BENCHMARKING PIECEWISE LINEAR RELAXATIONS FOR MINLPs: A COMPUTATIONAL STUDY BASED ON THE OPEN-SOURCE FRAMEWORK PWL-T-REX

KRISTIN BRAUN[1,2,*], ROBERT BURLACU[2]

[1] *Fraunhofer Institute for Integrated Circuits IIS, Nordostpark 84, D-90411 Nürnberg, Germany*

[2] *University of Technology Nuremberg, Nordostpark 93, D-90411 Nürnberg, Germany*

[*] *Corresponding author,* kristin.braun@iis.fraunhofer.de

Abstract. Solving mixed-integer nonlinear problems by means of piecewise linear approximations has emerged as a reasonable alternative to the commonly used spatial branch-and-bound. These approximations have been modeled by various mixed-integer linear models in recent decades. The idea is to exploit the availability of mature solvers for mixed-integer linear problems.

In this work, we implement a framework that reformulates mixed-integer nonlinear problems to eight commonly used different mixed-integer linear representations for piecewise linear relaxations where the reformulations can be compared in terms of behavior and runtime to determine which method to apply in practice. We use expression trees to reformulate all nonlinearities to one-dimensional functions and then compute a set of interpolation breakpoints for each function based on a given maximum error. This framework is made publicly available, see [7].

Further, we conduct an analysis on a benchmark set created from the MINLPLIB consisting of over 750 instances. It includes a comprehensive comparison of the number of problems solved, runtimes, and optimality gaps.

Keywords: Piecewise linear relaxations, mixed-integer nonlinear programming, mixed-integer linear programming, mixed-integer programming, discrete optimization

Acknowledgments

The authors gratefully acknowledge the scientific support and HPC resources provided by the Erlangen National High Performance Computing Center (NHR@FAU) of the Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU). The hardware is funded by the German Research Foundation (DFG). This work has been done within the joint project "TrinkXtrem" funded by the Federal Ministry of Education and Research (BMBF) under the project number 02WEE1625B in the funding "Wasser-Extremereignisse" (WaX) of

the Federal Program "Wasser:N" and as part of the announcement "Artificial Intelligence in Civil Security Research II" of the BMBF within the program "Research for Civil Security" of the Federal Government.

1. Introduction

To this day, general MINLPs remain very difficult to solve. Spatial branch-and-bound is still at the heart of most state-of-the-art solvers. However, over the last two decades, several methods have been presented that treat non-convex MINLPs by piecewise relaxations without direct branching on continuous variables, ranging from piecewise linear relaxations/approximations [28, 15, 32, 8, 1] to more general piecewise convex relaxations [25, 16, 24, 4]. While these approaches are sometimes quite different, they all need to tackle the following two problems: The construction of tight relaxations of the nonlinear functions and the incorporation of these relaxations into a mixed-integer linear program (MILP) or convex nonlinear program (NLP).

One approach to obtain such relaxations is to compute an optimal linearization of a nonlinear function in terms of the number of breakpoints and an a priori given accuracy as in [36, 35] and [37]. Therein, δ-underestimators and δ-overestimators are created – this is essentially what is needed for the creation of relaxations. This is supplemented by the construction of optimal polynomial relaxations of one-dimensional functions in [32]. The construction of optimal breakpoints for quadratic functions is presented in [34]. Specific approximation techniques for general nonlinear functions with dimensions smaller than three are proposed in [30]. The number of simplices in the approximation grows exponentially with the function's dimension. In this regard, we refer to the method of [38], which avoids this problem on the basis that the piecewise linear approximation does not have to interpolate the original function at the vertices of the triangulation.

This article covers a variety of different approaches existing for modeling PWL functions as a MILP: the disaggregated (as in [12, 39]) and aggregated convex combination models (as in [23, 33]) as well as their logarithmic variants presented by [42], the classical incremental method of [27], and the multiple choice model as in [2]. Moreover, we also consider the very recently introduced binary and integer Zig-Zag formulations, which are also logarithmic versions of the convex combination models; see [20]. In [21], these methods are clustered and extensively tested on transportation problems. For more details on the theoretical foundations of MILP models for PWL functions and general mixed-integer linear formulation techniques, we refer to the extensive survey by [41]. Further, recent work by [26] introduces ideal formulations for tight simultaneous upper and lower bounds. Because their approach relies on a fundamentally different modeling structure, a direct empirical comparison falls outside the scope of this baseline benchmark.

