and $n-1$ is odd, then $G+\{n+1\}$ is an odd bicycle wheel and we have exactly one odd bicycle wheel inequality for $\mathsf{CUT}(W_{n-1} + \{n+1\})$. Mapping this inequality to $QP(W_{n-1})$ using the well-known linear bijection between $QP(G)$ and $\mathsf{CUT}(G + \{v\})$ [Sim90] leads to the inequality

$$\binom{n}{2} x_n + 2\sum_{i=1}^{n-1} x_i - y(E) \leqslant n-1 \tag{25}$$

defining a facet of $QP(W_{n-1})$ when n is odd. When $n-1=5$, the inequalities (24) and (25) are sufficient to convexify f.

Observation (5-wheel)**.** *If $G = W_5$ and all edge weights are equal to 1, then $X(f) = \pi[f](P)$ where P is the polytope described by the McCormick inequalities (2) together with*

$$2x_6 + x_1 + \cdots + x_5 - y(E) \leqslant 2, \tag{26}$$

$$3x_6 + 2(x_1 + \cdots + x_5) - y(E) \leqslant 5, \tag{27}$$

which are are exactly inequalities (24) and (25) for $n = 6$.

We conjecture that this observation extends to any wheel W_{n-1} with $n \geqslant 6$, n even.

The open question on wheel graphs can be extended to a richer family of graphs called *Halin graphs* which was introduced by Halin [Hal71]. A Halin graph is a planar graph obtained from a tree without vertices of degree 2 by adding a cycle through all the leaves. Thus, a wheel is the simplest kind of a Halin graph. The insights gained from generalizing the above observation on wheel graphs might be useful to characterizing a polytope P that does not project onto $QP(G)$ yet has $X(f) = \pi[f](P)$ and is a minimal such extension of $X(f)$, when G is a Halin graph.

Acknowledgements. The first author thanks Jon Lee for bringing cactus graphs to his attention during stimulating discussions at the Dagstuhl Seminar on Designing and Implementing Algorithms for Mixed-Integer Nonlinear Optimization held at Schloss Dagstuhl in February 2018.

REFERENCES

[Ass+17] Benjamin Assarf, Ewgenij Gawrilow, Katrin Herr, Michael Joswig, Benjamin Lorenz, Andreas Paffenholz, and Thomas Rehn. "Computing convex hulls and counting integer points with polymake". In: *Mathematical Programming Computation* 9.1 (2017), pp. 1–38.

[AT15] David Avis and Hans Raj Tiwary. "On the extension complexity of combinatorial polytopes". In: *Mathematical Programming* 153.1 (2015), pp. 95–115.

[BCC93] Egon Balas, Sebastián Ceria, and Gérard Cornuéjols. "A lift-and-project cutting plane algorithm for mixed 0–1 programs". In: *Mathematical Programming* 58.1 (1993), pp. 295–324.

[BM14] Martin Ballerstein and Dennis Michaels. "Extended formulations for convex envelopes". In: *Journal of Global Optimization* 60.2 (2014), pp. 217–238.

[BM86] Francisco Barahona and Ali Ridha Mahjoub. "On the cut polytope". In: *Mathematical Programming* 36.2 (1986), pp. 157–173.

[Bol+17] Natashia Boland, Santanu S. Dey, Thomas Kalinowski, Marco Molinaro, and Fabian Rigterink. "Bounding the gap between the McCormick relaxation and the convex hull for bilinear functions". In: *Mathematical Programming* 162 (2017), pp. 523–535.

[BGL18] Pierre Bonami, Oktay Günlük, and Jeff Linderoth. "Globally solving nonconvex quadratic programming problems with box constraints via integer programming methods". In: *Mathematical Programming Computation* 10.3 (2018), pp. 333–382.

[BCH92] Endre Boros, Yves Crama, and Peter L. Hammer. "Chvátal cuts and odd cycle inequalities in quadratic 0–1 optimization". In: *SIAM Journal on Discrete Mathematics* 5.2 (1992), pp. 163–177.

[BL09] Samuel Burer and Adam N. Letchford. "On nonconvex quadratic programming with box constraints". In: *SIAM Journal on Optimization* 20.2 (2009), pp. 1073–1089.

[CCZ10] Michele Conforti, Gérard Cornuéjols, and Giacomo Zambelli. "Extended formulations in combinatorial optimization". In: *4OR* 8.1 (2010), pp. 1–48.

