Parity Polytopes and Binarization

Dominik Ermel & Matthias Walter

15th Cologne Twente Workshop on Graphs and Combinatorial Optimization,
Cologne 2017
Reformulating integer variables with binary ones:

- Consider an integer variable \( z \in \{0, 1, \ldots, n - 1, n\} \).
- Idea: Write \( z \) as the projection of some 0/1-polytope.
- Goals: Cutting planes or \textit{modeling} (e.g., to exclude holes in the domain)
Reformulating integer variables with binary ones:

- Consider an integer variable \( z \in \{0, 1, \ldots, n - 1, n\} \).
- Idea: Write \( z \) as the projection of some 0/1-polytope.
- Goals: Cutting planes or modeling (e.g., to exclude holes in the domain)
- Variants:

\[
k := \lfloor \log_2(n) \rfloor, \quad x \in \{0, 1\}^k, \quad \leadsto z = \sum_{i=0}^{k} 2^i x_i
\] (1)
Reformulating integer variables with binary ones:

- Consider an integer variable $z \in \{0, 1, \ldots, n - 1, n\}$.
- Idea: Write $z$ as the projection of some 0/1-polytope.
- Goals: Cutting planes or modeling (e.g., to exclude holes in the domain)
- Variants:

$$k := \lceil \log_2(n) \rceil, \quad x \in \{0, 1\}^k, \quad \sim z = \sum_{i=0}^{k} 2^i x_i$$ (1)

$$1 \geq x_1 \geq x_2 \geq \ldots \geq x_{n-1} \geq x_n \geq 0, \quad x \in \{0, 1\}^n, \quad \sim z = \sum_{k=1}^{n} x_k$$ (2)

- Variant (1) is more compact, but yields a weaker relaxation.
- Today: focus on (2): Let $X_{\text{ord}}^{n}$ be the set of $x$ with (2).
Application: Graphic TSP

- **Input:** Graph $G = (V, E)$
- **Output:** Minimum length closed walk visiting each node (at least once).
Application: Graphic TSP

- **Input:** Graph $G = (V, E)$
- **Output:** Minimum length closed walk visiting each node (at least once).
**Application: Graphic TSP**

- **Input:** Graph $G = (V, E)$
- **Output:** Minimum length closed walk visiting each node (at least once).

Can be solved as TSP with $O(|V|^2)$ variables (via metric closure).
Application: Graphic TSP

- **Input:** Graph $G = (V, E)$
- **Output:** Minimum length closed walk visiting each node (at least once).

Can be solved as TSP with $O(|V|^2)$ variables (via metric closure).

With only $O(|E| + |V|)$ variables:

$$\min \quad z(E)$$  \hspace{1cm} (3)

$$\text{s.t.} \quad z(\delta(S)) \geq 2 \quad \text{for all } \emptyset \neq S \subseteq V$$  \hspace{1cm} (4)

$$z_e \geq 0 \quad \text{for all } e \in E$$  \hspace{1cm} (5)

$$z(\delta(v)) = 2y_v \quad \text{for all } v \in V$$  \hspace{1cm} (6)

$$y_v \in \mathbb{Z} \quad \text{for all } v \in V$$  \hspace{1cm} (7)

$$z_e \in \mathbb{Z} \quad \text{for all } e \in E$$  \hspace{1cm} (8)
Binarization meets Parity

**Ordered binary vectors:**

- $X_{ord}^n$: set of all binary vectors $x$ of length $n$ of type $(1, \ldots, 1, 0, \ldots, 0)$. 
Binarization meets Parity

Ordered binary vectors:

- $X^n_{ord}$: set of all binary vectors $x$ of length $n$ of type $(1, \ldots, 1, 0, \ldots, 0)$.
- $P^n_{ord} := \text{conv}(X^n_{ord})$ is described by $1 \geq x_1 \geq x_2 \geq \ldots \geq x_{n-1} \geq x_n \geq 0$. 
Ordered binary vectors:
- $X^n_{ord}$: set of all binary vectors $x$ of length $n$ of type $(1, \ldots, 1, 0, \ldots, 0)$.
- $P^n_{ord} := \text{conv}(X^n_{ord})$ is described by $1 \geq x_1 \geq x_2 \geq \ldots \geq x_{n-1} \geq x_n \geq 0$.

