A NOVEL PARETO-OPTIMAL CUT SELECTION STRATEGY FOR BENDERS DECOMPOSITION

LUKAS GLOMB, FRAUKE LIERS, FLORIAN RÖSEL

All authors: FAU Erlangen-Nuremberg, Department Data Science, Cauerstraße 11, 91058 Erlangen, Germany

Abstract. Decomposition methods can be used to create efficient solution algorithms for a wide range of optimization problems. For example, Benders Decomposition can be used to solve scenario-expanded two-stage stochastic optimization problems effectively. Benders Decomposition iteratively generates Benders cuts by solving a simplified version of an optimization problem, the so-called subproblem. The choice of the generated cuts can impact the performance of this approach. Several approaches to making this choice are known.

The Magnanti-Wong method focuses on generating Pareto-optimal cuts. Cuts based on minimal infeasible subsystems of a modified version of the subproblem have been proven to be effective. Additionally, methods that use facets of the subproblem's value function's epigraph have been developed recently.

We have made a contribution to the field of cut selection strategies for Benders Decomposition by introducing a novel concept of Pareto-optimality, which leads to an efficient cut selection strategy. This strategy aims for cuts that exclude a large set of points from being optimal. Furthermore, we have established the algorithmic framework to fully leverage the potential of our cut selection strategy.

We have compared our cut selection strategy with several others on various instances, including instances from the MIPLib, network design problems, and randomly generated mixed-integer linear programs. The computational results indicate that our method solves problems faster than the benchmark approaches, particularly when combined as a hybrid selection strategy with the minimal infeasible subsystem cut selection. Moreover, the method clearly outperforms the other cut selection strategies in terms of the number of cuts required to solve a problem optimally. Therefore, this method is particularly effective in situations with limited memory or in cases where the subproblem is challenging to solve.

Keywords. Benders Decomposition, Cut Selection

1. Introduction and Literature Review

Initially published in Benders (1962) and developed for mixed-integer linear programs (**MILP** or **MIP**), the approach named Benders Decomposition (**BD**) works on a meta-level, as follows. It separates the set of variables of the MIP into two classes: the so-called master variables and the subvariables. The approach alternately solves two optimization problems, the master problem and the subproblem. It obtains a solution to the master problem in each iteration. It determines a solution to the subproblem by fixing the variables from the master problem to the obtained values. It derives a new valid constraint from the dual solution of the subproblem. It adds this constraint to the master problem.

BD is the backbone of various solution methodologies for various optimization problems. It works exceptionally well, if the subproblem has a separable structure. This means that the subproblem represents

E-mail address: lukas.glomb@fau.de, frauke.liers@fau.de, florian.roesel@fau.de (CA's email).

a collection of smaller optimization problems that can be solved separately (but are coupled if the master variables are not fixed).

1.1. **Literature on BD applications.** A large number of articles document the high practical relevance of the approach. BD has been applied to solve a large number of often stochastic optimization problems with the property that fixing the first stage variable decomposes the problem into independent subproblems, but BD has also successfully solved optimization problems without block structure.

In Baringo and Conejo (2011) the approach was applied to maximize profits from wind power investments. This problem is highly affected by uncertainty, and hence, the emerging scenario-expanded optimization model is decomposed into algorithmically tractable parts. BD can solve the otherwise intractable problem quickly. In Adulyasak et al. (2015), the approach is applied to solve a variant of the production routing problem under uncertainty. The scenario-expanded model, which is, in practice, difficult to solve due to its size, could be decomposed into smaller pieces, and the authors report the successful solution of instances of realistic size and considerable speedups of BD in comparison to standard methods. The authors also report on several algorithmic improvements like lower-bound lifting inequalities, Pareto-optimal cuts (in the sense of Magnanti and Wong (1981)), and cut aggregation. They report that Pareto-optimal cuts have the most significant effect on CPU time (factor 2 to 4 on their instance set). In Maheo et al. (2019), BD is applied to solve a variant of the network design problem to create a public transit system for Canberra. The problem structure allows for splitting the problem into multiple independent subproblems after fixing only a few variables. The authors used Pareto-optimal cuts in the sense of Magnanti and Wong (1981) and reported significant speedups compared to standard solution methods. Bärmann et al. (2015) could solve huge network design problems using an iterative aggregation procedure that dynamically expands subnetworks. It is shown that the expansion routines are closely related to Benders feasibility cuts. The method outperforms standard approaches by far when applied to railway network design problems. Abdolmohammadi and Kazemi (2013) applied BD to optimize the utilization of combined heat and power systems and showed that the approach is superior to several benchmark methods. In Nasri et al. (2015), the approach has been applied to solve a network-constrained AC unit commitment problem with uncertain wind power production. It is used to decompose the emerging intractable scenario-expanded optimization problems into tractable parts that can be solved independently. In Contreras et al. (2011), BD has been successfully applied to solve large-scale uncapacitated hub location problems. The problem as well decomposes into several subproblems if a small proportion of the variables is fixed. In Mansouri et al. (2020), the operational planning of energy hubs with demand response under uncertainty is optimized using BD, splitting the intractable scenario-expanded optimization model into tractable parts. The authors report considerable speedups in comparison to to standard methods. Grimm et al. (2019) have developed a generalized decomposition algorithm for a three-stage energy market problem that computes optimum electricity market price zones. They could show that welfare-maximum solutions can be computed within a reasonable time, although the problem is complex. In Ambrosius et al. (2020) the approach is specified for the German electricity market, considering network expansion and renewable energy. In Bayram and Yaman (2018), generalized BD has been applied to optimize the location of shelters and evacuation routes under uncertainty, allowing the intractable scenario-expanded model to be split into tractable parts. The authors applied several acceleration strategies, including Pareto-optimal cuts, and reported significant CPU time savings. You and Grossmann (2013) apply a multi-cut version of the approach to solve a supply chain planning problem under uncertainty. The authors report considerable CPU time savings compared to the single-cut version. Botton et al. (2013)

applied it to efficiently handle extended formulations of the hop-constrained survivable network design problem. The authors report a considerable speedup compared to standard methods. Qian et al. (2013) have successfully used it to optimize communication networks. Azad et al. (2013) applied it to optimize supply chain networks suffering from facility and transportation disruptions. Compared to standard solution methods, the authors report considerable CPU time savings, using the covering-cut-bundle method as presented in Saharidis et al. (2010) and maximum density cuts. In Fischetti et al. (2016), it is tuned to be applicable to capacitated facility location problems, while in Fischetti et al. (2017), it is applied to the uncapacitated version of this problem. The authors use several acceleration strategies and achieve considerable CPU time savings compared to alternative solution approaches for both problem versions. In Glomb et al. (2023a), a logic-based variant of BD has been applied to solve an integrated tail assignment and turnaround planning problem. The authors report considerable speedups compared to classical solution approaches. In Glomb et al. (2023b), it has been applied to solve the tail assignment problem suffering from part failure scenarios. Solving the huge scenario-expanded model with the decomposition led to significant CPU time savings compared to standard approaches.

1.2. **Known approaches to accelerate BD.** Several approaches and techniques to accelerate BD are known. Many have been used in Rahmaniani et al. (2018), where the approach has been applied to solve large-scale stochastic network design problems. For this reason, this article is a good starting point to acquire knowledge about acceleration techniques. The authors of Geoffrion and Graves (1974) proposed a strategy that avoids solving the master problem to optimality in each iteration to attain performance benefits. In Cote and Laughton (1984), it has been shown that heuristic solutions to the master problems generate valid cuts. This has been specified, e.g., in Poojari and Beasley (2009), where a genetic algorithm is applied to create solutions to the master problem. McDaniel and Devine (1977) show that valid cuts can also be calculated if relaxations of the master problems are used. Based on that, approaches to solve MIPs, often referred to as the two-phase-algorithm, have been proposed, which solve the linear programming relaxation of the master problem in the first phase and reestablish the integrality constraints of the master problem in the second phase. The authors of Costa et al. (2012) proposed several strategies that aim to determine of points in a superset of the master problem domain, for which it is beneficial to generate cuts. Several authors have observed that stabilizing the solution process of the master problem leads to performance gains. A concrete implementation of a stabilization procedure is Rei et al. (2009), where local branching has been used to limit the distance of subsequent master solutions. An alternative implementation of a stabilization approach has been set up by Santoso et al. (2005) by limiting the Hamming distance to a specified stabilization point. Our algorithmic framework uses modern MIP solvers being capable of inserting additional constraints to an optimization problem on the fly during a single Branch-and-Cut run. The approach is called Branch-and-Benders-Cut and is described e.g. in Rahmaniani et al. (2017), and applied e.g. in Gendron et al. (2016). The authors of Saharidis et al. (2011) point out that for several optimization problems, it is beneficial to initialize the master problem with several valid inequalities that can either be added upfront or treated as cutting planes that are added to the master problem during the solution process whenever it seems to be suitable. This strategy has also been applied in Fischetti et al. (2016). In Bonami et al. (2020), several crucial numerical aspects regarding BD have been addressed. The authors provide a precise guideline on how to implement the approach. The focus lies on an in-out procedure to derive cuts that tighten the LP-relaxation of the problem, on perceiving constraints in the subproblem that contain only one subvariable as variable bounds rather than as constraints, and on normalization of feasibility cuts. The authors provide an extensive benchmark of their implementation and report considerable speedups. Some authors have demonstrated the potential of making the approach more efficient by implementing problem-specific techniques to strengthen the cuts such as Van Roy (1986) or Wentges (1996) for some variants of the well-known facility location problem. In Wu and Shahidehpour (2010), the approach is accelerated using several acceleration strategies, like Pareto-optimal cuts or cuts with increased density, for network-constrained unit commitment problems. The authors report significant speedups compared to a standard implementation of the approach.

A vital element to improve the performance of BD is the cut selection strategy. Magnanti and Wong (1981) propose to generate so-called non-dominated Benders cuts. Whenever the dual of the subproblem has multiple solutions, each generated cut is readjusted by solving an auxiliary version of the subproblem that depends on a so-called core point. Adaptations to these methods have been proposed: In Papadakos (2008), it is shown that clever updates of the core point accelerate the procedure further, and Sherali and Lunday (2013) show that the solution of a perturbed version of the subproblem automatically provides a non-dominated cut, making the solution of a second optimization problem in each iteration unnecessary. For specific optimization problems, it seems to be beneficial to generate multiple cuts per master problem solution. Saharidis et al. (2010) propose generating a so-called covering cut bundle, i.e., a set of Benders cuts that have as many master variables as possible in the union of their supports. Fischetti et al. (2010) and Stursberg (2019) propose cut selection strategies based on slightly adapted subproblems. We will present their contribution in Section 2 more precisely.

1.3. **Contribution and Structure of this Article.** We develop a novel and original cut selection strategy in the sense of Magnanti and Wong (1981), who presented Pareto-optimal Benders cuts. To implement this, we first develop a novel notion of Pareto-optimality. In contrast to Magnanti and Wong (1981), our notion of Pareto-optimality is based on the sets of points in the domain of the master problem that has a chance to be the optimal solution of the overall problem after a Benders cut is inserted into the master problem. Based on this, we develop a novel cut selection strategy, i.e., it generates, under mild assumptions, Pareto-optimal cuts according to our notion. The strategy only requires the solution of one linear program per generated cut. This program is not considerably more difficult than the original subproblem. Further, we develop an algorithmic framework to apply this cut selection strategy. The framework specifies some details relevant to the proposed strategy's performance. We further conduct a computational study to demonstrate the good performance of our algorithmic framework in comparison to several other cut selection strategies: the Magnanti-Wong method, as presented in Magnanti and Wong (1981), cuts based on minimal infeasible subsystems (**MIS**) as given in Fischetti et al. (2010), and facet defining cuts as presented in Stursberg (2019). We consider this contribution relevant because BD is applied frequently, as shown in Section 1.1, and because many acceleration strategies and especially cut selection strategies have recently been developed, as demonstrated in Section 1.2. We want to acknowledge that the method we propose uses an idea of Conforti and Wolsey (2019) to generate cuts for BD, following a suggestion made in the final remarks of Conforti and Wolsey (2019). This article extends the main result of Conforti and Wolsey (2019) that applies only to polyhedra for which the outer description is explicitly known. We describe how the idea of Conforti and Wolsey (2019) can be used to generate Benders cuts.

In Section 2, we introduce BD and known cut selection strategies. In Section 3, we propose our novel notion of Pareto-optimality and our cut selection strategy and prove several vital statements about it. In Section 4, we present the algorithmic framework that enables the application of the cut selection strategy. In Section 5, we describe our computational study and its results. Section (6) summarizes our results and provides an outlook on future research directions.

2. Benders Decomposition and known Cut Selection Strategies

In this section, we want to present a new approach to selecting valid cutting planes for BD. First, BD, as initially presented in Benders (1962), is introduced. Afterward, we will present and discuss known cut selection strategies. Some of these are variants of the method initially presented in Magnanti and Wong (1981). Furthermore, we present a framework called MIS-cuts, initially presented by Fischetti et al. (2010). We also present an innovative cut selection strategy proposed by Stursberg (2019), capable of generating so-called facet-generating cuts under mild assumptions. For a more detailed overview, we refer to Rahmaniani et al. (2017), an excellent survey of recent developments in this field.

2.1. **Benders Decomposition.** Initially presented in Benders (1962), the approach has been developed for optimization problems containing a set of "complicating" variables, typically of the form

$$\min f^T x + c^T y \tag{1a}$$

$$\text{s.t. } Hx + Ay \geq b \tag{1b}$$

$$x \in X, y \geq 0, \tag{1c}$$

where $X \subset \mathbb{R}^n$ (the n-dimensional Euclidean space), $b \in \mathbb{R}^m$, $f \in \mathbb{R}^n$, $H \in \mathbb{R}^{m \times n}$, $c \in \mathbb{R}^p$ and $A \in \mathbb{R}^{m \times p}$.

The approach works straight-forwardly under the mild assumption that $\min_{x \in X} f^T x$ exists and a lower bound $\hat{\eta} \in \mathbb{R}$ for

$$\tilde{\eta} := \min_{x \in X} S(x),$$

where

$$S(x) := \min_{y \geq 0} \; c^T y \quad \text{s.t. } Ay \geq b - Hx,$$

exists and is easily obtainable for all $x \in \text{conv}(X)$. This is, e.g. the case, if c is non-negative. Then, the inner product $c^T y$ is also certainly non-negative, and the choice $\hat{\eta} := 0$ is hence valid.

An artificial variable η, along with the lower bound $\hat{\eta}$ is introduced, and afterward the so-called master problem is solved repeatedly, which is defined as

$$\min \; f^T x + \eta \tag{2a}$$

$$\text{s.t. } 0 \geq \omega(x) \text{ for all } \omega \in \Omega^{\text{feas}} \tag{2b}$$

$$\eta \geq \omega(x) \text{ for all } \omega \in \Omega^{\text{opt}} \tag{2c}$$

$$\eta \geq \hat{\eta}, x \in X. \tag{2d}$$

Constraints (2b) and (2c) are the so-called Benders cuts and are frequently updated, as described in the following. The set $\Omega := \Omega^{\text{feas}} \cup \Omega^{\text{opt}}$ is initialized as $\emptyset$. Hence, in the first iteration, the solution $(\bar{x}, \bar{\eta})$ of Problem (2) is generally $\bar{x} = \min_{x \in X} f^T x$, $\bar{\eta} = \hat{\eta}$.

After an optimal solution $(\bar{x}, \bar{\eta})$ of (2) is obtained, one solves the so-called dual subproblem, which is the linear programming dual of the primal subproblem,

$$\max \; \pi^T (b - H\bar{x}) \tag{3a}$$

$$\text{s.t. } \pi^T A \leq c^T \tag{3b}$$

$$\pi \geq 0. \tag{3c}$$

We note that (3) is a linear program that is defined, independently on $\bar{x}$, on the polyhedron

$$P := \{\pi \mid \pi^T A \leq c^T, \pi \geq 0\}.$$

The polyhedron P is non-empty, implied by the assumption that $S(x)$ is bounded from below.

If Problem (3) is unbounded, we insert a so-called feasibility cut to Ω^{feas}, which reads as

$$\bar{\pi}^T(b - Hx) \leq 0,$$

where $\bar{\pi}$ is an extreme ray of P with a positive objective function value, i.e.,

$$\bar{\pi}^T A \leq 0,\ \bar{\pi}^T(b - H\bar{x}) > 0.$$

Otherwise, we insert a so-called optimality cut to Ω^{opt}, which reads as

$$\bar{\pi}^T(b - Hx) \leq \eta,$$

where $\bar{\pi}$ is an optimal solution of (3).

The procedure is repeated until the optimal value of (3) coincides with $\bar{\eta}$. If extreme rays and optimal solutions of (3) are chosen out of a (necessarily finite) inner representation of P, this happens after a finite number of steps. We refer the interested reader to Benders (1962) for a detailed proof of convergence.

2.2. **Known Cut Selection Strategies.** If (3) has multiple optimal solutions in some iterations, the approach can be accelerated by choosing an appropriate one. In some cases, the appropriate solution implying an efficient cut does not have to be implied by an optimal solution of Problem (3). The following describes some approaches defining which cuts should be used.