[DG15] Santanu S Dey and Akshay Gupte. "Analysis of MILP techniques for the pooling problem". In: *Operations Research* 63.2 (2015), pp. 412–427.

[DL97] Michel Marie Deza and Monique Laurent. *Geometry of cuts and metrics*. Vol. 15. Algorithms and Combinatorics. Springer, 1997.

[Gup+13] Akshay Gupte, Shabbir Ahmed, Myun S. Cheon, and Santanu S Dey. "Solving mixed integer bilinear problems using MILP formulations". In: *SIAM Journal on Optimization* 23.2 (2013), pp. 721–744.

[Gup+17] Akshay Gupte, Shabbir Ahmed, Santanu S. Dey, and Myun Seok Cheon. "Relaxations and discretizations for the pooling problem". In: *Journal of Global Optimization* 67.3 (2017), pp. 631–669.

[Hal71] Rudolf Halin. "Studies on minimally n-connected graphs". In: *Combinatorial Mathematics and its Applications (Proc. Conf., Oxford, 1969)*. 1971, pp. 129–136.

[LS14] Adam N. Letchford and Michael M. Sørensen. "A new separation algorithm for the Boolean quadric and cut polytopes". In: *Discrete Optimization* 14 (2014), pp. 61–71.

[LS13] Marco Locatelli and Fabio Schoen. *Global optimization: theory, algorithms, and applications*. SIAM, 2013.

[LNL12] James Luedtke, Mahdi Namazifar, and Jeffrey T. Linderoth. "Some results on the strength of relaxations of multilinear functions". In: *Mathematical Programming* 136.2 (2012), pp. 325–351.

[McC76] Garth P. McCormick. "Computability of global solutions to factorable nonconvex programs: Part I – Convex underestimating problems". In: *Mathematical Programming* 10.1 (1976), pp. 147–175.

[Mic18] Carla Michini. "Forbidden minors for tight cycle relaxations". Optimization Online Preprint 5483. optimization-online:5483. 2018.

[MSF15] Ruth Misener, James B. Smadbeck, and Christodoulos A. Floudas. "Dynamically generated cutting planes for mixed-integer quadratically constrained quadratic programs and their incorporation into GloMIQO 2". In: *Optimization Methods and Software* 30.1 (2015), pp. 215–249.

[Pad89] Manfred W. Padberg. "The boolean quadric polytope: Some characteristics, facets and relatives". In: *Mathematical Programming* 45.1–3 (1989), pp. 139–172.

[Rik97] Anatoliy D. Rikun. "A Convex Envelope Formula for Multilinear Functions". In: *Journal of Global Optimization* 10.4 (1997), pp. 425–437.

[She97] Hanif D. Sherali. "Convex envelopes of multilinear functions over a unit hypercube and over special discrete sets". In: *Acta Mathematica Vietnamica* 22.1 (1997), pp. 245–270.

[SA90] Hanif D. Sherali and Warren P. Adams. "A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems". In: *SIAM Journal on Discrete Mathematics* 3.3 (1990), pp. 411–430.

[Sim90] Caterina De Simone. "The cut polytope and the Boolean quadric polytope". In: *Discrete Mathematics* 79.1 (1990), pp. 71–75.

[TRX12] Mohit Tawarmalani, Jean-Philippe P. Richard, and Chuanhui Xiong. "Explicit convex and concave envelopes through polyhedral subdivisions". In: *Mathematical Programming* 138.1 (2012), pp. 531–577.

[Zuc04] Mark Zuckerberg. "A Set Theoretic Approach to Lifting Procedures for 0, 1 Integer Programming". PhD thesis. Columbia University, 2004.

[Zuc16] Mark Zuckerberg. "Geometric proofs for convex hull defining formulations". In: *Operations Research Letters* 44.5 (2016), pp. 625–629.

APPENDIX A. MINIMALITY PROOF FOR THE POLYTOPE IN THEOREM 1

We want to show that each of the inequalities listed in Theorem 1 is necessary in the sense that omitting it leads to a polytope P' with $\pi[f](P') \supsetneq X(f)$.