With (even) parity constraint:
- Consider $k$ “blocks” of binarization variables, $i$’th one having length $r_i$.
- $P^r_{\text{even}} := \text{conv} \left\{ (x^{(1)}, \ldots, x^{(k)}) \in X^{r_1}_{ord} \times \ldots \times X^{r_k}_{ord} \mid \sum_{i=1}^{k} \sum_{j=1}^{r_i} x^{(i)}_{j} \text{ even} \right\}$
Binarization meets Parity

Ordered binary vectors:
- $X^n_{\text{ord}}$: set of all binary vectors $x$ of length $n$ of type $(1, \ldots, 1, 0, \ldots, 0)$.
- $P^n_{\text{ord}} := \text{conv}(X^n_{\text{ord}})$ is described by $1 \geq x_1 \geq x_2 \geq \ldots \geq x_{n-1} \geq x_n \geq 0$.

With (even) parity constraint:
- Consider $k$ “blocks” of binarization variables, $i$’th one having length $r_i$.
- $P^r_{\text{even}} := \text{conv}\left\{ (x^{(1)}, \ldots, x^{(k)}) \in X^{r_1}_{\text{ord}} \times \ldots \times X^{r_k}_{\text{ord}} \mid \sum_{i=1}^k \sum_{j=1}^{r_i} x_j^{(i)} \text{ even} \right\}$
- Note: Convexification of $(z_1, \ldots, z_k)$ with even $\sum_{i=1}^k z_i$ does not work:
  \[ 1 \in \text{conv}\{0, 2\} \]
Jeroslow, 1975: \( P_{\text{even}}^1 \) is described by \( 0 \leq x \leq 1 \) and

\[
\sum_{i \in [n] \setminus F} x_i + \sum_{i \in F} (1 - x_i) \geq 1 \text{ for all } F \subseteq [n] \text{ with } |F| \text{ odd}.
\]
Jeroslow, 1975: $P_{\text{even}}$ is described by $0 \leq x \leq 1$ and

$$\sum_{i \in [n] \setminus F} x_i + \sum_{i \in F} (1 - x_i) \geq 1$$

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

Observation 1: For $X_{\text{ord}}^n$, parity can be measured with a linear function $f$:

$$f(x) := x_1 - x_2 + x_3 - x_4 + \ldots \mp x_{n-1} \pm x_n$$

$$\begin{align*}
(0, & \quad 0, \quad 0, \quad 0, \ldots, \quad 0, \quad 0) \mapsto 0 \\
(1, & \quad 0, \quad 0, \quad 0, \ldots, \quad 0, \quad 0) \mapsto 1 \\
(1, & \quad 1, \quad 0, \quad 0, \ldots, \quad 0, \quad 0) \mapsto 0 \\
(1, & \quad 1, \quad 1, \quad 0, \ldots, \quad 0, \quad 0) \mapsto 1 \\
\vdots & \quad \vdots 
\end{align*}$$
Jeroslow, 1975: $P_{\text{even}}^1$ is described by $0 \leq x \leq 1$ and

$$\sum_{i \in [n] \setminus F} x_i + \sum_{i \in F} (1 - x_i) \geq 1 \text{ for all } F \subseteq [n] \text{ with } |F| \text{ odd}. $$

Observation 1: For $X_{\text{ord}}^n$, parity can be measured with a linear function $f$:

$$f(x) := x_1 - x_2 + x_3 - x_4 + \ldots \mp x_{n-1} \pm x_n$$

$$(0, 0, 0, 0, \ldots, 0, 0) \leadsto 0$$
$$(1, 0, 0, 0, \ldots, 0, 0) \leadsto 1$$
$$(1, 1, 0, 0, \ldots, 0, 0) \leadsto 0$$
$$(1, 1, 1, 0, \ldots, 0, 0) \leadsto 1$$

$$\vdots \quad \vdots$$

Observation 2: 0/1-polytopes can be glued together at a single coordinate.
Jeroslow, 1975: \( P_{\text{even}}^{\uparrow} \) is described by \( 0 \leq x \leq 1 \) and

\[
\sum_{i \in [n] \setminus F} x_i + \sum_{i \in F} (1 - x_i) \geq 1 \text{ for all } F \subseteq [n] \text{ with } |F| \text{ odd}.
\]