To the best of our knowledge, no comprehensive study has yet been conducted that demonstrates the performance of these MILP models on a wide and general range of MINLP problems. In most cases, some parts of the mentioned MILP models are compared only on some instances of very specific problems; see for instance [11, 19, 21, 34]. While these studies demonstrate that specialized formulations can be very strong in specific

TABLE 1.1. Overview of PWL models.

Full name	Short name	Presented in	Reference(s)
Disaggregated	(Disag)	Section 3.1	[42] and references therein
Log. Disaggregated	(LogDisag)	Section 3.1	[42]
Aggregated	(Ag)	Section 3.2	[42] and references therein
Log. Aggregated	(LogAg)	Section 3.2	[42]
Binary Zig-Zag	(BinZigZag)	Section 3.2	[20]
Integer Zig-Zag	(IntZigZag)	Section 3.2	[20]
Incremental	(Inc)	Section 3.3	[27]
Multiple Choice	(MC)	Section 3.4	[22]

contexts, it remains an open question how they perform across more general instance sets. Therefore, it becomes difficult to decide, in general, which MILP modeling is preferable, or if there is a best modeling at all. To this end, we perform an extensive computational study on over 750 instances from the MINLPLIB, cf. [9]. Using expression trees and the results from [3], we reformulate all MINLP instances to equivalent models that consist of only one-dimensional nonlinear functions. Based on these reformulations, we compare various MILP relaxations corresponding to the different PWL models. Note that, while some modern MINLP solvers like Gurobi provide internal piecewise linear approaches via adaptive refinement, our study focuses on the computational performance of different MILP encodings. This approach allows for a controlled comparison of the formulations themselves and remains applicable to general piecewise linear problems beyond the scope of specific MINLP solver features.

Overall, our study shows that the incremental method, although the oldest approach of all, generally performs best when MINLPs are solved by PWL relaxations that are modeled as MILPs, while the very recently introduced logarithmic integer Zig-Zag model by [20] is a reasonable alternative for very high-accuracy PWL relaxations. While the multiple choice method is rarely recommended in the literature, our findings offer compelling evidence in its efficiency in finding primal feasible solutions.

Contribution. In this work, we conduct a large-scale computational study comparing eight different MILP formulations based on PWL relaxations. Our key contributions include:

(1) A framework for generating and solving MINLPs via piecewise linear relaxations, with selectable MILP reformulations and error bounds. Our framework is made publicly available as an open-source implementation on GitHub, see [7].
(2) A detailed performance evaluation on over 750 instances from MINLPLIB, analyzing problem solvability, runtime, and optimality gaps.
(3) The creation and validation of a branching scheme for the logarithmic aggregated convex combination method, which is used in our computational study.

Structure. The rest of this article is structured as follows. We introduce all necessary definitions and basic ideas on how to solve MINLPs by PWL relaxations in Section 2. Section 3 provides a detailed overview of the various MILP models for PWL relaxations, see Table 1.1. After describing our framework in Section 4, we present the computational study in Section 5 that illustrates the practicability of the various MILP models for PWL relaxations. We discuss the numerical results in Section 6 and conclude this work in Section 7.

2. Preliminaries

We define a MINLP as an optimization problem of the following type:

$$\begin{aligned} \min_{x} \quad & c^\top x \\ \text{s.t.} \quad & Ax \leq b, \\ & f_i(x) \leq 0 \qquad \text{for all } i \in \{1, \dots, k\}, \\ & l \leq x \leq u, \\ & x \in \mathbb{R}^q \times \mathbb{Z}^p, \end{aligned} \tag{P}$$

where $A \in \mathbb{R}^{m \times (p+q)}, b \in \mathbb{R}^m, c \in \mathbb{R}^{p+q}$ and $k, q, p, m \in \mathbb{N}$. The variables x can be bounded from below and above by $l \in (\mathbb{R} \cup \{-\infty\})^{q+p}$ and $u \in (\mathbb{R} \cup \{\infty\})^{q+p}$. The inequalities $Ax \leq b$ denote the linear constraints, while the nonlinear constraints are denoted by $\mathcal{F} := \{f_1, \dots, f_k\}$. Each function $f_i \in \mathcal{F}$ is a nonlinear, real-valued function $f_i \colon D_{f_i} \to \mathbb{R}$ where $D_{f_i} \subset \mathbb{R}^{q+p}$ is the domain of f_i. Equality constraints, i.e., constraints of type $f_i(x) = 0$, are inherently contained in the description (P) by simply adding $f_i(x) \leq 0$ and $-f_i(x) \leq 0$. Further, we are not restricted to a linear objective function $c^\top x$, since we can include any nonlinear objective function $f_c \colon D_{f_c} \to \mathbb{R}$ by substituting $f_c(x)$ with a variable $y \in \mathbb{R}$ and adding $f_c(x) \leq y$ as a constraint to (P). As $\max c^\top x = -\min -c^\top x$, any maximization problem can be transformed to a minimization problem. Therefore, (P) serves as a comprehensive formal representation of a MINLP.