A.1. **The McCormick inequalities $y_{ij} \leqslant x_i$.** Let P' be the polytope obtained from P by omitting the inequality $y_{i^*j^*} \leqslant x_{i^*}$, and consider the point $\boldsymbol{x}$ with $x_{i^*} = 0$ and $x_i = 1$ for all $i \in V \setminus \{i^*\}$. Then $(\boldsymbol{x}, z) \in X(f)$ if and only if

$$z = \left|E(V \setminus \{i^*\})\right| = \begin{cases} \binom{n-1}{2} - 1 & \text{if } i^* \in V \setminus \{n-1, n\}, \\ \binom{n-1}{2} & \text{if } i^* \in \{n-1, n\}. \end{cases}$$

If $\{i^*, j^*\} \neq \{n-1, n\}$ then we obtain a point $(\boldsymbol{x}, \boldsymbol{y}) \in P'$ by setting $y_{ij} = 1$ for all $ij \in E(V \setminus \{i^*\})$, $y_{i^*j^*} = 1$ and $y_{i^*j} = 0$ for $j \neq j^*$. Then $\pi[f](\boldsymbol{x}, \boldsymbol{y}) \notin X(f)$ because $y(E) = \binom{n-1}{2} + 1$. For $i^* = n-1$, $j^* = n$, we obtain $(\boldsymbol{x}, \boldsymbol{y}) \in P'$ by setting $y_{in} = 1$ and $y_{i,n-1} = 0$ for all $i \in V \setminus \{n-1, n\}$, $y_{n-1,n} = 1$, and $y_{ij} = 1 - 1/\binom{n-2}{2}$ for all $ij \in E(V \setminus \{n-1, n\})$. Then $\pi[f](\boldsymbol{x}, \boldsymbol{y}) \notin X(f)$ because

$$\sum_{ij \in E} y_{ij} = \binom{n-2}{2} - 1 + (n-2) = \binom{n-1}{2} - 1.$$

A.2. **The inequalities (9).** Without loss of generality, $i = 1$. We obtain a point $(\boldsymbol{x}, \boldsymbol{y}) \in P'$ by setting

$$\begin{aligned} x_{n-1} = x_n = y_{n-1,n} &= 1/2, & x_1 = x_2 = y_{12} &= 5/6, & x_3 = x_4 = \dots = x_{n-2} = 0, \\ y_{1,n-1} = y_{1n} &= 0, & y_{2,n-1} = y_{2n} &= 1/2, \end{aligned}$$

and $y_{ij} = 0$ for all remaining $ij \in E$. Then $\pi[f](\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, 11/6) \notin X(f)$, because $(\boldsymbol{x}, z) \in X(f)$ implies

$$\begin{aligned} z \geqslant (x_1 + x_2 - 1) + (x_1 + x_{n-1} - 1) + (x_1 + x_n - 1) + (x_2 + x_{n-1} - 1) + (x_2 + x_n - 1) \\ = 2/3 + 4(1/3) = 2. \end{aligned}$$

A.3. **The inequalities (10).** Let $t \in [n-2]$, and consider the point $(\boldsymbol{x}, \boldsymbol{y})$ given by

$$\begin{aligned} & x_{n-1} = x_n = y_{n-1,n} = 1/2, \\ & x_i = \begin{cases} 1 & \text{for } 1 \leqslant i \leqslant t, \\ 0 & \text{for } t+1 \leqslant i \leqslant n-2, \end{cases} \\ & y_{i,n-1} = y_{in} = \begin{cases} 1/2 & \text{for } 1 \leqslant i \leqslant t, \\ 0 & \text{for } t+1 \leqslant i \leqslant n-2, \end{cases} \\ & y_{ij} = 1 - \frac{2}{t(t-1)} \qquad \text{for } 1 \leqslant i < j \leqslant t, \end{aligned}$$

and $y_{ij} = 0$ for all remaining $ij \in E$. Then $(\boldsymbol{x}, \boldsymbol{y}) \in P'$, where P' is the polytope obtained from P by omitting (10) for $s = t$, and $\pi[f](\boldsymbol{x}, \boldsymbol{y}) = (\boldsymbol{x}, \binom{s+1}{2} - 1) \notin X(f)$, because $(\boldsymbol{x}, z) \in X(f)$ implies

$$z \geqslant \binom{s}{2} + s = \binom{s+1}{2}.$$

A.4. **The inequalities (11).** The following lemma shows that (11) cannot be omitted for any $s < (n-1)/2$.