Observation 1: For \( X_{\text{ord}}^n \), parity can be measured with a linear function \( f \):

\[
f(x) := x_1 - x_2 + x_3 - x_4 + \ldots \mp x_{n-1} \pm x_n
\]

\[
(0, 0, 0, 0, \ldots, 0, 0) \sim 0
\]

\[
(1, 0, 0, 0, \ldots, 0, 0) \sim 1
\]

\[
(1, 1, 0, 0, \ldots, 0, 0) \sim 0
\]

\[
(1, 1, 1, 0, \ldots, 0, 0) \sim 1
\]

\[
\vdots
\]

Observation 2: 0/1-polytopes can be glued together at a single coordinate.

Main idea: Extend each binarization block with parity bit \( f(x) \) and glue all of them together at these bits.
Glueing at a single coordinate

Observation 2 more pictorially:

\[
\begin{array}{ccc}
\times & \lambda & y \\
\hline
\text{integral} & 0 & \lambda \in [0, 1] \\
\text{intersection integral} & 0 & \text{integral}
\end{array}
\]
Glueing at a single coordinate

Observation 2 more pictorially:

\[
\begin{array}{c|c|c|}
\times & \lambda & \gamma \\
\hline
0 & 0 & x \\
\hline
0 & \lambda & y \\
\end{array}
\]

\[\lambda \in [0,1]\]

\[\Rightarrow \text{intersection integral}\]

Observation 2 more formally: Let \(X_0, X_1 \subseteq \mathbb{R}^m\) and \(Y_0, Y_1 \subseteq \mathbb{R}^n\) be finite sets. Then

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

is equal to the intersection of

\[
\text{conv} ((X_0 \times \{0\}) \cup (X_1 \times \{1\})) \times \mathbb{R}^n
\]

and

\[
\mathbb{R}^m \times \text{conv} ((\{0\} \times Y_0) \cup (\{1\} \times Y_1)).
\]
Glueing at a single coordinate

Observation 2 more pictorially:

\[
\begin{array}{ccc}
\times & \lambda & y \\
\hline
0 & 0 & \lambda \\
\end{array}
\]

\[\lambda \in [0, 1]\]

integral

\[\Rightarrow \text{intersection integral}\]

integral

Observation 2 more formally: Let \(X_0, X_1 \subseteq \mathbb{R}^m\) and \(Y_0, Y_1 \subseteq \mathbb{R}^n\) be finite sets. Then

\[
\text{conv} \left( (X_0 \times \{0\} \times Y_0) \cup (X_1 \times \{1\} \times Y_1) \right)
\]

is equal to the intersection of

\[
\text{conv} \left( (X_0 \times \{0\}) \cup (X_1 \times \{1\}) \right) \times \mathbb{R}^n
\]

and

\[
\mathbb{R}^m \times \text{conv} \left( (\{0\} \times Y_0) \cup (\{1\} \times Y_1) \right)
\]

Proof: Convex multipliers in dimension 1 are unique.
Result for Description

Reminder:

- $X^n_{ord}$: binary vectors $x$ of length $n$ of type $(1, \ldots, 1, 0, \ldots, 0)$.
- $f(x) := x_1 - x_2 + x_3 - x_4 \ldots$ measures parity if $x \in X^n_{ord}$.
- $P^r_{even} := \text{conv} \left\{ (x^{(1)}, \ldots, x^{(k)}) \in X^{r_1}_{ord} \times \ldots \times X^{r_k}_{ord} \mid \sum_{i=1}^{k} \sum_{j=1}^{r_i} x_j^{(i)} \text{ even} \right\}$
Reminder:

- $X_{\text{ord}}^n$: binary vectors $x$ of length $n$ of type $(1, \ldots, 1, 0, \ldots, 0)$.
- $f(x) := x_1 - x_2 + x_3 - x_4 \ldots$ measures parity if $x \in X_{\text{ord}}^n$.
- $P_{\text{even}}^r := \text{conv} \left\{ (x^{(1)}, \ldots, x^{(k)}) \in X_{\text{ord}}^{r_1} \times \ldots \times X_{\text{ord}}^{r_k} \mid \sum_{i=1}^{k} \sum_{j=1}^{r_i} x_j^{(i)} \text{ even} \right\}$

