Parity polytopes and binarization

Dominik Ermel [a], Matthias Walter [b],*

[a] *Institute for Mathematical Optimization (IMO), Otto-von-Guericke-Universität Magdeburg, Universitätsplatz 2, 39106 Magdeburg, Germany*
[b] *Lehrstuhl für Operations Research, Kackertstraße 7, 52072 Aachen, Germany*

ABSTRACT

We consider generalizations of parity polytopes whose variables, in addition to a parity constraint, satisfy certain ordering constraints. More precisely, the variable domain is partitioned into k contiguous groups, and within each group, we require $x_i \geq x_{i+1}$ for all relevant i. Such constraints are used to break symmetry after replacing an integer variable by a sum of binary variables, so-called binarization. We provide extended formulations for such polytopes, derive a complete outer description, and present a separation algorithm for the new constraints. It turns out that applying binarization and only enforcing parity constraints on the new variables is often a bad idea. For our application, an integer programming model for the graphic traveling salesman problem, we observe that parity constraints do not improve the dual bounds, and we provide a theoretical explanation of this effect.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

The term *binarization* refers to techniques to reformulate mixed-integer linear programs by replacing general integer (bounded) variables by sets of binary variables. Roy compared several approaches with the goal to separate strong cutting planes in the reformulation and project it back [7]. Recently, Bonami and Margot continued this line of work using simple split disjunctions of ranks 1 and 2 to generate cutting planes [2]. The strongest bounds were obtained by a binarization in which an integer variable $z \in \{0, 1, \ldots, n\}$ is replaced by the sum of n binary variables $x_1, \ldots, x_n$, and symmetry-breaking constraints

$$1 \geq x_1 \geq \cdots \geq x_n \geq 0 \tag{1}$$

are added. We denote by X_{ord}^n the set of ordered binary vectors $x \in \{0, 1\}^n$ satisfying (1) and by P_{ord}^n their convex hull. Note that Inequalities (1) already describe P_{ord}^n since the corresponding matrix is totally unimodular (every row has at most one $+1$ and at most one -1, see Theorem 19.3 in [8]) and the right-hand side is integral. We call the vectors in X_{ord}^n *ordered binary vectors*.

In this article we consider the strong binarization in combination with parity constraints on the affected integer variables. For this we define the *ordered even parity polytope* for a vector $r \in \mathbb{N}^k$ as

$$P_{\mathrm{even}}^r := \mathrm{conv}\{(x^{(1)}, \ldots, x^{(k)}) \in X_{\mathrm{ord}}^{r_1} \times \cdots \times X_{\mathrm{ord}}^{r_k} \mid \sum_{i=1}^{k} \sum_{j=1}^{r_i} x_j^{(i)} \text{ even}\}.$$

* Corresponding author.
E-mail addresses: dominik.ermel@st.ovgu.de (D. Ermel), walter@or.rwth-aachen.de (M. Walter).

The integer points in P_{even}^r are precisely those binary vectors of length $n := r_1 + \cdots + r_k$ whose entry-wise sum is even and which are ordered within each of the groups defined by r. In the special case of $r = \mathbb{1}_n$ (the all-ones vector), P_{even}^r is the even parity polytope, which has the following outer description [4] (with $[n] := \{1, 2, \ldots, n\}$):

$$P_{\text{even}}^{\mathbb{1}_n} = \{x \in [0, 1]^n \mid \sum_{i \in [n] \setminus F} x_i + \sum_{i \in F} (1 - x_i) \geq 1 \text{ for all } F \subseteq [n] \text{ with } |F| \text{ odd}\} \tag{2}$$

Outline. We start by describing an extended formulation for P_{even}^r for arbitrary $r \in \mathbb{N}^k$, which implies that the associated optimization problem can be solved in polynomial time (in the dimension of P_{even}^r). In Section 3 we state and prove the main result of this paper, a complete outer description for P_{even}^r in the original space. After transferring the result to odd parities in Section 4 we derive in Section 5 a simple linear-time algorithm that solves the respective separation problems. Section 6 is dedicated to an application in which binarization is applied to an integer programming model for the graphic traveling salesman problem. During computations for this model we observed that the corresponding parity constraints did not improve the dual relaxation value. In Section 7 we provide a theoretical explanation of this effect.

2. Extended formulations

In this section we review well-known extended formulations of parity polytopes based on dynamic programming, and adapt them to the ordered case. The linear-size formulation for $P_{\text{even}}^{\mathbb{1}_n}$ introduced by Carr and Konjevod [3] works as follows. We define a digraph $D = (V, A)$ with nodes $V := \{0, 1, 2, \ldots, n\} \times \{0, 1\} \setminus \{(0, 1), (n, 1)\}$ and arcs $A := \{((i - 1, \alpha), (i, \beta)) \in V \times V \mid i \in [n], \alpha, \beta \in \{0, 1\}\}$. We call the nodes $s := (0, 0)$ the source and $t := (n, 0)$ the sink. Clearly, the set $P_{\text{flow}}^{s-t}(D)$ of s-t-flows $y \in \mathbb{R}_+^A$ of flow-value 1 in any such acyclic digraph is described by the linear constraints

