Parity Polytopes and Binarization

Dominik Ermel & Matthias Walter

15th Cologne Twente Workshop
on Graphs and Combinatorial Optimization,

Cologne 2017

Operations Research
RWTH AACHEN UNIVERSITY

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)

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:

$$k := \lfloor \log_2(n) \rfloor, \quad x \in \{0,1\}^k, \quad \rightsquigarrow z = \sum_{i=0}^{k} 2^i x_i \tag{1}$$

Reformulating integer variables with binary ones:

- Consider an integer variable $z \in \{0, 1, \dots, n-1, n\}$.
- 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:

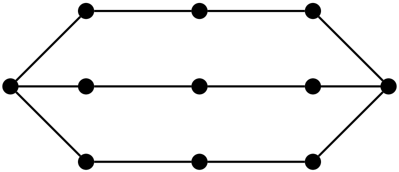
$$k := \lfloor \log_2(n) \rfloor, \quad x \in \{0,1\}^k, \quad \rightsquigarrow z = \sum_{i=0}^{k} 2^i x_i \tag{1}$$

$$1 \geqslant x_1 \geqslant x_2 \geqslant \dots \geqslant x_{n-1} \geqslant x_n \geqslant 0, \quad x \in \{0,1\}^n, \quad \rightsquigarrow z = \sum_{k=1}^{n} x_k \tag{2}$$

- Variant (1) is more compact, but yields a weaker relaxation.
- Today: focus on (2): Let X^n_{ord} be the set of x with (2).

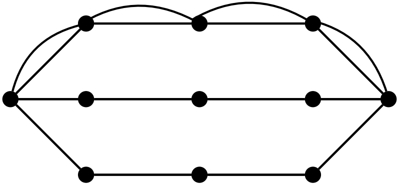
Application: Graphic TSP

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



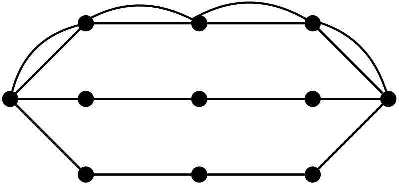
Application: Graphic TSP

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



Application: Graphic TSP

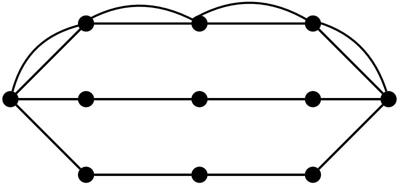
- **Input:** Graph $G = (V, E)$
- **Output:** Minimum length closed walk visiting each node (at least once).



- Can be solved as TSP with $\mathcal{O}\left(|V|^2\right)$ 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 $\mathcal{O}\left(|V|^2\right)$ variables (via metric closure).
- With only $\mathcal{O}\left(|E| + |V|\right)$ variables:

$$
\begin{aligned}
\min \quad & z(E) && && (3)\\
\text{s.t.} \quad z(\delta(S)) &\geqslant 2 && \text{for all } \emptyset \neq S \subsetneq V && (4)\\
z_e &\geqslant 0 && \text{for all } e \in E && (5)\\
z(\delta(v)) &= 2y_v && \text{for all } v \in V && (6)\\
y_v &\in \mathbb{Z} && \text{for all } v \in V && (7)\\
z_e &\in \mathbb{Z} && \text{for all } e \in E && (8)
\end{aligned}
$$

Ordered binary vectors:

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

Ordered binary vectors:

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

Ordered binary vectors:

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

With (even) parity constraint:

- Consider k "blocks" of binarization variables, i'th one having length r_i.
- $P^r_{\text{even}} := \text{conv}\left\{(x^{(1)}, \ldots, x^{(k)}) \in X^{r_1}_{\text{ord}} \times \ldots \times X^{r_k}_{\text{ord}} \mid \sum_{i=1}^{k} \sum_{j=1}^{r_i} x^{(i)}_j \text{ even}\right\}$

