AMSI VACATION RESEARCH SCHOLARSHIPS 2018-2019

# Convex Hulls of Graphs of Bilinear Functions

Mitchell Harris

Supervised by Dr. Thomas Kalinowski

University of New England

Vacation Research Scholarships are funded jointly by the Department of Education and Training and the Australian Mathematical Sciences Institute.

AMSI AUSTRALIAN MATHEMATICAL SCIENCES INSTITUTE

Australian Government Department of Education and Training

**Abstract**

Convex envelopes of non-convex functions play an important role in global optimisation. The project aims to find minimal characterisations of convex hulls of graphs of bilinear functions. The structure of a bilinear function can be described in terms of a graph with vertices corresponding to the variables and edges corresponding to the product terms. A new geometric method that has been developed recently can in special cases characterise the convex hull. We outline this method, and describe our progress towards a proof for a class of graphs known as wheels.

# Contents

# 1 Introduction

## 1.1 Optimisation

*Mathematical optimisation* concerns the theory and algorithms for computing maxima and minima of functions. Standardly we are interested in the minima of some function $f : \mathbb{R}^n \to \mathbb{R}$ over a feasible subset of $\mathbb{R}^n$. If the function and feasible region are convex, then any local minimum is also a global minimum. Moreover, there are efficient algorithms for solving such problems.

## 1.2 Global Optimisation

Considerably more difficult is the case where the functions are *not* convex. Local methods find solutions which are optimal in some subset of the feasible region. An example is *gradient descent*. The gradient descent algorithm follows the negative gradient of a differentiable objective function. If there is more than one local minimum, then the output is strongly dependent on the initialisation. In some applications, these local minima might be sufficient, but the aim of *global optimisation* is to find truly optimal solutions to non-convex problems. One strategy is to approximate by a convex underestimating function. The best such approximation comes from the *convex hull* of the graph of the function. This is a critical ingredient in all state-of-the-art global solvers.

# 2 Problem Formulation

## 2.1 Notation

We are interested in bilinear functions $f : [0, 1]^n \to \mathbb{R}$ on the unit $n$-cube of the form

$$f(\boldsymbol{x}) = \sum_{1 \leq i < j \leq n} a_{ij} x_i x_j$$

We want to compute the convex hull of the graph of $f$, which we denote

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

. By a result of Rikun [6] this hull is determined by the vertices.

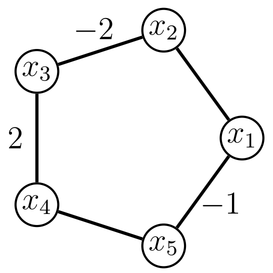
$$X(f) = \text{conv}\left\{(\boldsymbol{x}, z) : x \in \{0,1\}^n, z = f(\boldsymbol{x})\right\}$$

. This means that $X(f)$ is a polytope. In order to do computations, we must describe $X(f)$ in terms of the linear inequalities that determine it.

To this end, it will be helpful to introduce some graph theoretic notation. To the bilinear function $f$, we associate an edge weighted graph $G = (V, E)$, where $V = \{1, \ldots, n\}$. We have $ij \in E$ with weight $a_{ij}$ if and only if $a_{ij} \neq 0$. Let $m := |E|$. We call $G$ the *support graph* of $f$ and now express $f$ as

$$f(\boldsymbol{x}) = \sum_{ij \in E} a_{ij} x_i x_j \tag{1}$$

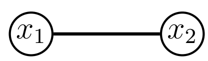
For example if $f(\boldsymbol{x}) = x_1x_2 - 2x_2x_3 + 2x_3x_4 + x_4x_5 - x_1x_5$:




An unlabelled edge is understood to have weight 1.

## The Simplest Example

Consider $f(x_1, x_2) = x_1x_2$. It has support graph:




Using Rikun $X(f)$ may be written

$$X(f) = \text{conv}\left\{\begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\0\\0\end{pmatrix}, \begin{pmatrix}0\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\1\end{pmatrix}\right\}$$

It is easy to see that these points satisfy the four linearly independent linear inequalities

$$x_3 \geq 0 \qquad\qquad x_3 \leq x_1$$

$$x_3 \geq x_1 + x_2 - 1 \qquad\qquad x_3 \leq x_2$$

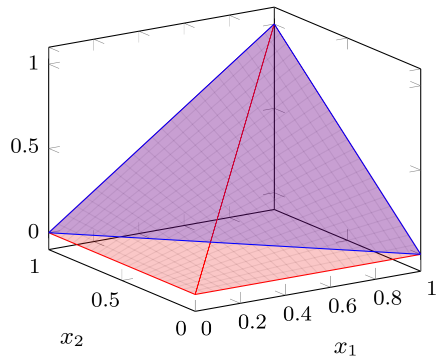



Figure 1: Convex hull of graph of $x_1x_2$.

Every point in the convex hull satisfies them as well. We now have

$$X(f)\left\{(x_1,x_2,z)\in[0,1]^2\times\mathbb{R}:\max\{0,x_1+x_2-1\}\leq z\leq\min\{x_1,x_2\}\right\}$$

These inequalities are called the *McCormick inequalities*, and each one yields a face of $X(f)$. See figure 2.1.

