Binary extended formulations and sequential convexification

Manuel Aprile, Michele Conforti, Marco Di Summa
University of Padua
Binarizations

Let \( x \) be a variable that ranges from 0 to \( k \). A binarization of \( x \) is a linear formulation with variables \( x \) and \( y_1, \ldots, y_d \) (between 0 and 1), so that integrality of \( x \) is implied by the integrality of \( y_1, \ldots, y_d \).

- **Unary:** \( x = \sum_{i=1}^{k} y_i \) with \( y_1 \geq \cdots \geq y_k \)  
  [Roy 07]

- **Full:** \( x = \sum_{i=1}^{k} i \cdot y_i \) with \( \sum_{i=1}^{k} y_i \leq 1 \)  
  [Sherali, Adams 13]
Binarizations

Let $x$ be a variable that ranges from 0 to $k$.
A binarization of $x$ is a linear formulation with variables $x$ and $y_1, \ldots, y_d$ (between 0 and 1), so that integrality of $x$ is implied by the integrality of $y_1, \ldots, y_d$.

- **Unary**: $x = \sum_{i=1}^{k} y_i$ with $y_1 \geq \cdots \geq y_k$ [Roy 07]
- **Full**: $x = \sum_{i=1}^{k} i \cdot y_i$ with $\sum_{i=1}^{k} y_i \leq 1$. [Sherali, Adams 13]
- **Logarithmic**: $x = \sum_{i=1}^{t} 2^{i-1} y_i$, with $k = 2^t - 1$ [Owen, Mehrotra 02]
Binarizations

Let $x$ be a variable that ranges from 0 to $k$. A binarization of $x$ is a linear formulation with variables $x$ and $y_1, \ldots, y_d$ (between 0 and 1), so that integrality of $x$ is implied by the integrality of $y_1, \ldots, y_d$.

- **Unary**: $x = \sum_{i=1}^{k} y_i$ with $y_1 \geq \cdots \geq y_k$ [Roy 07]
- **Full**: $x = \sum_{i=1}^{k} i \cdot y_i$ with $\sum_{i=1}^{k} y_i \leq 1$. [Sherali, Adams 13]
- **Logarithmic**: $x = \sum_{i=1}^{t} 2^{i-1} y_i$, with $k = 2^t - 1$ [Owen, Mehrotra 02]

A polytope $B \subseteq \{(x, y) : (x, y) \in \mathbb{R} \times [0, 1]^d\}$ is a binarization of $x$ in the range $\{0, \ldots, k\}$ if

$$\pi_x(\{(x, y) \in B : y \in \{0, 1\}^d\}) = \{0, \ldots, k\}.$$
Why binarizations, and which one?

IP solvers deal more easily with binary variables than general integer variables.

- Cutting planes generated from variables of a binarizations can be more effective. [Bonami, Margot 15]
- Unimodular (generalization of full and unary) are optimal in terms of split closure, but they have $k$ variables. [Dash, Gunluk, Hildebrand 18]

But...

- The logarithmic binarization has only $O(\log k)$ variables, but can lead to worse B&B trees than original formulation. [Owen, Mehrotra 02]
- “Although this substitution is valid, it should be avoided if possible.” [Optimization Modelling with LINGO]
Why binarizations, and which one?

- Cutting planes generated from variables of a binarizations can be more effective. [Bonami, Margot 15]
- Unimodular (generalization of full and unary) are optimal in terms of split closure, but they have $k$ variables. [Dash, Gunluk, Hildebrand 18]

But...

- The logarithmic binarization has only $O(\log k)$ variables, but can lead to worse B&B trees than original formulation. [Owen, Mehrotra 02]
- “Although this substitution is valid, it should be avoided if possible.” [Optimization Modelling with LINGO]

We propose a different way to compare binarizations inspired by a connection with sequential convexification.
Our contributions

• We propose a “natural” notion of binarizations and we characterize the vertices of formulations that use such binarizations.