$$\begin{aligned} y(\delta^{\text{out}}(v)) - y(\delta^{\text{in}}(v)) &= 0 \quad \text{for all } v \in V \setminus \{s, t\} \\ y(\delta^{\text{out}}(s)) &= 1 \\ y(\delta^{\text{in}}(t)) &= 1 \\ y_a &\geq 0 \quad \text{for all } a \in A, \end{aligned}$$

where we denote by $\delta^{\text{out}}(\cdot)$ resp. $\delta^{\text{in}}(\cdot)$ the set of outgoing resp. incoming arcs. Furthermore, again due to the total unimodularity of D's incidence matrix (see Application 19.2 in [8]), $P_{\text{flow}}^{s-t}(D)$ is an integral polytope. We claim that the following map projects $P_{\text{flow}}^{s-t}(D)$ onto $P_{\text{even}}^{\mathbb{1}_n}$:

$$\pi : \mathbb{R}^A \to \mathbb{R}^n \qquad \text{with} \qquad \pi(y)_i := \begin{cases} y_{((0,0),(1,1))} & \text{for } i = 1 \\ y_{((i-1,0),(i,1))} + y_{((i-1,1),(i-1,0))} & \text{for } 2 \leq i \leq n - 1 \\ y_{((n-1,1),(n,0))} & \text{for } i = n. \end{cases}$$

It is not hard to see that the integer points in $\pi(P_{\text{flow}}^{s-t}(D))$ are precisely the binary vectors with an even number of 1's, since an s-t-path in D must use an even number of diagonal arcs. The claim follows because $P_{\text{flow}}^{s-t}(D)$ is integral and because π maps integral points to integral points. Fig. 1 illustrates the constructed network and the projection (denoted by $r_i = 1$).

We can now modify the network such that a similar projection yields P_{even}^r for arbitrary vectors $r \in \mathbb{N}^k$. The new construction works as follows. The node set of the new digraph is the same as the old digraph for $n = k$, i.e., the node set of a $(2 \times (k + 1))$-grid graph without the lower-left and lower-right corner nodes. Again, s and t are the left-most and right-most nodes, respectively. The arcs again go from left to right, and the arcs from the k grid-layers correspond to the sets of variables $x^{(i)}$ for $i \in [k]$ in the sense that the projection of an arc variable only has impact on a subset of the corresponding x-variables. For every layer $i \in [k]$ and every node $(i - 1, \alpha)$ (for $\alpha = 0$ if $i = 1$ or $i = k$, and for $\alpha \in \{0, 1\}$ otherwise) on the left of this layer there are $r_i + 1$ outgoing arcs, each associated to one of the possible values $\ell \in \{0, 1, 2, \ldots, r_i\}$. The arcs with even ℓ all go to the same node (i, α), and the arcs with odd ℓ all go to node $(i, 1 - \alpha)$.

The projection map is defined by a 0/1-matrix. We describe it column-wise, i.e., we state for each arc in which rows (corresponding to x-variables) it contains a 1-entry. Consider an arc going from node $(i - 1, \alpha)$ to node (i, β) that is associated to $\ell \in \{0, 1, 2, \ldots, r_i\}$, and thus satisfies $\alpha + \beta \equiv \ell \pmod 2$. The column of the projection matrix that corresponds to this arc contains 1's in rows corresponding to the variables $x_1^{(i)}, x_2^{(i)}, \ldots, x_\ell^{(i)}$. The last part in Fig. 1 illustrates a layer for $r_i = 3$.

The correctness of the modification is justified similarly to the ordinary construction. First, the flow polytope is integral due to total unimodularity of its constraint matrix. Second, the projection matrix is integral. Third, the integral s-t-flows map to the correct vectors x. To see this, observe that for every $i \in [k]$, the corresponding arcs of the network correspond to the vectors in $X_{\text{ord}}^{r_i}$, and that the number of diagonal arcs in such a path must be an even number.

3. Outer description

We begin with the simple observation that for ordered binary vectors, parity can be measured using a linear function. For $n \in \mathbb{N}$ we define $f : P_{\text{ord}}^n \to \mathbb{R}$ via $f(x) := \sum_{i=1}^n (-1)^{i-1} x_i$. From