2.2.1. *The Magnanti-Wong Method.* A well-known cut selection strategy has been proposed by Magnanti and Wong (1981). The idea is to generate a non-dominated cut. A cut implied by a solution $\bar{\pi}$ of Problem (3) dominates another cut implied by $\tilde{\pi}$, if and only if for all feasible master points $x \in X$

$$\bar{\pi}^T(b - Hx) \geq \tilde{\pi}^T(b - Hx),$$

with strict inequality for at least one $x \in X$. We call a cut non-dominated if no other cut dominates it. According to Magnanti and Wong (1981), a non-dominated cut can be obtained if a so-called core point that is contained in the relative interior of the convex hull of the feasible master solutions is known, i.e., a point $\tilde{x} \in \operatorname{relint}(\operatorname{conv}(X))$. In this case, in each cut-generating step, Problem (3) is solved first, providing a solution $\hat{\pi}$. Then the constraint

$$\pi^T(b - H\bar{x}) \geq \hat{\pi}^T(b - H\bar{x})$$

is inserted, while the objective function is changed to

$$\pi^T(b - H\tilde{x}).$$

Then, Problem (3) is solved again to get a possibly new solution $\bar{\pi}$, which is guaranteed to define a non-dominated cut. We refer to Magnanti and Wong (1981) for details regarding the proof of the statement. The procedure accelerates the approach for various applications, e.g., Froyland et al. (2014). A drawback is the necessity to determine an appropriate core point, which is difficult for some problems. The authors of Papadakos (2008) suggest a strategy to determine an approximate core point on the fly with the help of an iterative procedure to overcome this drawback.

2.2.2. *Cuts defined by Minimal Infeasible Subsystems.* Fischetti et al. (2010) proposed a method to determine cuts using MIS of the slightly adapted variant of the subproblem, which is the feasibility problem

$$\begin{aligned} \text{Find } y &\geq 0 \\ \text{s.t. } c^T y &\leq \bar{\eta} \\ Ay &\geq b - H\bar{x}. \end{aligned}$$

Their computational study implies significant speedups. The approach is implemented by adding a variable and a constraint to Problem (3). The modified dual subproblem then reads

$$\max \ \pi^T(b - H\bar{x}) - \pi_0\bar{\eta} \tag{4a}$$

$$\text{s.t. } \pi^T A \leq \pi_0 c^T \tag{4b}$$

$$\pi^T w + \pi_0 w_0 = 1 \tag{4c}$$

$$\pi \geq 0. \tag{4d}$$

The program uses an appropriately dimensioned positive-valued weight vector w, which can be arbitrarily chosen. The authors of Fischetti et al. (2010) suggest choosing w_i to be 1 for $i = 0$ and for indices i for which the corresponding row of H is not the zero row, and 0 otherwise. The cut that is generated by a solution $(\bar{\pi}, \bar{\pi}_0)$ of the modified dual subproblem (4) would read

$$\bar{\pi}^T(b - Hx) \leq \bar{\pi}_0 \eta.$$

An advantage of the proposed framework is that the feasibility and optimality cuts can be derived in the same way as dual solutions instead of extreme rays. If $\bar{\pi}_0$ is 0, we get a cut that does not depend on the variable η. Hence, in this case, it is a feasibility cut.

The authors of Fischetti et al. (2010) report considerable computational benefits compared to alternative cut selection strategies. Even though the MIS approach is performing extraordinarily well in practice, it has several drawbacks. It might generate cuts that do not support the epigraph of the subproblem's value function, as defined later in (5), at any point. As a result, multiple cuts might have to be generated to cut off a single assignment of master variables. Furthermore, the scaling of the constraints influences the behavior of the MIS approach and the existence of unnecessary constraints in the sense that they are implied by other constraints. These statements are substantiated by Example A.1, which can be found in Appendix A.

Furthermore, the strategy cannot be applied straightforwardly to solve problems with equality constraints, leading to dual variables that are not restricted in sign. If the dual variables are sign-restricted, Fischetti et al. (2010)'s choice of w is substantiated by the following Lemma.

Lemma 2.1. *Let H_i denote the i-th row of H, and let*

$$w_i := \begin{cases} 0 & \text{if } H_i = 0 \\ 1 & \text{otherwise,} \end{cases} \qquad \text{for all } i = 1, ..., m.$$

Then, it holds that if Problem (4) is unbounded, then Problem (1) is infeasible.

Proof. If Problem (4) is unbounded, there is an extreme ray $\bar{\pi}$ of Problem (4) with a positive value. As $\bar{\pi} \geq 0$ and $w^T\bar{\pi} = 0$ holds, we know that $\bar{\pi}_i = 0$ for all i with $H_i \neq 0$. Furthermore, $\bar{\pi}_0 = 0$ holds. Hence, we get that

$$0 < \bar{\pi}^T(b - H\bar{x}) - \bar{\pi}_0\bar{\eta} = \sum_{i:H_i=0} \bar{\pi}_i b_i$$

and as $\bar{\pi}^T A \leq 0$ we can obtain the valid Benders cut

$$\sum_{i:H_i=0} \bar{\pi}_i b_i \leq 0.$$

As the LHS is a positive real, Problem (1) must be infeasible. □

Lemma 2.1 no longer holds if π_i is not sign-restricted for an index i with $H_i \neq 0$. This is a side result of Example A.2, found in the appendix. In this case, choosing w as proposed by Fischetti et al. (2010) can lead to incorrect results, as demonstrated in Example A.2. Optimizing directly over the alternative polyhedron proposed in Fischetti et al. (2010) does not solve the problem either, as shown in Example A.3. Hence, to choose the weight vector gets non-trivial in this case.

2.2.3. *Facet-Generating Cuts.* Stursberg (2019) recently proposed a cut selection framework that parallels MIS cut selection. A selection of the dissertation results has also been published in the journal article Brandenberg and Stursberg (2021). The idea is to aim for cuts that represent facets of the polyhedral epigraph of the subproblem's value function,

$$\mathrm{E} := \{(x, \eta) \mid \eta \geq \min c^T y \colon Ay \geq b - Hx, y \geq 0\}. \tag{5}$$

The theory developed by Stursberg (2019) is based on reverse polar sets of the epigraph. Since immediate representations of reverse polar sets are in general intractable, the author solves instead linear optimization problems with objective function vector $(w, w_0) \in \mathbb{R}^{m+1}$ over the so-called Relaxed Alternative Polyhedron

$$P^{\leq}(x, \eta) := \{\pi, \pi_0 \geq 0 \mid \pi^T A + \pi_0 c^T = 0,\ \pi^T(b - Hx) + \pi_0\eta \geq 1\}. \tag{6}$$

These are under mild assumptions equivalent to the optimization problems over the reverse polar set.

According to Theorem 3.32 of Stursberg (2019), there is an optimal extremal point of the Relaxed Alternative Polyhedron inducing a cut that supports the epigraph in one of its facets under mild assumptions on (w, w_0).

In the empirical study, it turns out that setting (w, w_0) to $(H(\tilde{x} - \bar{x}), \tilde{\eta} - \bar{\eta})$ is quite efficient. The core point $(\tilde{x}, \tilde{\eta})$ is a possibly frequently updated, feasible solution of the master problem and the corresponding subproblem objective function value. The vector $(\bar{x}, \bar{\eta})$ is the current iterate. The author calls this strategy "adaptive cuts". Since this seems to be an efficient choice of (w, w_0), we will use this strategy in our benchmarks, calling the cuts "facet generating" or simply "facet".

Furthermore, the author proposes several equivalent variants of subproblems that can be chosen when the "facet" cut selection is applied. One is the following, slightly adapted from Stursberg (2019).

$$\max\ \pi^T(b - H\bar{x}) - \pi_0\bar{\eta} \tag{7a}$$

$$\text{s.t. } \pi^T A \leq \pi_0 c^T \tag{7b}$$

$$\pi^T H(\tilde{x} - \bar{x}) + \pi_0(\tilde{\eta} - \bar{\eta}) = 1 \tag{7c}$$

$$\pi \geq 0. \tag{7d}$$

We note that the difference between the MIS subproblem and the facet-generating subproblem reduces to an adapted normalization constraint: While (4c) is the normalization constraint of Problem (4), Constraint (7c) is the normalization constraint of Problem (7). Apart from the normalization constraint, the two problems are identical.

Remark. While MIS cut selection relies on sign-restricted dual variables, the facet approach can also be applied when the dual variables are free. Furthermore, as discussed before, MIS cut selection depends, e.g., on the scaling of certain constraints. In contrast to that, the author of Stursberg (2019) claims that a cut that is generated using Problem (7) supports the epigraph of the subproblem value function at a point on the connection line of the iterate $(\bar{x}, \bar{\eta})$ and the core point $(\tilde{x}, \tilde{\eta})$. This implies that facet-generating cuts are independent of algebraic transformations of the optimization problems under consideration that leave the epigraph of the subproblem unaffected.

3. β-Dominance and our Cut Selection Strategy.

In the following, we propose a new cut selection strategy that incorporates knowledge about the master problem and solutions obtained in previous iterations of the algorithm. The observation is that it is beneficial to determine a solution of the dual subproblem (3) that leads to a cut that excludes as many points as possible from having a chance to be an optimal solution to Problem (1). This idea is illustrated with the help of an example in Section 3.1. The set of points excluded to be optimal by a Benders cut implies a partial order on the set of valid Benders cuts. In Section 3.2, it is demonstrated that based on this partial order, Pareto-optimal Benders cuts can be obtained. It is further shown that Pareto-optimal cuts are automatically non-dominated in the sense of Magnanti and Wong (1981). Section 3.3 introduces so-called *optimal line-shifting cuts* and shows that they are Pareto-optimal. In Section 3.4 it is described how optimal line-shifting cuts can be calculated.

3.1. An Introductory Example.

The fact that an optimal dual solution does not necessarily lead to a cut that excludes as many master solutions as possible from being optimal is demonstrated with the following example.

Example 3.1. *Consider the following optimization problem, which is defined on the Euclidean plane:*

$$\min -\frac{1}{10}x + y \tag{8a}$$

$$s.t.\ \ y \geq \frac{1}{4}x \tag{8b}$$

$$y \geq -x \tag{8c}$$

$$y \geq -1 - 2x \tag{8d}$$

$$x \in [-2, 2]. \tag{8e}$$

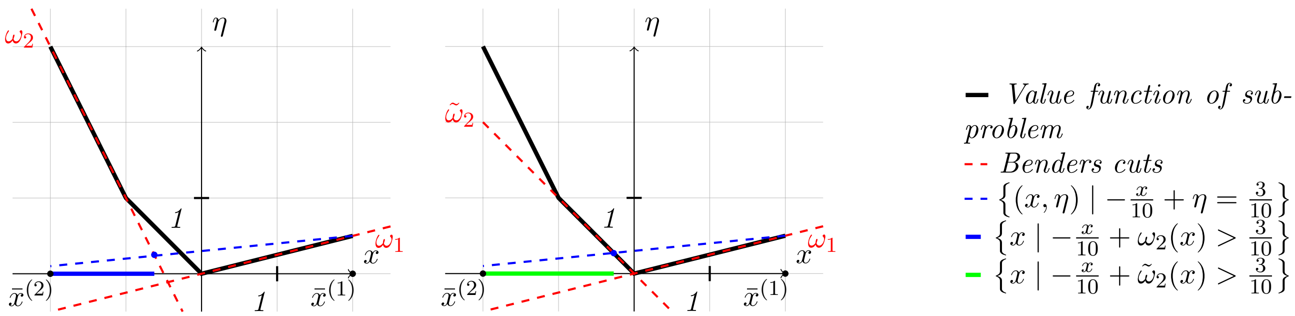
We consider x as the master variable and y as the subvariable.

We obtain $\bar{x}^{(1)} = 2$ by optimizing $\frac{-1}{10}x$ on the interval $[-2, 2]$ with a value of $-\frac{1}{5}$. The dual subproblem reads

$$\max \pi_1 \cdot \frac{1}{4}\bar{x} + \pi_2 \cdot (-\bar{x}) + \pi_3 \cdot (-1 - 2\bar{x}) \tag{9a}$$

$$\pi_1 + \pi_2 + \pi_3 = 1 \tag{9b}$$

$$\pi \geq 0. \tag{9c}$$



(A) Classical BD. (B) Minimizing C_ω.

FIGURE 1. Two different courses of BD.

Inserting $\bar{x}^{(1)}$, we obtain the solution $\bar{\pi} = (1,0,0)^T$ with value $\frac{1}{2}$ and derive the Benders cut that reads $\omega_1(x) = \frac{1}{4}x$. We note that we get $\frac{3}{10}$ as an upper bound for Problem (8). The master problem gets

$$\min -\frac{1}{10}x + \eta$$
$$\text{s.t. } \eta \geq \frac{1}{4}x.$$

Its unique solution is $\bar{x}^{(2)} = -2, \bar{\eta}^{(2)} = -\frac{1}{2}$, with a value of $-\frac{3}{10}$. The classical version of BD obtains the solution $\bar{\pi} = (0,0,1)$ to Problem (9), implying the cut $\omega_2(x) = -1 - 2x$ as shown in Figure 1a. An observation we make is that the solution of the next master problem is, in any case, in the set $C_{\tilde{\omega}_2}$ with

$$C_\omega := \left\{ x \mid -\frac{1}{10}x + \omega(x) \leq \frac{3}{10} \right\} \text{ for all valid Benders cuts } \omega,$$

as each solution that is not in this set would have a worse objective than $\frac{3}{10}$, and the point $(\tilde{x}, \tilde{\eta}) = (2, \frac{1}{2})$ is feasible and has an objective of exactly $\frac{3}{10}$. The complement of this set is shown as solid blue line in Figure 1a. The right border of the interval coincides with the x-coordinate of the intersection point of the dashed red line corresponding to the graph of ω_2 and the dashed blue line corresponding to the set $\{(x,\eta) \mid -\frac{x}{10} + \eta = \frac{3}{10}\}$.

The central idea of this article is now to aim for a Benders cut $\tilde{\omega}_2$ that has the property that C_ω is minimized over all valid cuts ω with regards to set inclusion. The Benders cut that minimizes this set is $\tilde{\omega}_2(x) = -x$, as shown in Figure 1b. The complement of $C_{\tilde{\omega}_2}$ is shown as the solid green line in Figure 1b. The right border of the interval coincides with the x-coordinate of the intersection point of the dashed red line corresponding to the graph of $\tilde{\omega}_2$ and the dashed blue line corresponding to the set $\{(x,\eta) \mid -\frac{x}{10} + \eta = \frac{3}{10}\}$.

We want to note that $\tilde{\omega}_2$ is the better choice, as it lets the master problem terminate immediately. In contrast, the algorithm needs another iteration if ω_2 is chosen. We also want to note that the inferior cut ω_2 is Pareto-optimal in the sense of Magnanti and Wong (1981), it is an MIS cut, and it is also facet-defining.

The example demonstrates that the existing cut selection strategies might fail to choose the clearly superior cut. Hence, we next present a new technique for generating Benders cuts, following the idea to minimize the set of potential solutions to the next master problem.

3.2. **Cut Selection Paradigm.** The cuts in the example aimed to minimize the size of the set of points in the master problem domain that are potentially better than an already obtained solution. This is the paradigm our cut selection strategy follows. Hence, we formally define a solution candidate set that contains all master solutions better than a threshold $\beta \in \mathbb{R}$ if the master problem contains either one or a set of Benders cuts. The value of β is often set to the incumbent value of Problem (1), i.e., $\beta = f^T x^* + S(x^*)$ for a solution $(x^*, S(x^*))$ of the master problem (2).

Definition 3.2 (Solution Candidate Set.)**.** *Given the master problem* (2) *with feasibility and optimality cuts* $\Omega = \Omega^{feas} \cup \Omega^{opt}$, *and* $\beta \in \mathbb{R}$, *we define the* solution candidate set (**SC-set**) *as*

$$C_{\Omega,\beta} := \left\{ x \in X \mid \exists \eta \colon f^T x + \eta \leq \beta, 0 \geq \omega(x) \text{ for } \omega \in \Omega^{feas}, \eta \geq \omega(x) \text{ for } \omega \in \Omega^{opt} \right\}.$$

We furthermore define $C_{\omega,\beta} := C_{\{\omega\},\beta}$. *For a feasible solution* $\tilde{x}$ *of Problem* (2) *with* $\tilde{\eta}$ *denoting the corresponding subproblem optimal value, we define the* SC-set *analogously as* $C_{\Omega,\tilde{x},\tilde{\eta}} := C_{\Omega,f^T\tilde{x}+\tilde{\eta}}$.

We first observe that the SC-set of a system of cuts and for a threshold, the value of the best solution found so far, is precisely the set where a potential optimal solution of (1) can be.

Observation 3.3. *Given Problem* (1) *and a system of valid cuts* Ω, *it holds that the set of points in* X *that can be completed to an optimal solution of* (1) *is a subset of* $C_{\Omega,\beta}$ *for all thresholds* β *that are not lower than the optimal value of* (1).

We can state a criterion that excludes a certain point in X from the SC-set for a certain threshold.

Observation 3.4. *Given a feasible master solution* $\bar{x} \in X$ *and a threshold* $\beta \in \mathbb{R}$. *Given an optimality cut* ω *induced by a feasible solution* $\bar{\pi}$ *of* (3), *it holds that*

$$\bar{x} \in C_{\omega,\beta} \Leftrightarrow f^T \bar{x} + \bar{\pi}^T (b - H\bar{x}) \leq \beta.$$

The observation implies that a version of the approach that generates a cut with a value greater than the current threshold at the point where it is generated has the property that the algorithm never visits a point in X twice before it terminates.