• We define the rank of a binarization, related to sequential convexification and the lift-and-project rank.

• We give formulas for the rank of the binarizations known in the literature, and show that
  • Unary is better than full
  • Logarithmic is optimal (among those with the same number of variables).
Sequential convexification

The **convexification** a polytope $Q$ with respect to a binary variable $x_i$ is

$$Q_{x_i} := \text{conv} \left( \{ x \in Q : x_i = 0 \} \cup \{ x \in Q : x_i = 1 \} \right).$$

If $Q \subset [0, 1]^p \times \mathbb{R}^{n-p}$, one has

$$\text{conv}\{ x \in Q : x_i \in \{0, 1\} \ \forall \ i \in [p] \} = (((Q_{x_1})_{x_2}) \ldots )_{x_p}.$$ 

[Balas Perregaard 02]
Sequential convexification

The convexification of a polytope $Q$ with respect to a binary variable $x_i$ is

$$Q_{x_i} := \text{conv} \left( \{ x \in Q : x_i = 0 \} \cup \{ x \in Q : x_i = 1 \} \right).$$

If $Q \subset [0, 1]^p \times \mathbb{R}^{n-p}$, one has

$$\text{conv}\{ x \in Q : x_i \in \{0, 1\} \ \forall i \in [p] \} = (((Q_{x_1})_{x_2}) \ldots)_{x_p}.$$

The lift-and-project rank of $Q$ is the minimum integer $k$ such that there are $i_1, \ldots, i_k \in [p]$ such that

$$\text{conv}\{ x \in Q : x_i \in \{0, 1\} \ \forall i \in [p] \} = ((((Q_{x_{i_1}}) \ldots)_{x_{i_k}}.$$  

One can see this as a hitting set problem: convexifying wrt $x_i$ we “get rid” of all vertices of $Q$ whose $x_i$-component is fractional. We need to pick $i_1, \ldots, i_k \in [p]$ so that each fractional vertex of $Q$ has a fractional component in some of $i_1, \ldots, i_k$. 


Sequential convexification converges in a finite number of steps to the integer hull, while general disjunctions do not converge. In this example, using split disjunctions does not converge if only $x_1, x_2$ are required to be integer. But, if we associate to $x_1, x_2$ a binarization, we obtain the integer hull by convexifying a small number of 0/1 variables.

[Cook, Kannan, Schrijver 90]
Natural binarizations and their vertices

We consider a polytope $P \subseteq [0, k]^n$ and a binary extended formulation

$$Q := \{(x, y) \in \mathbb{R}^n \times [0, 1]^{nk} : x \in P, (x_i, y_i) \in B_i \ \forall i \in [n]\}.$$ 

where $B_i$ is a binarization for $x_i$.

Convexifying all the $y$-variables leads to the integer hull $P_I = P \cap \mathbb{Z}^n$.

In order to study the lift-and-project rank of $Q$, we would like to understand its vertices...
Natural binarizations and their vertices

We consider a polytope $P \subseteq [0, k]^n$ and a binary extended formulation

$$Q := \{(x, y) \in \mathbb{R}^n \times [0, 1]^{nk} : x \in P, (x_i, y_i) \in B_i \forall i \in [n]\}.$$

where $B_i$ is a binarization for $x_i$.

Convexifying all the $y$-variables leads to the integer hull $P_I = P \cap \mathbb{Z}^n$.

In order to study the lift-and-project rank of $Q$, we would like to understand its vertices...

We can characterize exactly the vertices of $Q$, and their $x$-projections, under a natural assumption.