$$f(x) = \underbrace{x_1 - x_2}_{\geq 0} + \underbrace{x_3 - x_4}_{\geq 0} + x_5 - \cdots \pm x_n \qquad \text{and}$$
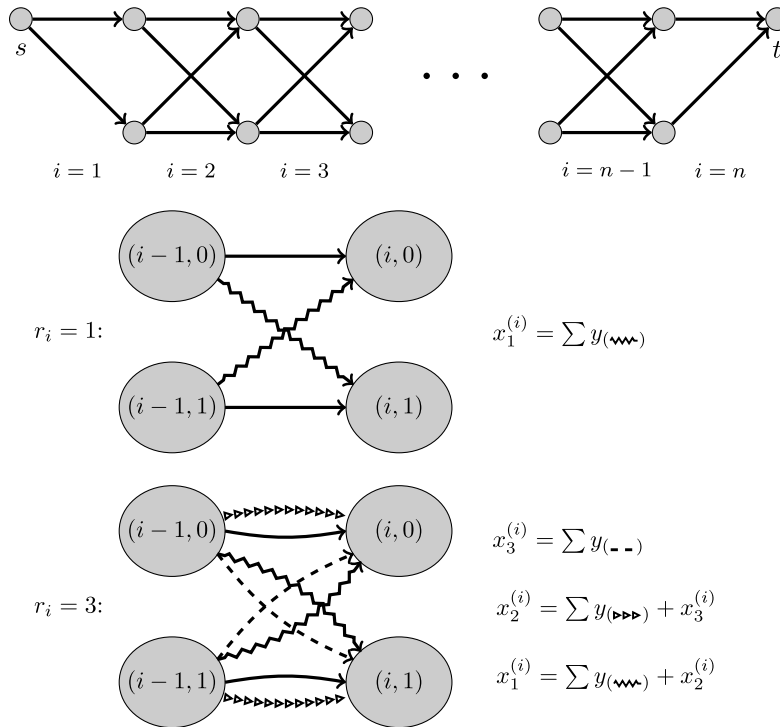
Fig. 1. Flow-extensions of (ordered) even parity polytopes.

$$f(x) = \underbrace{x_1}_{\leq 1} - \underbrace{(x_2 - x_3)}_{\geq 0} - \underbrace{(x_4 - x_5)}_{\geq 0} + \cdots \pm x_n$$

we obtain that f maps P_{ord}^n into $[0, 1]$. Thus, for $x \in X_{ord}^n$, the value $f(x)$ is binary, and it is easy to check that $f(x) = 1$ holds if and only if x has an odd number of 1's. Note that we will use f for different values of n, and assume that it is clear from the context. We can now state our main theorem.

Theorem 1. *Let $r \in \mathbb{N}^k$. The ordered even parity polytope P_{even}^r is equal to the set of those points $(x^{(1)}, \ldots, x^{(k)}) \in P_{ord}^{r_1} \times \cdots \times P_{ord}^{r_k}$ that satisfy the inequalities*

$$\sum_{i \in [k] \setminus F} f(x^{(i)}) + \sum_{i \in F} (1 - f(x^{(i)})) \geq 1 \qquad (3)$$

for all $F \subseteq [k]$ with $|F|$ odd.

In order to prove the theorem we will make use of the following lemma.

Lemma 2. *Let $X_0, X_1 \subseteq \mathbb{R}^m$ and $Y_0, Y_1 \subseteq \mathbb{R}^n$ be finite sets. Then the polytope*

$$\text{conv} ((X_0 \times \{0\} \times Y_0) \cup (X_1 \times \{1\} \times Y_1))$$

is equal to the polytope

$$\text{conv} ((X_0 \times \{0\}) \cup (X_1 \times \{1\})) \times \mathbb{R}^n \ \cap \ \mathbb{R}^m \times \text{conv} ((\{0\} \times Y_0) \cup (\{1\} \times Y_1)) .$$

Proof of Lemma 2. The inclusion "$\subseteq$" is simple since for $i \in \{0, 1\}$, every vector $(x, i, y) \in X_i \times \{i\} \times Y_i$ satisfies $(x, i) \in X_i \times \{i\}$ and $(i, y) \in \{i\} \times Y_i$.

To see "$\supseteq$", consider a point (x, λ, y) from the second set, which clearly satisfies $0 \leq \lambda \leq 1$. Thus (x, λ) is a convex combination of a point $(x^{(0)}, 0) \in \text{conv} (X_0) \times \{0\}$ and a point $(x^{(1)}, 1) \in \text{conv} (X_1) \times \{1\}$. From the λ-coordinate we obtain $x = \lambda x^{(0)} + (1 - \lambda) x^{(1)}$. Similarly, there exist points $y^{(0)} \in \text{conv} (Y_0)$ and $y^{(1)} \in \text{conv} (Y_1)$ such that $y = \lambda y^{(0)} + (1 - \lambda) y^{(1)}$ holds. This implies that $(x, \lambda, y) = \lambda(x^{(0)}, 0, y^{(0)}) + (1 - \lambda)(x^{(1)}, 1, y^{(1)})$ holds, which concludes the proof. □

In our situation this lemma implies that if we "glue together" arbitrary 0/1-polytopes at a *single* coordinate, then the outer description of the resulting polytope is obtained by taking the union of the outer descriptions of the two component polytopes. We can now prove Theorem 1.

Proof of Theorem 1. We start with the validity of Inequalities (3) for P_{even}^r. Let $F \subseteq [k]$ be of odd cardinality. Since $f(x^{(i)}) \in \{0, 1\}$ for every $x^{(i)} \in X_{\text{ord}}^{r_i}$, the inequality can only be violated if all summands are equal to 0, i.e., if $f(x^{(i)}) = 0$ for all $i \in [k] \setminus F$ and $f(x^{(i)}) = 1$ for all $i \in F$. This in turn means that $x^{(i)}$ contains an odd number of 1's if and only if $i \in F$ holds, contradicting the fact that the overall number of 1's is even because $|F|$ is odd.

