Benders cuts via corner polyhedra: an application to the stochastic vehicle routing problem

Matheus Jun Ota*, Ricardo Fukasawa, Aleksandr M. Kazachkov

Institution: University of Waterloo, Department of Combinatorics and Optimization; Email: mjota@uwaterloo.ca;

UNIVERSITY OF WATERLOO

Introduction

Consider the following mixed integer linear program.

$$
\begin{aligned}
\textbf{(MILP)}\quad \min\ & c^\top x + d^\top y \\
\text{s.t.}\ & Tx + Qy = h, && \text{Linking Constraints}\\
& Ax \geq b, && X := \{x \geq 0 : Ax \geq b\}\\
& Gy = g, && Y := \{y \geq 0 : Gy = g\} \neq \emptyset\\
& x \in \mathbb{Z}^{n_1} \times \mathbb{R}^{n_2},\\
& y \in \mathbb{R}^m_+,
\end{aligned}
$$

Benders decomposition: "project out" the y-variables.

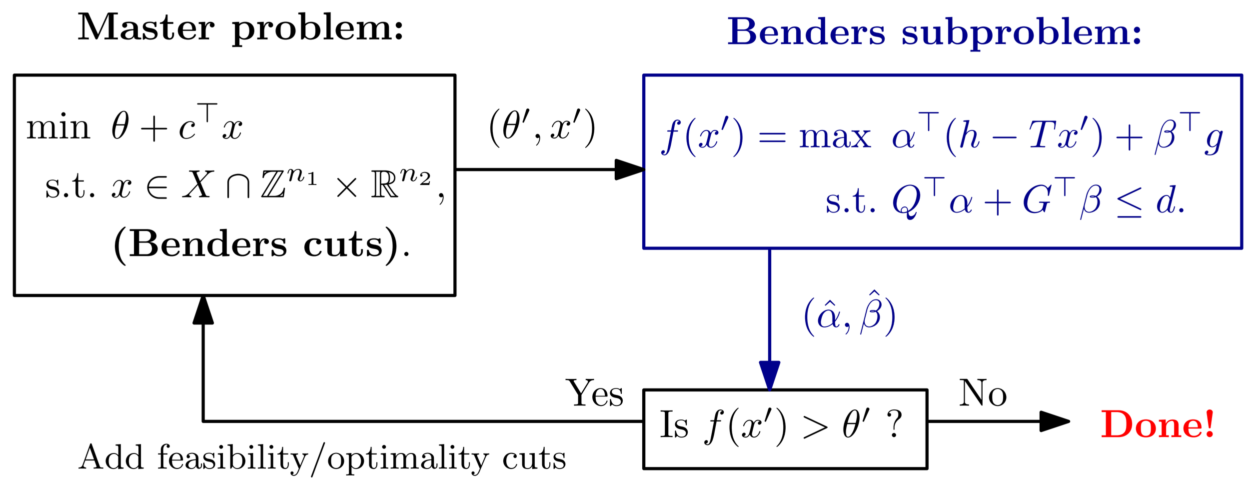


Figure 1. Flowchart of the Benders decomposition procedure.

Definition 1 (projection): For any $y \in \mathbb{R}^m$, $[d|Q](y) := (d^\top y, Qy)$. (The linking constraints map Qy to $(h - Tx)$, which is not (necessarily) x!) Since we allow $\theta \geq f(x)$, for any set of points $P \subseteq \mathbb{R}^m$, we also define $\mathcal{M}(P) := [d|Q](P) + \text{cone}(\{(1, \vec{0})\})$.

$\implies$ *Benders decomposition separates points from $\mathcal{M}(Y)$.*

Motivation: We consider situations where **solving the Benders subproblem is too expensive**; but on the other hand, we have access to an efficient oracle $\mathcal{O}$ that can **efficiently find an optimal basic feasible solution (BFS) of the problem** $\min_{y \in Y}\{s^\top y\}$, for any $s \in \mathbb{R}^m$.

Examples: Our investigation was motivated by cases where $\mathcal{O}$ solves a **shortest-path** or a **min-cost flow** problem over an extended state digraph.

- Example 1: **stochastic programming** models where the y variables correspond to **flows over an extended state digraph**.
- Example 2: mathematical formulations that use **decision diagrams** to strengthen relaxations.