## 2.2 Extended Formulation

Our approach is to introduce an *extended formulation*. Often polytopes arising in combinatorial optimisation have prohibitively many facets, but can be written as the projection of a higher dimensional polytope with significantly fewer facets. See [2]. We will descrie our polytope in a lifted space in a way that allows us to use as few inequalities as possible. For each term $x_ix_j$ we introduce a new variable $y_{ij}$, and impose the McCormick inequalities:

$$y_{ij}\geq 0 \qquad y_{ij}\geq x_i+x_j-1 \qquad y_{ij}\leq x_i \qquad y_{ij}\leq x_j \tag{2}$$

Now let $M$ be the polytope

$$M:=\left\{(\boldsymbol{x},\boldsymbol{y})\in[0,1]^{n(n+1)/2}:(2)\text{ for all }1\leq i<j\leq n\right\}$$

consisting of all $(\boldsymbol{x},\boldsymbol{y})$ that satisfy the McCormick inequalities. The *McCormick relaxation* of a bilinear function $f$ is the polytope

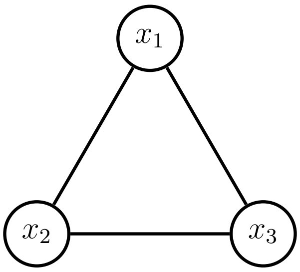
$$M(f)=\left\{(\boldsymbol{x},\sum_{ij\in E}a_{ij}y_{ij})\in[0,1]^n\times\mathbb{R}:(2)\text{ for all }ij\in E\right\}$$

Clearly, for the example in the previous section, we have:

$$X(f) = M(f)$$

THE TRIANGLE

Another example is the *triangle* $f(\boldsymbol{x}) = x_1x_2 + x_2x_3 + x_1x_3$ which has support graph:




Its vertex characterisation is

$$X(f) = \text{conv}\left\{ \begin{pmatrix} 0\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} 0\\1\\0\\0 \end{pmatrix}, \begin{pmatrix} 1\\0\\0\\0 \end{pmatrix}, \begin{pmatrix} 1\\1\\0\\1 \end{pmatrix}, \begin{pmatrix} 0\\0\\1\\0 \end{pmatrix}, \begin{pmatrix} 0\\1\\1\\0 \end{pmatrix}, \begin{pmatrix} 1\\0\\1\\0 \end{pmatrix}, \begin{pmatrix} 1\\1\\1\\3 \end{pmatrix} \right\}$$

Consider the point $(\boldsymbol{x}, \boldsymbol{y})$ with $x_i = 1/2$ and $y_{ij} = 0$ for all $i, j$. It is easy to see that it satisfies the McCormick inequalities, so that $(\boldsymbol{x}, \boldsymbol{y}) \in M(f)$. Clearly though, it is not in the convex hull $X(f)$, yielding the strict inclusion $X(f) \subset M(f)$. To completely describe $X(f)$ required one or more additional inequalities. In this case, it is enough to add the *triangle inequality*:

$$y_{12} + y_{23} + y_{13} \geq x_1 + x_2 + x_3 - 1$$

Notice that it indeed cuts off our problematic point:

$$0 + 0 + 0 < \frac{1}{2} + \frac{1}{2} + \frac{1}{2} - 1 = \frac{1}{2}$$

In general we have the following:

**Lemma 2.1.** $X(f) \subseteq M(f)$

*Proof.* First we show that the McCormick inequalities are valid for all $(\boldsymbol{x}, \boldsymbol{y}) \in \{0, 1\}^{n(n+1)/2}$ with $y_{ij} = x_ix_j$. The only non-trivial case is $y_{ij} \geq x_i + x_j - 1$. See that

$$\begin{aligned} x_i = x_j = 0 &\implies x_ix_j = 0 \geq 0 + 0 - 1 = -1 \\ x_i = 0, x_j = 1 &\implies x_ix_j = 0 \geq 0 + 1 - 1 = 0 \\ x_i = 1, x_j = 0 &\implies x_ix_j = 0 \geq 1 + 0 - 1 = 0 \\ x_i = 1, x_j = 1 &\implies x_ix_j = 1 \geq 1 + 1 - 1 = 1 \end{aligned}$$

This implies that (2) are valid for all $(\boldsymbol{x}, \boldsymbol{y}) \in \{(\boldsymbol{x}, \boldsymbol{y}) \in \{0,1\}^{n(n+1)/2} : y_{ij} = x_i x_j\}$. Since $x_i, x_j \in \{0,1\} \implies y_{ij} \in \{0,1\}$, (2) are valid for all $(\boldsymbol{x}, \boldsymbol{y}) \in \{(\boldsymbol{x}, \boldsymbol{y}) : \boldsymbol{x} \in \{0,1\}^n,\ y_{ij} = x_i x_j\}$. Moreover,

$$\begin{aligned} X(f) &= \operatorname{conv}\left\{(\boldsymbol{x}, \sum_{ij \in E} a_{ij} x_i x_j) : \boldsymbol{x} \in [0,1]^n\right\} \\ &= \operatorname{conv}\left\{(\boldsymbol{x}, \sum_{ij \in E} a_{ij} x_i x_j) : \boldsymbol{x} \in \{0,1\}^n\right\} \\ &= \operatorname{conv}\left\{(\boldsymbol{x}, \sum_{ij \in E} a_{ij} y_{ij}) : \boldsymbol{x} \in \{0,1\}^n, y_{ij} = x_i x_j\right\} \end{aligned}$$

which shows that (2) are valid in all of $X(f)$. Therefore $X(f) \subseteq M(f)$ as required. □

The general question is now: what inequalities do we need to add to the McCormick inequalities to characterise $X(f)$?