We now turn to the proof that Inequalities (3) cut off those points in $X_{\text{ord}}^{r_1} \times \cdots \times X_{\text{ord}}^{r_k}$ that have an odd number of 1's. Let $x = (x^{(1)}, \ldots, x^{(k)})$ be such a point and define F to be the set of indices $i \in [k]$ for which $x^{(i)}$ has an odd number of 1's. Since x has an odd number of 1's, we have that $|F|$ is odd. Furthermore, the left-hand side of (3) is equal to 0, i.e., x violates it.

It remains to prove that Inequalities (3) together with the inequalities that define each $P_{\text{ord}}^{r_i}$ for $i = 1, 2, \ldots, k$ yield an integral polytope. We show this by induction on the number ℓ of components of r that are strictly greater than 1. For $\ell = 0$ we have $r = \mathbb{1}_n$, and P_{even}^r is, according to (2), the even parity polytope.

Let $\ell \in \mathbb{N}$. W.l.o.g., we can assume that $r_k \geq 2$ holds since otherwise we permute the variables suitably. By induction hypothesis we know that

$$P_{\text{even}}^{(r_1, \ldots, r_{k-1}, 1)} = \left\{ (x^{(1)}, \ldots, x^{(k-1)}, \lambda) \in P_{\text{ord}}^{r_1} \times \cdots \times P_{\text{ord}}^{r_{k-1}} \times [0, 1] \mid (x^{(1)}, \ldots, x^{(k-1)}, \lambda) \text{ satisfies (3)} \right\}$$

holds, and is a 0/1-polytope. Furthermore, the polytope

$$Q := \left\{ (\lambda, x^{(k)}) \in [0, 1] \times P_{\text{ord}}^{r_k} \mid \lambda = f(x^{(k)}) \right\}$$

is affinely isomorphic to $P_{\text{ord}}^{r_k}$. Since the latter is a 0/1-polytope and since f maps $X_{\text{ord}}^{r_k}$ to $\{0, 1\}$, Q is also a 0/1-polytope. Applying Lemma 2 to $P_{\text{even}}^{(r_1, \ldots, r_{k-1}, 1)}$ and Q yields that

$$R := \{ (x^{(1)}, \ldots, x^{(k-1)}, \lambda, x^{(k)}) \in P_{\text{ord}}^{r_1} \times \cdots \times P_{\text{ord}}^{r_{k-1}} \times [0, 1] \times P_{\text{ord}}^{r_k} \mid \lambda = f(x^{(k)}) \text{ and }$$
$$(x^{(1)}, \ldots, x^{(k)}) \text{ satisfies (3)} \}$$

is an integral polytope. The orthogonal projection of R onto the x-variables is again an integral polytope and due to the equation $\lambda = f(x^{(k)})$ it is equal to the one in question. $\square$

4. Odd parities

In this section we derive the outer description of the ordered *odd* polytopes. The latter are defined as

$$P_{\text{odd}}^r := \text{conv}\{(x^{(1)}, \ldots, x^{(k)}) \in X_{\text{ord}}^{r_1} \times \cdots \times X_{\text{ord}}^{r_k} \mid \sum_{i=1}^{k} \sum_{j=1}^{r_i} x_j^{(i)} \text{ odd}\},$$

again for $r \in \mathbb{N}^k$. The result corresponding to Theorem 1 is the following:

Corollary 3. *The ordered odd parity polytope for $r \in \mathbb{N}^k$ is the set of $(x^{(1)}, \ldots, x^{(k)}) \in P_{\text{ord}}^{r_1} \times \cdots \times P_{\text{ord}}^{r_k}$ that satisfy Inequalities (3) for all $F \subseteq [k]$ with $|F|$ even.*

In principle we could repeat the proof for Theorem 1, and exchange "odd" and "even" suitably. Instead we present a proof that uses the previous result.

Proof. We claim that $P_{\text{odd}}^{(r_1, \ldots, r_k)}$ is the projection of the face defined by $x_1^{(k+1)} = 1$ of the polytope $P_{\text{even}}^{(r_1, \ldots, r_k, 1)}$ onto the variables $x^{(1)}, \ldots, x^{(k)}$. This is justified by the fact that fixing the last component to 1 exchanges the meaning of "odd" and "even" for the sum of the (remaining) components.

From Theorem 1 we know the outer description of $P_{\text{even}}^{(r_1, \ldots, r_k, 1)}$, so it remains to carry out the projection. The only constraints that involve the last component are the bounds $0 \leq x^{(k+1)} \leq 1$ and the Inequalities (3) for odd-cardinality sets $F \subseteq [k + 1]$. The former need not be considered, and if $k + 1 \notin F$ holds, then the latter inequalities are redundant since they read $\sum_{i \in [k] \setminus F} f(x^{(i)}) + 1 + \sum_{i \in F} (1 - f(x^{(i)})) \geq 1$, and for $x^{(i)} \in P_{\text{ord}}^{r_i}$ we know that $f(x^{(i)})$ and $1 - f(x^{(i)})$ are nonnegative. It remains to consider the case of $k + 1 \in F$, in which the inequality is $\sum_{i \in [k] \setminus F} f(x^{(i)}) + \sum_{i \in F \cap [k]} (1 - f(x^{(i)})) + (1 - 1) \geq 1$, which is equivalent to Inequality (3) for $F \cap [k]$ in the projection. Note that $F \cap [k]$ has even cardinality. $\square$

5. Separation

In this section we show how to solve the separation problem for Inequalities (3).