Contributions:

- A **new technique for generating Benders cuts** based on the projection of **corner polyhedra** in the space of the y-variables.
- In some cases, this technique gives an alternative to the method of [1] to **recover the Dantzig-Wolfe bound.**
- Test case: **vehicle routing problem with stochastic demands (VRPSD).**
- Computational experiments show that the proposed approach **can generate stronger cuts in significantly less computational time.**
- For the VRPSD, **combinatorial structure of the basis** found by $\mathcal{O}$ leads to **generalization/strengthening** of some known valid inequalities.

Preliminaries

Definition 2 (support): **For our purposes,** for any set of vectors P, we refer to the **support** of P as the function $\sigma_P : s \to \inf_{p \in P}\{s^\top p\}$.

Fact 1 (cone optimality): Let y^* be an optimal BFS for $\min_{y \in Y}\{s^\top y\}$ with basis B. Let $\{r\}_{r \in R}$ be the rays associated with the non-basic variables w.r.t. B. Define $C(y^*, R) := \{y^*\} + \text{cone}(R)$, then $\sigma_Y(s) = \sigma_{C(y^*,R)}(s)$.

Lemma 1 [OFK24]: For every $\alpha \in \mathbb{R}^p$ and every set of points $P \subseteq \mathbb{R}^m$, we have that $\sigma_P(d + Q^\top\alpha) = \sigma_{\mathcal{M}(P)}((1, \alpha))$.

Separating stronger Benders cuts via projected corner polyhedra

An initial idea: Use the **Lagrangian dual** to make use of the oracle $\mathcal{O}$.

$$
f(x) = \max_{\alpha \in \mathbb{R}^p}\left\{\min_{y \in Y}\{d^\top y + \alpha^\top(Qy - h + Tx)\}\right\} = \max_{\alpha \in \mathbb{R}^p}\left\{\alpha^\top(Tx - h) + \sigma_Y(d + Q^\top\alpha)\right\}. \tag{1}
$$

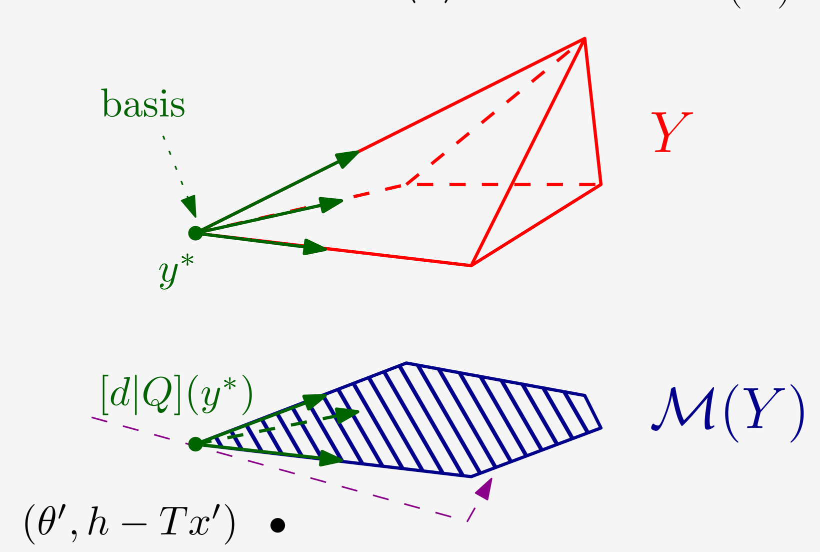
This gives a Benders optimality cut: for all $\hat\alpha \in \mathbb{R}^p$,

$$
\theta \geq \hat\alpha^\top(Tx - h) + \sigma_Y(d + Q^\top\hat\alpha) \quad\iff\quad \theta + \hat\alpha^\top(h - Tx) \geq \sigma_Y(d + Q^\top\hat\alpha) \overset{\text{Lemma 1}}{=} \sigma_{\mathcal{M}(Y)}((1, \hat\alpha)). \tag{2}
$$

However, in our preliminary experiments (using the subgradient/volume method), this approach still leads to **weak cuts with bad convergence.** So we need to do better...

Theorem 1 [OFK24]: Let (θ', x') be a candidate solution (see Figure 1) and let $\hat\alpha$ be such that (θ', x') violates the corresponding inequality (2). Let $C = C(y^*, R)$ be an optimal cone w.r.t. $\sigma_Y(d + Q^\top\hat\alpha)$, then $(\theta', h - Tx') \notin \mathcal{M}(C)$.