Next, we propose a dominance criterion similar to that of Magnanti and Wong (1981) based on SC-sets. One important advantage is that it also includes feasibility cuts, in contrast to Magnanti and Wong (1981)'s dominance notion.

Definition 3.5 (β-Dominance)**.** *Given the master problem* (2) *with valid cuts* ω_1 *and* ω_2, *cut* ω_1 β-*dominates cut* ω_2 *for a threshold* $\beta \in \mathbb{R}$, *if*

$$C_{\omega_1,\beta} \subsetneq C_{\omega_2,\beta}.$$

In the following we want to point out the connection between β-dominance and Magnanti and Wong (1981)'s dominance notion.

Theorem 3.6. *Given the master problem* (2) *with a convex feasible set* X *and two optimality cuts* ω_1 *and* ω_2 *with*

$$\omega_1(x) \geq \omega_2(x) \text{ for all } x \in X \text{ and } \omega_1(x) > \omega_2(x) \text{ for at least one } x \in X.$$

Then, ω_1 β-*dominates* ω_2 *for all thresholds* β *in the interval*

$$\mathcal{B} := \left(\inf \left\{ \tilde{\beta} \in \mathbb{R} \mid C_{\omega_2,\tilde{\beta}} \cap X \neq \emptyset \right\}, \sup \left\{ \tilde{\beta} \in \mathbb{R} \mid C_{\omega_1,\tilde{\beta}} \cap X \neq X \right\} \right).$$

Proof. Let $x' \in C_{\omega_1,\beta}$ for an arbitrary β. Then, there is an $\eta \in \mathbb{R}$ such that

$$f^T x' + \eta \leq \beta \text{ and } \eta \geq \omega_1(x').$$

Since $\omega_1(x') \geq \omega_2(x')$, this implies that for this choice of η

$$f^T x' + \eta \leq \beta \text{ and } \eta \geq \omega_2(x')$$

also applies, and hence $x' \in C_{\omega_2,\beta}$. Assume that there is a threshold β in $\mathcal{B}$ such that $C_{\omega_1,\beta} = C_{\omega_2,\beta}$. Let x' be a point for which $\omega_1(x') > \omega_2(x')$. If $x' \in C_{\omega_1,\beta}$, we choose $x'' \notin C_{\omega_1,\beta}$, if $x' \notin C_{\omega_1,\beta}$, we choose $x'' \in C_{\omega_1,\beta}$. The properties of $\mathcal{B}$ justify both choices.

We consider the case with $x' \in C_{\omega_1,\beta}$ first. This implies that

$$f^T x' + \omega_2(x') < f^T x' + \omega_1(x') \leq \beta.$$

On the other hand, we have

$$f^T x'' + \omega_1(x'') \geq f^T x'' + \omega_2(x'') > \beta.$$

We consider for $i \in \{1, 2\}$ the functions

$$g_i \colon \lambda \mapsto f^T(x' + \lambda(x'' - x')) + \omega_i(x' + \lambda(x'' - x')). \tag{10}$$

Since ω_i are affine functions, g_i is also an affine function and, therefore, continuous and either constant or strictly increasing or decreasing. We note that $g_2(\lambda) < g_1(\lambda)$ for all $\lambda \in [0, 1)$. We further note that $g_2(0) < \beta$ and $g_2(1) > \beta$. Since g_2 is continuous, there exists $\bar{\lambda} \in [0, 1)$ with $g_2(\bar{\lambda}) = \beta$. Since g_1 is strictly larger than g_2 on $[0, 1)$, we have that $g_1(\bar{\lambda}) > \beta$. Hence, $x' + \bar{\lambda}(x'' - x')$ is contained in $C_{\omega_2,\beta}$, but not in $C_{\omega_1,\beta}$. This is a contradiction to our assumption. The second case is analog. This proves the statement. ☐

We note that $\mathcal{B}$ contains the thresholds, for which at least one of the cuts ω_1, ω_2 has a candidate set that is neither X nor the empty set. Theorem 3.6 has the consequence that if a cut is generated that is not β-dominated for reasonable thresholds, it is also non-dominated in the sense of Magnanti and Wong (1981) under mild assumptions. The opposite direction is not valid, i.e., it is possible that a cut is non-dominated in the sense of Magnanti and Wong (1981) but is β-dominated, as already demonstrated in Example 3.1.

The following theorem states that a valid cut that β-dominates another cut does this for smaller thresholds under certain conditions. This further justifies choosing the dominating cut over the dominated one if these conditions apply to a pair of cuts. In this case, the dominating cut has smaller candidate sets over the whole remaining course of an algorithmic implementation of BD, for that β attains the best solution value found so far, which is non-increasing.

Theorem 3.7. *Given the master problem* (2) *with a convex feasible set X, a point $\bar{x} \in X$, a threshold $\bar{\beta}$ and two optimality cuts ω_1 and ω_2 with*

$$C_{\omega_1,\bar{\beta}} \subseteq C_{\omega_2,\bar{\beta}} \text{ and } \omega_2(\bar{x}) > \omega_1(\bar{x}) \geq \bar{\beta} - f^T \bar{x}.$$

Then it holds that

$$C_{\omega_1,\beta} \subsetneq C_{\omega_2,\beta}$$

for all $\beta \in \mathcal{B}$, where the interval $\mathcal{B}$ is defined as

$$\mathcal{B} := \left(\inf\left\{\tilde{\beta} \in \mathbb{R} \mid C_{\omega_1,\tilde{\beta}} \neq \emptyset\right\}, \bar{\beta}\right).$$

Proof. Assume there is a $\beta \in \mathcal{B}$ for which

$$C_{\omega_1,\beta} \nsubseteq C_{\omega_2,\beta}, \text{ i.e., } \exists x' \in X \text{ with } \omega_1(x') + f^T x' \leq \beta,\ \omega_2(x') + f^T x' > \beta.$$

We first note that $\omega_1(\bar{x}) + f^T\bar{x} > \bar{\beta}$, since otherwise $\bar{x} \in C_{\omega_1,\bar{\beta}}$ and $\bar{x} \notin C_{\omega_2,\bar{\beta}}$, which is a contradiction to our assumption. This implies that $x' \neq \bar{x}$. We further note that $\omega_2(x') + f^T x' \leq \bar{\beta}$ since otherwise $x' \in C_{\omega_1,\bar{\beta}}$ and $x' \notin C_{\omega_2,\bar{\beta}}$, which is a contradiction to our assumption. This implies that

$$\bar{\beta} < f^T\bar{x} + \omega_1(\bar{x}) < f^T\bar{x} + \omega_2(\bar{x}) \text{ and } f^T x' + \omega_1(x') < f^T x' + \omega_2(x') \leq \bar{\beta}.$$

We consider for $i \in \{1, 2\}$ the affine functions g_i as defined in Equation (10). We note that $g_2(\lambda) > g_1(\lambda)$ for all $\lambda \in [0, 1]$. We further note that $g_1(0) > \bar{\beta}$ and $g_1(1) < \bar{\beta}$. Since g_1 is continuous, there exists $\bar{\lambda} \in [0, 1]$ with $g_1(\bar{\lambda}) = \bar{\beta}$. Since g_2 is strictly larger than g_1 on $[0, 1]$, we have that $g_2(\bar{\lambda}) > \bar{\beta}$. Hence, $\bar{x} + \bar{\lambda}(x' - \bar{x})$ is contained in $C_{\omega_1,\bar{\beta}}$, but not in $C_{\omega_2,\bar{\beta}}$. This is a contradiction to our assumption, and we can conclude that $C_{\omega_1,\beta} \subseteq C_{\omega_2,\beta}$.

Assume there is a $\beta \in \mathcal{B}$ for which

$$C_{\omega_1,\beta} = C_{\omega_2,\beta}.$$

Then there exists $x' \in X$ such that

$$\omega_1(x') + f^T x' = \beta.$$

This holds because $\omega_2(\bar{x}) + f^T\bar{x} \geq \bar{\beta}$ and $C_{\omega_1,\bar{\beta}} \subseteq C_{\omega_2,\bar{\beta}}$ imply that $\bar{x} \in C_{\omega_1,\beta}$, and this implies that $\emptyset \neq X \setminus C_{\omega_1,\beta}$. Furthermore, $\beta \in \mathcal{B}$ implies directly that $\emptyset \neq C_{\omega_1,\beta}$. It further holds that

$$\omega_2(x') + f^T x' = \beta$$

as well, since $\omega_2(x') + f^T x' > \beta$ would imply that the candidate sets are not equal. If $\omega_2(x') + f^T x' < \beta$ would hold, we could define g_i as in Equation (10) for $i = 1, 2$ and could observe that there is $\bar{\lambda}$ close to 1 such that $g_1(\bar{\lambda}) > \beta$, but $g_2(\bar{\lambda}) < \beta$. This would also imply that the candidate sets are not equal.

So, to conclude our argument, we take g_i as defined in Equation (10) for $i = 1, 2$. We note that $g_2(\lambda) > g_1(\lambda)$ for $\lambda \in [0, 1)$. Since g_1 is continuous, this interval contains a point $\bar{\lambda}$ for which $g_1(\bar{\lambda}) = \bar{\beta}$, and this implies that $g_2(\bar{\lambda}) > \bar{\beta}$. This implies that $\bar{x} + \bar{\lambda}(x' - \bar{x}) \in C_{\omega_1,\bar{\beta}}$, but not in $C_{\omega_2,\bar{\beta}}$. This is a contradiction to our assumption and concludes the proof. ☐

Remark. For points $\bar{x}$ with $S(\bar{x}) = \infty$ and certain thresholds β, the only non-dominated cuts may be optimality cuts. Example A.4 in Appendix A shows a situation where this happens. Nevertheless, it may also happen that only feasibility cuts are not β-dominated, as demonstrated in Example A.5 in Appendix A.

In the following section, we describe how cuts can be obtained that are not β-dominated.

3.3. **Line-Shifting Cuts.** We introduce an approach to generate Benders cuts that are, under certain conditions, non-dominated in the sense of Definition 3.5. The idea is to exclude as many points on a line from the current master solution $\bar{x} \in X$ to a core point $\tilde{x} \in \text{conv}(X)$ from further consideration. We call each valid Benders cut that excludes at least $\bar{x}$ from further consideration a "line-shifting cut". Formally, line-shifting cuts are given by Definition 3.8.

Remark. Even though it is formally not necessary, we think that it could simplify to understand the following contents if one interprets $(\bar{x}, \bar{\eta})$ as the current master solution or "iterate", $\tilde{x}$ as core point, x^*

as the incumbent solution. It might also be helpful to interpret the value of β as the incumbent value, i.e., $\beta = f^T x^* + S(x^*)$. We note that the incumbent can always serve as a core point.

Definition 3.8 (Line-shifting cut)**.** *Given the master problem* (2) *with a feasible set* X*, and points* $\tilde{x} \in conv(X)$*,* $\bar{x} \in X$*, and* $\beta \in \mathbb{R}$ *with the property that*

$$f^T\tilde{x} + S(\tilde{x}) \leq \beta \quad \textit{and } f^T\bar{x} + S(\bar{x}) \geq \beta,$$

we define a line-shifting cut *induced by* $(\bar{x}, \tilde{x}, \beta)$ *as an optimality cut* $\bar{\omega}$*, for which a parameter* $\mu \in [0, 1]$ *exists, such that*

$$\begin{aligned} f^T(\bar{x} + \mu(\tilde{x} - \bar{x})) + \bar{\omega}(\bar{x} + \mu(\tilde{x} - \bar{x})) &= \beta, \\ f^T\bar{x} + \bar{\omega}(\bar{x}) &\geq \beta, \\ f^T(\bar{x} + \tilde{\mu}(\tilde{x} - \bar{x})) + \bar{\omega}(\bar{x} + \tilde{\mu}(\tilde{x} - \bar{x})) &< \beta \textit{ for all } \tilde{\mu} \in (\mu, 1), \end{aligned}$$

or a feasibility cut, $\bar{\omega}$*, for which a parameter* $\mu \in [0, 1]$ *exists, such that*

$$\begin{aligned} \bar{\omega}(\bar{x} + \mu(\tilde{x} - \bar{x})) &= 0, \\ \bar{\omega}(\bar{x}) &> 0, \\ \bar{\omega}(\bar{x} + \tilde{\mu}(\tilde{x} - \bar{x})) &< 0 \textit{ for all } \tilde{\mu} \in (\mu, 1). \end{aligned}$$

In both cases, we call μ *the* depth *of the line-shifting cut.*

We observe that for all points $\bar{x} \in X$, arbitrary feasibility cuts or optimality cuts $\bar{\omega}$ with $f^T\bar{x} + \bar{\omega}(\bar{x}) \geq \beta$ are line-shifting cuts with a certain depth.

We call a line-shifting cut with maximum depth optimal line-shifting (**OLS**) cut.

Definition 3.9 (OLS cut)**.** *Under the conditions of Definition 3.8, we define an* OLS cut *induced by* $(\bar{x}, \tilde{x}, \beta)$ *as an optimal solution to the optimization problem*

$$\max\{\mu(\omega) \mid \omega \textit{ is a line-shifting cut induced by } (\bar{x}, \tilde{x}, \beta)\},$$

where μ *maps line-shifting cuts to their depth.*

The concepts of the depth of a line-shifting cut and OLS cuts are illustrated in Figure 2, showing the two cuts generated in Example 3.1 for classical BD.

Continuance of Example 3.1. *We recall that the classical version of BD obtained the cut* $\omega_2 = -1 - 2x$*, and the improved version of BD obtained the cut* $\tilde{\omega}_2 = -x$*. After one iteration, the incumbent and core point of Problem* (8) *is* $\bar{x}^{(1)} = 2$ *with a value of* $\frac{3}{10} := \beta$*. The iterate of iteration* 2 *is* $\bar{x}^{(2)} = -2$*. Both,* ω_2 *and* $\tilde{\omega}_2$ *are line-shifting cuts induced by* $(\bar{x}^{(2)}, \bar{x}^{(1)}, \beta)$*, as*

$$f^T\bar{x}^{(2)} + \omega_2(\bar{x}^{(2)}) = \frac{2}{10} + 3 > \frac{3}{10} = \beta \textit{ and } f^T\bar{x}^{(2)} + \tilde{\omega}_2(\bar{x}^{(2)}) = \frac{2}{10} + 2 > \frac{3}{10} = \beta.$$

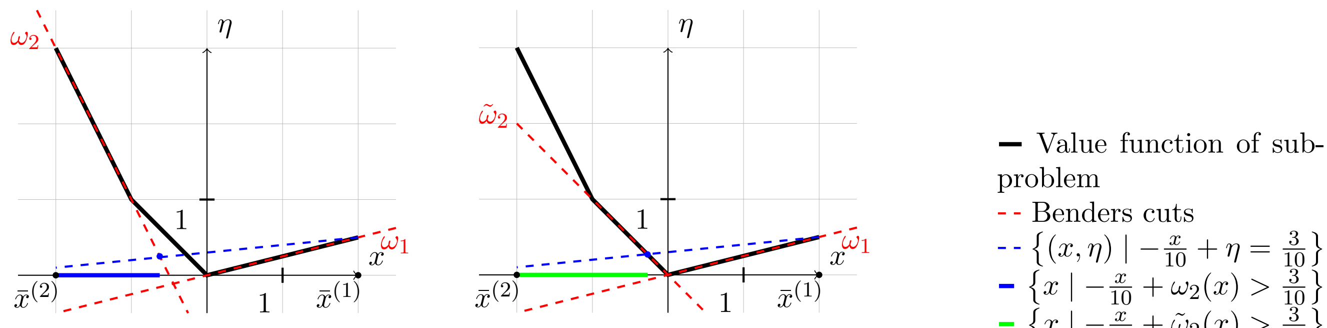
The depth of ω_2 *is* $\mu = \frac{29}{84}$*, as*

$$f^T(\bar{x}^{(2)} + \mu(\bar{x}^{(1)} - \bar{x}^{(2)})) + \omega_2(\bar{x}^{(2)} + \mu(\bar{x}^{(1)} - \bar{x}^{(2)})) = \frac{-1}{10}(-2 + \frac{29}{84}(2 - (-2))) - 1 - 2(-2 + \frac{29}{84}(2 - (-2))) = \frac{3}{10} = \beta.$$

The depth of $\tilde{\omega}_2$ *is* $\tilde{\mu} = \frac{19}{44}$*, as*

$$f^T(\bar{x}^{(2)} + \tilde{\mu}(\bar{x}^{(1)} - \bar{x}^{(2)})) + \tilde{\omega}_2(\bar{x}^{(2)} + \tilde{\mu}(\bar{x}^{(1)} - \bar{x}^{(2)})) = \frac{-1}{10}(-2 + \frac{19}{44}(2 - (-2))) - (-2 + \frac{19}{44}(2 - (-2))) = \frac{3}{10} = \beta.$$

There is no line-shifting cut induced by $(\bar{x}^{(2)}, \bar{x}^{(1)}, \beta)$ *with a higher depth. Hence,* $\tilde{\omega}_2$ *is an OLS cut.*



(A) Benders cut ω_2 with a depth of $\mu = \frac{29}{84}$, illustrated by the thick blue line that covers a proportion of μ of the distance between $\bar{x}^{(2)}$ and $\bar{x}^{(1)}$.

(B) Benders cut $\tilde{\omega}_2$ with a depth of $\tilde{\mu} = \frac{19}{44}$, illustrated by the thick green line that covers a proportion of $\tilde{\mu}$ of the distance between $\bar{x}^{(2)}$ and $\bar{x}^{(1)}$.