Theorem 4. *Let $r \in \mathbb{N}^k$ and let $\hat{x} = (\hat{x}^{(1)}, \ldots, \hat{x}^{(k)}) \in P_{\text{ord}}^{r_1} \times \cdots \times P_{\text{ord}}^{r_k}$. We can decide in linear time if $\hat{x} \in P_{\text{even}}^r$ holds, and if this is not the case, obtain an odd set $F \subseteq [k]$ whose associated Inequality (3) is violated by $\hat{x}$.*

Proof. It is easy to see that it suffices to compute $\hat{\lambda}_i := f(\hat{x}^{(i)}) \in [0, 1]$ for every $i \in [k]$ and then to minimize $\sum_{i \in [k] \setminus F} \hat{\lambda}_i + \sum_{i \in F} (1 - \hat{\lambda}_i)$ over all $F \subseteq [k]$ of odd cardinality. This is done by computing the set $F' := \left\{ i \in [k] \mid \hat{\lambda}_i > \frac{1}{2} \right\}$. If $|F'|$ is odd, then $F := F'$ is the minimizer. Otherwise, we let $F := F' \triangle \{\hat{i}\}$ (by $\triangle$ we denote the symmetric difference, i.e., $A \triangle B := (A \cup B) \setminus (A \cap B)$) where $\hat{i} \in [k]$ is the index for which $|\hat{\lambda}_{\hat{i}} - \frac{1}{2}|$ is minimum.

The computation of $\hat{\lambda}$ can obviously be done in linear time. Furthermore, the construction of F' and the search for $\hat{i}$ can be carried out in time $\mathcal{O}(k)$ which results in a linear total running time. $\quad\square$

Directly from Corollary 3 we obtain the separation algorithm for the ordered *odd* parity polytopes. The only difference to the even case is that $|F|$ must be even, i.e., we set $F := F' \triangle \{\hat{i}\}$ (for the same $\hat{i}$) if and only if $|F'|$ is odd.

Corollary 5. *Let $r \in \mathbb{N}^k$ and let $\hat{x} = (\hat{x}^{(1)}, \ldots, \hat{x}^{(k)}) \in P_{ord}^{r_1} \times \cdots \times P_{ord}^{r_k}$. We can decide in linear time if $\hat{x} \in P_{odd}^r$ holds, and if this is not the case, obtain an even set $F \subseteq [k]$ whose associated Inequality (3) is violated by $\hat{x}$.*

6. Strengthened Blossom Inequalities for the graphic TSP

In this section we consider the *graphic traveling salesman problem (GTSP)*, defined as follows. The input consists of an undirected graph $G = (V, E)$, and the goal is to find a minimum-length closed walk in G that visits every node at least once. Recently, Sebő and Vygen developed a 7/5-approximation algorithm for this problem [9].

When modeling it as an integer program (IP) one has different options. The first is based on the observation that the problem is equivalent to the traveling salesman problem on the complete graph with $|V|$ nodes and edge weights $c \in \mathbb{R}_+^E$ where $c_{u,v}$ is equal to the (combinatorial) distance from u to v in G. This model has $\binom{|V|}{2}$ binary variables which is much greater than $|E|$ if G is sparse.

In order to model the problem with fewer variables we can start with the IP

$$\min \sum_{e \in E} z_e \tag{4}$$

$$\text{s.t.} \quad z(\delta(S)) \geq 2 \quad \text{for all } \emptyset \neq S \subsetneq V \tag{5}$$

$$z_e \geq 0 \quad \text{for all } e \in E \tag{6}$$

$$z_e \in \mathbb{Z} \quad \text{for all } e \in E, \tag{7}$$

where by $\delta(\cdot)$ we denote the cut-sets. The solutions to this problem correspond to the 2-edge-connected spanning subgraphs of G. Since not every such subgraph is a closed walk, this IP does not model the GTSP. The necessary additional requirement is the parity condition $z(\delta(v)) \in 2\mathbb{Z}$ for every $v \in V$. Such a constraint cannot be modeled directly, i.e., by adding linear inequalities. To obtain a correct model, one may add integer variables $y_v = \frac{1}{2} z(\delta(v))$ for every $v \in V$. Unfortunately, this does not contribute to the strength of the LP relaxation since relaxing y_v's integrality constraints essentially makes them redundant.