Theorem: Let $r \in \mathbb{N}^k$. Then $P_{\text{even}}^r$ is described by

- $1 \geq x_1^{(i)} \geq x_2^{(i)} \geq \ldots \geq x_{r_i}^{(i)} \geq 0$ for each $i \in [k]$.
- $\sum_{i \in [k] \setminus F} f(x^{(i)}) + \sum_{i \in F} (1 - f(x^{(i)})) \geq 1$ for all $F \subseteq [k]$ with $|F|$ odd.
Reminder:

- $X^n_{\text{ord}}$: binary vectors $x$ of length $n$ of type $(1, \ldots, 1, 0, \ldots, 0)$.
- $f(x) := x_1 - x_2 + x_3 - x_4 \ldots$ measures parity if $x \in X^n_{\text{ord}}$.
- $P^r_{\text{even}} := \text{conv} \left\{ (x^{(1)}, \ldots, x^{(k)}) \in X^{r_1}_{\text{ord}} \times \ldots \times X^{r_k}_{\text{ord}} \mid \sum_{i=1}^{k} \sum_{j=1}^{r_i} x^{(i)}_j \text{ even} \right\}$

Theorem: Let $r \in \mathbb{N}^k$. Then $P^r_{\text{even}}$ is described by

- $1 \geq x_1^{(i)} \geq x_2^{(i)} \geq \ldots \geq x_{r_i}^{(i)} \geq 0$ for each $i \in [k]$.
- $\sum_{i \in [k] \setminus F} f(x^{(i)}) + \sum_{i \in F} (1 - f(x^{(i)})) \geq 1$ for all $F \subseteq [k]$ with $|F|$ odd.

Proof:

- Add parity variables for each block: $(x^{(1)}, y_1, x^{(2)}, y_2, \ldots, x^{(k)}, y_k)$.
- Isomorphism via $y_i := f(x^{(i)})$ (by linearity of $f$).
- Enforce parity polytope constraints on $y$-variables.
- Interaction of blocks with these is limited to the single $y$-variable per block.
- Apply Observation 2 (glueing trick).
Separation problem:

Given \((\hat{x}^{(1)}, \ldots, \hat{x}^{(k)}) \in X_{\text{ord}}^r \times \ldots \times X_{\text{ord}}^r\), is there an \(F \subseteq [k]\) with \(|F|\) odd and

\[
\sum_{i \in [k] \setminus F} f(\hat{x}^{(i)}) + \sum_{i \in F} (1 - f(\hat{x}^{(i)})) < 1?
\]
Separation problem:

- Given \((\hat{x}^{(1)}, \ldots, \hat{x}^{(k)}) \in X^{r_1}_{\text{ord}} \times \ldots \times X^{r_k}_{\text{ord}}\), is there an \(F \subseteq [k]\) with \(|F|\) odd and

  \[
  \sum_{i \in [k]\setminus F} f(\hat{x}^{(i)}) + \sum_{i \in F} (1 - f(\hat{x}^{(i)})) < 1?
  \]

- Easy: Compute \(\hat{y}_i := f(\hat{x})\) for all \(i \in [k]\) and call (linear-time) parity polytope separation.
Separation problem:

- Given \((\hat{x}^{(1)}, \ldots, \hat{x}^{(k)}) \in X_{\text{ord}}^{r_1} \times \ldots \times X_{\text{ord}}^{r_k}\), is there an \(F \subseteq [k]\) with \(|F|\) odd and

\[
\sum_{i \in [k] \setminus F} f(\hat{x}^{(i)}) + \sum_{i \in F} (1 - f(\hat{x}^{(i)})) < 1?
\]

- Easy: Compute \(\hat{y}_i := f(\hat{x})\) for all \(i \in [k]\) and call (linear-time) parity polytope separation.