In our work, we have two limitations for (P):

(1) The variables are finite, i.e., $l, u \in \mathbb{R}$.
(2) The nonlinear functions f are continuous and bounded through the choice of D_f.

By (2), the domain D_f is defined as a d-dimensional box with $d \leq q + p$, whose edges are parallel to the coordinate axes, making it a compact set.

To solve problems of type (P) with box-constrained domains, we employ piecewise linear (PWL) relaxations. Their practical applicability depends on the dimensionality of the nonlinear functions to be approximated, since the number of simplices required for a given approximation quality grows exponentially with this dimension. Therefore, PWL relaxations are particularly attractive when the nonlinear part can be decomposed into low-dimensional factorable terms. As is also common in spatial branch-and-bound methods for factorable MINLPs, we first reformulate the nonlinear expressions via expression trees into low-dimensional building blocks. By using PWL relaxations, we permit solutions to be infeasible within a strictly specified error bound

ϵ, effectively controlling the approximation quality. However, MINLPs are typically solved only within given feasibility and optimality tolerances, this approach remains practical, provided ϵ is chosen appropriately.

A common approach to PWL relaxations is to first construct a PWL approximation of f by interpolating the function over a d-dimensional simplex. The corresponding approximation error is then incorporated into the relaxation to ensure validity. While this construction yields a mathematically valid relaxation for any error bound ϵ, we note that larger values naturally lead to weaker bounds in the optimization process. Relaxations are preferable to approximations because they preserve all feasible solutions, thereby guaranteeing valid bounds, whereas approximations may exclude feasible solutions, leading to missing upper or lower bounds. To achieve this, we first construct a triangulation of D_f. However, the number of simplices in a triangulation of a d-dimensional box grows exponentially with d for a fixed approximation quality. As a result, this method is only practical for low-dimensional cases.

For a wide range of practically interesting MINLPs, the nonlinear functions are factorable, and can be represented by the following functions:

$$\begin{aligned} &f(x_1, x_2) = x_1 \cdot x_2,\ f(x_1, x_2) = \frac{x_1}{x_2},\ f(x) = \ln x,\ f(x) = \log_{10} x, \\ &f(x) = x^a \text{ (including } x^{-1} \text{ and } \sqrt{x}),\ f(x) = a^x \text{ (including } e^x), \\ &f(x) = \sin x,\ f(x) = \cos x,\ f(x) = \tanh x,\ f(x) = |x|, \end{aligned} \tag{2.1}$$

as described by [5]. The frequent use of this leads to an equivalent formulation of the MINLP problem that contains only univariate nonlinearities and, if necessary, the coupling bivariate functions $f(x_1, x_2) = x_1 x_2$ and $f(x_1, x_2) = \frac{x_1}{x_2}$. Following the results of [3], we can further reformulate $f(x_1, x_2)$ to a one-dimensional representation via

$$f(x_1, x_2) = \frac{1}{2}\left(p^2 - x_1^2 - x_2^2\right) \tag{2.2}$$

and

$$p = x_1 + x_2 \tag{2.3}$$

using a new variable $p \in \mathbb{R}$. The term $f(x_1, x_2) = \frac{x_1}{x_2}$ is reformulated similarly, as it equals $f(x_1, x_2) = x_1 \cdot \frac{1}{x_2}$. The authors show that in practice it is more favorable to use (2.2) and (2.3) instead of a bivariate product when dealing with PWL approximations. More precisely, the results in [3] indicate that bivariate PWL formulations are mainly attractive for very coarse approximations, whereas for the broader range of approximation accuracies considered in this paper, univariate reformulations are generally preferable when strengthened by McCormick inequalities as in [29]. We therefore use the reformulation (2.2) and (2.3) together with the following inequalities:

$$\frac{1}{2}\left(p^2 - x_1^2 - x_2^2\right) \geq l_1 \cdot x_2 + l_2 \cdot x_1 - l_1 \cdot l_2, \tag{2.4a}$$

$$\frac{1}{2}\left(p^2 - x_1^2 - x_2^2\right) \geq u_1 \cdot x_2 + u_2 \cdot x_1 - u_1 \cdot u_2, \tag{2.4b}$$

$$\frac{1}{2}\left(p^2 - x_1^2 - x_2^2\right) \leq u_1 \cdot x_2 + l_2 \cdot x_1 - u_1 \cdot l_2, \tag{2.4c}$$

$$\frac{1}{2}\left(p^2 - x_1^2 - x_2^2\right) \leq l_1 \cdot x_2 + u_2 \cdot x_1 - l_1 \cdot u_2, \tag{2.4d}$$

where l_1, u_1, l_2, and u_2 are the lower and upper bounds of x_1 and x_2, respectively. We use this reformulation as it tightens the relaxation, leading to improved solver performance.