FIGURE 2. Comparison of cut depth for two different line-shifting cuts.

The following theorem establishes a connection between OLS cuts and β-dominance. It states that OLS cuts are not β-dominated under certain conditions.

Theorem 3.10. *Given the master problem* (2) *with a convex feasible set* X*, and points* $\tilde{x} \in X$*,* $\bar{x} \in X$*, and* $\beta \in \mathbb{R}$*. An OLS cut* $\bar{\omega}$ *induced by* $(\bar{x}, \tilde{x}, \beta)$ *with depth* $\bar{\mu} < 1$ *is not* β*-dominated, if*

$$x' := \bar{x} + \bar{\mu}(\tilde{x} - \bar{x}) \in \mathit{relint}(X).$$

Proof. We prove statement that a line-shifting cut that is an optimality cut is not β-dominated by another optimality cut. This proof can be easily adapted for feasibility cuts. For this reason, we omit analog reasoning.

Assume there is a valid Benders cut $\tilde{\omega}$ that β-dominates $\bar{\omega}$, i.e.,

$$C_{\tilde{\omega},\beta} \subsetneq C_{\bar{\omega},\beta}.$$

We first note that $\bar{\mu} < 1$ implies that $f^T\bar{x} + \bar{\omega}(\bar{x}) > \beta$. This in turn implies that $f^T\bar{x} + \tilde{\omega}(\bar{x}) > \beta$ holds, since otherwise $\bar{x} \in C_{\tilde{\omega},\beta}$, but $\bar{x} \notin C_{\bar{\omega},\beta}$. This implies that $\tilde{\omega}$ is a line-shifting cut. Hence, it has a depth. The depth of $\tilde{\omega}$ is $\bar{\mu}$ as well: As $\bar{\omega}$ has been a line-shifting cut with maximum depth, $\tilde{\mu} \leq \bar{\mu}$ has to hold for the depth $\tilde{\mu}$ of $\tilde{\omega}$. If $\tilde{\mu} < \bar{\mu}$ would hold, then $x' \in C_{\tilde{\omega},\beta}$, but $x' \notin C_{\bar{\omega},\beta}$.

As we assumed that $\tilde{\omega}$ β-dominates $\bar{\omega}$, there exists a point $x'' \in X$ with $f^Tx'' + \tilde{\omega}(x'') > \beta$, but $f^Tx'' + \bar{\omega}(x'') \leq \beta$. Since $x' \in \text{relint}(X)$, there exists a parameter $\theta > 1$, such that $x'' + \theta(x' - x'') \in X$. This in turn implies that $f^Tx'' + \bar{\omega}(x'') = \beta$, as well as $f^T(x'' + \theta(x' - x'')) + \bar{\omega}(x'' + \theta(x' - x'')) = \beta$.

We note that $f^Tx + \tilde{\omega}(x) > \beta$ has to hold for all x in a relatively open neighborhood of $\frac{x''+x'}{2}$. Using the same argument as before, on this open neighborhood, it holds that $f^T(x) + \bar{\omega}(x) = \beta$, i.e., the affine function $f^T(x) + \bar{\omega}(x)$ is constant on this relatively open neighborhood. As X is convex, we can infer that $f^Tx + \bar{\omega}(x)$ is constant on X. This implies that $\bar{\omega}$ is a line-shifting cut with depth 1, which contradicts the assumptions made upfront. □

Remark. If X is non-convex, it is possible to calculate OLS cuts, but they might be β-dominated. Nevertheless, OLS cuts are helpful for solution methods that work with convex supersets of X rather than with X itself, e.g., Branch-and-Benders-Cut, as they are non-dominated on these convex supersets.

3.4. **Computing OLS Cuts.** This section describes how OLS cuts can be obtained. The most important results are the following: Lemma 3.12 gives a linear program whose solution implies an optimality cut that is line-shifting with maximum depth. Lemma 3.14 gives a linear program whose solution implies a line-shifting feasibility cut with maximum depth. Theorem 3.16 combines the results of the two lemmas and gives a linear program whose solution is an OLS cut, i.e., a valid line-shifting Benders cut with maximum depth. The two lemmas and the theorem have as conditions that the current iterate does not improve the incumbent's value. We recall that the idea of generating non-dominated Benders cuts is to get the value of as many points as possible in X over the old incumbent value in the next iteration. This goal is desirable if the current iterate does not improve the incumbent value. If it improves the incumbent value, a standard Benders cut would be preferable, as it admits to calculating the new incumbent value and gives a cut that excludes the current iterate from being chosen in subsequent iterations. From an algorithmic point of view, it is desirable not to have the obligation to check if the current iterate generates a better incumbent value. Hence, we proved Lemma 3.17, which states that the linear program given in Theorem 3.16 attains a non-positive value or is infeasible if the conditions of Theorem 3.16 are violated.

3.4.1. *Computation of optimality cuts.* The following proposition states that an optimality cut that is a line-shifting cut with maximum depth is the solution of a fractional linear program.

Proposition 3.11. *Consider Problem* (1) *and the corresponding master problem* (2) *with feasible set* X*, a threshold* $\beta \in \mathbb{R}$ *and points* $\tilde{x} \in conv(X)$*,* $\bar{x} \in X$*, with* $f^T\bar{x} + S(\bar{x}) \geq \beta$*,* $f^T\tilde{x} + S(\tilde{x}) \leq \beta$*. If no optimality cut* $\bar{\omega}$ *exists for which* $\bar{\omega}(\tilde{x}) + f^T\tilde{x} = \beta$ *and* $\bar{\omega}(\bar{x}) + f^T\bar{x} = \beta$*, then a solution* $(\bar{\pi}, \bar{\mu})$ *of the following optimization problem defines an optimality cut that is a line-shifting cut induced by* $(\bar{x}, \tilde{x}, \beta)$ *with depth* $\bar{\mu} \leq 1$*, and no optimality cut that is a line-shifting cut induced by this triple has a higher depth.*

$$\max \quad \mu \tag{11a}$$

$$\text{s.t. } \pi^T(b - H(\bar{x} + \mu(\tilde{x} - \bar{x}))) + f^T(\bar{x} + \mu(\tilde{x} - \bar{x})) = \beta \tag{11b}$$

$$\pi^T A \leq c^T \tag{11c}$$

$$\pi^T(b - H\bar{x}) + f^T\bar{x} \geq \beta \tag{11d}$$

$$\pi \geq 0, \ \mu \geq 0. \tag{11e}$$

The optimization problem is not unbounded and has at least one feasible solution.

Containing products of variables, Problem (11) is no longer a linear program, since it contains the bilinear constraint (11b). Nevertheless, Problem (11) can be reformulated to a fractional linear program. Under some mild assumptions the fractional linear program can be reformulated to an equivalent linear program with one variable that is restricted to be strictly positive, applying the method presented in Charnes and Cooper (1962). This is the statement of Lemma 3.12.

Lemma 3.12. *Consider Problem* (1) *and the corresponding master problem* (2) *with feasible set* X*, a threshold* $\beta \in \mathbb{R}$ *and points* $\tilde{x} \in conv(X)$*,* $\bar{x} \in X$*, with* $f^T\bar{x} + S(\bar{x}) \geq \beta \geq f^T\tilde{x} + S(\tilde{x})$*.*

i) If no optimality cut $\bar{\omega}$ exists for which $\bar{\omega}(\tilde{x}) + f^T\tilde{x} = \beta$ and $\bar{\omega}(\bar{x}) + f^T\bar{x} = \beta$, then each optimal solution $(\bar{\rho}, \bar{\alpha})$ of

$$\max \quad \rho^T(b - H\bar{x}) + \alpha(f^T\bar{x} - \beta) \tag{12a}$$

$$\rho^T H(\tilde{x} - \bar{x}) - \alpha f^T(\tilde{x} - \bar{x}) = 1 \tag{12b}$$

$$\rho^T A \leq \alpha c^T \tag{12c}$$

$$\alpha \geq 0, \ \ \rho \geq 0. \tag{12d}$$

has the property that $\bar{\pi} := \frac{\bar{\rho}}{\bar{\alpha}}$, $\bar{\mu} := \bar{\rho}^T(b - H\bar{x}) + \bar{\alpha}(f^T\bar{x} - \beta)$ is an optimal solution of Problem (11)*, if $\bar{\alpha} > 0$.*

ii) If $S(\bar{x}) < \infty$, then $\bar{\alpha} > 0$.

Remark. As a consequence of Lemma 3.14 and Theorem 3.16, presented later in this paper, it follows that the condition $S(\bar{x}) < \infty$ is not restrictive. Even if it does not hold, Problem (12) can be a surrogate for the dual subproblem.

Proof. For i), it is easy to check that Constraint (11b) is equivalent to

$$\mu = \frac{\pi^T(b - H\bar{x}) + f^T\bar{x} - \beta}{\pi^T H(\tilde{x} - \bar{x}) - f^T(\tilde{x} - \bar{x})}$$

if we can prove that the denominator is nonzero over the feasible set of Problem (11). Hence, μ can be replaced in the objective function of Problem (11). The result is an optimization problem with linear constraints and a fractional objective function. We use the approach of Charnes and Cooper (1962) to reformulate this optimization problem to obtain an equivalent linear program.

To apply the reformulation method, one has to ensure that the denominator does not change its sign over the set of feasible points of the program to be reformulated. This requirement is fulfilled for Problem (11): If we assume that there is a feasible solution $\bar{\pi}$ of Problem (11), for which the denominator is not positive, we can derive

$$\begin{aligned} 0 &\geq \bar{\pi}^T H(\tilde{x} - \bar{x}) - f^T(\tilde{x} - \bar{x}) \\ &= \bar{\pi}^T(b - H\bar{x}) + f^T\bar{x} - (\bar{\pi}^T(b - H\tilde{x}) + f^T\tilde{x}) \\ &> \beta - \beta = 0. \end{aligned}$$

The last inequality is implied by Constraint (11d) and the condition $\beta \geq f^T\tilde{x} + S(\tilde{x})$, implying $\beta \geq f^T\tilde{x} + \bar{\pi}^T(b - H\tilde{x})$. It is strict because we assumed that no cuts exist for which $\bar{\omega}(\tilde{x}) + f^T\tilde{x} = \beta$ and $\bar{\omega}(\bar{x}) + f^T\bar{x} = \beta$.

Hence, we can reformulate Problem (11) by introducing a new variable $\alpha > 0$ and the constraint

$$\alpha = \frac{1}{\pi^T H(\tilde{x} - \bar{x}) - f^T(\tilde{x} - \bar{x})} \Leftrightarrow \alpha(\pi^T H(\tilde{x} - \bar{x}) - f^T(\tilde{x} - \bar{x})) = 1.$$

The objective function gets

$$\alpha(\pi^T(b - H\bar{x}) + f^T\bar{x} - \beta).$$

Further, we can multiply Constraints (11b) to (11d) with α. We substitute $\alpha\pi = \rho$, relax $\alpha > 0$ to $\alpha \geq 0$, and drop the reformulation of Constraint (11d). This yields

$$\rho^T(b - H\bar{x}) \geq \alpha(\beta - f^T\bar{x}).$$

This constraint bounds the objective function from below and is, hence, not necessary. Adding everything together, we get Problem (12).

For ii), it remains to show that an optimal solution $(\bar{\rho}, \bar{\alpha})$ of Problem (12) has the property $\bar{\alpha} > 0$. Assume, $\bar{\alpha} = 0$. Then, we know that

$$\bar{\rho}^T(b - H\bar{x}) \leq \begin{cases} \max \pi^T(b - H\bar{x}) \\ \text{s.t. } \pi^T A \leq 0 \\ \pi \geq 0. \end{cases}$$

The optimal value is 0 since we assumed $S(\bar{x}) < \infty$. Nevertheless, there is always a solution $\tilde{\pi}$ of Problem (3) implying a cut $\tilde{\omega}$ that has the property that $\tilde{\omega}(\bar{x}) = S(\bar{x})$. If we define

$$\tilde{\alpha} := \frac{1}{\tilde{\pi}^T H(\tilde{x} - \bar{x}) - f^T(\tilde{x} - \bar{x})}, \ \tilde{\rho} := \tilde{\alpha}\tilde{\pi},$$

This is a feasible solution of Problem (12) with a positive objective function value - a contradiction. □

3.4.2. *Computation of feasibility cuts.* One can directly obtain results for the line-shifting cut procedure including only feasibility cuts, which are presented next.

Proposition 3.13. *Given Problem* (1) *and the corresponding master problem* (2) *with feasible set and points* $\tilde{x} \in conv(X)$, $\bar{x} \in X$, *with* $S(\tilde{x}) < \infty$ *and* $S(\bar{x}) = \infty$. *A solution* $(\bar{\pi}, \bar{\mu})$ *of the following optimization problem is a line-shifting cut with depth* $\bar{\mu}$ *induced by* $(\tilde{x}, \bar{x}, \beta)$ *for all* $\beta \in \mathbb{R}$, *and no feasibility cut induced by this triple has a higher depth.*

$$\max \quad \mu \tag{13a}$$
$$\text{s.t. } \pi^T(b - H(\bar{x} + \mu(\tilde{x} - \bar{x}))) = 0 \tag{13b}$$
$$\pi^T A \leq 0 \tag{13c}$$
$$\pi^T(b - H\bar{x}) = 1 \tag{13d}$$
$$\pi \geq 0, \ \mu \geq 0. \tag{13e}$$

The optimization problem is not unbounded and has at least one feasible solution.

The following lemma states that Problem (13) can be reformulated similarly to Problem (11).

Lemma 3.14. *Given Problem* (1) *and the corresponding master problem* (2) *with feasible set and points* $\tilde{x} \in conv(X)$, $\bar{x} \in X$, *with* $S(\tilde{x}) < \infty$ *and* $S(\bar{x}) = \infty$. *Then, each optimal solution* $\bar{\rho}$ *of*

$$\max \quad \rho^T(b - H\bar{x}) \tag{14a}$$
$$\rho^T H(\tilde{x} - \bar{x}) = 1 \tag{14b}$$
$$\rho^T A \leq 0 \tag{14c}$$
$$\rho \geq 0 \tag{14d}$$

has the property that $\bar{\pi} := \frac{\bar{\rho}}{\bar{\rho}^T(b - H\bar{x})}$, $\bar{\mu} := \bar{\rho}^T(b - H\bar{x})$ *is an optimal solution of Problem* (13).

Proof. It is easy to check that Constraint (13b) is equivalent to

$$\mu = \frac{\pi^T(b - H\bar{x})}{\pi^T H(\tilde{x} - \bar{x})}.$$

Hence, μ can be replaced in the objective function of Problem (13). The result is an optimization problem with linear constraints and a fractional objective function. We use the approach of Charnes and Cooper (1962) to reformulate this optimization problem to obtain an equivalent linear program.

For the reformulation method to be applicable, it has to be guaranteed that the denominator does not change its sign over the whole feasible set of the problem to be reformulated. This requirement is fulfilled for Problem (13): If we assume that there is a feasible solution $\bar{\pi}$ of Problem (13), for which the denominator is not positive, we can derive

$$\begin{aligned} 0 &\geq \bar{\pi}^T H(\tilde{x} - \bar{x}) \\ &= \bar{\pi}^T (b - H\bar{x}) - \bar{\pi}^T (b - H\tilde{x}) \\ &\geq 1 - 0 = 1. \end{aligned}$$

The last inequality can be derived from Constraint (13d) and the conditions of the Lemma.

Hence, we can reformulate Problem (13) by introducing a new variable $\alpha > 0$ and the constraint

$$\alpha = \frac{1}{\pi^T H(\tilde{x} - \bar{x})} \Leftrightarrow \alpha \pi^T H(\tilde{x} - \bar{x}) = 1.$$

The objective function gets

$$\alpha \pi^T (b - H\bar{x}) = \alpha.$$

Furthermore, we can multiply Constraints (13b) to (13d) with α. Substituting $\alpha\pi = \rho$ and substituting α by $\rho^T(b - H\bar{x})$ according to Constraint (13d) leads to Problem (14). An optimal solution to Problem (14) has a positive optimal value. □

3.4.3. *Joint computation of optimality and feasibility cuts.* We note that Problem (14) is Problem (12) with α restricted to be 0. Hence, the following lemma connects the results regarding feasibility cuts and optimality cuts.

Lemma 3.15. *Given Problem* (1) *and the corresponding master problem* (2) *with feasible set X, a threshold $\beta \in \mathbb{R}$ and points $\tilde{x} \in conv(X)$, $\bar{x} \in X$, with $f^T\bar{x} + S(\bar{x}) \geq \beta$, $f^T\tilde{x} + S(\tilde{x}) \leq \beta$. If there is no optimality cut $\bar{\omega}$ with $\bar{\omega}(\tilde{x}) + f^T\tilde{x} = \beta$ and $\bar{\omega}(\bar{x}) + f^T\bar{x} = \beta$ and if $\bar{\alpha} = 0$ holds for an optimal solution $(\bar{\rho}, \bar{\alpha})$ to Problem* (12)*, then the optimal value of Problem* (11) *is not greater than the optimal value of Problem* (13)*.*

Proof. Assume the optimal value of Problem (11) is $\bar{\mu}$, attained at $\bar{\pi}$, and the optimal value of Problem (13) is $\tilde{\mu}$ with $\bar{\mu} > \tilde{\mu}$. Then, the constraint

$$\pi^T (b - H\bar{x}) \leq \bar{\pi}^T (b - H\bar{x})$$

can be added to Problem (11) without changing the optimal value. This corresponds to an alternative dual subproblem that has no unbounded rays. Lemma 3.12 ascertains that in this case, for an optimal solution $(\tilde{\rho}, \tilde{\alpha})$ of the modified subproblem (12), it holds that $\tilde{\alpha} > 0$, with an optimal value of $\bar{\mu}$. Since $(\bar{\rho}, \bar{\alpha})$ is an optimal solution to the original Problem (12), its value is hence at least $\bar{\mu}$. Since $\bar{\rho}$ is feasible for Problem (14) and the optimal values of Problem (13) and Problem (14) coincide, we obtain

$$\tilde{\mu} \geq \bar{\mu} > \tilde{\mu},$$

a contradiction. □

We get the following result by merging Lemmas 3.12, 3.14, and 3.15.