Ordered binary vectors:

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

With (even) parity constraint:

- Consider k "blocks" of binarization variables, i'th one having length r_i.
- $P^r_{\text{even}} := \text{conv}\left\{(x^{(1)},\ldots,x^{(k)}) \in X^{r_1}_{\text{ord}} \times \ldots \times X^{r_k}_{\text{ord}} \mid \sum_{i=1}^{k}\sum_{j=1}^{r_i} x^{(i)}_j \text{ even}\right\}$
- Note: Convexification of $(z_1,\ldots,z_k)$ with even $\sum_{i=1}^k z_i$ does not work:

$$1 \in \text{conv}\,\{0,2\}$$

Parity

Jeroslow, 1975: $P^{\mathbb{1}}_{\text{even}}$ is described by $\mathbb{0} \leqslant x \leqslant \mathbb{1}$ and

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

Parity

Jeroslow, 1975: $P^{\mathbb{1}}_{\text{even}}$ is described by $\mathbb{O} \leqslant x \leqslant \mathbb{1}$ and

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

Observation 1: For X^n_{ord}, parity can be measured with a linear function f:

$$
\begin{array}{lllllllll}
f(x) := x_1 - x_2 + x_3 - x_4 + \ldots \mp x_{n-1} \pm x_n \\
(0, & 0, & 0, & 0, & \ldots, & 0, & 0) \rightsquigarrow 0 \\
(1, & 0, & 0, & 0, & \ldots, & 0, & 0) \rightsquigarrow 1 \\
(1, & 1, & 0, & 0, & \ldots, & 0, & 0) \rightsquigarrow 0 \\
(1, & 1, & 1, & 0, & \ldots, & 0, & 0) \rightsquigarrow 1 \\
\vdots & & & & & & \vdots
\end{array}
$$

Jeroslow, 1975: $P^{\mathbb{1}}_{\text{even}}$ is described by $\mathbb{0} \leqslant x \leqslant \mathbb{1}$ and

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

Observation 1: For X^n_{ord}, parity can be measured with a linear function f:

$$\begin{array}{rl} f(x) := & x_1 - x_2 + x_3 - x_4 + \ldots \mp x_{n-1} \pm x_n \\ & (0, \quad 0, \quad 0, \quad 0, \quad \ldots, \quad 0, \quad 0) \rightsquigarrow 0 \\ & (1, \quad 0, \quad 0, \quad 0, \quad \ldots, \quad 0, \quad 0) \rightsquigarrow 1 \\ & (1, \quad 1, \quad 0, \quad 0, \quad \ldots, \quad 0, \quad 0) \rightsquigarrow 0 \\ & (1, \quad 1, \quad 1, \quad 0, \quad \ldots, \quad 0, \quad 0) \rightsquigarrow 1 \\ & \quad \vdots \qquad\qquad\qquad\qquad\qquad\qquad \vdots \end{array}$$

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

Jeroslow, 1975: $P^{\mathbb{1}}_{\text{even}}$ is described by $\mathbb{0} \leqslant x \leqslant \mathbb{1}$ and

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

Observation 1: For X^n_{ord}, parity can be measured with a linear function f:

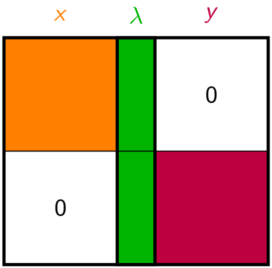
$$\begin{array}{rl} f(x) := & x_1 - x_2 + x_3 - x_4 + \ldots \mp x_{n-1} \pm x_n \\ & (0, \quad 0, \quad 0, \quad 0, \quad \ldots, \quad 0, \quad 0) \rightsquigarrow 0 \\ & (1, \quad 0, \quad 0, \quad 0, \quad \ldots, \quad 0, \quad 0) \rightsquigarrow 1 \\ & (1, \quad 1, \quad 0, \quad 0, \quad \ldots, \quad 0, \quad 0) \rightsquigarrow 0 \\ & (1, \quad 1, \quad 1, \quad 0, \quad \ldots, \quad 0, \quad 0) \rightsquigarrow 1 \\ & \quad \vdots \qquad\qquad\qquad\qquad\qquad \vdots \end{array}$$