### The Projection Map

To make these notions explicit, we define a function from $(\boldsymbol{x}, \boldsymbol{y})$-space into $(\boldsymbol{x}, z)$-space by

$$\pi[f] : \mathbb{R}^n \times \mathbb{R}^{n(n-1)/2} \to \mathbb{R}^n \times \mathbb{R}, (\boldsymbol{x}, \boldsymbol{y}) \mapsto \left(\boldsymbol{x}, \sum_{ij \in E} a_{ij} y_{ij}\right)$$

For any bilinear function $f$, the goal is to find a polytope $P$ in our extended space, such that

$$X(f) = \pi[f](P)$$

### The Boolean Quadric Polytope

In [5] Padberg introduced the *Boolean Quadric Polytope*. It is the convex hull of the binary $x_i, x_j$ and $y_{ij}$ satisfying the McCormick inequalities

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

Padberg proved that all of the McCormick inequalities define non-trivial facets of $\mathsf{QP}^n$, and introduced several additional families of valid inequalities (including the triangle inequality from above). Given a bilinear function $f$ with support graph $G = (V, E)$ we define

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

Clearly, $\mathsf{QP}^n = \mathsf{QP}^n(K_n)$. In general, $\mathsf{QP}^n(G)$ is a projection of $\mathsf{QP}^n$ where we omit the $y_{ij}$ for which $ij \neq E$. The McCormick inequalities are exact at the vertices. That is, if $x_i, x_j \in \{0,1\}$, then (2) implies $y_{ij} = x_i x_j$. This implies that

$$X(f) = \pi[f](\mathsf{QP}^n(G))$$

Naïvely, we are done. $\mathsf{QP}^n$ is precisely a polytope as sought above. However, $\mathsf{QP}^n$ has exponentially many facets as $n$ grows, not all of which are known. This motivates us to search sufficient subsets of the known facets of $\mathsf{QP}^n(G)$ that are *enough*, in the sense that they describe a polytope $P$ with polynomially many facets, for which $\pi[f](P) = X(f)$. The known families of valid inequalities for the boolean quadric polytope are a good starting place for particular problems.

### Convex and Concave Envelopes

We denote the convex and concave envelopes of $f$ by

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

respectively (see A.4), so that $X(f)$ may be written as

$$X(f) = \{(\boldsymbol{x}, z) \in [0,1]^n \times \mathbb{R} : \operatorname{vex}[f](\boldsymbol{x}) \leq z \leq \operatorname{cav}[f](\boldsymbol{x})\} \tag{3}$$

Now let $P$ be a polytope. Define lower and upper bound functions $\mathrm{LB}_P[f]$ and $\mathrm{UB}_P[f]$ on $[0,1]^n$ by

$$\mathrm{LB}_P[f](\boldsymbol{x}) := \min\left\{\sum_{ij\in E} a_{ij} y_{ij} : (\boldsymbol{x}, \boldsymbol{y}) \in P\right\} = \min\{z : (\boldsymbol{x}, z) \in \pi[f](P)\}$$

$$\mathrm{UB}_P[f](\boldsymbol{x}) := \max\left\{\sum_{ij\in E} a_{ij} y_{ij} : (\boldsymbol{x}, \boldsymbol{y}) \in P\right\} = \max\{z : (\boldsymbol{x}, z) \in \pi[f](P)\}$$

It should be noted that for fixed $\boldsymbol{x}$, $\mathrm{LB}_P[f](\boldsymbol{x})$ and $\mathrm{UB}_P[f](\boldsymbol{x})$ are the optimal solutions to *linear programs*. Comparing $\mathrm{LB}_P[f]$ and $\mathrm{UB}_P[f]$ with $\operatorname{vex}[f]$ and $\operatorname{cav}[f]$ immediately proves the following.

**Theorem 2.2.** *Let $f(\boldsymbol{x}) = \sum_{ij\in E} a_{ij} x_i x_j$ be a bilinear function, and $P$ a polytope in $(\boldsymbol{x}, \boldsymbol{y})$-space. Then $X(f) = \pi[f](P)$ if and only if $\mathrm{LB}_P[f](\boldsymbol{x}) = \operatorname{vex}[f](\boldsymbol{x})$, and $\mathrm{UB}_P[f](\boldsymbol{x}) = \operatorname{cav}[f](\boldsymbol{x})$ for all $\boldsymbol{x} \in [0,1]^n$.*

So long as $\mathsf{QP}^n \subseteq P \subseteq M$, by Lemma 2.1 $X(f) \subseteq M(f)$, which means that $\mathrm{LB}_P[f](\boldsymbol{x}) \leq \mathrm{vex}[f](\boldsymbol{x})$ and $\mathrm{cav}[f](\boldsymbol{x}) \leq \mathrm{UB}_P[f](\boldsymbol{x})$. Therefore, in practice, we only need to show the reverse inequalities.

# 3 Method

## 3.1 Zuckerberg's Geometric Method

In [7, 8], Zuckerberg developed a geometric method for proving convex hull characterizations of subsets of $\{0,1\}^n$. Any set $\mathcal{F} \subset \{0,1\}^n$ has a representation as a finite combination of unions, intersections, and complements of the sets

$$A_i := \{\boldsymbol{\xi} \in \{0,1\}^n : \xi_i = 1\} \text{ for all } i \in [n]$$

where $[n] = \{1, \dots, n\}$. We write such a representation as $F(A_1, \dots, A_n)$. The main theorem of [8] states that for every $\boldsymbol{x} \in [0,1]^n$, we have $\boldsymbol{x} \in \mathrm{conv}(\mathcal{F})$ if and only if there exists:

1. A set $U$ and a collection $\mathcal{L}$ of subsets of $U$, such that $\emptyset \in \mathcal{L}$, $U \in \mathcal{L}$, and $\mathcal{L}$ is closed under complements and finite unions,
2. a function $\mu : \mathcal{L} \to \mathbb{R}$ with $\mu(U) = 1$, and $\mu\left(\bigcup_{i=1}^n L_i\right) = \sum_{i=1}^k \mu(L_i)$ for any pairwise disjoint collection $L_1, \dots, L_k \in \mathcal{L}$, and
3. sets $X_1, \dots, X_n \in \mathcal{L}$ such that $\mu(X_i) = x_i$ for all $i \in [n]$ and $\mu(F(X_1, \dots, X_n)) = 1$.

The power of this condition is that the choice of $U$, $\mathcal{L}$, and $\mu$ are up to us. We may effectively understand an $n$-dimensional problem in terms of subsets of a lower dimensional space. Indeed, the examples in [8] all exploit the properties of $\mathbb{R}$.