This gives an equivalent reformulation of the MINLP that consists solely of one-dimensional nonlinearities. Hence, in the following, we consider only one-dimensional MILP models for PWL relaxations of nonlinear functions.

For tackling a MINLP by PWL relaxations, an adaptive refinement of the PWL relaxations is usually crucial for the performance of the approach. In this article, we omit this algorithmic overhead since we are primarily interested in comparing various MILP models for PWL relaxations with different approximation errors.

3. One-dimensional MILP models for piecewise linear relaxations

In this chapter, we describe in more detail the MILP representations of the piecewise linear models that we use in this work. For all representations, we show the MILP formulation as well as a visual representation.

3.1. **Disaggregated convex combination model.**

The first model that we discuss is the disaggregated convex combination model. Here, each feasible point $(x, \bar{f}(x))$ is represented as a convex combination of its two neighboring breakpoints. Assuming that x lies in the i-th segment, i.e., $\bar{x}_i \leq x \leq \bar{x}_{i+1}$ holds, then x can be represented by the equality

$$x = \lambda_i^1 \bar{x}_i + \lambda_i^2 \bar{x}_{i+1} \tag{3.1}$$

with $\lambda_i^1 + \lambda_i^2 = 1$ and $\lambda_i^1, \lambda_i^2 \geq 0$. Analogously, $\bar{f}(x)$ is then given by

$$\bar{f}(x) = \lambda_i^1 f(\bar{x}_i) + \lambda_i^2 f(\bar{x}_{i+1}) \tag{3.2}$$

using the same two variables λ_i^1, λ_i^2.

Since a piecewise linear representation consists of multiple segments, a binary variable $y_i \in \{0, 1\}$ is introduced for each segment $i \in [n]$. Exactly one of these variables must be one, the rest of the values is zero – this is also called one-hot encoded vector. The position of the one-entry indicates that x lies in segment i, i.e. segment i is active. The resulting model consists of n binary and $2n$ continuous variables. We describe all constraints in (Disag). Further, Figure 3.1 depicts a visual representation.

Model 1. Disaggregated convex combination model

$$\begin{aligned}
\sum_{i=1}^{n} \left(\lambda_i^1 \bar{x}_{i-1} + \lambda_i^2 \bar{x}_i\right) &= x, \\
\sum_{i=1}^{n} \left(\lambda_i^1 f(\bar{x}_{i-1}) + \lambda_i^2 f(\bar{x}_i)\right) &= z, \\
\lambda_i^1 + \lambda_i^2 &= y_i, \qquad i \in \{1, \dots, n\}, \qquad \text{(Disag)} \\
\sum_{i=1}^{n} y_i &= 1, \\
\lambda_i^1, \lambda_i^2 &\geq 0, \qquad i \in \{1, \dots, n\}, \\
y_i &\in \{0, 1\}, \quad i \in \{1, \dots, n\}.
\end{aligned}$$