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 pictorically:



$\lambda \in [0, 1]$

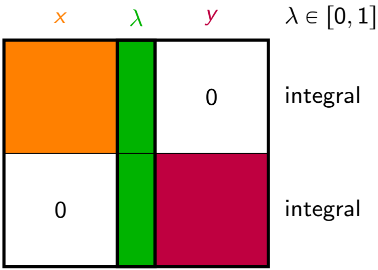
integral

integral

$\Rightarrow$ intersection integral

Glueing at a single coordinate

Observation 2 more pictorically:



$\Rightarrow$ 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

$$\operatorname{conv}\left((X_0 \times \{0\} \times Y_0) \cup (X_1 \times \{1\} \times Y_1)\right)$$

is equal to the intersection of

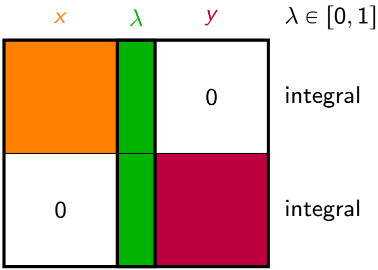
$$\operatorname{conv}\left((X_0 \times \{0\}) \cup (X_1 \times \{1\})\right) \times \mathbb{R}^n$$

and

$$\mathbb{R}^m \times \operatorname{conv}\left((\{0\} \times Y_0) \cup (\{1\} \times Y_1)\right).$$

Glueing at a single coordinate

Observation 2 more pictorially:



$\Rightarrow$ 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

$$\operatorname{conv}\left((X_0 \times \{0\} \times Y_0) \cup (X_1 \times \{1\} \times Y_1)\right)$$

is equal to the intersection of

$$\operatorname{conv}\left((X_0 \times \{0\}) \cup (X_1 \times \{1\})\right) \times \mathbb{R}^n$$

and

$$\mathbb{R}^m \times \operatorname{conv}\left((\{0\} \times Y_0) \cup (\{1\} \times Y_1)\right).$$

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 $x \in X^n_{\text{ord}}$.
- $P^r_{\text{even}} := \text{conv}\left\{ (x^{(1)}, \ldots, x^{(k)}) \in X^{r_1}_{\text{ord}} \times \ldots \times X^{r_k}_{\text{ord}} \;\middle|\; \sum_{i=1}^{k} \sum_{j=1}^{r_i} x^{(i)}_j \text{ even} \right\}$

Result for Description

Reminder:

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

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

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

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 $x \in X^n_{\text{ord}}$.
- $P^r_{\text{even}} := \operatorname{conv}\left\{ (x^{(1)},\ldots,x^{(k)}) \in X^{r_1}_{\text{ord}} \times \ldots \times X^{r_k}_{\text{ord}} \mid \sum_{i=1}^{k}\sum_{j=1}^{r_i} x^{(i)}_j \text{ even} \right\}$

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

- $1 \geqslant x^{(i)}_1 \geqslant x^{(i)}_2 \geqslant \ldots \geqslant x^{(i)}_{r_i} \geqslant 0$ for each $i \in [k]$.
- $\sum_{i\in[k]\setminus F} f(x^{(i)}) + \sum_{i \in F}(1 - f(x^{(i)})) \geqslant 1$ 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 / Odd parities

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

Separation Problem / Odd parities

Separation problem:

- Given $(\hat{x}^{(1)}, \dots, \hat{x}^{(k)}) \in X_{\text{ord}}^{r_1} \times \dots \times X_{\text{ord}}^{r_k}$, 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.