## 3.2 Refinement of Zuckerberg's Method

In [3] is a refinement and simplification of Zuckerberg's method. They showed that it is always enough to use subsets of the half-open unit interval $[0,1)$, and the Lebesgue measure.

**Theorem 3.1** (Theorem 4 in [3])**.** *Let $\mathcal{F} \subseteq \{0,1\}^n$, $\boldsymbol{x} \in [0,1]^n$. Let $U = [0,1)$, and $\mathcal{L}$ be the collection of unions of finitely many half-open intervals*

$$\mathcal{L} = \{[a_1, b_1) \cup \cdots \cup [a_k, b_k) : 0 \leq a_1 < b_1 < \cdots < a_k < b_k \leq 1, k \in \mathbb{N}\}$$

*and $\mu$ be the Lebesgue measure restricted to $\mathcal{L}$. Then $\boldsymbol{x} \in \operatorname{conv}(\mathcal{F})$ if and only if there exist sets $X_1, \ldots, X_n \in \mathcal{L}$ 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 $t \in [0, 1)$.*

We will illustrate the application of this theorem by returning to the simple example of $f(\boldsymbol{x}) = x_1 x_2$. Consider the point $(0.5, 0.4, 0.1)$. Put $\mathcal{F} := \left\{ \left(\begin{smallmatrix}0\\0\\0\end{smallmatrix}\right), \left(\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right), \left(\begin{smallmatrix}0\\1\\0\end{smallmatrix}\right), \left(\begin{smallmatrix}1\\1\\1\end{smallmatrix}\right) \right\}$ and recall that $X(f) = \operatorname{conv} \mathcal{F}$. We choose the sets $X_1 = [0, 0.5)$, $X_2 = [0.4, 0.8)$, and $X_3 = [0.4, 0.5)$ as in the following figure, where we use non-overlapping bricks to represent them.

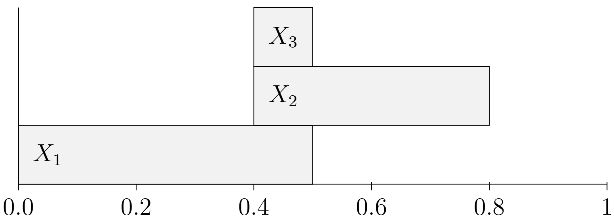



Figure 2: Choice of subsets

By construction, $X_i = x_i$ for each $i$, and every $t \in [0, 1)$ corresponds to al element of $\mathcal{F}$. The interval $[0, 0.4)$ intersects only $X_1$, and so is associated with the vector $\left(\begin{smallmatrix}1\\0\\0\end{smallmatrix}\right)$. The interval $[0.4, 0.5)$ intersects $X_1, X_2, X_3$, and so is associated with the vector $\left(\begin{smallmatrix}0\\1\\0\end{smallmatrix}\right)$. In the same manner $[0.5, 0.8)$ is associated with $\left(\begin{smallmatrix}0\\1\\0\end{smallmatrix}\right)$, and $[0.8, 1)$ with $\left(\begin{smallmatrix}0\\0\\0\end{smallmatrix}\right)$. That this choice is possible shows that $(0.5, 0.4, 0.1) \in \operatorname{conv} \mathcal{F}$.

### Generalisation to Graphs of Functions

In [3] is also a generalisation of Theorem 3.1 to the graphs of arbitrary functions $\psi : \{0, 1\}^n \to \mathbb{R}$. In particular, they prove the following characterisation for the graphs of bilinear functions.

**Theorem 3.2** (Corollary 1 in [3])**.** *Let $f(x)$ be a bilinear function, and $G$ be its support graph. Then*

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

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

This theorem has the following practical consequence. Let $f$ be a bilinear function and $G$ its support graph. By Theorem 2.2, if $P \subseteq \mathbb{R}^{n(n+1)/2}$ is a polytope with $P \subseteq X(f)$, then $\pi[f](P) = X(f)$ if and only if for every $\boldsymbol{x} \in [0,1]^n$, there exists $X_1, \dots, X_n \in \mathcal{L}$ and $X_1', \dots, X_n' \in \mathcal{L}$ with $\mu(X_i) = \mu(X_i') = x_i$ for all $i \in [n]$ and

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

Thus our problem has been reduced to the construction of sets $X_i, \dots, X_n$ such that pairwise intersections have certain measures. In the case that all edge weights are positive we also have:

**Corollary 3.3** (Corollary 2 in [3])**.** *If all edge weights are positive, then*

$$\mathrm{cav}[f](\boldsymbol{x}) = \sum_{ij \in E} a_{ij} \min\{x_i, x_j\}$$

This is enough, because $y_{ij} \leq x_i$ and $y_{ij} \leq x_j$ together imply $y_{ij} \leq \min\{x_i, x_j\}$ giving

$$\mathrm{cav}[f](\boldsymbol{x}) = \sum_{ij \in E} a_{ij} \min\{x_i, x_j\} \geq \sum_{ij \in E} a_{ij} y_{ij}$$

Therefore

$$\mathrm{cav}[f](\boldsymbol{x}) \geq \mathrm{UB}_P(\boldsymbol{x}) \text{ for all } \boldsymbol{x}$$

as required. It is generally harder to show that $\mathrm{vex}[f](\boldsymbol{x}) \leq \mathrm{LB}_P[f](\boldsymbol{x})$, as the choice of sets $X_1, \dots, X_n$ is not obvious. In the following section we will also assume that all edge weights are equal to 1. We summarise this situation with the following theorem.

