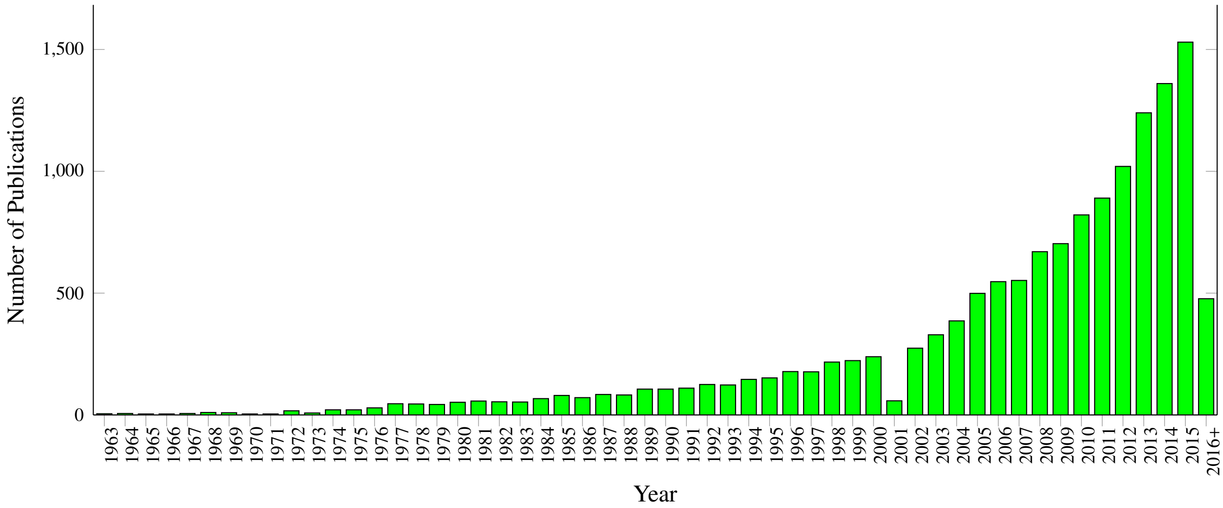

Figure 1: Publication–year distribution of BD research according to https://scholar.google.ca/.

taxonomy, used to survey the acceleration strategies in Sections 4 to 7. Section 8 presents Benders-type heuristics, and Section 9 describes extensions of the classical algorithm. Finally, Section 10 provides concluding remarks and describes promising research directions.

2. The Benders Decomposition Method

We present in this section the classical version of the Benders algorithm (Benders, 1962). We review its extensions to broader range of optimization problems in Section 9.

2.1. *The classical version*

We consider an MILP of the form

$$\begin{aligned} \text{Minimize} \quad & f^T y + c^T x && (1)\\ \text{subject to} \quad & Ay = b && (2)\\ & By + Dx = d && (3)\\ & x \geq 0 && (4)\\ & y \geq 0 \quad \text{and integer,} && (5) \end{aligned}$$

with complicating variables $y \in \mathfrak{R}^{n_1}$, which must take positive integer values and satisfy the constraint set $Ay = b$, where $A \in \mathfrak{R}^{m_1 \times n_1}$ is a known matrix and $b \in \mathfrak{R}^{m_1}$ is a given vector. The continuous variables $x \in \mathfrak{R}^{n_2}$, together with the y variables, must satisfy the linking constraint set $By + Dx = d$, with $B \in \mathfrak{R}^{m_2 \times n_1}$, $D \in \mathfrak{R}^{m_2 \times n_2}$, and $d \in \mathfrak{R}^{m_2}$. The objective function minimizes the total cost with the cost vectors $f \in \mathfrak{R}^{n_1}$ and $c \in \mathfrak{R}^{n_2}$.

Model (1–5) can be re-expressed as

$$\min_{\bar{y}\in Y}\{f^T\bar{y} + \min_{x\geq 0}\{c^T x : Dx = d - B\bar{y}\},\} \tag{6}$$

where $Y = \{y|Ay = b, y \geq 0 \text{ and integer}\}$. The inner minimization is a continuous linear problem that can be dualized by means of dual variables π associated with the constraint set $Dx = d - B\bar{y}$:

$$\max_{\pi\in\Re^{m_2}}\{\pi^T(d - B\bar{y}) : \pi^T D \leq c\} \tag{7}$$

Based on duality theory, the primal and dual formulations can be interchanged to extract the following equivalent formulation:

$$\min_{\bar{y}\in Y}\{f^T\bar{y} + \max_{\pi\in\Re^{m_2}}\{\pi^T(d - B\bar{y}) : \pi^T D \leq c\}\} \tag{8}$$