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

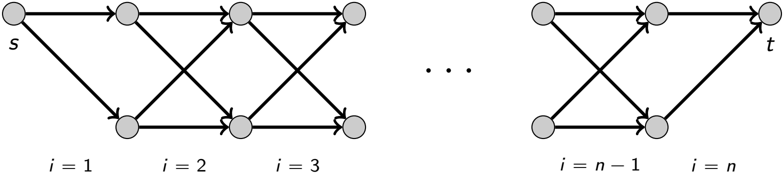
Separation problem:

- Given $(\hat{x}^{(1)}, \ldots, \hat{x}^{(k)}) \in X^{r_1}_{\text{ord}} \times \ldots \times X^{r_k}_{\text{ord}}$, 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.

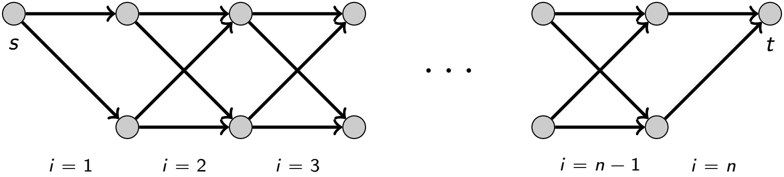
Odd parities:

- P^r_{odd} is projection of face of $P^{r'}_{\text{even}}$ for $r' = (r_1, \ldots, r_k, 1)$ with $x^{(k+1)} = 1$.
- Similar result ($|F|$ even), obtained by projecting (Fourier-Motzkin).

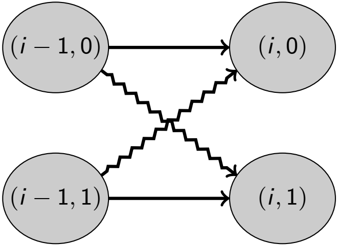
Extended Formulations



Extended Formulations

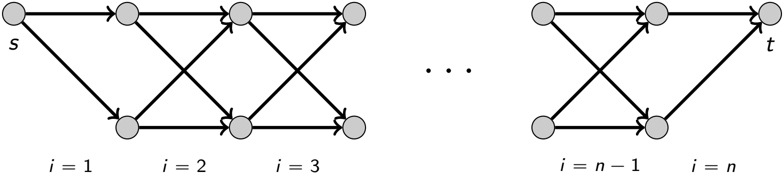


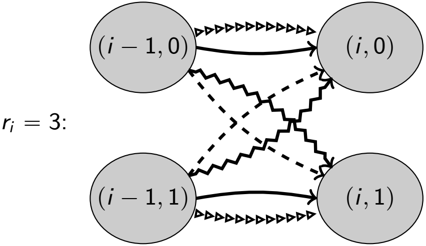
$r_i = 1$:



$x_1^{(i)} = \sum y_{(\rightsquigarrow)}$

Extended Formulations





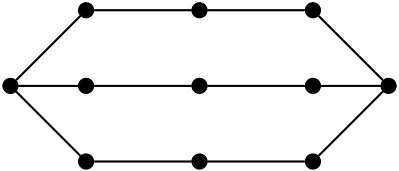
$$x_3^{(i)} = \sum y_{(- -)}$$

$$x_2^{(i)} = \sum y_{(\triangleright\triangleright\triangleright)} + x_3^{(i)}$$

$$x_1^{(i)} = \sum y_{(\rightsquigarrow)} + x_2^{(i)}$$

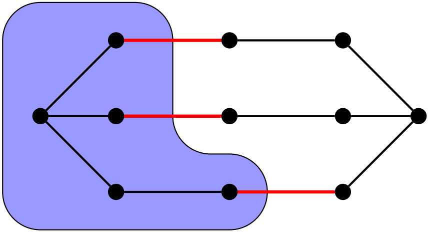
Back to Graphic TSP

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



Back to Graphic TSP

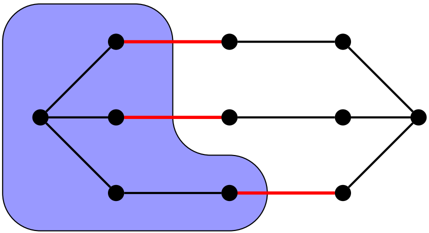
- **Input:** Graph $G = (V, E)$
- **Output:** Minimum length closed walk visiting each node (at least once).