Odd parities:

- \(P_{\text{odd}}^{r}\) is projection of face of \(P_{\text{even}}^{r'}\) for \(r' = (r_1, \ldots, r_k, 1)\) with \(x^{(k+1)} = 1\).
Separation problem:

- Given \((\hat{x}^{(1)}, \ldots, \hat{x}^{(k)}) \in X_{\text{ord}}^{r_1} \times \ldots \times X_{\text{ord}}^{r_k}\), is there an \(F \subseteq [k]\) with \(|F|\) odd and
\[
\sum_{i \in [k] \setminus F} f(\hat{x}^{(i)}) + \sum_{i \in F} (1 - f(\hat{x}^{(i)})) < 1?
\]

- Easy: Compute \(\hat{y}_i := f(\hat{x})\) for all \(i \in [k]\) and call (linear-time) parity polytope separation.

Odd parities:

- \(P_{\text{odd}}^r\) is projection of face of \(P_{\text{even}}^{r'}\) for \(r' = (r_1, \ldots, r_k, 1)\) with \(x^{(k+1)} = 1\).

- Similar result (\(|F|\) even), obtained by projecting (Fourier-Motzkin).
Extended Formulations

\[ i = 1 \quad i = 2 \quad i = 3 \quad \ldots \quad i = n - 1 \quad i = n \]
Extended Formulations

\[ r_i = 1: \]
\[ x_1^{(i)} = \sum y_{(\ldots)} \]

\[ \begin{array}{ccc}
(i - 1, 0) & (i, 0) & (i, 1) \\
(i - 1, 1) & (i, 1) & \\
\end{array} \]
Extended Formulations

\[
\begin{align*}
(i - 1, 0) & \quad (i, 0) \\
(i - 1, 1) & \quad (i, 1) \\
\end{align*}
\]

\[
x_3^{(i)} = \sum y_{(i-1,0,0)}
\]

\[
x_2^{(i)} = \sum y_{(i-1,1,0)} + x_3^{(i)}
\]

\[
x_1^{(i)} = \sum y_{(i-1,0,1)} + x_2^{(i)}
\]

\[
\begin{align*}
\ & \quad i = 1 \\
\ & \quad i = 2 \\
\ & \quad i = 3 \\
\ & \quad i = n - 1 \\
\ & \quad i = n
\end{align*}
\]
Back to Graphic TSP

- **Input:** Graph $G = (V, E)$
- **Output:** Minimum length closed walk visiting each node (at least once).
Back to Graphic TSP

- **Input:** Graph $G = (V, E)$
- **Output:** Minimum length closed walk visiting each node (at least once).

Every feasible solution intersects every cut in an **even** number of edges.
Back to Graphic TSP

- **Input:** Graph $G = (V, E)$
- **Output:** Minimum length closed walk visiting each node (at least once).

Every feasible solution intersects every cut in an **even** number of edges.

**Generalized $T$-join inequality:**

$$\sum_{e \in \delta(S) \setminus F} f(x_e) + \sum_{e \in \delta(S) \cap F} (1 - f(x_e)) \geq 1$$

for all $\emptyset \neq S \subseteq V$ and all $F \subseteq \delta(S)$ with $|S|$ odd.
Back to Graphic TSP

- **Input**: Graph $G = (V, E)$
- **Output**: Minimum length closed walk visiting each node (at least once).

Every feasible solution intersects every cut in an even number of edges.

Generalized $T$-join inequality:

$$\sum_{e \in \delta(S) \setminus F} f(x_e) + \sum_{e \in \delta(S) \cap F} (1 - f(x_e)) \geq 1$$

for all $\emptyset \neq S \subseteq V$ and all $F \subseteq \delta(S)$ with $|S|$ odd.

- **Separation algorithm** by Letchford, Reinelt & Theis can be reused.
Observations:

- In practice, lots of cuts are separated, but no bound improvement over subtour inequalities!
- Characteristic property: Parity constraint is the only one that depends on binary variables. Remaining inequalities can be formulated in integer variables.
Observations:

- In practice, lots of cuts are separated, but no bound improvement over subtour inequalities!
- Characteristic property: Parity constraint is the only one that depends on binary variables. Remaining inequalities can be formulated in integer variables.