**Definition**

A binarization $B$ is **natural** if, for each vertex $(x, y)$ of $B$, $x$ is integer.
Let $P \subseteq [0, k]^n$ be a polytope and let $Q$ be a binary extended formulation of $P$ with natural binarizations. Then $\bar{x} \in \mathbb{R}^n$ is a point in $\pi_x(V(Q))$ if and only if there exist $I \subseteq [n]$, $\alpha_i \in \mathbb{Z}$ for $i \in I$, and a face $F$ of $P$ of dimension $|I|$ such that

$$F \cap \{x_i = \alpha_i \; \forall i \in I\} = \{\bar{x}\}.$$

Projections of vertices are exactly the 0-dimensional intersections of faces of $P$ with the integer grid.
Theorem

Let $P \subseteq [0, k]^n$ be a polytope and let $Q$ be a binary extended formulation of $P$ with natural binarizations. Then $\bar{x} \in \mathbb{R}^n$ is a point in $\pi_x(V(Q))$ if and only if there exist $I \subseteq [n]$, $\alpha_i \in \mathbb{Z}$ for $i \in I$, and a face $F$ of $P$ of dimension $|I|$ such that

$$F \cap \{x_i = \alpha_i \ \forall \ i \in I\} = \{\bar{x}\}.$$ 

In particular, projections of vertices of $Q$ do not depend on the binarizations used!
Let $P \subseteq [0, k]^n$ be a polytope and let $Q$ be a binary extended formulation of $P$ with natural binarizations. Then $(\bar{x}, \bar{y}) \in \mathbb{R}^n \times [0, 1]^{nd}$ is a vertex of $Q$ if and only if there exist $I \subseteq [n]$, $\alpha_i \in \mathbb{Z}$ for $i \in I$, and a face $F$ of $P$ of dimension $|I|$ such that:

- $F \cap \{x_i = \alpha_i \ \forall i \in I\} = \{\bar{x}\}$;
- $(\bar{x}_i, \bar{y}_i) \in V(B_i) \ \forall i \in I$;
- $(\bar{x}_i, \bar{y}_i) \in V(B_i \cap \{x_i = \bar{x}_i\}) \ \forall i \in [n] \setminus I$. 
\[ P = \{(x_1, x_2, x_3) \in [0, 2]^2 \times \mathbb{R} : \quad hx_1 + hx_2 + x_3 \leq 2h \]
\[ \quad x_3 \leq 2hx_1 \]
\[ \quad x_3 \leq 2hx_2 \]
\[ \quad x_3 \geq 0 \}\]

For \( i = 1, 2 \), \( B_i = \{(x_i, y_{i1}, y_{i2}) \in \mathbb{R} \times [0, 1]^2 : x_i = y_{i1} + y_{i2}, \ y_{i1} \geq y_{i2}\}\)
For $i = 1, 2$, $B_i = \{(x_i, y_{i1}, y_{i2}) \in \mathbb{R} \times [0, 1]^2 : x_i = y_{i1} + y_{i2}, y_{i1} \geq y_{i2}\}$

$$Q = \{(x_1, x_2, x_3) \in [0, 2]^2 \times \mathbb{R}, (y_{11}, y_{12}, y_{21}, y_{22}) \in [0, 1]^4 : \begin{align} hx_1 + hx_2 + x_3 &\leq 2h \\ x_3 &\leq 2hx_1 \\ x_3 &\leq 2hx_2 \\ x_3 &\geq 0 \end{align} \}$$

$(x_i, y_{i1}, y_{i2}) \in B_i$ for $i = 1, 2$
$V(Q)$ consists of the following points:

| $x_1$ | $x_2$ | $x_3$ | $y_{11}$ | $y_{12}$ | $y_{21}$ | $y_{22}$ |
|-------|-------|-------|----------|----------|----------|----------|
| 0     | 0     | 0     | 0        | 0        | 0        | 0        |
| 2     | 0     | 0     | 1        | 1        | 0        | 0        |
| 0     | 2     | 0     | 0        | 0        | 1        | 1        |
| 1/2   | 1/2   | $h$   | 1/2      | 0        | 1/2      | 0        |
| 1/2   | 1/2   | $h$   | 1/2      | 0        | 1/4      | 1/4      |
| 1/2   | 1/2   | $h$   | 1/4      | 1/4      | 1/2      | 0        |
| 1/2   | 1/2   | $h$   | 1/4      | 1/4      | 1/4      | 1/4      |
| 1     | 0     | 0     | 1        | 0        | 0        | 0        |
| 0     | 1     | 0     | 0        | 0        | 1        | 0        |
| 1     | 1     | 0     | 1        | 0        | 1        | 0        |
| 1     | 1     | 0     | 1/2      | 1/2      | 1        | 0        |
| 1     | 1     | 0     | 1/2      | 1/2      | 1/2      | 1/2      |
| 1     | 1/3   | $2h/3$ | 1      | 0        | 1/3      | 0        |
| 1     | 1/3   | $2h/3$ | 1      | 0        | 1/6      | 1/6      |
| 1/3   | 1     | $2h/3$ | 1/3    | 0        | 1        | 0        |
| 1/3   | 1     | $2h/3$ | 1/6    | 1/6      | 1        | 0        |
\textbf{\(V(Q)\) consists of the following points:}

| \(x_1\) | \(x_2\) | \(x_3\) | \(y_{11}\) | \(y_{12}\) | \(y_{21}\) | \(y_{22}\) |
|--------|--------|--------|--------|--------|--------|--------|
| 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 2 | 0 | 0 | 1 | 1 | 0 | 0 |
| 0 | 2 | 0 | 0 | 0 | 1 | 1 |
| 1/2 | 1/2 | \(h\) | 1/2 | 0 | 1/2 | 0 |
| 1/2 | 1/2 | \(h\) | 1/2 | 0 | 1/4 | 1/4 |
| 1/2 | 1/2 | \(h\) | 1/4 | 1/4 | 1/2 | 0 |
| 1/2 | 1/2 | \(h\) | 1/4 | 1/4 | 1/4 | 1/4 |
| 1 | 0 | 0 | 1 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| 1 | 1 | 0 | 1 | 0 | 1 | 0 |
| 1 | 1 | 0 | 1/2 | 1/2 | 1 | 0 |
| 1 | 1 | 0 | 1 | 0 | 1/2 | 1/2 |
| 1 | 1/3 | \(2h/3\) | 1 | 0 | 1/3 | 0 |
| 1 | 1/3 | \(2h/3\) | 1 | 0 | 1/6 | 1/6 |
| 1/3 | 1 | \(2h/3\) | 1/3 | 0 | 1 | 0 |
| 1/3 | 1 | \(2h/3\) | 1/6 | 1/6 | 1 | 0 |

Convexifying variables \(y_{11}, y_{21}\) is enough to obtain the integer hull.
The structure of the hitting set problem of a binary extended formulation depends on both $P$ and the binarizations and can be complex. However, the situations simplifies if we restrict to a single $x_i$ and $B$.

Let $\alpha \in \mathbb{Z}$. What is the minimum number of variables $y_{ij}$ to convexify in order to get rid of all vertices $(x, y) \in Q$ with $\alpha < x_i < \alpha + 1$?

Thanks to our characterization of vertices of $Q$, it turns out that, if $B$ is natural, the answer only depends on $B$ and $\alpha$, and not on $P$!
Given any binary extended formulation where natural binarization $B$ is associated to variable $x_i$, and $\alpha \in \mathbb{Z}$, the rank $r_k_B(\alpha)$ is the minimum number of variables $y_{ij}$ of $B$ to convexify in order to get rid of all vertices $(x, y) \in Q$ with $\alpha < x_i < \alpha + 1$.