Theorem 3.16. *Given Problem* (1) *and the corresponding master problem* (2) *with feasible set X, a threshold $\beta \in \mathbb{R}$ and points $\tilde{x} \in conv(X)$, $\bar{x} \in X$, with $f^T\bar{x} + S(\bar{x}) \geq \beta$, $f^T\tilde{x} + S(\tilde{x}) \leq \beta$. If no optimality cut $\bar{\omega}$ exists for which $\bar{\omega}(\tilde{x}) + f^T\tilde{x} = \beta$ and $\bar{\omega}(\bar{x}) + f^T\bar{x} = \beta$, then each optimal solution $(\bar{\rho}, \bar{\alpha})$ of Problem* (12) *defines a line-shifting cut with maximal depth. The depth equals the optimal value of the problem, and the cut added to Problem* (2) *reads*

$$\bar{\rho}^T(b - Hx) \leq \bar{\alpha}\eta.$$

Proof. A line-shifting cut with maximal depth is either a feasibility cut, captured by Problem (13), or an optimality cut, captured by Problem (11), depending on which problem has the higher optimal value. If we consider a solution $(\bar{\rho}, \bar{\alpha})$ of Problem (12) with $\bar{\alpha} = 0$, Lemma 3.15 guarantees that the optimal value of Problem (13) is not lower than the optimal value of Problem (11). Furthermore, $\bar{\rho}$ is feasible for Problem (14), and we can apply Lemma 3.14 to transform this solution into a feasibility cut that reads

$$\bar{\rho}^T(b - Hx) \leq 0.$$

If $\bar{\alpha} > 0$, the solution is also a solution to Problem (12) with α restricted to be positive. We further know that the optimal value of this problem is not lower than the solution of Problem (14), since it is a relaxation of the latter. The solution $(\bar{\rho}, \bar{\alpha})$ can be transformed by applying Lemma 3.12 to a solution of Problem (11) with the same objective function value, and this can be transformed into a valid optimality cut that reads

$$\bar{\rho}^T(b - Hx) \leq \bar{\alpha}\eta.$$

Taking everything together, we get the desired result. □

The following lemma states that Problem (12) is infeasible or has a non-positive value if the requirements of Lemma 3.12 are not met, i.e., if $f^Tx^* + S(x^*) > f^T\bar{x} + S(\bar{x})$ for the current iterate $\bar{x}$ and incumbent x^*.

Lemma 3.17. *Given Problem* (1) *and the corresponding master problem* (2) *with feasible set X, a threshold $\beta \in \mathbb{R}$ and points $\tilde{x} \in conv(X)$, $\bar{x} \in X$, with $f^T\bar{x} + S(\bar{x}) < \beta$, $f^T\tilde{x} + S(\tilde{x}) \leq \beta$. Then, Problem* (12) *is infeasible or has a non-positive optimal value.*

Proof. Assume that Problem (12) has an optimal value that is positive, or is unbounded. Then, it has a feasible solution $(\bar{\rho}, \bar{\alpha})$ with a positive value. If $\bar{\alpha} = 0$, we can derive that $\bar{\rho}^T A \leq 0$ and $\bar{\rho}^T(b - H\bar{x}) > 0$. This translates into a valid feasibility cut that cuts off $\bar{x}$, a contradiction. If $\bar{\alpha} > 0$, we can derive that $\frac{1}{\bar{\alpha}}\bar{\rho}^T A \leq c^T$, and $\frac{1}{\bar{\alpha}}\bar{\rho}^T(b - H\bar{x}) + f^T\bar{x} > \beta$. This translates into a valid optimality cut that guarantees that $f^T\bar{x} + S(\bar{x}) > \beta$, which is also a contradiction. □

The construction of OLS cuts relies on knowing good feasible solutions to the original problem (1). We observed that it is beneficial if the feasible solution the approach relies on is varied in every iteration. This is the motivation of the following lemma, stating how a new core point in the convex hull of the master problem's feasible space can be constructed after an OLS cut is generated.

Lemma 3.18. *Given Problem* (1) *and the corresponding master problem* (2) *with feasible set X, a threshold $\beta \in \mathbb{R}$ and points $\tilde{x} \in conv(X)$, $\bar{x} \in X$, with $f^T\bar{x} + S(\bar{x}) \geq \beta$, $f^T\tilde{x} + S(\tilde{x}) \leq \beta$. Let furthermore $\bar{\mu} < 1$ denote the optimal value of Problem* (12) *corresponding to this setting, and let $x' := \bar{x} + \bar{\mu}(\tilde{x} - \bar{x})$.*

If no optimality cut $\bar{\omega}$ exists for which $\bar{\omega}(\tilde{x}) + f^T\tilde{x} = \beta$ and $\bar{\omega}(\bar{x}) + f^T\bar{x} = \beta$, then it holds that

$$f^T x' + S(x') \leq \beta,$$
$$f^T x + S(x) < \beta \quad \textit{for all } x \in \{x' + \lambda(\tilde{x} - x'), \lambda \in (0,1)\}.$$

Proof. First, we show $f^T x' + S(x') \leq \beta$. To do this, assume first that $S(x') = \infty$. Then, there exists a valid feasibility cut $\tilde{\omega}$, such that $\tilde{\omega}(x') > 0$. Since the cut is valid, it holds that $\tilde{\omega}(\tilde{x}) \leq 0$. Hence, there is $\hat{\mu} \in (0,1)$ for which $\tilde{\omega}(\bar{x} + \hat{\mu}(\tilde{x} - \bar{x})) = 0$ and $\tilde{\omega}(\bar{x} + \mu(\tilde{x} - \bar{x})) > 0$ for all $\mu \in (0, \hat{\mu})$. Hence, $\tilde{\omega}$ implies a solution of Problem (12) with value $\hat{\mu} > \bar{\mu}$, and this is a contradiction to our assumption.

Assume that $\infty > f^T x' + S(x') > \beta$. Then, there exists a valid optimality cut $\tilde{\omega}$ such that $f^T x' + \tilde{\omega}(x') > \beta$. Since the cut is valid, it holds that $f^T\tilde{x} + \tilde{\omega}(\tilde{x}) \leq \beta$. Hence, there is $\hat{\mu} \in (0,1)$, such that $f^T(\bar{x} + \hat{\mu}(\tilde{x} - \bar{x})) + \tilde{\omega}(\bar{x} + \hat{\mu}(\tilde{x} - \bar{x})) = \beta$ and $f^T(\bar{x} + \mu(\tilde{x} - \bar{x})) + \tilde{\omega}(\bar{x} + \mu(\tilde{x} - \bar{x})) > \beta$ for all $\mu \in (0, \hat{\mu})$. Hence, $\tilde{\omega}$ implies a solution of Problem (12) with value $\hat{\mu} > \bar{\mu}$, and this is a contradiction.

Second, we show $f^T x + S(x) < \beta$ for all $x \in \{x' + \lambda(\tilde{x} - x'), \lambda \in (0,1)\}$. Assume that there exists $x'' \in \{x' + \lambda(\tilde{x} - x'), \lambda \in (0,1)\}$, i.e., $x'' = x' + \check{\lambda}(\tilde{x} - x')$ for a $\check{\lambda} \in (0,1)$, with $f^T x'' + S(x'') \geq \beta$. Then there is a valid optimality cut $\tilde{\omega}$ with $f^T x'' + \tilde{\omega}(x'') \geq \beta$. Since this cut is valid, it holds that $f^T x' + \tilde{\omega}(x') \leq \beta$ and $f^T\tilde{x} + \tilde{\omega}(\tilde{x}) \leq \beta$. Since one condition of the lemma was that no optimality cut $\bar{\omega}$ exists for which $\bar{\omega}(\tilde{x}) + f^T\tilde{x} = \beta$ and $\bar{\omega}(\bar{x}) + f^T\bar{x} = \beta$, we can derive that

$$\beta > \check{\lambda}(f^T x' + \tilde{\omega}(x')) + (1 - \check{\lambda})(f^T\tilde{x} + \tilde{\omega}(\tilde{x})) = f^T x'' + S(x'') \geq \beta,$$

and this is a contradiction. □

3.5. **Concluding Remarks.**

3.5.1. *On the conditions of Theorem 3.16.* As already stated at the beginning of Section 3.4, none of the conditions of Theorem 3.16 has to be checked before Problem (12) is set up and solved. Lemma 3.17 ascertains that whenever $f^T\bar{x} + S(\bar{x}) < \beta$, i.e., $\bar{x}$ improves the currently best-known objective function value of Problem (1), Problem (12) has a non-positive optimal value or is infeasible. In the rare case that a valid optimality cut $\bar{\omega}$ exists, for which $\bar{\omega}(\tilde{x}) + f^T\tilde{x} = \beta$ and $\bar{\omega}(\bar{x}) + f^T\bar{x} = \beta$, Problem (12) can be set up and solved as well. If it has a positive optimal value, it implies a cut that the iterate violates and can be used. If not, a standard Benders cut can be calculated. We will provide a more detailed explanation in Section 4.

3.5.2. *Connection of facet cuts and optimal line-shifting cuts.* The optimal line-shifting strategy can also be considered as a facet-generating strategy, with the difference that instead of the current iterate $(\bar{x}, \bar{\eta})$ the point $(\bar{x}, \beta - f^T\bar{x})$ is separated from the epigraph. We recall that the facet strategy solves Problem (7) to generate a cut violated by $(\bar{x}, \bar{\eta})$. If we want to generate a cut that is violated by $(\bar{x}, \beta - f^T\bar{x})$ instead, we replace $\bar{\eta}$ with $\beta - f^T\bar{x}$ in Problem (7) and get

$$\max \ \pi^T(b - H\bar{x}) - \pi_0(\beta - f^T\bar{x}) \tag{15a}$$
$$\text{s.t. } \pi^T H(\tilde{x} - \bar{x}) + \pi_0(\tilde{\eta} - \beta + f^T\bar{x}) = 1 \tag{15b}$$
$$\pi^T A \leq \pi_0 c^T \tag{15c}$$
$$\pi_0, \pi \geq 0. \tag{15d}$$

As $\tilde{\eta} = \beta - f^T\tilde{x}$, we can reformulate the left-hand side of Constraint (15b) into

$$\pi^T H(\tilde{x} - \bar{x}) + \pi_0(\beta - f^T\tilde{x} - \beta + f^T\bar{x}) = \pi^T H(\tilde{x} - \bar{x}) + \pi_0 f^T(\bar{x} - \tilde{x}).$$

We observe that Problem (15) and Problem (12) are identical.

3.5.3. *Subproblems with block-diagonal structure.* BD works exceptionally well if the subproblem (3) decomposes into K independent optimization problems. This is the case if the subproblem has a so-called block-diagonal structure. The independent subproblem parts, or "blocks", can then be solved separately, which leads to a considerably reduced computational effort. In this case, the master problem (2) contains not only one variable η to estimate the value function of the subproblem but one variable η_k for each block $k = 1, ..., K$ to estimate the value function of each subproblem part.

We note that Problem (12) does not decompose into blocks, even if Problem (3) decomposes into blocks, as Constraint (12b) has one non-zero coefficient for at least one variable in each block for all reasonably stated optimization problems.

Nevertheless, block-diagonal structure can be exploited within the OLS approach. The last remark states that the OLS approach separates a point $(\bar{x}, \hat{\eta})$ with $\hat{\eta} > \bar{\eta}$ from the epigraph of S in the sense of Stursberg (2019), instead of the iterate $(\bar{x}, \bar{\eta})$. This principle can be directly transferred to the block-diagonal case, i.e., we aim to separate a point $(\bar{x}, \hat{\eta}_k)$ with $\hat{\eta}_k > \bar{\eta}_k$ from the epigraph of the k-th block's value function S_k instead of separating the iterate $(\bar{x}, \bar{\eta}_k)$ from it. We expect that an algorithm based on this principle will perform well.

The following chapter specifies the algorithmic framework we propose to solve optimization problems with the approach provided in this chapter.

4. The algorithmic framework.

In this section, we discuss how the OLS procedure is implemented. We note that Theorem 3.16 can only be applied if the current master solution $(\bar{x})$ has a value that is not below the value of the currently best-known solution (x^*), i.e., if $f^T\bar{x} + S(\bar{x}) \geq \beta = f^T x^* + S(x^*)$. Furthermore, the theorem has as the prerequisite that no valid optimality cut $\bar{\omega}$ exists that has the property that $f^T\bar{x}+\bar{\omega}(\bar{x}) = f^T\tilde{x}+\bar{\omega}(\tilde{x}) = \beta$ for a core point $\tilde{x} \in \text{conv}(X)$. The latter is automatically fulfilled if $f^T\tilde{x} + S(\tilde{x}) < \beta$. Otherwise, we can check if the optimal value of Problem (3) with the additional constraint

$$\pi^T(b - H\tilde{x}) = \beta - f^T\tilde{x} \tag{16}$$

is lower than $\beta - f^T\bar{x}$.

The first prerequisite can be checked by solving Problem (12). Lemma 3.17 guarantees that if the prerequisite is violated, the problem has a non-positive value or is infeasible. We can now state an algorithm that calculates OLS cuts.

Since we test BD on mixed-integer linear programs, with the integer variables in the master problem and the continuous variables in the subproblem, we implemented a Branch-and-Benders-Cut procedure that starts a single Branch-and-Cut framework and generates valid cuts whenever a new integer solution of the master problem is generated. The Branch-and-Cut framework maintains values for $x^* \in X$ (incumbent), $\tilde{x} \in \text{conv}(X)$ (core point), $\beta \in \mathbb{R}$ (incumbent value), and a lower bound $l \in \mathbb{R}$ on the value of Problem (1). Algorithm 1 is called whenever the Branch-and-Cut framework detects a new integral solution $(\bar{x}, \bar{\eta})$ with a value lower than the incumbent value.

Algorithm 1 gets all the parameters described in the last paragraph as input and a parameter $\gamma \in (0, 1)$ that coordinates the update of the core point. Algorithm 1 checks first if the incumbent value is sufficiently close to the lower bound provided by the Branch-and-Cut framework. If it is close, it returns optimal.

Otherwise, if the last Benders cut has not been obtained from a solution of Problem (12), it solves Problem (3) + (16) and terminates with the implied Benders cut if the value of the problem is greater or equal to β, leaving threshold, core point and incumbent unchanged. Otherwise, it solves Problem (12) and determines its optimal value $\bar{\mu}$, if possible. An OLS cut is desirable if the iterate does not improve the incumbent value. The algorithm exploits Lemma 3.17: If the iterate improves the incumbent, Problem (12) is infeasible, or $\bar{\mu}$ is non-positive. The algorithm solves Problem (3) and determines its optimal value $S(\bar{x})$. A Benders cut ω is derived from the solution of Problem (3), β is set to $f^T\bar{x} + S(\bar{x})$, and the core point and the incumbent are set to the iterate. If Problem (12) has a positive optimal value, a Benders cut ω is derived from the solution of Problem (12). Theorem 3.16 ascertains that the derived cut is an OLS cut. The core point $\tilde{x}$ is updated to $\bar{x} + (\bar{\mu} + \gamma(1 - \bar{\mu}))(\tilde{x} - \bar{x})$. This step is justified by Lemma 3.18. The lemma also guarantees that the value of the new core point is lower than the incumbent value. Afterward, the Benders cut, the new core, the new incumbent, and the new incumbent value are returned. The Branch-and-Cut framework inserts the returned cut ω and updates the values of $\tilde{x}, x^*$, and β afterward. It terminates as soon as "optimal" is returned by Algorithm 1, with a solution x^* that is part of a solution of Problem (1), at most $tol \in \mathbb{R}$ worse than an optimal one.

Remark. It is possible to check the condition $f^T\bar{x} + S(\bar{x}) \geq \beta$ by solving either Problem (12) as in Algorithm 1 or by solving Problem (3). If the condition holds, Problem (12) has to be solved to generate the OLS cut. Otherwise, Problem (3) has to be solved to evaluate $S(\bar{x})$ and to generate the Benders cut. Hence, solving Problem (12) first is more efficient if one expects that the condition holds for most iterates. In some situations, solving Problem (3) first is more efficient. One situation is if the condition does not hold for more than 50% of the iterates. Another situation is if Problem (3) is easier to solve, e.g., due to the presence of many so-called "Generalized Bound Constraints", as used in Bonami et al. (2020). If Problem (3) is supposed to be solved first, Lines 9 to 17 must be changed accordingly. Empirical tests showed that the variant presented in Algorithm 1 is the more efficient for the instance set we used for our computational experiments.