Idea:

- We fix the integer variables $z \in [0, 2]^E$.
- Try to modify $x \in [0, 1]^{E \times \{1, 2\}}$ such that it satisfies all parity constraints?

$$\sum_{e \in \delta(S) \setminus F} f(x_e) + \sum_{e \in \delta(S) \cap F} (1 - f(x_e)) \geq 1$$

- Try to modify such that $f(x_e)$ and $1 - f(x_e)$ are large enough, i.e., $f(x_e) \approx \frac{1}{2}$. 
Betraying with fractional solutions

Setup:

- Consider (integer) variable $z \in [0, r]$ and
- the binarization $z = \sum_{i=1}^{r} x_i$ with $1 \geq x_1 \geq x_2 \geq \ldots \geq x_r \geq 0$. 
Betraying with fractional solutions

Setup:

- Consider (integer) variable $z \in [0, r]$ and
- the binarization $z = \sum_{i=1}^{r} x_i$ with $1 \geq x_1 \geq x_2 \geq \ldots \geq x_r \geq 0$.
- We fix $z$ but allow changes to $x_1, \ldots, x_r$. 

Auxiliary problem:

$$\min |f_p x_q|$$
subject to $x_P P$, $P$ and $r$ $\sum_{i=1}^{r} x_i = z$ (9)

Lemma: The optimal value of (9) is 

- $z$ if $z \leq 1/2$
- $\frac{r - z}{2}$ if $z \leq \frac{r}{2}$
- otherwise

Consequence:

- Satisfied if only two variables participating in the parity constraint are $1/2$ away from their respective bounds.
Betraying with fractional solutions

Setup:

- Consider (integer) variable $z \in [0, r]$ and
- the binarization $z = \sum_{i=1}^{r} x_i$ with $1 \geq x_1 \geq x_2 \geq \ldots \geq x_r \geq 0$.
- We fix $z$ but allow changes to $x_1, \ldots, x_r$.

Auxiliary problem:

$$\min |f(x) - \frac{1}{2}| \text{ subject to } x \in P_{ord}^r \text{ and } \sum_{i=1}^{r} x_i = z$$

(9)
Betraying with fractional solutions

Setup:

- Consider (integer) variable $z \in [0, r]$ and
- the binarization $z = \sum_{i=1}^{r} x_i$ with $1 \geq x_1 \geq x_2 \geq \ldots \geq x_r \geq 0$.
- We fix $z$ but allow changes to $x_1, \ldots, x_r$.

Auxiliary problem:

\[
\min |f(x) - \frac{1}{2}| \quad \text{subject to } x \in P_{\text{ord}}^r \text{ and } \sum_{i=1}^{r} x_i = z
\]  

(9)

Lemma: The optimal value of (9) is

\[
\begin{cases} 
  z & \text{if } z \leq \frac{1}{2} \\
  r - z & \text{if } z \geq r - \frac{1}{2} \\
  \frac{1}{2} & \text{otherwise}
\end{cases}
\]
Betraying with fractional solutions

Setup:
- Consider (integer) variable $z \in [0, r]$ and
- the binarization $z = \sum_{i=1}^{r} x_i$ with $1 \geq x_1 \geq x_2 \geq \ldots \geq x_r \geq 0$.
- We fix $z$ but allow changes to $x_1, \ldots, x_r$.

Auxiliary problem:

$$\min |f(x) - \frac{1}{2}| \text{ subject to } x \in P_{\text{ord}}^r \text{ and } \sum_{i=1}^{r} x_i = z \quad (9)$$

Lemma: The optimal value of (9) is

$$\begin{cases} 
z & \text{if } z \leq \frac{1}{2} \\
r - z & \text{if } z \geq r - \frac{1}{2} \\
\frac{1}{2} & \text{otherwise}
\end{cases}$$

Consequence:
- Satisfied if only two variables participating in the parity constraint are $\frac{1}{2}$ away from their respective bounds.
Not useful at all!

- Parity is wrong: odd number of cut edges used.
- But: All edge variables are $1 \in [0, 2]$, so lemma from previous slide applies!
Not useful at all!

- Parity is wrong: odd number of cut edges used.
- But: All edge variables are $1 \in [0, 2]$, so lemma from previous slide applies!

**What did we learn:**
- Complete description of binarization plus parity constraint.
- Not useful to binarize in order to add parity constraints!
Not useful at all!

- Parity is wrong: odd number of cut edges used.
- But: All edge variables are $1 \in [0, 2]$, so lemma from previous slide applies!

What did we learn:

- Complete description of binarization plus parity constraint.
- Not useful to binarize in order to add parity constraints!

Thank you for your attention!