**Theorem 3.4.** *Let $f$ be a bilinear function with support graph $G$ and unit coefficients. We have $\pi[f](P) = X(f)$ if and only if $P \supseteq \mathsf{QP}^n(G)$, and for every $\boldsymbol{x} \in [0,1]^n$ there exist sets $X_1, \dots, X_n \subseteq [0,1]$ such that*

1. *$\mu(X_i) = x_i$ for all $i \in [n]$, and*

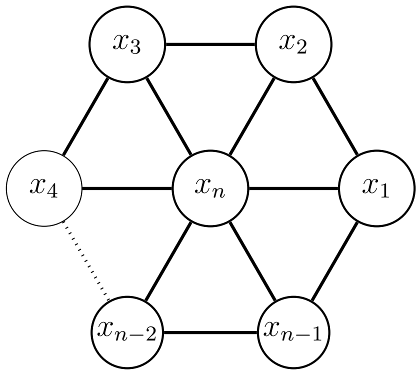
2. *$\mathrm{LB}_P(\boldsymbol{x}) \geq \sum_{ij \in E} \mu(X_i \cap X_i)$.*

# 4 Results

The contribution of this project was to make progress on the following special case. Let $G = W_{n-1}$ be the $(n-1)$-wheel, that is, the graph with vertex set $V = [n]$ and edge set $E = E_0 \cup E_1$

where

$$E_0 = \{ij : i \in [n-1],\, j \equiv i+1 \pmod{n-1}\}, \qquad E_1 = \{in : i \in [n-1]\}.$$




We are interested in the convex hull $X(f)$ when the support graph of $f$ is a wheel. An important observation is that each of the cycle edges is contained in precisely one triangle.

**Definition 4.1.** *The* triangle relaxation *of the* $\mathsf{QP}^n(G)$ *for a graph* $G = (V, E)$ *is the polytope in* $[0,1]^{|V|+|E|}$ *described by the McCormick inequalities for each* $ij \in E$ *and the triangle inequalities*

$$y_{ij} + y_{jk} + y_{jk} \geq x_i + x_j + x_k - 1$$

*for every triangle* $ijk$ *in* $G$.

We conjecture the following for even and odd wheels respectively.

**Conjecture 4.2.** *Let* $G = W_{n-1}$ *for even* $n-1$*, and let* $P \subseteq [0,1]^{3n-2}$ *be its triangle relaxation. Then* $\pi[f](P) = X(f)$.

**Conjecture 4.3.** *Let* $G = W_{n-1}$ *for odd* $n-1$*, and let* $P \subseteq [0,1]^{3n-2}$ *be defined by the McCormick inequalities and*

$$\left\lfloor \frac{n-1}{2} \right\rfloor x_n + x_1 + \cdots + x_{n-1} - \sum_{ij \in E} y_{ij} \leq \left\lfloor \frac{n-1}{2} \right\rfloor,$$
$$\frac{n}{2} x_n + 2(x_1 + \cdots + x_n) - \sum_{ij \in E} y_{ij} \leq n - 1.$$

*Then* $\pi[f](P) = X(f)$.

The following theorem gives a condition under which the triangle inequalities are sufficient.

**Theorem 4.4** (Sufficiency of the Triangle Inequalities)**.** *Let $G = (V, E)$ be a graph, with all edge weights equal to 1, and $f$ be the associated bilinear function. Suppose that for all $\boldsymbol{x} \in [0, 1]^n$, there exists a collection $T^* \subseteq T$ of edge-disjoint triangles $ijk$, and sets $X_i \subset [0, 1), i = 1, \dots, n$, such that*

1. *$\mu(X_i) = x_i$ for all $i$,*
2. *$X_i \cap X_j \neq \emptyset$ implies either $X_i \cap X_j = [0, 1)$, or that there exists a $k$ such that $ijk \in T^*$,*
3. *For all $ijk \in T^*$, $X_i \cup X_j \cup X_k = [0, 1)$, and $X_i \cap X_j \cap X_k = \emptyset$.*

*Then $X(f) = \pi[f](P)$ where $P$ is the triangle relaxation of $G$.*

*Proof.* We need $\mathrm{vex}[f](\boldsymbol{x}) \leq \mathrm{LB}_P[f](\boldsymbol{x})$ for every $\boldsymbol{x} \in [0, 1]^n$. The relevant inequalities are

$$y_{ij} \geq x_i + x_j - 1 \text{ for all } ij \in E \tag{4}$$

$$y_{ij} + y_{ik} + y_{jk} \geq x_i + x_j + x_k - 1 \text{ for all } ijk \in T \tag{5}$$

$$y_{ij} \geq 0 \text{ for all } ij \in E \tag{6}$$

By Theorem 3.4, we show that

$$\sum_{ij \in E} \mu(X_i \cap X_j) = \min \left\{ \sum_{ij \in E} y_{ij} : (4), (5), (6) \right\}$$

Our approach involves the dual linear program. We have dual variables $z_{ij}$ for each McCormick inequality $x_i + x_j - 1 \leq y_{ij}$, and $w_{ijk}$ for each triangle inequality $y_{ij} + y_{ik} + y_{jk} \geq x_i + x_j + x_k - 1$. The dual constraints are

$$z_{ij} + \sum_{k : ijk \in T^*} w_\tau \leq 1 \tag{7}$$

$$z_{ij} \geq 0 \tag{8}$$

$$w_{ijk} \geq 0 \tag{9}$$

By the *weak duality theorem* (A.3) is enough to find fesible $z_{ij}$ and $w_{ijk}$ such that

$$\sum_{ij \in E} \mu(X_i \cap X_j) = \sum_{ij \in E} (x_i + x_j - 1) z_{ij} + \sum_{ijk \in T} (x_i + x_j + x_j - 1) w_{ijk}$$

For $z_{ij}$ and $w_{ijk}$ we choose

$$z_{ij} = \begin{cases} 0 & \exists k \text{ such that } ijk \in T^* \\ 1 & \nexists k \text{ such that } ijk \in T^* \text{ and } X_i \cap X_j \neq \emptyset \\ 0 & \nexists k \text{ such that } ijk \in T^* \text{ and } X_i \cap X_j = \emptyset \end{cases}$$

$$w_{ijk} = \begin{cases} 1 & ijk \in T^* \\ 0 & \text{otherwise} \end{cases}$$

For each $ij \in E$ we have the following:

**Case 1:** If $X_i \cap X_j = \emptyset$. Then $\mu(X_i \cap X_j) = 0$ and $z_{ij} = 0$.

**Case 2:** Suppose $X_i \cap X_j \neq \emptyset$, and $X_i \cup X_j = [0,1)$, and $\nexists k$ such that $ijk \in T^*$. Then $z_{ij} = 1$, and by 1, $\mu(X_i \cap X_j) = \mu(X_i) + \mu(X_j) - \mu(X_i \cup X_j) = x_i + x_j - 1$.