Remark. OLS cuts can be incorporated into the framework presented by Bonami et al. (2020). We suggest using OLS cuts whenever Problem (3) is unbounded, i.e., instead of a subproblem with "L^1-normalization" or "CW-normalization". Preliminary computational tests suggest that this has a positive influence on running times.

In the following section, the results of our computational study are presented.

5. Computational Results

In this section, we present our computational results. This includes a description of our computational setup and the presentation of results for different classes of instances: instances taken from the MIPLib, multicommodity-flow network design instances, and instances of randomly generated MIPs. All instances and our code are publicly available; see Glomb et al. (2024). We benchmarked our cut selection strategy against several existing solution approaches.

5.1. **Computational Setup.** In this section, the results of our computational study are presented. All programs have been written using the programming language Python, version 3.10. Optimization problems have been solved in single-thread mode using the Gurobi optimizer, version 10.0.2, see Gurobi Optimization (2020). The programs have been executed on nodes of a high-performance computing cluster using a Xeon E3-1240 v6 CPU with four cores at a 3.7GHz base frequency and a total memory of 32 GB. We solved four instances on one node at a time. All calculations are terminated after one hour,

Algorithm 1 OLS cut generation

Input $\bar{x}, x^* \in X$, $\bar{\eta} \in \mathbb{R}$, $\tilde{x} \in \text{conv}(X)$ with $\infty \neq \beta := f^T x^* + S(x^*) \geq f^T \tilde{x} + S(\tilde{x})$, $(\bar{x}, \bar{\eta})$ feasible solution of (2) with $f^T \bar{x} + \bar{\eta} < \beta$, a lower bound l on (2), and ´$\gamma \in (0,1)$.
Output Message that (1) is solved, **or** a valid cut and updated $\tilde{x}, x^*, \beta$.

if $l \leq \beta - tol$ **then**
    if $\tilde{x} = x^*$ **then**
        Solve (3) + (16). Let α denote its optimal value.
        if $f^T \bar{x} + \alpha \geq \beta$ **then**
            $\omega \leftarrow$ Benders cut implied by (3) + (16).
            return $\tilde{x}, x^*, \beta, \omega$.
        end if
    end if
    Solve (12). Let $\bar{\mu}$ denote its optimal value.
    if Problem (12) is infeasible or $\bar{\mu} \leq 0$ **then**
        Solve Problem (3) to determine $S(\bar{x})$.
        $\omega \leftarrow$ Benders cut implied by (3).
        $\beta \leftarrow f^T \bar{x} + S(\bar{x})$.
        $\tilde{x}, x^* \leftarrow \bar{x}$.
    else
        $\omega \leftarrow$ Benders cut implied by (12).
        $\tilde{x} \leftarrow \bar{x} + (\bar{\mu} + \gamma(1 - \bar{\mu}))(\tilde{x} - \bar{x})$.
    end if
    return $\tilde{x}, x^*, \beta, \omega$.
else
    return optimal.
end if

leaving the processes 180 seconds for non-optimization tasks like setting up the problems or writing out the solution files. The duration of these tasks has not been included in the reported solution times. Whenever a relative optimality gap of 10^{-4} has been reached, the calculation has been terminated as well. Since we only consider minimization problems with a positive optimal value, the optimality gap is defined depending on a known upper bound $u > 0$ on the optimal value and a known lower bound $u \geq l \geq 0$ on the optimal value as $\frac{u-l}{u} \in [0,1]$. All instances have been solved directly with Gurobi without a decomposition approach (**van**). Furthermore, we tried BD using five different cut selection approaches: the Magnanti-Wong strategy (**mwb**), the MIS strategy (**mis**), the Facet strategy (**fcb**), the OLS strategy (**ols**), and a hybrid strategy (**hyb**) combining MIS and OLS. The hybrid strategy uses MIS selection for at least 100 cuts and at least until the optimality gap reaches 10%. Then, it uses the OLS selection. The hybrid strategy has the additional advantage that it can be started without knowing a feasible solution.

For the cut selection methods that require a feasible solution of the optimization problem to get started, i.e., facet and OLS, a feasible solution calculated using the Gurobi optimizer, see Gurobi Optimization (2020), terminating as soon as an optimality gap of 25% is reached. The time required to execute this heuristic is negligible, i.e., far below 1 s for all instances except 24 instances of the MIPLib instance set and far below the running times of the solution algorithms we tested for these. Hence, we do not consider the running time of the heuristic as a part of the solution time of the methods.

All instances are mixed-integer linear programs, and all integer variables have been retained in the master problem, while all continuous variables have been put into the subproblem. Hence, we embedded

Instance set	Facet	Magnanti-Wong	MIS	OLS	MIP-Solver	Hybrid	Benders	Total
MIPLib	17	17	17	17	42	17	17	44
MCF-NWD	24	16	26	26	35	26	27	50
Fischetti et al. (2010)	18	16	18	19	9	19	19	20
Random MIP	45	45	46	47	48	46	47	48

TABLE 1. Instances per instance set that could be solved to optimality by different solution approaches within the time limit. The right column denotes the size of the instance set.

all cut selection strategies into a Branch-and-Cut framework, analogous to the framework described in Algorithm 1.

To establish a proper graphical illustration of the algorithms' performance, performance plots for the algorithms' solution time, number of required cuts and optimality gap after one hour have been generated. Cut and time performance plots include only instances that could be solved to optimality within the time limit by at least one algorithm based on BD. This prevents the evaluations from being strongly biased by assigning the algorithm that produces new cuts at the slowest rate overly high cut performance values. Table 1 shows how many instances per instance set could be solved to optimality by the different solution approaches. The second column from the right denotes how many instances could be solved to optimality by BD for at least one of the proposed cut selection strategies. These instances have been used to generate the performance plots for the algorithms' runtime. The runtime has been set to the time limit whenever an instance has not been solved to optimality. Hence, the performance plots shown might overestimate the true performance plots by $1 - \frac{\text{\# Instances solved by algorithm}}{\text{\# Instances solved by Benders}}$. The same applies to the cut efficiency plots. Whenever a significant proportion of instances could not be solved to optimality, we additionally presented a performance plot for the optimality gap of these instances. For the other instance sets, the gaps for instances that could not be solved using BD are reported in tables whenever suitable.

The following sections will present our results on MIPLIB instances, network design problems, and randomly generated MIPs. The results are depicted as so-called performance plots. Performance plots are the graphs of the monotone functions

$$\phi_a : [1, \infty) \to [0, 1], \quad a \in A,$$
$$\rho \mapsto \frac{|\{i \in I : p(a, i) \leq \rho \min_{\tilde{a} \in A} p(\tilde{a}, i)\}|}{|I|},$$

where I is a set of instances, A is a set of algorithms, and p is a performance measure, depending on the algorithm and the instance, like, as in our case, "time the algorithm needs to solve the instance", "cuts the algorithm needs to solve the instance" or "optimality gap after termination". The graphs of algorithms with high performance are in the top-left regions of the figures.

5.2. **Results for MIPLib Instances.** We benchmarked our cut selection strategy against MIS, Facet and Magnanti-Wong cuts on a selection of decomposable instances from the MIPLib (2017) Collection set Gleixner et al. (2021). We identified 55 instances with the following properties:

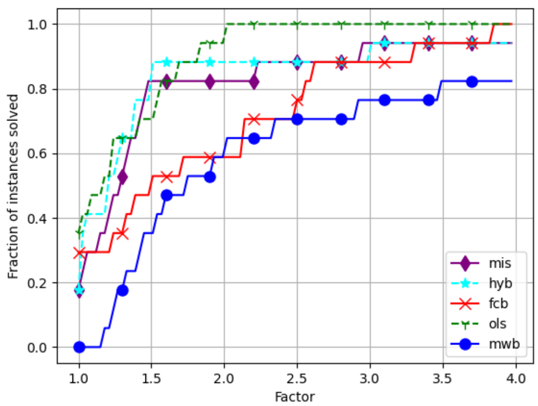


FIGURE 3. Fraction of MIPLIB instances that could be solved within a multiple of fastest running time.

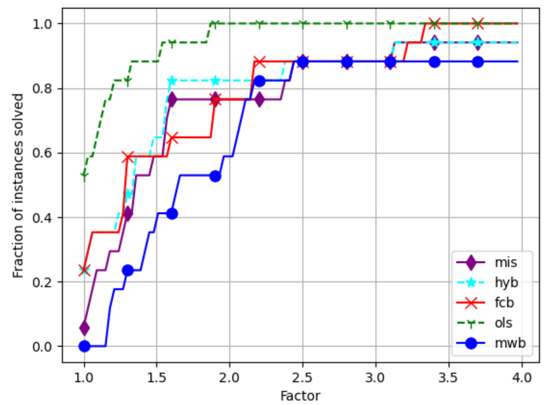


FIGURE 4. Fraction of MIPLIB instances that could be solved using a multiple of the cuts needed by the algorithm with the fewest cuts.

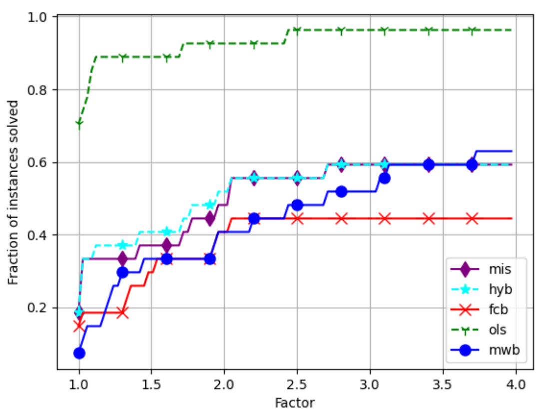


FIGURE 5. Fraction of MIPLIB instances achieving a multiple of the gap the algorithm with the lowest gap achieves after one hour of calculation.

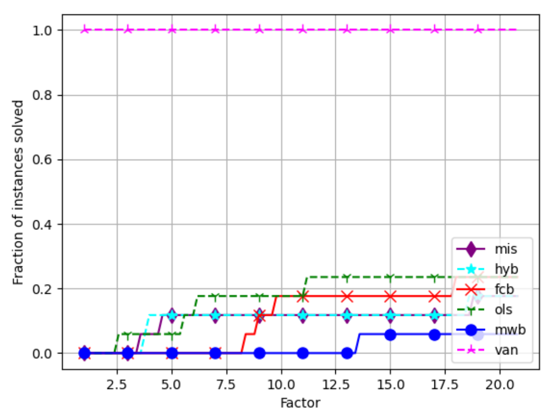


FIGURE 6. Fraction of MIPLIB instances that could be solved within a multiple of fastest running time, including the state-of-the-art MIP solver.

they contain at least 1 and at most 1000 integer/binary variables, at least one and at most 100000 continuous variables, and at most 10000 constraints. They contain more continuous than integer variables and have a positive optimal value. We only selected instances labeled as "easy" and not as "numerics".

Of these 55 instances, we omitted the 11 instances b2c1s1, bienst2, binkar10_1, cost266-UUE, fastxgemm-n2r6s0t2, n6-3, neos-3072252-nete, neos-3627168-kasai, neos-3665875-lesum, neos-480878, and rentacar due to numerical problems. The results for the remaining 44 instances are shown in Figures 3 to 6.

19 of 44 MIPLIB instances could be solved to optimality using BD. Figure 3 shows that, measured in running time, hybrid, MIS, and OLS selection are competitive.

Comparing the number of cuts needed to solve instances to optimality, as shown in Figure 4, OLS outperforms the other cut selection strategies, needing the fewest cuts for over 50% of all instances. For no instance, it needs more than 80% more cuts to solve it to optimality, compared to the best selection strategy for this instance. We want to note that the hybrid strategy, which has competitive solution times, is the second-best regarding cut efficiency.

Figure 5 shows that for the remaining 24 instances that could not be solved using BD, OLS achieves the lowest optimality gap for over 70% of all instances.

We are not very surprised that Figure 6 demonstrates that BD is not competitive against one of the best state-of-the-art MIP solvers on MIPLIB instances regardless of the cut selection strategy.

5.3. **Results for Network Design Problems.** We tested the algorithms on 50 newly created instances and on 20 of the original instances used in Fischetti et al. (2010) (those with positive flow costs, i.e., the "optimality" instances). The instances we created ourselves are network design problems, with all combinations of 20, 40, 60, 80, or 100 commodities, graphs that have either 5×4 or 6×4 nodes, and edges that are created either as a grid, as Erdös-Renyi-Graph, as a random 5-regular graph, as a graph that is initialized as empty and adds random edges until each node has a degree between 2 and 6, and as a graph that is derived of a random 4-regular graph, removing each edge with a probability of 0.06 and adding 13 random edges afterward. Setup costs for an arc have been set to 5, while costs for a unit of flow along an arc have been set to 1. Each commodity has one origin node, one destination node, and a demand of 1. The algorithmic performance is shown in Figures 7 - 10.

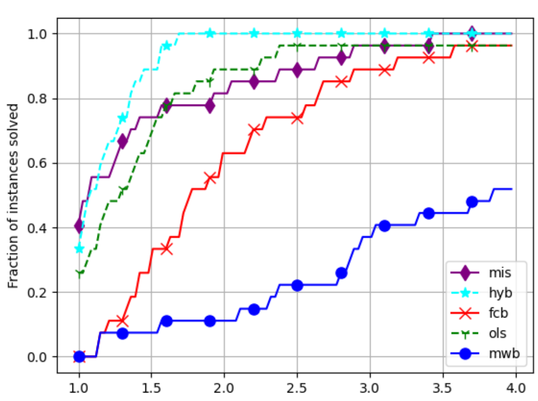


FIGURE 7. Fraction of network design instances that could be solved within a multiple of fastest running time.

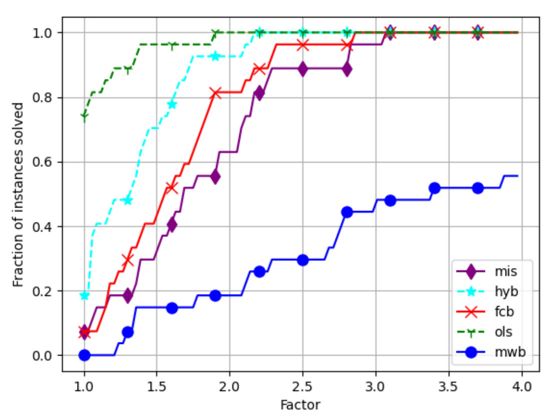


FIGURE 8. Fraction of network design instances that could be solved using a multiple of the cuts needed by the algorithm with the fewest cuts.

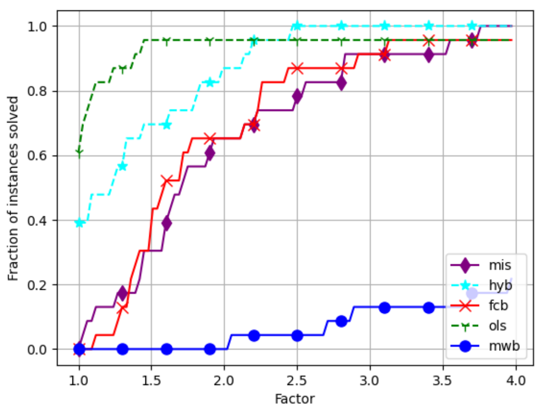


FIGURE 9. Fraction of network design instances achieving a multiple of the gap the algorithm with the lowest gap achieves after one hour of calculation.

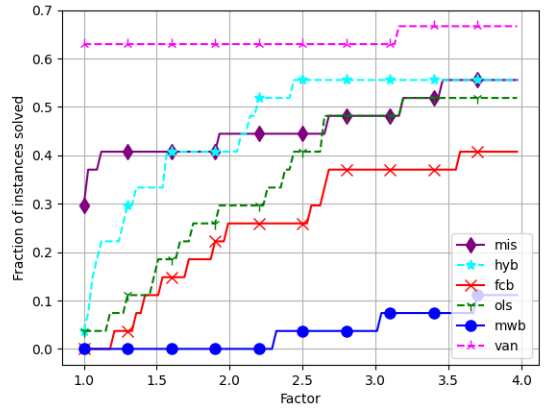


FIGURE 10. Fraction of network design instances that could be solved within a multiple of fastest running time, including the state-of-the-art MIP solver.

Figure 7 shows that regarding solution time, the hybrid strategy is the best one. The hybrid strategy has the lowest running time for approximately 40% of the instances and it solves all instances within an additional 70% of the best algorithm's running time.

Figure 8 shows that the OLS strategy needs the fewest cuts for the instances that could be solved to optimality by BD.

Figure 9 shows that for the instances that could not be solved to optimality by one of the cut selection strategies, OLS selection achieves the best optimality gaps after one hour.

Figure 10 shows that a state-of-the-art MIP solver is the best choice for more than 60% of all network design instances we created, while the MIS approach is the best for 30% of all instances tested. Pure OLS selection is the best approach for around 5% of the instances.

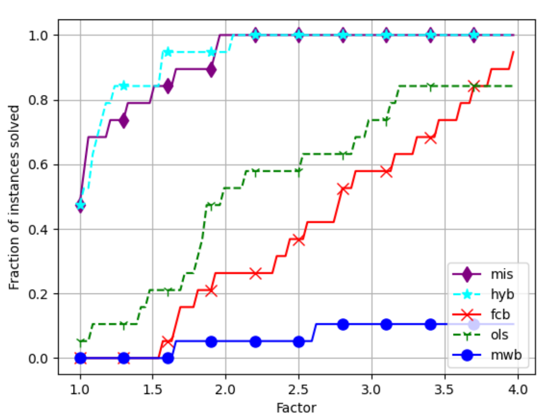


FIGURE 11. Fraction of Fischetti et al. (2010) instances that could be solved within a multiple of fastest running time.