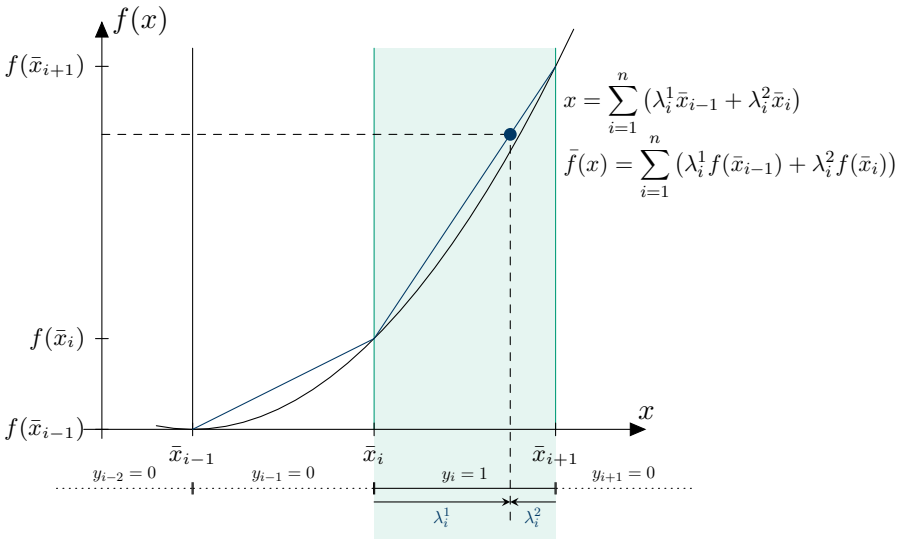


FIGURE 3.1. Representation of (Disag). The active segment ($y_i = 1$) is shaded in green. The continuous variables λ_i^1 and λ_i^2 define the point $(x, \bar{f}(x))$ as a convex combination of the segment's endpoints.

Given that it is possible to encode an n-digit number using $\lceil \log_2 n \rceil$ binary variables, the idea to create a binary representation of the segments 1 to n is obvious. To this end, new binary variables y_l with $l \in \{1, ..., \lceil \log_2 n \rceil\}$ and a binary encoding $B : [n] \to \{0, 1\}^{\lceil \log_2 n \rceil}$ are introduced. The binary encoding is not fixed; in our framework, we simply use the usual conversion from decimal to binary system. Using this encoding, one can define a branching scheme by

$$\mathscr{P}^0(B, l) := \{i \in [n] \mid B(i)_l = 0\} \tag{3.3}$$

and

$$\mathscr{P}^+(B, l) := \{i \in [n] \mid B(i)_l = 1\}. \tag{3.4}$$

Now, as an arbitrary number of binary variables can be nonzero, the representation has to be changed to ensure that $\lambda_i^1 + \lambda_i^2 = 1$ holds for exactly

one $i \in [n]$ and $\lambda_j^1 + \lambda_j^2 = 0$ holds for all $j \in [n]$ with $i \neq j$. These adjustments result in (LogDisag) that uses the same number of continuous variables as in (Disag), but now only $\lceil \log_2 n \rceil$ binary variables.

Model 2. Logarithmic disaggregated convex combination model

$$
\begin{aligned}
\sum_{i=1}^{n} \left(\lambda_i^1 \bar{x}_{i-1} + \lambda_i^2 \bar{x}_i \right) &= x, \\
\sum_{i=1}^{n} \left(\lambda_i^1 f(\bar{x}_{i-1}) + \lambda_i^2 f(\bar{x}_i) \right) &= z, \\
\sum_{i=1}^{n} \left(\lambda_i^1 + \lambda_i^2 \right) &= 1, \\
\sum_{i \in \mathscr{P}^+(B,l)} \left(\lambda_i^1 + \lambda_i^2 \right) &\leq y_l, \qquad l \in \{1, \dots, \lceil \log_2 n \rceil\}, \\
\sum_{i \in \mathscr{P}^0(B,l)} \left(\lambda_i^1 + \lambda_i^2 \right) &\leq 1 - y_l, \quad l \in \{1, \dots, \lceil \log_2 n \rceil\}, \\
\lambda_i^1, \lambda_i^2 &\geq 0, \qquad i \in \{1, \dots, n\}, \\
y_l &\in \{0,1\}, \quad l \in \{1, \dots, \lceil \log_2 n \rceil\}.
\end{aligned}
\tag{LogDisag}
$$

3.2. **Aggregated convex combination model.** For introducing the aggregated convex combination model, the previous model just needs to be adjusted slightly.

In (Disag), variables λ_{i-1}^2 and λ_i^1 represent the same tuple $(\bar{x}_i, f(\bar{x}_i))$ and, thus, are aggregated. To this end, new continuous variables $\lambda_0, \dots, \lambda_n$ are introduced. The new variable λ_0 replaces λ_1^1, λ_n replaces λ_n^2, and, for $i = 1, \dots, n-1$, λ_i replaces λ_{i-1}^2 and λ_i^1. Finally, one only needs $n+1$ instead of $2n$ continuous variables. Now, a variable λ_i is allowed to be nonzero if either segment i or segment $i+1$ is used, i.e., $\lambda_i \leq y_i + y_{i+1}$. The constraints are given in (Ag), and, again, Figure 3.2 provides a visual representation.