**Case 3:** Suppose $X_i \cap X_j \neq \emptyset$, and $\exists k$ such that $ijk \in T^*$. Then, $w_{ijk} = 1$, $z_{ij} = 0$, and

$$X_i \cup X_j \cup X_k = [0,1), \text{ and } X_i \cap X_j \cap X_k = \emptyset$$
$$\implies \mu(X_i \cap X_j) + \mu(X_j \cap X_k) + \mu(X_i \cap X_k) = x_i + x_j + x_k - 1$$

Our $z_{ij}$ and $w_{ijk}$ satisfy (8) and (9) by construction, (7) because the triangles $T^*$ are edge disjoint, and so are feasible. □

We assume without loss of generality that $\max\{0, x_2+x_3-1\}+\cdots+\max\{0, x_{n-1}+x_1-1\} \geq \max\{0, x_1+x_2-1\}+\cdots+\max\{x_{n-1}+x_{n-2}-1\}$. Then we choose $T^* := \{12n, 34n, \ldots, (n-2, n-1, n)\}$. Let $\boldsymbol{x} = (x_1, \ldots, x_{n-1}, x_n) \in [0,1]^n$. We introduce $2n-2$ new variables $x_i', x_i''$, $i \in [n-1]$. They will reflect the components of $X_i$ that intersect $X_n$ and $\overline{X}_n$ respectively. Constructibility of the sets $X_1, \ldots, X_n$ reduces to the feasibility of the following system of inequalities.

**Lemma 4.5.** *Suppose there exists a feasible solution* $(\boldsymbol{x}', \boldsymbol{x}'')$ *for the system*

$$x_i' + x_i'' = x_i \qquad i \in [n-1] \tag{10}$$

$$x_i' \le 1 - x_n \qquad i \in [n-1] \tag{11}$$

$$x_i'' \le x_n \qquad i \in [n-1] \tag{12}$$

$$x_i' + x_{i+1}' \ge 1 - x_n \qquad i \in \{1, 3, \dots, n-2\} \tag{13}$$

$$x_i'' + x_{i+1}'' \le x_n \qquad i \in \{1, 3, \dots, n-2\} \tag{14}$$

$$x_i' + x_{i+1}' \begin{cases} \le 1 - x_n & \text{if } x_i + x_{i+1} \le 1 \\ \ge 1 - x_n & \text{if } x_i + x_{i+1} > 1 \end{cases} \qquad i \in \{2, 4, \dots, n-1\} \tag{15}$$

$$x_i'' + x_{i+1}'' \begin{cases} \le x_n & \text{if } x_i + x_{i+1} \le 1 \\ \ge x_n & \text{if } x_i + x_{i+1} > 1 \end{cases} \qquad i \in \{2, 4, \dots, n-1\}, \tag{16}$$

$$x_i', x_i'' \ge 0 \qquad i \in [n-1]. \tag{17}$$

*Then the sets* $X_n = [0, x_n)$ *and*

$$X_i = \begin{cases} [x_n, x_n + x_i') \cup [0, x_i'') & \text{for odd } i \in [n-1], \\ [1 - x_i', 1) \cup [x_n - x_i') & \text{for even } i \in [n-1] \end{cases}$$

*satisfy the conditions in Theorem 3.4.*

*Proof.* Constraints (11) and (12) ensure that $X_i \subseteq [0, 1)$ for all $i$.

Now, if $i$ is even, then

$$\mu(X_i) = 1 - (1 - x_i') + x_n - (x_n - x_i'') = x_i' + x_i''$$

and if $i$ is odd, then

$$\mu(X_i) = x_n + x_i' - x_n + x_i'' = x_i' + x_i''$$

In either case, by (10), $\mu(X_i) = x_i$ as required for condition 1.

Secondly, suppose $X_i \cap X_{i+1} \neq \emptyset$, and $(i, i+1, n) \notin T^*$. By our choice of $T^*$, $i$ is even. Furthermore, $X_i \cap X_{i+1} \neq \emptyset$ implies $x_i + x_{i+1} > 1$. Now

$$\begin{aligned} X_i \cup X_{i+1} &= [1 - x_i', 1) \cup [x_n - x_i'', x_n) \cup [x_n, x_n + x_{i+1}') \cup [0, x_{i+1}'') \\ &= ([x_n, x_n + x_{i+1}')) \cup [1 - x_i', 1) \cup ([0, x_{i+1}'') \cup [x_n - x_i'', x_n)) \end{aligned}$$

That $x_i + x_{i+1} > 1$ implies that $x_i + x'_{i+1} \geq 1 - x_n$ by (15), which gives $x_n + x'_{i+1} \geq 1 - x'_i$. It also implies that $x''_i + x''_{i+1} \geq x_n$ by (16), which gives $x''_{i+1} \geq x_n - x''_i$. Thus we have

$$X_i \cup X_{i+1} = [x_n, 1) \cup [0, x_n) = [0, 1)$$

as required for condition 2.