Intuitively, $r_k_B(\cdot)$ measures the progress made towards ensuring the integrality of $x_i$ via application of sequential convexification.
Given any binary extended formulation where natural binarization $B$ is associated to variable $x_i$, and $\alpha \in \mathbb{Z}$, the rank $r_{kB}(\alpha)$ is the minimum number of variables $y_{ij}$ of $B$ to convexify in order to get rid of all vertices $(x, y) \in Q$ with $\alpha < x_i < \alpha + 1$.

Intuitively, $r_{kB}(\cdot)$ measures the progress made towards ensuring the integrality of $x_i$ via application of sequential convexification.

For $\alpha_1, \ldots, \alpha_\ell \in \mathbb{Z}$, the rank $r_{kB}(\alpha_1, \ldots, \alpha_\ell)$ is the minimum number of variables $y_{ij}$ of $B$ to convexify in order to get rid of all vertices $(x, y) \in Q$ with $\alpha_j < x_i < \alpha_j + 1$ for any $j = 1, \ldots, \ell$. 
\( \text{rk}_B(\alpha) \) = minimum number of variables \( y_{ij} \) of \( B \) to convexify in order to get rid of all vertices \((x, y) \in Q \) with \( \alpha < x_i < \alpha + 1 \).

\[
B = \{ (x_i, y) \in \mathbb{R} \times [0, 1]^3 : x_i = \sum_{j=1}^d y_j, \ 1 \geq y_1 \geq y_2 \geq y_3 \geq 0 \};
\]

\( \text{rk}_B(\alpha) = 1 \) for \( \alpha = 0, 1, 2 \)
Given a natural binarization \( B \subseteq [0, k] \times [0, 1]^d \) and \( \alpha \in \{0, \ldots, k-1\} \), we say that edge \( ((x^u, y^u), (x^v, y^v)) \) of \( B \) is an \( \alpha \)-edge if \( x^u \leq \alpha \) and \( x^v \geq \alpha + 1 \), or vice versa.

The **indicator set** of edge \( ((x^u, y^u), (x^v, y^v)) \) is the set of indices \( i \in [d] \) for which \( y^u_i \neq y^v_i \).

\[
\begin{array}{c|c}
0\text{-edges} & \text{sets} \\
0 - 1 & \{1\} \\
0 - 2 & \{1, 2\} \\
0 - 3 & \{1, 2, 3\}
\end{array}
\]
Given a natural binarization $B \subseteq [0, k] \times [0, 1]^d$ and $\alpha \in \{0, \ldots, k - 1\}$, we say that edge $((x^u, y^u), (x^v, y^v))$ of $B$ is an $\alpha$-edge if $x^u \leq \alpha$ and $x^v \geq \alpha + 1$, or vice versa.

The indicator set of edge $((x^u, y^u), (x^v, y^v))$ is the set of indices $i \in [d]$ for which $y^u_i \neq y^v_i$.

**Lemma**

$$\text{rk}_B(\alpha_1, \ldots, \alpha_\ell) = \min |I| : I \subseteq [d] \text{ hits the indicator sets of all } \alpha_j\text{-edges of } B, \text{ for } j \in [\ell].$$

Proof idea: the rank is also equal to the lift-and-project rank of a certain polytope inside $B$. 


Unary binarization

\[ B^U = \{(x, y) \in \mathbb{R} \times [0, 1]^k : x = \sum_{i=1}^k y_i, 1 \geq y_1 \geq \cdots \geq y_k \geq 0\}; \]

\[
\begin{array}{c|c}
1\text{-edges} & \text{sets} \\
0 - 2 & \{1, 2\} \\
0 - 3 & \{1, 2, 3\} \\
1 - 2 & \{2\} \\
1 - 3 & \{2, 3\}
\end{array}
\]

\[ \text{rk}_{B^U}(\alpha_1, \ldots, \alpha_\ell) = \ell. \]
Full binarization

\[ B^F = \{(x, y) \in \mathbb{R} \times [0, 1]^k : x = \sum_{i=1}^k i \cdot y_i, \sum_{i=1}^k y_i \leq 1\}; \]

\[
\begin{array}{c|c}
0\text{-edges} & \text{sets} \\
0 - 1 & \{1\} \\
0 - 2 & \{2\} \\
0 - 3 & \{3\} \\
\end{array}
\]

\[ \text{rk}_{B^F}(\alpha_1, \ldots, \alpha_\ell) = k - \min_{j \in [\ell]} \alpha_j. \]

\[ k - \min_{j \in [\ell]} \alpha_j \geq k - (k - \ell) = \ell, \text{ hence:} \]