Proof: We show that (2) is valid for $\mathcal{M}(C)$: $\sigma_Y(d + Q^\top\hat\alpha) \overset{\text{Fact 1}}{=} \sigma_C(d + Q^\top\hat\alpha) \overset{\text{Lemma 1}}{=} \sigma_{\mathcal{M}(C)}((1, \hat\alpha))$. ∎



$$
\begin{aligned}
\textbf{(PRLP)}\quad \min\ & \alpha_\theta^\top\theta' + \alpha_x^\top(h - Tx') - \beta \\
\text{s.t.}\ & \beta = (\alpha_\theta, \alpha_x)^\top[d|Q](y^*), && \text{(P1)}\\
& (\alpha_\theta, \alpha_x)^\top[d|Q](r) \geq 0, \quad \forall r \in R. && \text{(P2)}\\
& \alpha_\theta \geq 0. && \text{(P3)}
\end{aligned}
$$

Figure 2. In the left, we have an illustration of $C = C(y^*, R)$ and $\mathcal{M}(C)$. The point $(\theta', h - Tx')$ is not in $\mathcal{M}(C)$ and violates inequality (2) (dashed purple line). The LP in the right is a cut-generating linear program w.r.t. the point-ray representation of $\mathcal{M}(C)$. Feasible solutions to (PRLP) correspond to valid inequalities for $\mathcal{M}(C)$ and extreme rays of (PRLP) correspond to facets of $\mathcal{M}(C)$.

Recovering strong bounds with a single cone: Let us now assume that T has an inverse. In this case, we may rewrite the linking constraints as $x = T^{-1}h - T^{-1}Qy$. By calling $Q^\circ = -T^{-1}Q$ and $x^\circ = x - T^{-1}h$, we assume WLOG that the linear relaxation of **(MILP)** is

$$
\begin{aligned}
z_{\text{LP}} &= \min_{x,y \in Y}\{c^\top x + d^\top y : x = Qy, Ax \geq b, x \geq q\}\\
&= \min_{y \in Y}\{c^\top(Qy) + d^\top y : A(Qy) \geq b, Qy \geq q\}\\
&= \max_{\rho \geq 0,\, \gamma \geq 0}\left\{\min_{y \in Y}\{c^\top(Qy) + d^\top y + \rho^\top(b - A(Qy)) + \gamma^\top(q - (Qy))\}\right\}.
\end{aligned} \tag{3}
$$

Theorem 2 [OFK24]: Let $(\hat\rho, \hat\gamma)$ be optimal for the outer problem in (3) and let $C = C(y^*, R)$ be an optimal cone for the inner problem in (3) (with $\rho = \hat\rho$ and $\gamma = \hat\gamma$). Then $\min_{(\theta,x) \in \mathcal{M}(C) \cap (\mathbb{R}_+ \times X)}\{\theta + c^\top x\} \geq z_{\text{LP}}$.

Proof: Let $(\theta', x') \in \mathcal{M}(C) \cap (\mathbb{R}_+ \times X)$, so $\exists y' \in C$ such that $\theta' \geq d^\top y'$ and $x' = Qy'$. Then

$$
\theta' + c^\top x' \geq c^\top(Qy') + d^\top y' \geq c^\top(Qy') + d^\top y' + \underbrace{\hat\rho^\top(b - A(Qy)) + \hat\gamma^\top(q - (Qy))}_{\leq 0 \text{ since } x' \in X} \geq z_{\text{LP}}. \quad \blacksquare
$$

When the **pricing problem** of a column-generation formulation is solved with **dynamic-programming (DP)**, one can construct a formulation equivalent to **(MILP)** where the y-variables correspond to flows over a DP state digraph. In these cases, **Theorem 2 let us recover the Dantzig-Wolfe bound with cuts.**

The vehicle routing problem with stochastic demands

Input
Complete graph $G = (V = \{0\} \cup V_+, E)$
k: Number of vehicles
B: Vehicle capacity
$c \in \mathbb{R}^E_+$: Edge costs
$d \sim \mathbb{P}$: Demands prob. distribution
Q: Recourse function