Finally, suppose $(i, i+1, n) \in T^*$. By our choice of $T^*$, $i$ is odd. By (13), $x'_i + x'_{i+1} \geq 1 - x_n$, so that $x_n + x'_{i+1} \geq 1 - x'_i$. Thus $[x_n, x_n + x'_{i+1}) \cup [1 - x'_i, 1) = [x_n, 1)$. By (14), $x''_i + x''_{i+1} \leq x_n$, so that $x''_{i+1} \leq x_n - x''_i$. thus $[0, x''_{i+1}) \cap [x_n - x''_i, x_n) = \emptyset$. This gives

$$X_n \cup X_i \cup X_{i+1} = [0, x_n) \cup [x_n, 1) = [0, 1)$$

and

$$X_n \cap X_i \cap X_{i+1} = \emptyset$$

as required for condition 3. □

We illustrate this with an example point and choice of sets. Consider the 6 wheel, and the point $(0.7, 0.5, 0.6, 0.4, 0.7, 0.3, 0.7)$. We choose the edge-disjoint triangles $127, 347$, and $567$, and sets $X_1, \ldots, X_6, X_7$ as in the following figure.

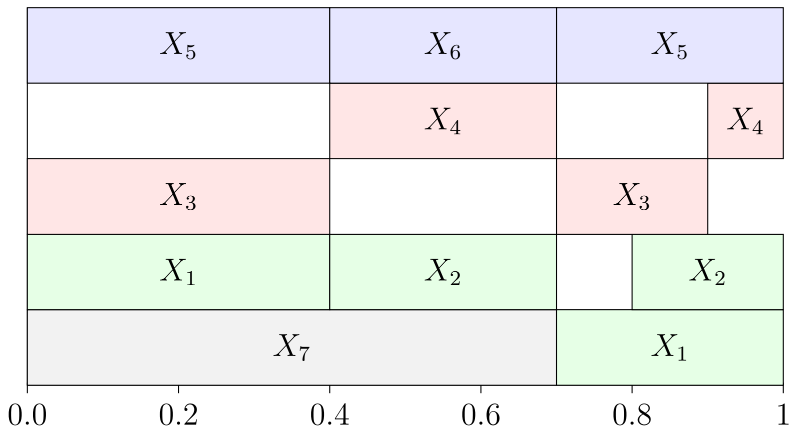



Figure 3: Choice of sets for 6-wheel.

We can see, for example, that $\mu(X_1 \cap X_2 \cap X_7) = x_1 + x_2 + x_7 - 1 = 0.9$, where we have used the triangle inequalities. And also, that $\mu(X_2 \cap X_3) = x_2 + x_3 - 1 = 0.1$ where we have used the McCormick inequalities. For the general proof, we need to show that this choice is always possible. Using a variant of Farkas' Lemma (A.1), we can say that the system in Lemma 4.5 is feasible if the dual system has a bounded objective function. Showing this would be sufficient to complete the proof for the even wheel, and would shed light on the case of odd wheels.

# References

[1] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

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

[3] Akshay Gupte, Thomas Kalinowski, Fabian Rigterink, and Hamish Waterer. "Extended formulations for convex hulls of graphs of bilinear functions". ArXiv:1702.04813.

[4] Marco Locatelli and Fabio Schoen. *Global Optimization: Theory, Algorithms, and Applications*. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2013.

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

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

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

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

# A Appendices

## A.1 Convexity

Proceeding according to [1], a subset $C \subseteq \mathbb{R}^n$ is called *convex* if for any $\boldsymbol{x}_1, \boldsymbol{x}_2 \in C$ and any $\theta$ with $0 \leq \theta \leq 1$, we have $\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2 \in C$. Geometrically this means that the line segment between any two points of $C$ is contained entirely within $C$. We extend this, and say that a *convex combination* is a linear combination of the form $\theta_1\boldsymbol{x}_1 + \cdots + \theta_p\boldsymbol{x}_p$, where $\boldsymbol{x}_i \in C$, $\theta_i \geq 0$, for $i = 1, \ldots, p$, and $\theta_1 + \cdots + \theta_p = 1$. Indeed, a set $C \subseteq \mathbb{R}^n$ is convex if and only if it is closed under convex combinations. The *convex hull* of a set $C \subseteq \mathbb{R}^n$ is the smallest convex set that contains it. It is the set of all convex combinations of elements of $C$ and we denote it as

$$\operatorname{conv} C := \left\{ \sum_{i=1}^{n} \theta_i \boldsymbol{x}_i \ : \ \boldsymbol{x}_i \in C,\ \theta_i \geq 0,\ i = 1, \ldots, k,\ \sum_{i=1}^{n} \theta_i = 1 \right\}$$

For example, a *hyperplane* is the solution set to a linear equality, and a *half-space* is the solution set to a linear inequality. A *polyhedron* is the solution set to finitely many linear equalities and inequalities, and a bounded polyhedron is called a *polytope*.

A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be *convex* if its domain is a convex set, and for all $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \operatorname{dom} f$ and $\theta$ with $0 \leq \theta \leq 1$ we have $f(\theta\boldsymbol{x}_1 + (1-\theta)\boldsymbol{x}_2) \leq \theta f(\boldsymbol{x}_1) + (1-\theta) f(\boldsymbol{x}_2)$. Geometrically this means that the secant between any two points on the graph of $f$ lies entirely above $f$.