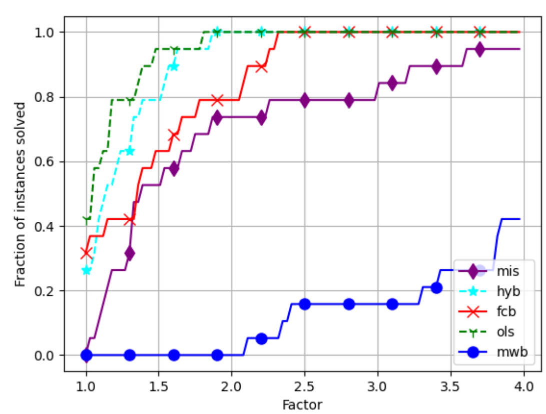


FIGURE 12. Fraction of Fischetti et al. (2010) instances that could be solved using a multiple of the cuts needed by the algorithm with the fewest cuts.

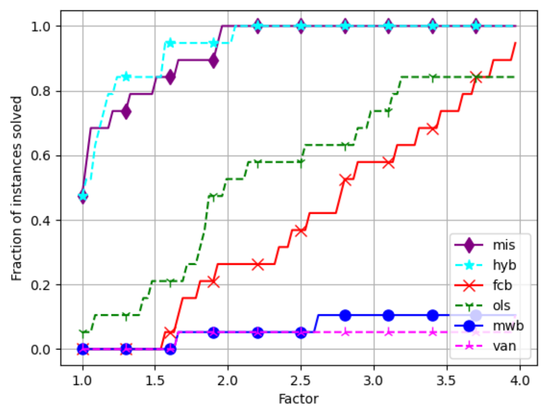


FIGURE 13. Fraction of Fischetti et al. (2010) instances that could be solved within a multiple of fastest running time, including the state-of-the-art MIP solver.

The results for the instances from Fischetti et al. (2010) are graphically summarized in Figures 11 - 13. As before, the hybrid strategy and MIS are competitive regarding time (Figure 11), and OLS and hybrid selection are competitive regarding cuts (Figure 12). For all instances from Fischetti et al. (2010) that have been tested, the state-of-the-art MIP solver is outperformed, as demonstrated in Figure 13.

5.4. **Results for randomly generated MIPs.** We present our computational results for randomly generated MIPs. The problems are generated randomly of all combinations of 50, 100, or 150 integer variables, 200 or 400 continuous variables, 50 or 100 inequality constraints containing only integer variables, 100 or 200 inequality constraints containing integer and continuous variables, and 200 or 400 inequality constraints containing only continuous variables. Additionally, each instance contains five equality constraints containing integer and continuous variables.

Integer variables have an objective function coefficient between 0.5 and 5. Continuous variables have an objective function coefficient between 0 and 10 times the ratio of the number of integer variables and the number of continuous variables. The objective function coefficients are drawn from these intervals following a uniform distribution. Each integer variable has a nonzero coefficient in constraints that contain integer variables with a probability that equals the ratio of 10 and the number of integer variables. Each continuous variable has a nonzero coefficient in constraints that contain continuous variables with a probability that equals the ratio of 20 and the number of integer variables.

The results can be taken from Figures 14 - 16.

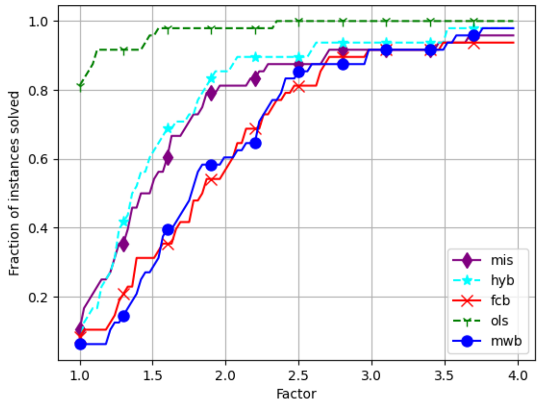


FIGURE 14. Fraction of Random MIP instances that could be solved within a multiple of fastest running time.

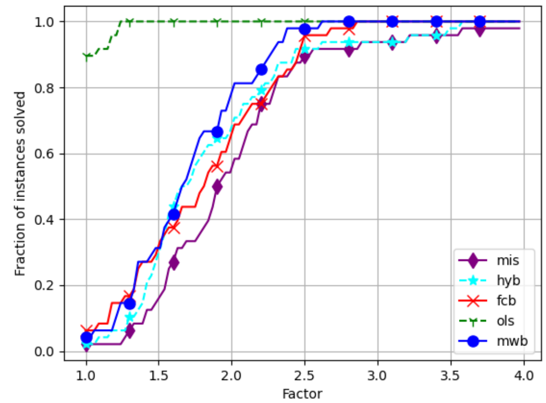


FIGURE 15. Fraction of Random MIP instances that could be solved using a multiple of the cuts needed by the algorithm with the fewest cuts.

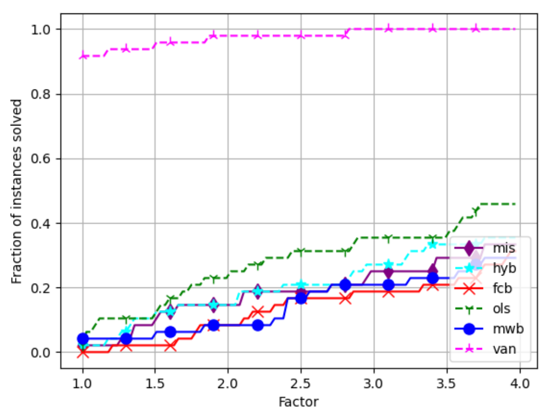


FIGURE 16. Fraction of Random MIP instances that could be solved within a multiple of fastest running time, including the state-of-the-art MIP solver.

Figure 14 implies that OLS cut selection is competitive in terms of calculation time on this instance set. OLS is the fastest selection strategy for over 80% of all instances and solves all instances in not more than 130% more time than the fastest algorithm.

Figure 15 shows that OLS cut selection is competitive when comparing the number of needed cuts. For over 90% of all instances, OLS needs the fewest cuts to solve the instance. Hence, OLS outperforms all other cut selection strategies.

Figure 16 demonstrates that for the random MIPs we generated, using a state-of-the-art MIP solver is superior to choosing BD as a solution approach.

We want to conclude with a summary of what we consider to be the advantages and disadvantages of the three cut selection strategies MIS, Facet, and OLS. The summary can be found in Table 2.

	MIS	**Facet**	**OLS**
Adv.	• Excellent practical performance • Easy to implement • No update of normalization constraint is necessary • Works without knowledge of a feasible solution	• High practical performance • Independent on constraint scaling and redundant constraints	• Excellent practical performance • Immediately recognizes improvements • Independent on constraint scaling and redundant constraints • Always cuts off the iterate point
Disadv.	• Depends on constraint scaling • Affected by redundant constraints • Difficult to interpret • Does not immediately recognize improvements • Needs sign-restricted dual variables • Does not necessarily cut off the iterate point	• Knowledge of a feasible solution is necessary • Does not immediately recognize improvements • Update of normalization constraint is necessary in each iteration • Does not necessarily cut off iterate	• Knowledge of feasible solution is necessary • Two subproblems have to be solved if the incumbent improves • Update of normalization constraint in each iteration is necessary

Table 2. Properties of MIS, OLS, and Facet selection strategies.

6. Summary and Outlook

In this article, an innovative notion of Pareto-optimality for Benders cuts has been first developed. This notion is based on so-called solution-candidate sets, which describe the set of feasible points to the master problem that are potentially an optimal solution to the original optimization problem. We showed that cuts that are non-dominated in our sense are also non-dominated in the sense of Magnanti and Wong (1981), but the opposite is not necessarily true.

Based on our notion of Pareto-optimality, we developed a novel cut selection strategy for BD that is capable of calculating Pareto-optimal cuts under mild assumptions.

Further, we developed the algorithmic framework necessary to optimally exploit the potential of the cut selection strategy. The algorithm has been benchmarked against other known cut selection strategies (Magnanti-Wong, MIS, Facet) on various instance classes. For all instance classes (MIPLib, multicommodity flow network design problems, randomly generated mixed-integer linear programming problems), the computational results show that the developed method is competitive, measured in CPU seconds needed to solve a problem to optimality, and the results showed that the developed method needs to generate a smaller number of cuts than the benchmark approaches. Hence, the method is effective in situations with scarce memory or a difficult-to-solve subproblem.

Possible future research directions are the transfer of the method to more general versions of the approach, such as Generalized BD, as published in Geoffrion (1972). The approach (as well as the MIS selection strategy as published in Fischetti et al. (2010) and Facet cuts as published in Stursberg (2019)) can be exploited to accelerate various algorithms based on BD that have been applied to solve real-world optimization problems, of which many are mentioned in Section 1, since in many of these articles, the approach is either applied without considering a cut selection strategy at all or using the Magnanti-Wong method. Considering the high performance of MIS, Facet, and OLS cuts for the optimization problems investigated in this article, this could also lead to significant speedups for real-world applications.

Acknowledgements

This research was conducted within the framework of the OPs-TIMAL research project, financed by the German Federal Ministry of Economic Affairs and Energy (BMWi). Furthermore we want to thank three anonymous reviewers for their helpful feedback, which helped to substantially improve the article.

References

Abdolmohammadi, H. R. and Kazemi, A. (2013). A benders decomposition approach for a combined heat and power economic dispatch. *Energy conversion and management*, 71:21–31.

Adulyasak, Y., Cordeau, J.-F., and Jans, R. (2015). Benders decomposition for production routing under demand uncertainty. *Operations Research*, 63(4):851–867.

Ambrosius, M., Grimm, V., Kleinert, T., Liers, F., Schmidt, M., and Zöttl, G. (2020). Endogenous price zones and investment incentives in electricity markets: An application of multilevel optimization with graph partitioning. *Energy Economics*, 92:104879.

Azad, N., Saharidis, G. K., Davoudpour, H., Malekly, H., and Yektamaram, S. A. (2013). Strategies for protecting supply chain networks against facility and transportation disruptions: an improved benders decomposition approach. *Annals of Operations Research*, 210:125–163.

Baringo, L. and Conejo, A. J. (2011). Wind power investment: A benders decomposition approach. *IEEE Transactions on Power Systems*, 27(1):433–441.

Bärmann, A., Liers, F., Martin, A., Merkert, M., Thurner, C., and Weninger, D. (2015). Solving network design problems via iterative aggregation. *Mathematical Programming Computation*, 7:189–217.

Bayram, V. and Yaman, H. (2018). Shelter location and evacuation route assignment under uncertainty: A benders decomposition approach. *Transportation science*, 52(2):416–436.

Benders, J. (1962). Partitioning procedures for solving mixed-variables programming problems. *Numerische mathematik*, 4(1):238–252.

Bonami, P., Salvagnin, D., and Tramontani, A. (2020). Implementing automatic benders decomposition in a modern mip solver. In *Integer Programming and Combinatorial Optimization: 21st International Conference, IPCO 2020, London, UK, June 8–10, 2020, Proceedings*, pages 78–90. Springer.

Botton, Q., Fortz, B., Gouveia, L., and Poss, M. (2013). Benders decomposition for the hop-constrained survivable network design problem. *INFORMS journal on computing*, 25(1):13–26.

Brandenberg, R. and Stursberg, P. (2021). Refined cut selection for benders decomposition: applied to network capacity expansion problems. *Mathematical Methods of Operations Research*, 94(3):383–412.

Charnes, A. and Cooper, W. W. (1962). Programming with linear fractional functionals. *Naval Research logistics quarterly*, 9(3-4):181–186.

Conforti, M. and Wolsey, L. A. (2019). "facet" separation with one linear program. *Mathematical Programming*, 178:361–380.

Contreras, I., Cordeau, J.-F., and Laporte, G. (2011). Benders decomposition for large-scale uncapacitated hub location. *Operations research*, 59(6):1477–1490.

Costa, A. M., Cordeau, J.-F., Gendron, B., and Laporte, G. (2012). Accelerating benders decomposition with heuristic master problem solutions. *Pesquisa Operacional*, 32:03–20.

Cote, G. and Laughton, M. A. (1984). Large-scale mixed integer programming: Benders-type heuristics. *European Journal of Operational Research*, 16(3):327–333.

Fischetti, M., Ljubić, I., and Sinnl, M. (2016). Benders decomposition without separability: A computational study for capacitated facility location problems. *European Journal of Operational Research*, 253(3):557–569.

Fischetti, M., Ljubić, I., and Sinnl, M. (2017). Redesigning benders decomposition for large-scale facility location. *Management Science*, 63(7):2146–2162.

Fischetti, M., Salvagnin, D., and Zanette, A. (2010). A note on the selection of benders' cuts. *Mathematical Programming*, 124(1):175–182.

Froyland, G., Maher, S. J., and Wu, C.-L. (2014). The recoverable robust tail assignment problem. *Transportation Science*, 48(3):351–372.

Gendron, B., Scutellà, M. G., Garroppo, R. G., Nencioni, G., and Tavanti, L. (2016). A branch-and-benders-cut method for nonlinear power design in green wireless local area networks. *European Journal of Operational Research*, 255(1):151–162.

Geoffrion, A. M. (1972). Generalized benders decomposition. *Journal of optimization theory and applications*, 10:237–260.

Geoffrion, A. M. and Graves, G. W. (1974). Multicommodity distribution system design by benders decomposition. *Management science*, 20(5):822–844.

Gleixner, A., Hendel, G., Gamrath, G., Achterberg, T., Bastubbe, M., Berthold, T., Christophel, P. M., Jarck, K., Koch, T., Linderoth, J., Lübbecke, M., Mittelmann, H. D., Ozyurt, D., Ralphs, T. K., Salvagnin, D., and Shinano, Y. (2021). MIPLIB 2017: Data-Driven Compilation of the 6th Mixed-Integer Programming Library. *Mathematical Programming Computation*.

Glomb, L., Liers, F., and Rösel, F. (2023a). Optimizing integrated aircraft assignment and turnaround handling. *European Journal of Operational Research*, 310(3):1051–1071.

Glomb, L., Liers, F., and Rösel, F. (2023b). A stochastic optimization approach for optimal tail assignment with knowledge-based predictive maintenance. *CEAS Aeronautical Journal*, pages 1–14.

Glomb, L., Liers, F., and Rösel, F. (2024). Compare css. `https://github.com/DiscreteOptimizer/compare_css`.

Grimm, V., Kleinert, T., Liers, F., Schmidt, M., and Zöttl, G. (2019). Optimal price zones of electricity markets: a mixed-integer multilevel model and global solution approaches. *Optimization methods and software*, 34(2):406–436.

Gurobi Optimization, L. (2020). Gurobi optimizer reference manual.

Magnanti, T. L. and Wong, R. T. (1981). Accelerating benders decomposition: Algorithmic enhancement and model selection criteria. *Operations research*, 29(3):464–484.

Maheo, A., Kilby, P., and Van Hentenryck, P. (2019). Benders decomposition for the design of a hub and shuttle public transit system. *Transportation Science*, 53(1):77–88.

Mansouri, S., Ahmarinejad, A., Ansarian, M., Javadi, M., and Catalao, J. (2020). Stochastic planning and operation of energy hubs considering demand response programs using benders decomposition approach. *International Journal of Electrical Power & Energy Systems*, 120:106030.

McDaniel, D. and Devine, M. (1977). A modified benders' partitioning algorithm for mixed integer programming. *Management Science*, 24(3):312–319.

Nasri, A., Kazempour, S. J., Conejo, A. J., and Ghandhari, M. (2015). Network-constrained ac unit commitment under uncertainty: A benders' decomposition approach. *IEEE transactions on power systems*, 31(1):412–422.

Papadakos, N. (2008). Practical enhancements to the Magnanti–Wong method. *Operations Research Letters*, 36(4):444–449.

Poojari, C. A. and Beasley, J. E. (2009). Improving benders decomposition using a genetic algorithm. *European Journal of Operational Research*, 199(1):89–97.

Qian, L. P., Zhang, Y. J. A., Wu, Y., and Chen, J. (2013). Joint base station association and power control via benders' decomposition. *IEEE Transactions on Wireless Communications*, 12(4):1651–1665.

Rahmaniani, R., Crainic, T. G., Gendreau, M., and Rei, W. (2017). The benders decomposition algorithm: A literature review. *European Journal of Operational Research*, 259(3):801–817.

Rahmaniani, R., Crainic, T. G., Gendreau, M., and Rei, W. (2018). Accelerating the benders decomposition method: Application to stochastic network design problems. *SIAM Journal on Optimization*, 28(1):875–903.

Rei, W., Cordeau, J.-F., Gendreau, M., and Soriano, P. (2009). Accelerating benders decomposition by local branching. *INFORMS Journal on Computing*, 21(2):333–345.

Saharidis, G. K., Boile, M., and Theofanis, S. (2011). Initialization of the benders master problem using valid inequalities applied to fixed-charge network problems. *Expert Systems with Applications*, 38(6):6627–6636.

Saharidis, G. K., Minoux, M., and Ierapetritou, M. G. (2010). Accelerating benders method using covering cut bundle generation. *International Transactions in Operational Research*, 17(2):221–237.