It is well-known that an optimum solution will never use an edge more than twice (otherwise we can decrease the value by 2 and obtain a better feasible solution). Hence we can restrict z_e to the set $\{0, 1, 2\}$ which allows us to perform binarization, i.e., to replace, for every $e \in E$, the variable z_e by $x_e^{(1)} + x_e^{(2)}$ with $x_e^{(1)}, x_e^{(2)} \in \{0, 1\}$. This again leads to the same strength of the LP relaxation, but now allows us to enforce parity constraints: since every closed walk traverses each cut $\delta(S)$ an even number of times, the following inequalities are valid for every $S \subseteq V$ and every $F \subseteq \delta(S) \times \{0, 1\}$ with $|F|$ odd.

$$\sum_{(e,i) \in (\delta(S) \times \{0,1\}) \setminus F} x_e^{(i)} + \sum_{(e,i) \in F} (1 - x_e^{(i)}) \geq 1 \tag{8}$$

If we identify the variables $x_e^{(1)}$ and $x_e^{(2)}$ with the doubled edges of G, then these inequalities are the well-known *Blossom Inequalities*. These enforce that, for every cut $\delta(S)$, the projection of x onto the variables corresponding to $\delta(S)$ is in the even parity polytope.

We will soon strengthen these inequalities. In order to argue later why the separation problem for the strengthened inequalities can be solved efficiently, we review the separation problem for the unstrengthened ones. The first separation algorithm for Inequalities (8) was developed by Padberg and Rao [6]. The asymptotically fastest separation algorithm is by Letchford et al. [5]. More precisely, they describe how to efficiently solve (a generalization of) the following problem. Given a graph $\overline{G} = (\overline{V}, \overline{E})$ and weight vectors $c, c' \in \mathbb{R}_+^{\overline{E}}$, their algorithm can minimize the objective

$$\beta(S, F) := \sum_{e \in \delta(S) \setminus F} c_e + \sum_{e \in F} c'_e \tag{9}$$

over all sets $S \subseteq \overline{V}$ and all odd sets $F \subseteq \delta(S)$. Clearly, by letting $\overline{G}$ be the graph obtained from G by doubling every edge and identifying the two edges $e_1, e_2 \in \overline{E}$ with the variables $x_e^{(1)}$ and $x_e^{(2)}$ for $e \in E$, respectively, we can solve the separation problem for a point $\hat{x} \in [0, 1]^{2E}$ by minimizing $\beta(S, F)$ with the following weight vectors:

$$c_{e_i} := \hat{x}_e^{(i)} \quad \text{and} \quad c'_{e_i} := 1 - \hat{x}_e^{(i)} \quad \text{for all } i = 1, 2 \text{ and all } e \in E.$$

In fact we do not have to double the graph's edges for the computation: the cases where F contains none or both edges contribute the same to the parity constraint on F. Hence, we only have to consider the case that has lower weight. Similarly, only one of the two cases of F containing precisely one of the two edges needs to be considered. This can be implemented using the weights

$$c_e := \min(\hat{x}_e^{(1)} + \hat{x}_e^{(2)}, 2 - \hat{x}_e^{(1)} - \hat{x}_e^{(2)}) \quad \text{and} \quad c_e' := \min(1 - \hat{x}_e^{(1)} + \hat{x}_e^{(2)}, 1 + \hat{x}_e^{(1)} - \hat{x}_e^{(2)})$$

for all $e \in E$.

Strengthened constraints. To break the symmetry of $x_e^{(1)}$ and $x_e^{(2)}$ we add $x_e^{(1)} \geq x_e^{(2)}$ for each edge $e \in E$. Thus, we can now enforce that, for every cut $\delta(S)$, the projection of x onto the variables corresponding to $\delta(S)$ is in the ordered even parity polytope P_{even}^r, where $r = (2, 2, \ldots, 2) \in \mathbb{Z}^{|\delta(S)|}$. By Theorem 1, the strengthened version of Constraint (8) reads

$$\sum_{e \in \delta(S) \setminus F} (x_e^{(1)} - x_e^{(2)}) + \sum_{e \in F}(1 - x_e^{(1)} + x_e^{(2)}) \geq 1 \tag{10}$$

for every $S \subseteq V$ and every $F \subseteq \delta(S)$ with $|F|$ odd. Using almost the same separation procedure we can solve the separation problem by working in the original graph G, and defining the weight vectors via

$$c_e := \hat{x}_e^{(1)} - \hat{x}_e^{(2)} \quad \text{and} \quad c_e' := 1 - \hat{x}_e^{(1)} + \hat{x}_e^{(2)} \quad \text{for all } e \in E .$$

Note that the computational costs for separating Constraints (8) and (10) are almost equal.

7. Strength of the relaxation

We carried out computational experiments on the GTSP, and it turned out that the lower bound obtained by the linear relaxation (4)–(6) was never improved by binarization of the variables and addition of (strengthened) parity constraints. In fact, in the root node many (strengthened) Blossom Inequalities were added, but with no effect on the dual objective.

We investigated this effect, and observed the following weakness of the approach of binarization and addition of parity constraints. The decisive property of our model is that *only* parity constraints actually use the binarization variables $x_e^{(i)}$: all other variables consider the original variables z_e (or, equivalently, the sum of the binarization variables $x_e^{(i)}$).

In a more abstract setting we consider arbitrary (integer) variables $z_1, \ldots, z_k$ with domains $z_i \in [0, r_k] \cap \mathbb{Z}$ for all $i \in [k]$, and apply binarization, i.e., introduce variables $x_j^{(i)} \in P_{\text{ord}}^{r_i}$ and linking constraints $z_i = \sum_{j=1}^{r_i} x_j^{(i)}$ for each $i \in [k]$. We can assume that further (arbitrary) constraints that link the z-variables are also present.