Model 3. Convex combination model

$$
\begin{aligned}
\sum_{i=0}^{n} \lambda_i \bar{x}_i &= x, \\
\sum_{i=0}^{n} \lambda_i f(\bar{x}_i) &= z, \\
\sum_{i=0}^{n} \lambda_i &= 1, \\
\lambda_0 &\leq y_1, \\
\lambda_i &\leq y_i + y_{i+1}, \quad i \in \{1, \dots, n-1\}, \\
\lambda_n &\leq y_n, \\
\sum_{i=1}^{n} y_i &= 1, \\
\lambda_i &\geq 0, \qquad i \in \{0, \dots, n\}, \\
y_i &\in \{0,1\}, \qquad i \in \{1, \dots, n\}.
\end{aligned}
\tag{Ag}
$$

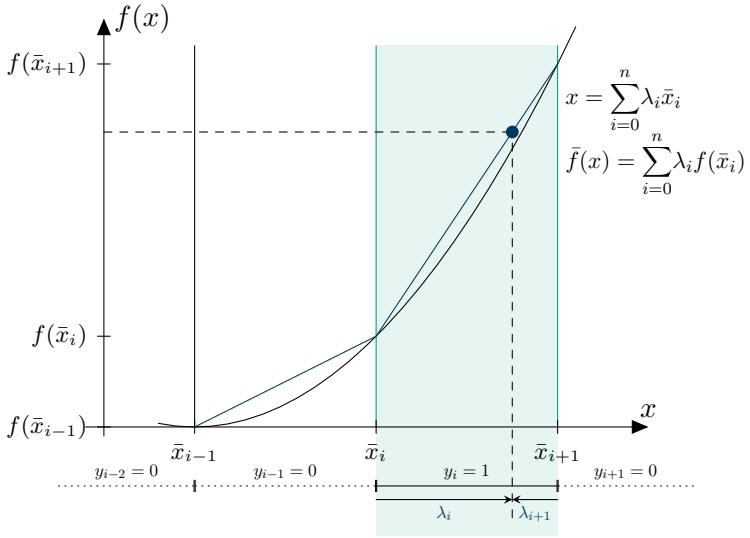


FIGURE 3.2. Representation of (Ag). The active segment ($y_i = 1$) is shaded in green. The continuous variables λ_i and λ_{i+1} define the point $(x, \bar{f}(x))$ as a convex combination of the segment's endpoints.

As before, one wants to reduce the number of binary variables in our model. For this, we present different approaches: On the one hand, a formulation using a binary branching scheme, and, on the other hand, two new formulations introduced by [20].

In (Disag) and (LogDisag), each continuous variable λ_i belongs to two different segments. The binary branching scheme of (Disag) fixes all variables λ_i^1, λ_i^2 to zero if segment i is not used.

If we use the same branching scheme as in the aggregated version here, this would result in fixing all continuous variables to zero and then no segment is usable anymore. Thus, we need to define a branching scheme $\left(\mathcal{L}^S, \mathcal{R}^S\right)$ with $\mathcal{L}^S := \left\{L_1^S, \ldots, L_S^S\right\}$ and $\mathcal{R}^S := \left\{R_1^S, \ldots, R_S^S\right\}$ for $S := \lceil \log_2 n \rceil$ such that, for all $i \in \{1, \ldots, n\}$, there exists a series $T^i = [T_1^i, \ldots, T_S^i]$ with

$$\{i-1, i\} = \bigcap_{s=1}^{S} \left([n]_0 \setminus T_s^i\right), \tag{3.5}$$

where $T_s^i \in \left\{L_s^S, R_s^S\right\}$ and $[n]_0 := [n] \cup \{0\} = \{0, \ldots, n\}$. We defined such a branching scheme for this work and describe it in the following. Logarithmic encodings for convex combinations were already introduced by [42], we include an explicit implementation as a baseline to computationally evaluate the advantages of further formulations.

Definition 3.1. For a given $S \geq 1$, our branching scheme $\left(\mathcal{L}^S, \mathcal{R}^S\right)$ is defined by

$$\begin{aligned} \mathcal{L}^S &:= \left\{L_1^S|_{[n]_0}, \ldots, L_S^S|_{[n]_0}\right\} \text{ and} \\ \mathcal{R}^S &:= \left\{R_1^S|_{[n]_0}, \ldots, R_S^S|_{[n]_0}\right\}, \end{aligned} \tag{3.6}$$

where

$$\begin{aligned} L_S^S &:= \left\{0, \dots, 2^{S-1} - 1\right\}, \\ R_S^S &:= \left\{2^{S-1} + 1, \dots, 2^S\right\}, \end{aligned} \tag{3.7}$$

and, for all $s \in [S-1]$,

$$\begin{aligned} L_s^S &:= \left(L_s^{S-1} \cup \left\{2^S - j : j \in L_s^{S-1}\right\}\right), \\ R_s^S &:= \left(R_s^{S-1} \cup \left\{2^S - j : j \in R_s^{S-1}\right\}\right). \end{aligned} \tag{3.8}$$

In (3.8), L_s^{S-1} and R_s^{S-1} are recursively calculated.