Lemma 4. *Let t be an integer with $1 \leqslant t < (n-1)/2$, and consider the point $(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^{n(n+1)/2}$ given by $x_{n-1} = 1$, $x_n = y_{n-1,n} = 0$,*

$$x_i = \frac{t-1/2}{n-2} \qquad \text{for } 1 \leqslant i \leqslant n-2,$$
$$y_{i,n-1} = y_{in} = 0 \qquad \text{for } i \leqslant n-2,$$
$$y_{ij} = \frac{t^2-1}{(n-2)(n-3)} \qquad \text{for } 1 \leqslant i < j \leqslant n-2.$$

Then $(\boldsymbol{x}, \boldsymbol{y})$ satisfies all the constraints describing P except (11) for $s = t$. Moreover,

$$\pi[f](\boldsymbol{x}, \boldsymbol{y}) = \Big(\boldsymbol{x},\, (t^2-1)/2\Big) \notin X(f).$$

Proof. All y-variables are non-negative so $y(E) \geqslant 0$. For the inequalities $y_{ij} \leqslant \min\{x_i, x_j\}$, we just need to check that

$$\frac{t^2-1}{(n-2)(n-3)} \leqslant \frac{t-1/2}{n-2},$$

or equivalently, $\phi(t) \leqslant 0$ where ϕ is the quadratic function $\phi(t) = (t^2-1) - (t-1/2)(n-3)$. This follows from $\phi(1) = (3-n)/2 < 0$ and $\phi((n-1)/2) = -(n-3)(n-5)/4 \leqslant 0$.

For (9), we use the assumption $2t < n-1$:

$$2x_i + x_{n-1} + x_n - y_{i,n-1} - y_{in} = \frac{2t-1}{n-2} + 1 < 2.$$

For (10), we use

$$x(V \setminus \{n-1, n\}) + \frac{x_{n-1} + x_n}{2} = t,$$
$$y(E(V \setminus \{n-1, n\})) + \frac{1}{2}\sum_{i=1}^{n-2}\big(y_{i,n-1} + y_{in}\big) = \frac{t^2-1}{2}.$$

Now (10) follows from

$$st - \frac{t^2-1}{2} - \binom{s+1}{2} = -\frac{1}{2}\left[\left(s - \frac{2t-1}{2}\right)^2 + t - \frac{5}{4}\right] \leqslant 0$$

for all integers s.

For (11), we we have

$$sx(V) - y(E) - y_{n-1,n} - \binom{s+1}{2} = s\left(t + \frac{1}{2}\right) - \frac{t^2-1}{2} - \binom{s+1}{2}$$
$$= -\frac{1}{2}(s-t+1)(s-t-1) \begin{cases} \leqslant 0 & \text{for } s \neq t, \\ = 1/2 & \text{for } s = t. \end{cases}$$

Finally, $\big(\boldsymbol{x},\, (t^2-1)/2\big) \notin X(f)$ because

$$tx(V) - z - x_n - \binom{t+1}{2} = t\left(t + \frac{1}{2}\right) - \frac{t^2-1}{2} - \binom{t+1}{2} = \frac{1}{2}$$

while

$$tx(V) - z - x_n - \binom{t+1}{2} \leqslant 0$$

is a valid inequality for $X(f)$. □

To conclude the proof, the next lemma shows that (11) cannot be omitted for any $s \geqslant (n-1)/2$.

Lemma 5. *Let t be an integer with $(n-1)/2 \leqslant t \leqslant n-2$, and consider the point $(\boldsymbol{x},\boldsymbol{y}) \in \mathbb{R}^{n(n+1)/2}$ given by $x_{n-1}=1$, $x_n = y_{n-1,n} = 1/2$,*

$$\begin{aligned} x_i &= \frac{t-1}{n-2} && \text{for } 1 \leqslant i \leqslant n-2,\\ y_{i,n-1} &= \frac{4t-n-2}{3(n-2)} \text{ and } y_{i,n} = \frac{4t-n-2}{6(n-2)} && \text{for } i \leqslant n-2,\\ y_{ij} &= \frac{2t^2-8t+2n+1}{2(n-2)(n-3)} && \text{for } 1 \leqslant i < j \leqslant n-2. \end{aligned}$$

Then $(\boldsymbol{x},\boldsymbol{y})$ satisfies all the constraints describing P except (11) for $s=t$. Moreover,

$$\pi[f](\boldsymbol{x},\boldsymbol{y}) = \left(\boldsymbol{x},\, (2t^2-3)/4\right) \notin X(f).$$