The feasible space of the inner maximization, i.e., $F = \{\pi|\pi^T D \leq c\}$, is independent of the choice of $\bar{y}$. Thus, if F is not empty, the inner problem can be either unbounded or feasible for any arbitrary choice of $\bar{y}$. In the former case, there is a direction of unboundedness $r_q, q \in Q$ for which $r_q^T(d - B\bar{y}) > 0$; this must be avoided because it indicates the infeasiblity of the $\bar{y}$ solution. We add a cut

$$r_q^T(d - B\bar{y}) \leq 0 \qquad \forall q \in Q \tag{9}$$

to the problem to restrict movement in this direction. In the latter case, the solution of the inner maximization is one of the extreme points $\pi_e, e \in E$. If we add all the cuts of the form (9) to the outer minimization problem, the value of the inner problem will be one of its extreme points. Consequently, problem (8) can be reformulated as:

$$\begin{aligned} \min_{\bar{y}\in Y} \quad & f^T\bar{y} + \max_{e\in E}\{\pi_e^T(d - B\bar{y})\} && (10)\\ \text{subject to} \quad & r_q^T(d - B\bar{y}) \leq 0 \qquad \forall q \in Q && (11) \end{aligned}$$

This problem can easily be linearized via a continuous variable $\eta \in \Re^1$ to give the following equivalent formulation to problem (1–5), which we refer to as the Benders *Master Problem* (*MP*):

$$\begin{aligned} \min_{y,\eta} \quad & f^T y + \eta && (12)\\ \text{subject to} \quad & Ay = b && (13)\\ & \eta \geq \pi_e^T(d - By) \qquad \forall e \in E && (14)\\ & 0 \geq r_q^T(d - By) \qquad \forall q \in Q && (15)\\ & y \geq 0 \quad \text{and integer} && (16) \end{aligned}$$

Constraints (14) and (15) are referred to as *optimality* and *feasibility* cuts, respectively. The complete enumeration of these cuts is generally not practical. Therefore, Benders (1962) proposed a relaxation of the feasibility and optimality cuts and an iterative approach. Thus, the BD algorithm repeatedly solves the MP, which includes only a subset of constraints (14) and (15), to obtain a trial value for the y variables, i.e., $\bar{y}$. It then solves subproblem (7) with $\bar{y}$. If the subproblem is feasible and bounded, a cut of type (14) is produced. If the subproblem is unbounded, a cut of type (15) is produced. If the cuts are violated by the current solution, they are inserted into the current MP and the process repeats.

Figure 2 illustrates the BD algorithm. After deriving the initial MP and subproblem, the algorithm alternates between them (starting with the MP) until an optimal solution is found. To confirm the convergence, the optimality gap can be calculated at each iteration. The objective function of the MP gives a valid lower bound on the optimal cost because it is a relaxation of the equivalent Benders reformulation. On the other hand, combining the $\bar{y}$ solution with the objective value of the subproblem, which is equivalent to fixing $\bar{y}$ in the original formulation, yields a valid upper bound on the optimal cost.

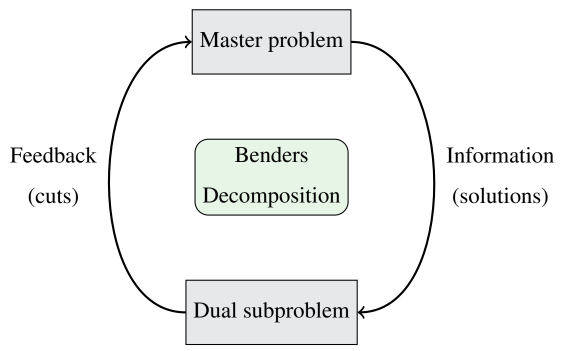

Figure 2: Schematic representation of Benders decomposition method.

2.2. *Model selection for Benders decomposition*

A given problem can usually be modeled with different but equivalent formulations. However, from a computational point of view, the various formulations may not be equivalent. Geoffrion and Graves (1974) observed that the formulation has a direct impact on the performance of the BD. Magnanti and Wong (1981) demonstrated that a formulation with a stronger LP relaxation will have better performance. This is because of the tighter root node and the smaller number of fractional variables and also because the generated cuts are provably stronger. Sahinidis and Grossmann (1991) proved that the BD method applied to a mixed integer nonlinear programming (NLP) formulation with a zero NLP relaxation gap requires only the cut corresponding to the optimal solution to converge. Cordeau et al. (2006) studied a stochastic logistic network design problem. They found that when the original formulation was strengthened with a set of valid inequalities (VIs), the performance of the BD method was considerably improved.

These observations confirm the importance of tight formulations in the context of the BD method. However, tighter formulations are often obtained by adding additional constraints. This may result in a more time-consuming subproblem, which may also exhibit a higher degree of degeneracy. Therefore, there must be a trade-off between the reduction in the number of iterations and the additional difficulty of the subproblem.