The following Lemma 3.2 shows that Definition 3.1 fulfills the necessary conditions.

Lemma 3.2. *Let a partition of $[\bar{x}_0, \bar{x}_n]$ into n segments with $n+1$ breakpoints $\bar{x}_0, \dots, \bar{x}_n$ be given. Further, let $\left(\mathcal{L}^S, \mathcal{R}^S\right)$ for $S := \lceil \log_2(n) \rceil$ be given as in Definition 3.1. Then, for all segments $[\bar{x}_{i-1}, \bar{x}_i]$ with $1 \leq i \leq n$, there exists a series of sets $T^i := [T_1^i, \dots, T_S^i]$ with $T_s^i \in \left\{L_s^S, R_s^S\right\}$, $s \in [S]$, such that*

$$[n]_0 \setminus \{i-1, i\} = \bigcup_{s=1}^{S} T_s^i, \tag{3.9}$$

i.e., we can represent the vertices $\bar{x}_{i-1}$ and $\bar{x}_i$ of each segment by a disjunction of $\lceil \log_2(n) \rceil$ pre-defined sets.

Proof. We can assume w.l.o.g. that n is a power of two, i.e., there is some $S \in \mathbb{N}^+$ with $S = \log_2 n$ and proof the lemma via induction in the following. We use the notation $[n]_0 := \{0, \dots, n\}$. For $n = 2$, i.e., $S = 1$, let $T_1^1 := R_1^1$ and $T_1^2 := L_1^1$. Then, we have

$$\bigcup_{s=1}^{S} T_s^i = \begin{cases} R_1^1 = \{2\} = [2]_0 \setminus \{0, 1\} & \text{if } i = 1, \\ L_1^1 = \{0\} = [2]_0 \setminus \{1, 2\} & \text{if } i = 2. \end{cases} \tag{3.10}$$

Now, let $n > 2$, i.e., $S > 1$: We assume that, for any $i \in \left\{1, \dots, 2^{S-1}\right\}$, there is a series $\bar{T}^i := \left[\bar{T}_1^i, \dots, \bar{T}_{S-1}^i\right]$ with $\bar{T}_s^i \in \left\{L_s^{S-1}, R_s^{S-1}\right\}$ such that

$$[2^{S-1}]_0 \setminus \{i-1, i\} = \bigcup_{s=1}^{S-1} \bar{T}_s^i. \tag{3.11}$$

It further holds that, for any $i \in \left\{1, \dots, 2^{S-1}\right\}$,

$$\begin{aligned} \bigcup_{s=1}^{S-1} T_s^i &= \bigcup_{s=1}^{S-1} \left(\bar{T}_s^i \cup \left\{2^S - j : j \in \bar{T}_s^i\right\}\right) \\ &= \bigcup_{s=1}^{S-1} \bar{T}_s^i \cup \bigcup_{s=1}^{S-1} \left\{2^S - j : j \in \bar{T}_s^i\right\} \\ &= \bigcup_{s=1}^{S-1} \bar{T}_s^i \cup \left\{2^S - j : j \in \bigcup_{s=1}^{S-1} \bar{T}_s^i\right\}. \end{aligned} \tag{3.12}$$

Using (3.11), we can reformulate this to

$$\begin{aligned}\bigcup_{s=1}^{S-1} T_s^i &= \left([2^{S-1}]_0 \setminus \{i-1,i\}\right) \cup \left(\left\{2^S - j : j \in [2^{S-1}]_0 \setminus \{i-1,i\}\right\}\right) \\ &= \left([2^{S-1}]_0 \setminus \{i-1,i\}\right) \cup \left(\left\{2^{S-1},\dots,2^S\right\} \setminus \left\{2^S - i, 2^S - (i-1)\right\}\right) \\ &= [2^S]_0 \setminus \left\{i-1, i, 2^S - i, 2^S - (i-1)\right\} \\ &= [n]_0 \setminus \left\{i-1, i, 2^S - i, 2^S - (i-1)\right\}.\end{aligned} \tag{3.13}$$