Proof. All y-variables are non-negative so $y(E) \geqslant 0$. For the inequalities $y_{ij} \leqslant \min\{x_i, x_j\}$, we need to check that

$$\frac{2t^2-8t+2n+1}{2(n-2)(n-3)} \leqslant \frac{t-1}{n-2}, \qquad \frac{4t-n-2}{3(n-2)} \leqslant \frac{t-1}{n-2}.$$

The first inequality is equivalent to $\phi(t) \leqslant 0$ where $\phi(t) = 2t^2-8t+2n+1-2(n-3)(t-1)$, and this follows from

$$\phi\left(\frac{n-1}{2}\right) = -\frac{1}{2}(n^2-6n+7) \leqslant 0, \qquad \phi(n-2) = 7-2n \leqslant 0.$$

The second inequality follows from $4t-n-2-3(t-1) = t+1-n < 0$.

For (9), we have

$$2x_i + x_{n-1} + x_n - y_{i,n-1} - y_{in} = \frac{2t-2}{n-2} + \frac{3}{2} - \frac{4t-n-2}{2(n-2)} = 2.$$

For (10), we use

$$\begin{aligned} x(V\setminus\{n-1,n\}) + \frac{x_{n-1}+x_n}{2} &= t - \frac{1}{4},\\ y(E(V\setminus\{n-1,n\})) + \frac{1}{2}\sum_{i=1}^{n-2}\left(y_{i,n-1}+y_{in}\right) &= \frac{2t^2-8t+2n+1}{4} + \frac{4t-n-2}{4}\\ &= \frac{2t^2-4t+n-1}{4}. \end{aligned}$$

Now (10) follows from

$$\begin{aligned} s\left(t-\frac{1}{4}\right) - \frac{2t^2-4t+n-1}{4} - \binom{s+1}{2} &= -\frac{1}{2}\left[\left(s-\frac{4t-3}{4}\right)^2 + \frac{n-t}{2} - \frac{17}{16}\right]\\ &\leqslant -\frac{1}{2}\left[\left(s-\frac{4t-3}{4}\right)^2 - \frac{1}{16}\right] \leqslant 0. \end{aligned}$$

for all integers s.

For (11), we have

$$y(E) + y_{n,n-1} = \frac{2t^2-8t+2n+1}{4} + \frac{4t-n-2}{2} + \frac{1}{2} = \frac{2t^2-1}{4},$$

and then

$$\begin{aligned} sx(V) - y(E) - y_{n-1,n} - \binom{s+1}{2} &= s\left(t+\frac{1}{2}\right) - \frac{2t^2-1}{4} - \binom{s+1}{2}\\ &= -\frac{1}{2}\left[(s-t)^2 - \frac{1}{2}\right] \begin{cases} \leqslant 0 & \text{for } s \neq t,\\ = 1/4 & \text{for } s = t. \end{cases} \end{aligned}$$

Finally, $\big(\boldsymbol{x},\,(2t^2-3)/4\big)\notin X(f)$ because

$$tx(V)-z-x_n-\binom{t+1}{2}=t\left(t+\frac{1}{2}\right)-\frac{2t^2-3}{4}-\binom{t+1}{2}=\frac{3}{4}$$

while

$$tx(V)-z-x_n-\binom{t+1}{2}\leqslant 0$$

is a valid inequality for $X(f)$. □

Appendix B. Detailed case analysis in the proof of Lemma 2

Case 1: $n-1,n\in E^+$. We may assume $\mu(X_1\cap X_{n-1})<x_n<x_1+x_{n-1}-\mu(X_1\cap X_{n-1})$ since otherwise $\delta_n=0$.

Case 1.1: $x_{n-1}\leqslant x_1$. In this case $\delta_{n-1}=\mu(X_{n-1}\setminus X_1)$.

If $x_n\leqslant x_{n-1}$ then

$$\begin{aligned}\delta_n&=\mu(X_n\setminus X_1)=\delta_{n-1}-(x_{n-1}-x_n)\\&=\gamma-\mu_{n-1}-\mu_n+x_n+\min\{x_{n-1},\,x_1\}-(x_{n-1}-x_n)\\&=\gamma-x_n-x_n+x_n+x_{n-1}-(x_{n-1}-x_n)=\gamma.\end{aligned}$$

If $x_{n-1}\leqslant x_n\leqslant x_1$ then

$$\begin{aligned}\delta_n&=\mu(X_n\setminus X_1)=\delta_{n-1}\\&=\gamma-\mu_{n-1}-\mu_n+x_n+\min\{x_{n-1},\,x_1\}\\&=\gamma-x_{n-1}-x_n+x_n+x_{n-1}=\gamma.\end{aligned}$$