## A.2 Convex Optimisation

An *optimisation problem* in *standard form* is written

$$\begin{aligned} \min \quad & f_0(\boldsymbol{x}) \\ \text{subject to} \quad & f_i(\boldsymbol{x}) \leq 0,\ i = 1, \ldots, m \\ & h_i(\boldsymbol{x}) = 0,\ i = 1, \ldots, p \end{aligned} \tag{18}$$

where $f_0 : \mathbb{R}^n \to \mathbb{R}$ is the *objective*, $f_i : \mathbb{R}^n \to \mathbb{R},\ i = 1, \ldots, n$, and $h_i : \mathbb{R}^n \to \mathbb{R},\ i = 1, \ldots, p$. The inequalities $f_i(\boldsymbol{x}) \leq 0$ and equalities $h_i(\boldsymbol{x}) = 0$ are called the *constraints*. A point $\boldsymbol{x} \in \mathbb{R}^n$ is called *feasible* if it satisfies all of the constraints and is in the domain of $f_0$. The optimal value $p$ of the problem is defined by $p := \inf\{f_0(\boldsymbol{x}) : f_i(\boldsymbol{x}) \leq 0, h_i(\boldsymbol{x}) = 0\}$. A point $\boldsymbol{x}$ is called *optimal* if it is feasible, and $f_0(\boldsymbol{x}) = p$. If there exists a optimal point, then the problem is *solvable*. We say that $\hat{\boldsymbol{x}}$ is *locally optimal* if there exists a neighbourhood $U$ of $\hat{\boldsymbol{x}}$ in which it is optimal. A *maximisation problem* can be transformed into a minimisation problem, because maximising $f_0$ is equivalent to minimising $-f_0$.

The problem (18) is a *linear program* if the functions $f_0, \dots, f_n$ and $h_1, \dots, h_p$ are linear functions. It is a *convex optimisation problem* if the functions $f_0, \dots, f_n$ are convex, and the functions $h_1, \dots, h_p$ are affine. The set of feasible points of a convex optimisation problem is convex, and importantly, if $\boldsymbol{x}$ is a locally optimal point of a convex problem, then it is also globally optimal.

For a more complete treatment of convex optimisation, see [1].

## A.3 Linear Programming Duality

Suppose we have a linear program of the form

$$\begin{aligned} \min_{\boldsymbol{y}} \quad & \sum_{j=1}^{n} b_j y_j \\ \text{subject to} \quad & \sum_{j=1}^{n} a_{ji} y_j \geq c_i, \; i = 1, \dots, m \\ & y_j \geq 0, \; j = 1, \dots, n \end{aligned} \tag{19}$$

which we will call the *primal* program. To it is associated another linear program, called the *dual program*. For each variable in the primal problem, we have a constraint in the dual program, and for each constraint, a variable. The dual program has the form

$$\begin{aligned} \max_{\boldsymbol{x}} \quad & \sum_{i=1}^{m} c_i x_i \\ \text{subject to} \quad & \sum_{i=1}^{m} a_{ji} x_i \leq b_j, \; j = 1, \dots, n \\ & x_i \geq 0, \; i = 1, \dots, m \end{aligned} \tag{20}$$

The *weak duality theorem* states that if $\boldsymbol{y}$ is feasible for the primal (minimisation) problem, and $\boldsymbol{x}$ is feasible for the dual (maximisation) problem, then $\sum_{i=1}^{m} c_i x_i \leq \sum_{j=1}^{n} b_j y_j$.

The *strong duality theorem* states that if $\boldsymbol{y}$ is optimal for the primal problem, and $\boldsymbol{x}$ is optimal for the dual problem, then $\sum_{i=1}^{m} c_i x_i = \sum_{j=1}^{n} b_j y_j$.

Another important result is known as *Farkas' Lemma*.

**Lemma A.1** (Farkas' Lemma)**.** *Let* $A \in \mathbb{R}^{m \times n}$ *and* $b \in \mathbb{R}^m$*. Then exactly one of the following is satisfied:*

1. $\exists \boldsymbol{x}$ *such that* $A\boldsymbol{x} = b$.

2. $\exists \boldsymbol{y}$ *such that* $A^T\boldsymbol{y} \geq 0$ *and* $b^T\boldsymbol{y} < 0$.

## A.4 Convex Relaxations

As in [4], we define a *relaxation* as follows.

**Definition A.2.** *Let* $\min_{\boldsymbol{x} \in S} f(\boldsymbol{x})$ *be an optimisation problem. A relaxation of this problem is another problem*

$$\min_{\boldsymbol{x} \in \hat{S}} \hat{f}(\boldsymbol{x}) \tag{21}$$

*such that* $S \subseteq \hat{S}$*, and* $\boldsymbol{x} \in S$ *implies* $\hat{f}(\boldsymbol{x}) \leq f(\boldsymbol{x})$.

A solution of a relaxation of an optimisation problem is a lower bound on the solution set of the original problem. If $\hat{S}$ is a convex set and $\hat{f}$ is a convex function, then (21) is called a *convex relaxation*.

An important notion is that of a *convex envelope*.

**Definition A.3.** *The convex envelope of a function* $f : \mathbb{R}^n \to \mathbb{R}$ *over a non-empty convex set* $C \subseteq \operatorname{dom} f$ *is the unique function* $\operatorname{vex}[f] : C \to \mathbb{R}$ *for which*

1. $\operatorname{vex}[f]$ *is a convex on* $C$*,*

2. $\operatorname{vex}[f](\boldsymbol{x}) \leq f(\boldsymbol{x})$ *for all* $\boldsymbol{x} \in C$*, and*

3. *If* $g : C \to \mathbb{R}$ *is a convex function that satisfies* $g(\boldsymbol{x}) \leq f(x)$ *for all* $\boldsymbol{x} \in C$*, then* $g(\boldsymbol{x}) \leq \operatorname{vex}[f](\boldsymbol{x})$ *for all* $\boldsymbol{x} \in C$.

Geometrically, $\operatorname{vex}[f]$ is the point-wise supremum of all affine underestimating functions for $f$. We may similarly define the *concave envelope* of $f$ as the point-wise infimum of all affine overestimating functions for $f$. Since the graph of a function is a set, it has a convex hull. Importantly, the convex hull of the graph of $f$ is bounded below and above by the convex and concave envelopes respectively.