Now consider a parity constraint on (a subset of) the z-variables, which is of course stated in terms of the corresponding x-variables. Suppose we have a certain fractional relaxation solution $(\hat{z}, \hat{x})$ that may violate this parity constraint, but satisfies all other constraints. If we can modify the x-variables such that the linking constraints are still satisfied (i.e., that the sums $\sum_{j=1}^{r_i} \hat{x}_j^{(i)}$ remain constant) and such that the parity constraint is satisfied, then we obtain a feasible solution of the same value. Unfortunately, such a modification is possible under very mild conditions. In fact, the modifications are very general in the sense that they can be done independently for every set of binarization variables, that is, for all variables $x_j^{(i)}$ for a *single* $i \in [k]$. Furthermore, they do neither depend on the parity inequality that is actually violated nor on the parity. Hence, even if there exist many parity constraints whose variable sets overlap, such modifications can still exist, as it is often the case for Blossom Inequalities (note that there actually exist $2^{|V|-1}$ such parity constraints). We will now describe the modifications, and later discuss why the conditions are often satisfied for solutions of the linear relaxation (4)–(6).

Lemma 6. *Let $n \in \mathbb{N}$, and $z \in [0, n]$. Then the maximum value of $\min(f(x), 1 - f(x))$ over all $x \in P_{\text{ord}}^n$ with $z = \sum_{i=1}^n x_i$ is equal to the minimum of z, $n - z$, and $\frac{1}{2}$.*

Proof. We denote the maximum in the lemma by γ^*, and by definition we have $\gamma^* \in [0, \frac{1}{2}]$. We distinguish three cases, depending on where the minimum of z, $n - z$, and $\frac{1}{2}$ is attained.

Case 1: $z < \frac{1}{2}$. Clearly, since $f(x) = x_1 - x_2 + x_3 - \cdots \pm x_n \leq x_1 + x_2 + x_3 + \cdots + x_n = z$ holds, we have $\gamma^* \leq z$. This value is obtained by $x = (z, 0, \ldots, 0, 0) \in P_{\text{ord}}^n$ which satisfies $\sum_{i=1}^n x_i = z$ and $f(x) = z$.

Case 2: $n - z < \frac{1}{2}$. This implies $x_i \geq \frac{1}{2}$ and hence $x_i \leq 2 - x_i$ for all $i \in [n]$. For even n we obtain $f(x) \leq (2 - x_1) - x_2 + (2 - x_3) - x_4 + \cdots + (2 - x_{n-1}) - x_n \leq n - z$. For odd n, $1 - f(x) \leq 1 - (x_1 - (2 - x_2) + x_3 - (2 - x_4) + \cdots + x_{n-1}) \leq n - z$. Together, this proves $\gamma^* \leq n - z$. This value is obtained by $x = (1, 1, \ldots, 1, z - n + 1) \in P_{\text{ord}}^n$ which satisfies $\sum_{i=1}^n x_i = z$. Depending on the parity of n it also satisfies $f(x) = z - n + 1$ or $1 - f(x) = z - n + 1$, i.e., we have $\gamma^* = n - z$.

Case 3: $\min(z, n - z) \geq \frac{1}{2}$. Since $\gamma^* \leq \frac{1}{2}$ always holds, it remains to construct a solution of that value. We define $k \in \mathbb{N}$ to be an integer whose distance to z is at most $\frac{1}{2}$, breaking ties such that k is not equal to 0 or equal to n (but arbitrary otherwise). We furthermore define the values $z_k^- := \frac{2z - 2k + 1}{4} \in [0, \frac{1}{2}]$, and $z_k^+ := \frac{2z - 2k + 3}{4} \in [\frac{1}{2}, 1]$. We again consider two cases, depending on k:

Case 3 (a): $1 \leq k \leq n - 2$. We consider the vector x defined via

$$x_i := \begin{cases} 1 & \text{for } i \leq k - 1 \\ \dfrac{1}{2} & \text{for } i = k \\ z_k^- & \text{for } i \in \{k+1, k+2\} \\ 0 & \text{for } i \geq k + 3, \end{cases}$$

and observe that $x \in P_{\text{ord}}^n$ holds, and that the first $(k-1)$ indices have 1's. Furthermore, we have that $\sum_{i=1}^n x_i = (k-1) + \frac{1}{2} + 2z_k^- = k - 1 + \frac{1}{2} + \frac{2z - 2k + 1}{2} = z$ and $f(x) = \frac{1}{2}$ hold.

Case 3 (b): $k = n - 1$. We consider the vector $x = (1, 1, \ldots, 1, z_{n-1}^+, z_{n-1}^+, \frac{1}{2})$ and observe that $x \in P_{\text{ord}}^n$ holds. Again, we have that $\sum_{i=1}^n x_i = (n-3) + 2z_{n-1}^+ + \frac{1}{2} = n - 3 + \frac{2z - 2(n-1) + 3}{2} + \frac{1}{2} = z$ and $f(x) = \frac{1}{2}$ hold. This concludes the proof. □

Motivated by Lemma 6 we define $\gamma(z) := \min(z, r - z, \frac{1}{2})$ for a variable $z \in [0, r]$. In the following theorem we binarize a set of variables and consider parity constraints on subsets of them.