Finally, if $x_n\geqslant x_1$ then

$$\begin{aligned}\delta_n&=\mu(X_1\setminus X_n)=\delta_{n-1}-(x_n-x_1)\\&=\gamma-\mu_{n-1}-\mu_n+x_n+\min\{x_{n-1},\,x_1\}-(x_n-x_1)\\&=\gamma-x_{n-1}-x_1+x_n+x_{n-1}-(x_n-x_1)=\gamma.\end{aligned}$$

Case 1.2: $x_{n-1}>x_1$. In this case $\delta_{n-1}=\mu(X_1\setminus X_{n-1})$.

If $x_n\leqslant x_1$ then

$$\begin{aligned}\delta_n&=\mu(X_n\setminus X_1)=\delta_{n-1}-(x_1-x_n)\\&=\gamma-\mu_{n-1}-\mu_n+x_n+\min\{x_{n-1},\,x_1\}-(x_1-x_n)\\&=\gamma-x_n-x_n+x_n+x_1-(x_1-x_n)=\gamma.\end{aligned}$$

If $x_1\leqslant x_n\leqslant x_{n-1}$ then

$$\begin{aligned}\delta_n&=\mu(X_1\setminus X_n)=\delta_{n-1}\\&=\gamma-\mu_{n-1}-\mu_n+x_n+\min\{x_{n-1},\,x_1\}\\&=\gamma-x_n-x_1+x_n+x_1=\gamma.\end{aligned}$$

Finally, if $x_n\geqslant x_{n-1}$ then

$$\begin{aligned}\delta_n&=\mu(X_1\setminus X_n)=\delta_{n-1}-(x_n-x_{n-1})\\&=\gamma-\mu_{n-1}-\mu_n+x_n+\min\{x_{n-1},\,x_1\}-(x_n-x_{n-1})\\&=\gamma-x_{n-1}-x_1+x_n+x_1-(x_n-x_{n-1})=\gamma.\end{aligned}$$

Case 2: $n-1,n\in E^-$. We may assume $1-\mu(X_1\cup X_{n-1})<x_n<1-\mu(X_1\cap X_{n-1})$ since otherwise $\delta_n=0$.

Case 2.1: $x_{n-1}\leqslant x_1$. In this case $\delta_{n-1}=\mu(X_{n-1}\setminus X_1)=\mu(X_1\cup X_{n-1})-x_1$.

If $x_n\leqslant 1-x_1$ then

$$\begin{aligned}\delta_n&=\mu(X_n\cap X_1)=x_n-(1-\mu(X_1\cup X_{n-1}))=\delta_{n-1}-(1-x_1-x_n)\\&=\gamma+\eta_{n-1}+\eta_n-x_n-x_{n-1}-x_1+\min\{x_{n-1},\,x_1\}+1-(1-x_1-x_n)\\&=\gamma+0+0-x_n-x_{n-1}-x_1+x_{n-1}+1-(1-x_1-x_n)=\gamma.\end{aligned}$$

If $1-x_1 \leqslant x_n \leqslant 1-x_{n-1}$ then

$$\begin{aligned}\delta_n &= \mu([0,1]\setminus(X_n\cup X_1)) = \mu(X_{n-1}\setminus X_1) = \delta_{n-1}\\ &= \gamma+\eta_{n-1}+\eta_n-x_n-x_{n-1}-x_1+\min\{x_{n-1},\,x_1\}+1\\ &= \gamma+0+(x_n+x_1-1)-x_n-x_{n-1}-x_1+x_{n-1}+1=\gamma.\end{aligned}$$

Finally, if $x_n \geqslant 1-x_{n-1}$ then

$$\begin{aligned}\delta_n &= \mu([0,1]\setminus(X_n\cup X_1)) = \delta_{n-1}-(x_n+x_{n-1}-1)\\ &= \gamma+\eta_{n-1}+\eta_n-x_n-x_{n-1}-x_1+\min\{x_{n-1},\,x_1\}+1-\eta_{n-1}\\ &= \gamma+\eta_{n-1}+(x_n+x_1-1)-x_n-x_{n-1}-x_1+x_{n-1}+1-\eta_{n-1}=\gamma.\end{aligned}$$