Unary has smaller rank than Full:

\[ \text{rk}_{B^F}(\cdots) \geq \text{rk}_{B^U}(\cdots). \]
Logarithmic binarization

\[ B^L = \{(x, y) \in \mathbb{R} \times [0, 1]^d : x = \sum_{i=1}^{d} 2^{i-1} y_i \}. \]

Observation: indicator sets of \( \alpha \)-edges are singletons, and parallel edges have the same indicator set.
Logarithmic binarization

\[ B^L = \{(x, y) \in \mathbb{R} \times [0, 1]^d : x = \sum_{i=1}^{d} 2^{i-1} y_i \}. \]

| 3-edges | sets       |
|---------|------------|
| 0 – 4   | \{3\}      |
| 1 – 5   | \{3\}      |
| 2 – 6   | \{3\}      |
| 3 – 7   | \{3\}      |

\( \text{rk}_{B^L}(0) = 3, \ \text{rk}_{B^L}(1) = 2, \ \text{rk}_{B^L}(3) = 1. \)
Logarithmic binarization

\[ B^L = \{ (x, y) \in \mathbb{R} \times [0, 1]^d : x = \sum_{i=1}^{d} 2^{i-1} y_i \}. \]

**Lemma**

\[ \text{rk}_{B^L}(\alpha) = d - f(\alpha). \]

where \( f(\alpha) \) is the largest \( t \) such that \( 2^t \) divides \( \alpha + 1 \).
Logarithmic binarization

\[ B^L = \{(x, y) \in \mathbb{R} \times [0, 1]^d : x = \sum_{i=1}^{d} 2^{i-1} y_i \}. \]

### Lemma

\[ \text{rk}_{B^L}(\alpha_1, \ldots, \alpha_\ell) = d - f(\alpha_1, \ldots, \alpha_\ell). \]

where \( f(\alpha_1, \ldots, \alpha_\ell) = \max\{t : 2^t \text{ divides } \alpha_j + 1 \ \forall j \in [\ell]\}. \)
Hypercube binarizations

The logarithmic binarization has \( \lceil \log_2(k) \rceil \) variables, but large rank.

Is there any binarization with the same number of variables, but with lower rank?
Hypercube binarizations

The logarithmic binarization has $\lceil \log_2(k) \rceil$ variables, but large rank.

Is there any binarization with the same number of variables, but with lower rank? No!

**Definition**

Binarization $B \subseteq [0, k] \times [0, 1]^d$ is a hypercube binarization if $\pi_y(B) = [0, 1]^d$ ( $\implies d = \lceil \log_2(k) \rceil$).

**Theorem**

For any hypercube binarization $B$,

$$\text{rk}_B(\alpha_1, \ldots, \alpha_\ell) \geq \text{rk}_{B^L}(\alpha_1, \ldots, \alpha_\ell).$$

The logarithmic binarization is optimal among hypercube binarizations.
Open questions

- What is the trade-off between the number of variables in a binarization and its rank?

- Is the unary binarization optimal among the “simplex” binarizations?

- Is there a binarization with \( O(\log k) \) variables that is better than the logarithmic?
Open questions

- What is the trade-off between the number of variables in a binarization and its rank?

- Is the unary binarization optimal among the “simplex” binarizations?

- Is there a binarization with $O(\log k)$ variables that is better than the logarithmic?

Thank you for your attention.