Theorem 7. *Let $r \in \mathbb{N}^k$ and let $z \in [0, r_1] \times \cdots \times [0, r_k]$. Let $\mathcal{I}$ be a family of subsets $I \subseteq [k]$. If every $I \in \mathcal{I}$ satisfies $\sum_{i \in I} \gamma(z_i) \geq 1$, then there exist $x^{(i)} \in P_{\text{ord}}^{r_i}$ with $z_i = \sum_{j=1}^{r_i} x_j^{(i)}$ for all $i \in [k]$ such that for every $I \in \mathcal{I}$ the vector $(x^{(i)})_{i \in I}$ is contained in the even and odd parity polytopes corresponding to $(r_i)_{i \in I}$.*

Proof. Let $x^{(i)}$ be the maximizer from Lemma 6 for all $i \in [k]$, that is, $\min(f(x^{(i)}), 1 - f(x^{(i)})) = \gamma(z_i)$ holds for every $i \in [k]$. Now consider one of the sets $I \in \mathcal{I}$ and any subset $F \subseteq I$. From

$$1 \leq \sum_{i \in I} \gamma(z_i) = \sum_{i \in I} \min(f(x^{(i)}), 1 - f(x^{(i)})) \leq \sum_{i \in I \setminus F} f(x^{(i)}) + \sum_{i \in F} (1 - f(x^{(i)}))$$

we obtain that Inequality (3) is satisfied, which concludes the proof. □

The theorem essentially states sufficient conditions for the case that after binarization and enforcing of several parity constraints, the values of the original variables remain feasible. Note that in order to satisfy $\sum_{i \in I} \gamma(z_i) \geq 1$ for a constraint on variable set I, it suffices that two of the participating variables have a distance to their respective bounds of at least $\frac{1}{2}$ (see Lemma 6), which is not very restrictive for nonbinary variables.

Implications for the graphic TSP. We consider an optimum solution $\hat{z} \in [0, 2]^E$ of the LP relaxation (4)–(6) of the model introduced in Section 6. The requirements for Theorem 7 are satisfied if and only if for every nontrivial cut $\delta(S)$ ($S \subsetneq V$, $S \neq \emptyset$), the inequality $\sum_{e \in \delta(S)} \gamma(\hat{z}_e) \geq 1$ is satisfied. This is in particular the case if every cut contains two edges e, f with $\frac{1}{2} \leq \hat{z}_e, \hat{z}_f \leq \frac{3}{2}$. Note that integrality of the z-variables does not play a role here, i.e., even if $\hat{z}$ is integral, and $\hat{z}(\delta(S))$ has the wrong parity for some set S, then there may exist a (fractional) assignment for x-variables that is feasible for Constraints (10). For the IP model this means that the Blossom Inequalities only become useful if relevant z-variables are near their bounds or if branching or cutting restricted the x-variables further.

Acknowledgments

References

[1] D. Avis, A Revised Implementation of the Reverse Search Vertex Enumeration Algorithm, in: G. Kalai, G.M. Ziegler (Eds.), Polytopes−Combinatorics and Computation, Birkhäuser Basel, 2000, pp. 177–198.

[2] P. Bonami, F. Margot, Cut Generation through Binarization, in: J. Lee, J. Vygen (Eds.), Integer Programming and Combinatorial Optimization: 17th International Conference, IPCO 2014, Bonn, Germany, June 23–25, 2014. Proceedings, Springer International Publishing, 2014, pp. 174–185.

[3] R.D. Carr, G. Konjevod, Polyhedral combinatorics, in: H.J. Greenberg (Ed.), Tutorials on Emerging Methodologies and Applications in Operations Research, in: International Series in Operations Research & Management Science, vol. 76, Springer, 2005, pp. 1–46 (Chapter 2).

[4] R. Jeroslow, On defining sets of vertices of the hypercube by linear inequalities, Discrete Math. 11 (2) (1975) 119–124.

[5] A.N. Letchford, G. Reinelt, D.O. Theis, Odd minimum cut sets and b-matchings revisited, SIAM J. Discrete Math. 22 (4) (2008) 1480–1487.

[6] M.W. Padberg, M.R. Rao, Odd minimum cut-sets and b-matchings, Math. Oper. Res. 7 (1) (1982) 67–80.

[7] J.-S. Roy, "Binarize and Project" to generate cuts for general mixed-integer programs, Algorithmic Oper. Res. 2 (1) (2007) 37–51.

[8] A. Schrijver, Theory of Linear and Integer Programming, John Wiley & Sons, Inc., New York, NY, USA, 1986.

[9] A. Sebő, J. Vygen, Shorter tours by nicer ears: 7/5-approximation for the graph-TSP, 3/2 for the path version, and 4/3 for two-edge-connected subgraphs, Combinatorica (2014) 1–34.