Case 2.2: $x_{n-1} > x_1$. In this case $\delta_{n-1} = \mu(X_1\setminus X_{n-1}) = \mu(X_1\cup X_{n-1}) - x_{n-1}$.
If $x_n \leqslant 1-x_{n-1}$ then

$$\begin{aligned}\delta_n &= \mu(X_n\cap X_1) = x_n-(1-\mu(X_1\cup X_{n-1})) = \delta_{n-1}-(1-x_n-x_{n-1})\\ &= \gamma+\eta_{n-1}+\eta_n-x_n-x_{n-1}-x_1+\min\{x_{n-1},\,x_1\}+1-(1-x_n-x_{n-1})\\ &= \gamma+0+0-x_n-x_{n-1}-x_1+x_1+1-(1-x_n-x_{n-1})=\gamma.\end{aligned}$$

If $1-x_{n-1} \leqslant x_n \leqslant 1-x_1$ then

$$\begin{aligned}\delta_n &= \mu(X_n\cap X_1) = \mu(X_1\setminus X_{n-1}) = \delta_{n-1}\\ &= \gamma+\eta_{n-1}+\eta_n-x_n-x_{n-1}-x_1+\min\{x_{n-1},\,x_1\}+1\\ &= \gamma+(x_{n-1}+x_n-1)+0-x_n-x_{n-1}-x_1+x_1+1=\gamma.\end{aligned}$$

Finally, if $x_n \geqslant 1-x_1$ then

$$\begin{aligned}\delta_n &= \mu([0,1]\setminus(X_n\cup X_1)) = \delta_{n-1}-(x_n+x_1-1)\\ &= \gamma+\eta_{n-1}+\eta_n-x_n-x_{n-1}-x_1+\min\{x_{n-1},\,x_1\}+1-\eta_n\\ &= \gamma+(x_{n-1}+x_n-1)+\eta_n-x_n-x_{n-1}-x_1+x_1+1-\eta_n=\gamma.\end{aligned}$$

Case 3: $n-1\in E^+$, $n\in E^-$. We may assume $\mu(X_{n-1}\setminus X_1) < x_n < 1-\mu(X_1\setminus X_{n-1})$ since otherwise $\delta_n = 0$.

Case 3.1: $x_1+x_{n-1}\leqslant 1$. In this case $\delta_{n-1} = \mu(X_1\cap X_{n-1})$.
If $x_n \leqslant x_{n-1}$ then

$$\begin{aligned}\delta_n &= \delta_{n-1}-(x_{n-1}-x_n)\\ &= \gamma-\mu_{n-1}+\eta_n+x_{n-1}-\max\{0,\,x_{n-1}+x_1-1\}-(x_{n-1}-x_n)\\ &= \gamma-x_n+0+x_{n-1}-0-(x_{n-1}-x_n)=\gamma.\end{aligned}$$

If $x_{n-1}\leqslant x_n \leqslant 1-x_1$ then

$$\begin{aligned}\delta_n &= \delta_{n-1}\\ &= \gamma-\mu_{n-1}+\eta_n+x_{n-1}-\max\{0,\,x_{n-1}+x_1-1\}\\ &= \gamma-x_{n-1}+0+x_{n-1}-0=\gamma.\end{aligned}$$

Finally, if $x_n\geqslant 1-x_1$ then

$$\begin{aligned}\delta_n &= \delta_{n-1}-(x_1+x_n-1)\\ &= \gamma-\mu_{n-1}+\eta_n+x_{n-1}-\max\{0,\,x_{n-1}+x_1-1\}-(x_1+x_n-1)\\ &= \gamma-x_{n-1}+(x_1+x_n-1)+x_{n-1}-0-(x_1+x_n-1)=\gamma.\end{aligned}$$

Case 3.2: $x_1+x_{n-1} > 1$. In this case $\delta_{n-1} = \mu([0,1]\setminus(X_1\cup X_{n-1}))$.
If $x_n\leqslant 1-x_1$ then

$$\begin{aligned}\delta_n &= \delta_{n-1}-(1-x_1-x_n)\\ &= \gamma-\mu_{n-1}+\eta_n+x_{n-1}-\max\{0,\,x_{n-1}+x_1-1\}-(1-x_1-x_n)\\ &= \gamma-x_n+0+x_{n-1}-(x_{n-1}+x_1-1)-(1-x_1-x_n)=\gamma.\end{aligned}$$