Santoso, T., Ahmed, S., Goetschalckx, M., and Shapiro, A. (2005). A stochastic programming approach for supply chain network design under uncertainty. *European Journal of Operational Research*, 167(1):96–115.

Sherali, H. D. and Lunday, B. J. (2013). On generating maximal nondominated benders cuts. *Annals of Operations Research*, 210:57–72.

Stursberg, P. M. (2019). *On the mathematics of energy system optimization*. PhD thesis, Technische Universität München.

Van Roy, T. J. (1986). A cross decomposition algorithm for capacitated facility location. *Operations Research*, 34(1):145–163.

Wentges, P. (1996). Accelerating benders' decomposition for the capacitated facility location problem. *Mathematical Methods of Operations Research*, 44(2):267–290.

Wu, L. and Shahidehpour, M. (2010). Accelerating the benders decomposition for network-constrained unit commitment problems. *Energy Systems*, 1(3):339–376.

You, F. and Grossmann, I. E. (2013). Multicut benders decomposition algorithm for process supply chain planning under uncertainty. *Annals of Operations Research*, 210:191–211.

Appendix A. Examples

Example A.1 (MIS cuts are affected by unnecessary constraints, constraint scaling, and might need more than one cut to cut off a single assignment of master variables)**.** *Consider the optimization problem*

$$\min \quad \frac{7}{10}x + y \tag{17a}$$

$$s.t. \quad y \geq 3 - x \tag{17b}$$

$$10y \geq 20 - 5x \tag{17c}$$

$$x \in [-1, 1], \ y \geq 0. \tag{17d}$$

We note that Constraint (17c) *does not influence the feasible set of Problem* (17) *but strongly influences the course of BD using MIS cut selection, as we will see.*
The first Benders master problem evaluates to

$$\begin{aligned} \min \quad & \frac{7}{10}x + \eta \\ s.t. \quad & x \in [-1, 1], \ \eta \geq 0. \end{aligned}$$

Its solution is $\bar{x}^{(1)} = -1, \bar{\eta}^{(1)} = 0$*, with a value of* $-\frac{7}{10}$*. Problem* (4) *reads*

$$\begin{aligned} \max \quad & 4\pi_1 + 25\pi_2 \\ s.t. \quad & \pi_1 + 10\pi_2 \leq \pi_0 \\ & \pi_1 + \pi_2 + \pi_0 = 1 \\ & \pi \geq 0. \end{aligned}$$

Its solution is $\bar{\pi} = (0, \frac{1}{11}, \frac{10}{11})^T$*, with a value of* $\frac{25}{11}$*. It generates the cut*

$$\frac{10}{11}\eta \geq \frac{1}{11}(20 - 5x).$$

This cut does not support the epigraph of the subproblem's value function at $\bar{x}^{(1)}$*. We further note that the updated master problem,*

$$\begin{aligned} \min \quad & \frac{7}{10}x + \eta \\ s.t. \quad & \eta \geq 2 - 0.5x \\ & x \in [-1, 1], \ \eta \geq 0, \end{aligned}$$

has the solution $\bar{x}^{(2)} = -1$*,* $\bar{\eta}^{(2)} = \frac{5}{2}$*. We note that* $\bar{x}^{(1)} = \bar{x}^{(2)}$*.*
If Constraint (17c) *is replaced by the equivalent constraint*

$$y \geq 2 - \frac{1}{2}x,$$

then Problem (4) *gets*

$$\begin{aligned} \max \quad & 4\pi_1 + \frac{5}{2}\pi_2 \\ s.t. \quad & \pi_1 + \pi_2 \leq \pi_0 \\ & \pi_1 + \pi_2 + \pi_0 = 1 \\ & \pi \geq 0. \end{aligned}$$

This is solved by $\bar{\pi} = (\frac{1}{2}, 0, \frac{1}{2})^T$*. The cut reads*

$$\frac{1}{2}\eta \geq \frac{3}{2} - \frac{1}{2}x.$$

We further note that leaving away Constraint (17c) *does not change the feasible set of Problem* (17)*. In this case, Problem* (4) *gets*

$$\begin{aligned} \max \quad & 4\pi_1 \\ s.t. \quad & \pi_1 \leq \pi_0 \\ & \pi_1 + \pi_0 = 1 \\ & \pi \geq 0. \end{aligned}$$

It is solved by $\bar{\pi} = (\frac{1}{2}, \frac{1}{2})^T$*. The cut reads*

$$\frac{1}{2}\eta \geq \frac{3}{2} - \frac{1}{2}x$$

as well. The updated master problem reads

$$\begin{aligned} \min \quad & \frac{7}{10}x + \eta \\ s.t. \quad & \eta \geq 3 - x \\ & x \in [-1, 1], \; \eta \geq 0 \end{aligned}$$

in both cases and is solved by $\bar{x}^{(2)} = 1$, $\bar{\eta}^{(2)} = 2$*. As* $(1, 2)^T$ *is the (unique) optimal solution to Problem* (17)*, the algorithm terminates here.*

Example A.2 (MIS cuts for problems with equality constraints)**.** *Alternatively, we can investigate what happens if Constraint* (17c) *is replaced by*

$$y = 1 - \frac{1}{4}x.$$

We note that this makes Problem (17) *infeasible. With* $\bar{x}^{(1)} = -1$ *and* $\bar{\eta}^{(1)} = 0$ *as before, Problem* (4) *gets*

$$\begin{aligned} \max \quad & 4\pi_1 + \frac{5}{4}\pi_2 \\ s.t. \quad & \pi_1 + \frac{1}{2}\pi_2 \leq \pi_0 \\ & \pi_1 + \pi_2 + \pi_0 = 1 \\ & \pi_2 \in \mathbb{R}, \pi_1, \pi_0 \geq 0. \end{aligned}$$

The feasible set of this problem has one vertex, $(0, \frac{2}{3}, \frac{1}{3})^T$*, and two extreme rays,* $(1, \frac{-4}{3}, \frac{1}{3})^T$ *and* $(0, -1, 1)^T$*. The vertex has a positive objective, and the first extreme ray has a positive objective. Hence, the problem* *is unbounded. The cut generated by the vertex reads*

$$\eta \geq 2 - \frac{1}{2}x.$$

Inserting this cut into the master problem generates $\bar{x}^{(2)} = -1$ *again and* $\bar{\eta}^{(2)} = \frac{5}{2}$*. Problem* (4) *gets*

$$\begin{aligned} \max \quad & 4\pi_1 + \frac{5}{4}\pi_2 - \frac{5}{2}\pi_0 \\ s.t. \quad & \pi_1 + \frac{1}{2}\pi_2 \leq \pi_0 \\ & \pi_1 + \pi_2 + \pi_0 = 1 \\ & \pi_2 \in \mathbb{R}, \pi_1, \pi_0 \geq 0. \end{aligned}$$

We cannot use the cut generated by the vertex again, as this would not lead to progress. The cut generated by the extreme ray reads

$$\frac{1}{3}\eta \geq 1(3 - x) - \frac{4}{3}(1 - \frac{1}{4}x) \Leftrightarrow \eta \geq 9 - 3x - 4 + x \Leftrightarrow \eta \geq 5 - 2x.$$

Inserting this cut into the master problem generates $\bar{x}^{(3)} = 1$ *and* $\bar{\eta}^{(3)} = 3$*. Problem* (4) *gets*

$$\begin{aligned} \max \quad & 2\pi_1 + \frac{3}{4}\pi_2 - 3\pi_0 \\ s.t. \quad & \pi_1 + \frac{1}{2}\pi_2 \leq \pi_0 \\ & \pi_1 + \pi_2 + \pi_0 = 1 \\ & \pi_2 \in \mathbb{R}, \pi_1, \pi_0 \geq 0. \end{aligned}$$

As the optimal value of this problem is negative, it cannot generate a cut that cuts off $(\bar{x}, \bar{\eta})$*. The method terminates with a wrong result.*

Example A.3 (MIS cuts for problems with equality constraints, derived from the alternative polyhedron)**.** *Alternatively, we can investigate what happens if Constraint* (17c) *is replaced by the equation*

$$y = 1 - \frac{1}{4}x$$

as before, but we solve a dual subproblem over the alternative problem described in Fischetti et al. (2010), and with a weight vector of $(w_1, w_2, w_0)^T = (1, 1, 1)^T$*. This means that we exchange objective function and the left-hand side of the normalization constraint in Problem* (4)*. With* $\bar{x}^{(1)} = -1$ *and* $\bar{\eta}^{(1)} = 0$ *as before, modified Problem* (4) *gets*

$$\begin{aligned} \max \quad & \pi_1 + \pi_2 + \pi_0 \\ s.t. \quad & \pi_1 + \frac{1}{2}\pi_2 \leq \pi_0 \\ & 4\pi_1 + \frac{5}{4}\pi_2 = 1 \\ & \pi_2 \in \mathbb{R}, \pi_1, \pi_0 \geq 0. \end{aligned}$$

The feasible set of this problem has two vertices, $(0, \frac{4}{5}, \frac{2}{5})^T$ *and* $(\frac{2}{3}, -\frac{4}{3}, 0)^T$*, and two extreme rays,* $(1, \frac{-16}{5}, 0)^T$ *and* $(0, 0, 1)^T$*. The first vertex has a positive objective, and the second extreme ray has a* *positive objective. Hence, the problem is unbounded. The cut generated by the vertex reads*

$$\eta \geq 2 - \frac{1}{2}x.$$

Inserting this cut into the master problem generates $\bar{x}^{(2)} = -1$ again and $\bar{\eta}^{(2)} = \frac{5}{2}$. Modified Problem (4) gets

$$\begin{aligned} \max \quad & \pi_1 + \pi_2 + \pi_0 \\ s.t. \quad & \pi_1 + \frac{1}{2}\pi_2 \leq \pi_0 \\ & 4\pi_1 + \frac{5}{4}\pi_2 - \frac{5}{2}\pi_0 = 1 \\ & \pi_2 \in \mathbb{R}, \pi_1, \pi_0 \geq 0. \end{aligned}$$

The feasible set of this problem has one vertex $(\frac{2}{3}, -\frac{4}{3}, 0)^T$ and two extreme rays $(1, \frac{-16}{5}, 0)^T$ and $(0, 2, 1)^T$. Only the second ray has a positive objective value and generates the cut

$$\eta \geq 2 - \frac{1}{2}x.$$

This is the same cut as before, and the method terminates with a wrong result.

Example A.4 (Optimality cuts can cut off infeasible points)**.** *Consider the optimization problem*

$$\min \quad \frac{1}{2}x + y \tag{18a}$$
$$s.t. \quad y \geq 3 - x \tag{18b}$$
$$y \geq 3 + x \tag{18c}$$
$$-y \geq -4 + x \tag{18d}$$
$$x \in [-1, 1], \; y \geq 0. \tag{18e}$$

The first Benders master problem evaluates to

$$\begin{aligned} \min \quad & \frac{1}{2}x + \eta \\ s.t. \quad & x \in [-1, 1], \; \eta \geq 0. \end{aligned}$$

Its solution is $\bar{x}^{(1)} = -1, \bar{\eta}^{(1)} = 0$, with a value of $-\frac{1}{2}$. Problem (3) reads

$$\begin{aligned} \max \quad & 4\pi_1 + 2\pi_2 - 5\pi_3 \\ s.t. \quad & \pi_1 + \pi_2 - \pi_3 \leq 1 \\ & \pi \geq 0. \end{aligned}$$

Its solution is $\bar{\pi} = (1, 0, 0)^T$, with a value of 4. We get $\frac{7}{2}$ as an upper bound for the value of Problem (18). It generates the cut

$$\eta \geq -x =: \omega_1(x).$$

The Benders master problem gets

$$\begin{aligned} \min \quad & \frac{1}{2}x + \eta \\ s.t. \quad & \eta \geq 3 - x \\ & x \in [-1, 1], \;\; \eta \geq 0. \end{aligned}$$

Its solution is $\bar{x}^{(2)} = 1, \bar{\eta}^{(2)} = 2$*, with a value of* $\frac{5}{2}$*. Problem* (3) *reads*

$$\begin{aligned} \max \quad & 2\pi_1 + 4\pi_2 - 3\pi_3 \\ s.t. \quad & \pi_1 + \pi_2 - \pi_3 \leq 1 \\ & \pi \geq 0. \end{aligned}$$

This problem is unbounded, with ray $\bar{\pi} = (0, 1, 1)$*. It generates the cut*

$$0 \geq 3 + x - 4 + x \Leftrightarrow 0 \geq 2x - 1 =: \omega_2(x).$$

The SC-set of this cut is $C_{\omega_2, \frac{7}{2}} = [-1, \frac{1}{2}]$*. We note that* $\bar{\pi} = (0, 1, 0)$ *is a feasible solution of the dual subproblem as well, and the cut*

$$\eta \geq 3 + x =: \tilde{\omega}_2(x)$$

is also violated by $(\bar{x}^{(2)}, \bar{\eta}^{(2)}) = (1, 2)$*. The SC-set of this cut is* $C_{\tilde{\omega}_2, \frac{7}{2}} = [-1, \frac{1}{3}]$*.*

Example A.5 (Feasibility cuts can have minimal SC-sets)**.** *Consider the optimization problem*

$$\begin{aligned} \text{(19a)} \quad \min \quad & \frac{1}{10}x + y \\ \text{(19b)} \quad s.t. \quad & y \geq 3 - x \\ \text{(19c)} \quad & y \geq 3 + x \\ \text{(19d)} \quad & -y \geq -4 + x \\ \text{(19e)} \quad & x \in [-1, 1], \;\; y \geq 0. \end{aligned}$$

The first Benders master problem evaluates to

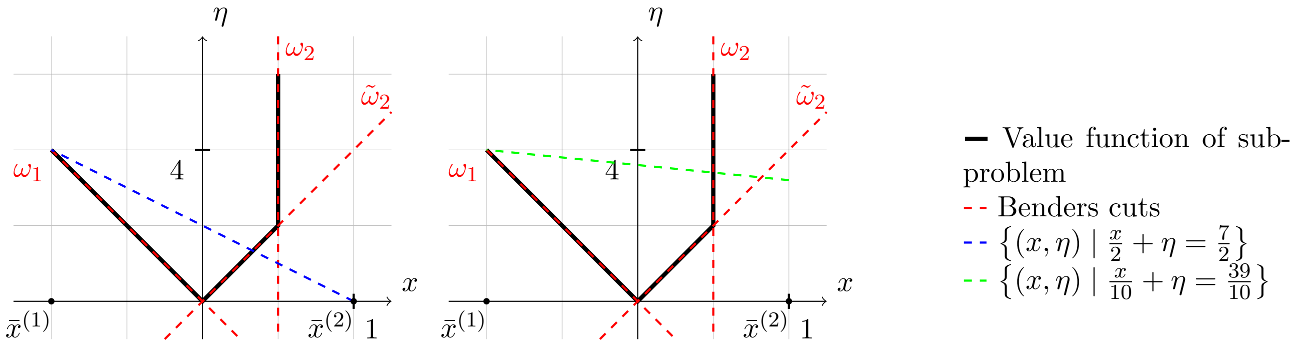
$$\begin{aligned} \min \quad & \frac{1}{10}x + \eta \\ s.t. \quad & x \in [-1, 1], \;\; \eta \geq 0. \end{aligned}$$

Its solution is $\bar{x}^{(1)} = -1, \bar{\eta}^{(1)} = 0$*, with a value of* $-\frac{1}{10}$*. Problem* (3) *reads*

$$\begin{aligned} \max \quad & 4\pi_1 + 2\pi_2 - 5\pi_3 \\ s.t. \quad & \pi_1 + \pi_2 - \pi_3 \leq 1 \\ & \pi \geq 0. \end{aligned}$$

Its solution is $\bar{\pi} = (1, 0, 0)^T$*, with a value of* 4*. We note that we get* $\frac{39}{10}$ *as an upper bound for the value of Problem* (18)*. It generates the cut*

$$\eta \geq -x =: \omega_1(x).$$



(A) Value function of Problem (3) and Benders cuts for Example A.4.

(B) Value function of Problem (3) and Benders cuts for Example A.5.

FIGURE 17. Value function of Problem (3), and Benders cuts for Example A.4 and A.5.

The Benders master problem gets

$$\begin{aligned} \min \quad & \frac{1}{10}x + \eta \\ s.t. \quad & \eta \geq 3 - x \\ & x \in [-1, 1], \ \eta \geq 0. \end{aligned}$$

Its solution is $\bar{x}^{(2)} = 1, \bar{\eta}^{(2)} = 2$, *with a value of* $\frac{21}{10}$. *Problem* (3) *reads*

$$\begin{aligned} \max \quad & 2\pi_1 + 4\pi_2 - 3\pi_3 \\ s.t. \quad & \pi_1 + \pi_2 - \pi_3 \leq 1 \\ & \pi \geq 0. \end{aligned}$$

This problem is unbounded, with ray $\bar{\pi} = (0, 1, 1)$. *It generates the cut*

$$0 \geq 3 + x - 4 + x \Leftrightarrow 0 \geq 2x - 1 =: \omega_2(x).$$

The SC-set of this cut is $C_{\omega_2, \frac{39}{10}} = [-1, \frac{1}{2}]$. *Each optimality cut* $\tilde{\omega}_2$ *has to fulfill* $\tilde{\omega}_2(\frac{1}{2}) \leq \frac{7}{2} < \frac{77}{20}$. *This implies that the point* $\frac{1}{2}$ *is in the interior of its SC-set, which is hence not minimal.*