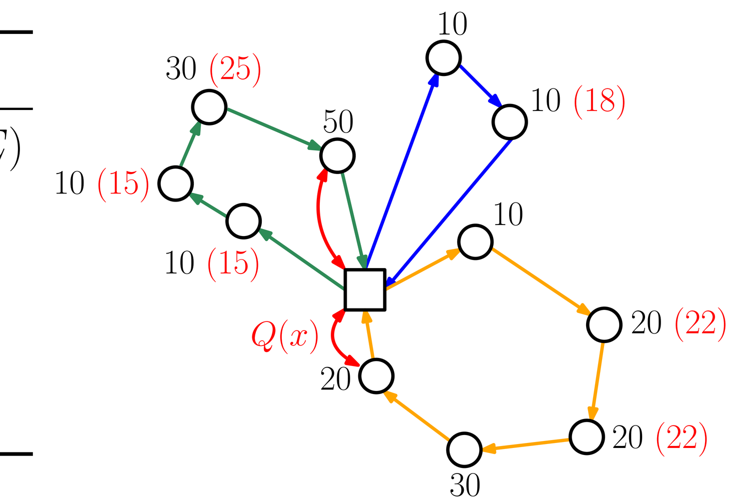


Figure 3: Example with $B = 100$. Numbers in black are the expected demands and numbers in red are realizations that differ from their expectations. This forces some vehicles to unload at the depot before continuing the route, incurring an additional recourse cost.

Goal: find k feasible routes $R_1, \ldots, R_k$ that visit every customer exactly once and that minimizes $\sum_{i \in [k]}(c(R_i) + \mathbb{E}[Q(R_i)])$.

Formulation: $x \in \mathbb{Z}^E$: edges in the routes; $y \in \mathbb{R}^{\mathcal{A}}_+$: flows over $\mathcal{N} = (\mathcal{V} \cup \{r, t\}, \mathcal{A})$.

Network $\mathcal{N}$: each state $s \in \mathcal{V}$ correspond to a tuple (customer, acc. demand) (we assume that the demands are Poisson and independently distributed). We set $w \in \mathbb{Q}^{\mathcal{A}}_+$ to encode the recourse costs. Each arc $a \in \mathcal{A}$ corresponds to an edge edge$(a) \in E$. The linking constraints are then $x_{ij} = \sum(y_a : \text{edge}(a) = ij)$, for all $ij \in E$.

Combinatorial structure: We model the problem so that Y is a flow polytope associated with $r - t$ flows of value k (without capacities on the arcs). Thus, **(basis $\iff$ spanning tree $\mathcal{T}$ of $\mathcal{N}$)** and **(rays $\iff$ cycles w.r.t. $\mathcal{T}$.)**

$\implies$ *Examining these cycles allow us to improve special cases of integer L-shaped cuts previously proposed by Spliet and Hoogendoorn (2023)*

Computational experiments

Standard Benders based on LP duality (Figure 1) and Lagrangian duality (Equation (1)): Cannot solve root node in 30 minutes.

Algorithms: ILS: our implementation of a state-of-the-art integer L-shaped (ILS) algorithm [2]; **+Cone:** addition of our cuts (Theorem 2).

- For gaps, we ignored instances that an algorithm was unable to find a feasible solution within the time limit.
- For times, the numbers outside parentheses are averages over all instances; numbers inside parentheses ignore instances that were not solved within the time limit.

Table 1. Computational experiments on instances from Jabali et al. (2014)

		Root Gap		Time (s)		Solved	
$\lvert V\rvert$	k	ILS	Cone+ILS	ILS	Cone+ILS	ILS	Cone+ILS
40	4	4.3%	**1.8%**	208 **(31)**	**95** (52)	9 / 10	**10 / 10**
50	4	3.2%	**1.7%**	**406 (232)**	482 (299)	8 / 10	8 / 10
40	6	5.4%	**2.4%**	913 (319)	**806 (271)**	6 / 10	**7 / 10**
50	6	4.5%	**2.4%**	1362 (326)	**1061 (203)**	3 / 10	**6 / 10**

References

[1] Rui Chen, Oktay Günlük, and Andrea Lodi.
Recovering dantzig-wolfe bounds by cutting planes.
Operations Research, 2024.

[2] Lucas Parada, Robin Legault, Jean-François Côté, and Michel Gendreau.
A disaggregated integer l-shaped method for stochastic vehicle routing problems with monotonic recourse.
European Journal of Operational Research, 2024.