If $1 - x_1 \leqslant x_n \leqslant x_{n-1}$ then

$$\begin{aligned}\delta_n &= \delta_{n-1}\\ &= \gamma - \mu_{n-1} + \eta_n + x_{n-1} - \max\{0,\, x_{n-1} + x_1 - 1\}\\ &= \gamma - x_n + (x_n + x_1 - 1) + x_{n-1} - (x_{n-1} + x_1 - 1) = \gamma.\end{aligned}$$

Finally, if $x_n \geqslant x_{n-1}$ then

$$\begin{aligned}\delta_n &= \delta_{n-1} - (x_n - x_{n-1})\\ &= \gamma - \mu_{n-1} + \eta_n + x_{n-1} - \max\{0,\, x_{n-1} + x_1 - 1\} - (x_n - x_{n-1})\\ &= \gamma - x_{n-1} + (x_n + x_1 - 1) + x_{n-1} - (x_{n-1} + x_1 - 1) - (x_n - x_{n-1}) = \gamma.\end{aligned}$$

Case 4: $n - 1 \in E^-$, $n \in E^+$. We may assume $\mu(X_{n-1} \setminus X_1) < x_n < 1 - \mu(X_{n-1} \setminus X_1)$ since otherwise $\delta_n = 0$.

Case 4.1: $x_1 + x_{n-1} \leqslant 1$. In this case $\delta_{n-1} = \mu(X_1 \cap X_{n-1})$.

If $x_n \leqslant x_1$ then

$$\begin{aligned}\delta_n &= \delta_{n-1} - (x_1 - x_n)\\ &= \gamma + \eta_{n-1} - \mu_n + x_1 - \max\{0,\, x_{n-1} + x_1 - 1\} - (x_1 - x_n)\\ &= \gamma + 0 - x_n + x_1 - 0 - (x_1 - x_n) = \gamma.\end{aligned}$$

If $x_1 \leqslant x_n \leqslant 1 - x_{n-1}$ then

$$\begin{aligned}\delta_n &= \delta_{n-1}\\ &= \gamma + \eta_{n-1} - \mu_n + x_1 - \max\{0,\, x_{n-1} + x_1 - 1\}\\ &= \gamma + 0 - x_1 + x_1 - 0 = \gamma.\end{aligned}$$

Finally, if $x_n \geqslant 1 - x_{n-1}$ then

$$\begin{aligned}\delta_n &= \delta_{n-1} - (x_{n-1} + x_n - 1)\\ &= \gamma + \eta_{n-1} - \mu_n + x_1 - \max\{0,\, x_{n-1} + x_1 - 1\} - (x_{n-1} + x_n - 1)\\ &= \gamma + (x_{n-1} + x_n - 1) - x_1 + x_1 - 0 - (x_{n-1} + x_n - 1) = \gamma.\end{aligned}$$

Case 4.2: $x_1 + x_{n-1} > 1$. In this case $\delta_{n-1} = \mu([0,1] \setminus (X_1 \cup X_{n-1}))$.

If $x_n \leqslant 1 - x_{n-1}$ then

$$\begin{aligned}\delta_n &= \delta_{n-1} - (1 - x_{n-1} - x_n)\\ &= \gamma + \eta_{n-1} - \mu_n + x_1 - \max\{0,\, x_{n-1} + x_1 - 1\} - (1 - x_{n-1} - x_n)\\ &= \gamma + 0 - x_n + x_1 - (x_{n-1} + x_1 - 1) - (1 - x_{n-1} - x_n) = \gamma.\end{aligned}$$

If $1 - x_{n-1} \leqslant x_n \leqslant x_1$ then

$$\begin{aligned}\delta_n &= \delta_{n-1}\\ &= \gamma + \eta_{n-1} - \mu_n + x_1 - \max\{0,\, x_{n-1} + x_1 - 1\}\\ &= \gamma + (x_{n-1} + x_n - 1) - x_n + x_1 - (x_{n-1} + x_1 - 1) = \gamma.\end{aligned}$$

Finally, if $x_n \geqslant x_1$ then

$$\begin{aligned}\delta_n &= \delta_{n-1} - (x_n - x_1)\\ &= \gamma + \eta_{n-1} - \mu_n + x_1 - \max\{0,\, x_{n-1} + x_1 - 1\} - (x_n - x_1)\\ &= \gamma + (x_{n-1} + x_n - 1) - x_1 + x_1 - (x_{n-1} + x_1 - 1) - (x_n - x_1) = \gamma.\end{aligned}$$