- 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).



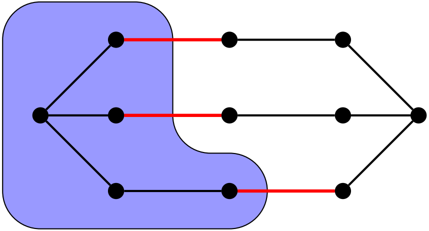
- Every feasible solution intersects every cut in an even number of edges.
- Generalized T-join inequality:

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

for all $\emptyset \neq S \subsetneq V$ 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:

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

for all $\emptyset \neq S \subsetneq V$ 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.

Bad news

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?

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

- 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 $1 \geqslant x_1 \geqslant x_2 \geqslant \ldots \geqslant x_r \geqslant 0$.

Betraying with fractional solutions

Setup:

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

Betraying with fractional solutions

Setup:

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

Auxiliary problem:

$$\min |f(x) - \tfrac{1}{2}| \text{ subject to } x \in P^r_{\text{ord}} \text{ and } \sum_{i=1}^{r} x_i = z \tag{9}$$

Setup:

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

Auxiliary problem:

$$\min \left|f(x) - \tfrac{1}{2}\right| \text{ subject to } x \in P^r_{\text{ord}} \text{ and } \sum_{i=1}^{r} x_i = z \tag{9}$$

Lemma: The optimal value of (9) is $\begin{cases} z & \text{if } z \leqslant \frac{1}{2} \\ r - z & \text{if } z \geqslant r - \frac{1}{2} \\ \frac{1}{2} & \text{otherwise} \end{cases}.$

Betraying with fractional solutions

Setup:

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

Auxiliary problem:

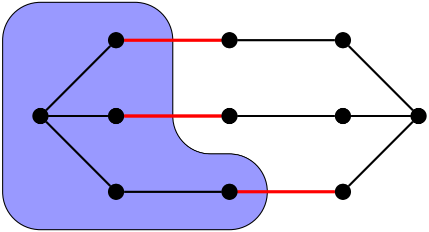
$$\min \left|f(x) - \tfrac{1}{2}\right| \text{ subject to } x \in P^r_{\text{ord}} \text{ and } \sum_{i=1}^{r} x_i = z \tag{9}$$

Lemma: The optimal value of (9) is $\begin{cases} z & \text{if } z \leqslant \frac{1}{2} \\ r - z & \text{if } z \geqslant r - \frac{1}{2} \\ \frac{1}{2} & \text{otherwise} \end{cases}$.

Consequence:

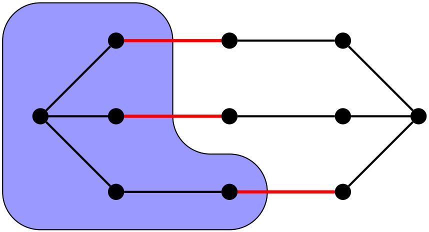
- Satisfied if <u>only two</u> variables participating in the parity constraint are $\frac{1}{2}$ away from their respective bounds.

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!

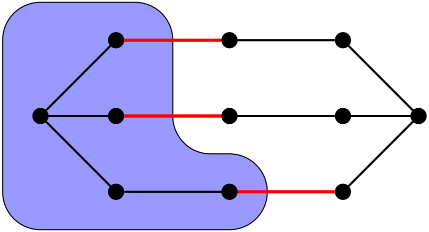


- Parity is wrong: odd number of cut edges used.
- But: All edge variables are $1 \in [0, 2]$, so lemma from previous slide applies!

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!