Now, we need to prove that, using the new sets L_S^S and R_S^S, we can find a series of sets $T^i := [T_1^i, \dots, T_S^i]$ such that (3.9) holds for all $i \in \{1,\dots,n\}$. If $i \le 2^{S-1}$, we choose $T_s^i = \bar{T}_s^i$ for $s \in \{1,\dots,S-1\}$ and $T_S^i = R_S^S$. Then,

$$\begin{aligned}\bigcup_{s=1}^{S} T_s^i = T_S^i \cup \bigcup_{s=1}^{S-1} T_s^i &= R_S^S \cup \left([n]_0 \setminus \left\{i-1, i, 2^S - i, 2^S - (i-1)\right\}\right) \\ &= \left\{2^{S-1}+1,\dots,n\right\} \cup \left([n]_0 \setminus \left\{i-1, i, 2^S - i, 2^S - (i-1)\right\}\right) \\ &= [n]_0 \setminus \{i-1, i\}.\end{aligned} \tag{3.14}$$

For $i > 2^{S-1}$, the proof works similarly with $T_s^i = L_S^S$:

$$\begin{aligned}\bigcup_{s=1}^{S} T_s^i = T_S^i \cup \bigcup_{s=1}^{S-1} T_s^i &= L_S^S \cup \left([n]_0 \setminus \left\{i-1, i, 2^S - i, 2^S - (i-1)\right\}\right) \\ &= \left\{0,\dots 2^{S-1}-1\right\} \cup \left([n]_0 \setminus \left\{i-1, i, 2^S - i, 2^S - (i-1)\right\}\right) \\ &= [n]_0 \setminus \{i-1, i\}.\end{aligned}$$

□

Using this branching scheme, (LogAg) gives an explicit logarithmic version of (Ag).

Model 4. Logarithmic branching convex combination model

$$\begin{aligned}\sum_{i=0}^{n} \lambda_i \bar{x}_i &= x, \\ \sum_{i=0}^{n} \lambda_i f(\bar{x}_i) &= z, \\ \sum_{i=0}^{n} \lambda_i &= 1, \\ \sum_{i \in L_s} \lambda_i &\le y_s, && s \in S, \\ \sum_{i \in R_s} \lambda_i &\le 1 - y_s, && s \in S, \\ \lambda_i &\ge 0, && i \in \{0,\dots,n\}, \\ y_s &\in \{0,1\}, && s \in S.\end{aligned} \tag{LogAg}$$

Additionally, two new reformulations are provided by [20]. For these reformulations, one needs the encoding C^r that is defined recursively as

$$\begin{aligned} C^1 &= \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\ C^{k+1} &= \begin{pmatrix} C^k & 0^d \\ C^k + 1^d C_d^{k\,T} & 1^d \end{pmatrix} \text{ for } k = 1, \dots, r-1, \end{aligned} \tag{3.15}$$

where $d = 2^k$, $C_d^k \in \mathbb{R}^k$ is the d-th row of C^k and $0^d, 1^d \in \mathbb{R}^d$ contain only zeroes or ones, respectively. Further, for the sake of simplicity, they set $C_0^k = C_1^k$ and $C_{d+1}^k = C_d^k$. This encoding helps to provide the following equations:

$$\sum_{v=0}^{n} C_{v,k}^r \lambda_v \le y_k + \sum_{l=k+1}^{r} 2^{l-k-1} y_l \le \sum_{v=0}^{n} C_{v+1,k}^r \lambda_v, \tag{3.16}$$

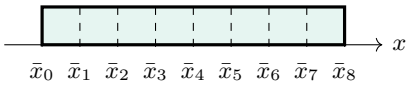
$$\sum_{v=0}^{n} C_{v,k}^r \lambda_v \le y_k \le \sum_{v=0}^{n} C_{v+1,k}^r \lambda_v. \tag{3.17}$$

Using each of these two equations, one can then define two new models, a binary formulation in (BinZigZag) and an integer one in (IntZigZag). They are called Zig-Zag formulations as that describes the behavior of the single vectors. A proof for the correctness of these equations is given by [20].

Model 5. Binary Zig-Zag Formulation

$$\begin{aligned} \sum_{i=0}^{n} \lambda_i \bar{x}_i &= x, \\ \sum_{i=0}^{n} \lambda_i f(\bar{x}_i) &= z, \\ \sum_{i=0}^{n} \lambda_i &= 1, \\ \sum_{v=0}^{n} C_{v,k}^r \lambda_v &\le y_k + \sum_{l=k+1}^{r} 2^{l-k-1} y_l, && k \in \{1, \dots, r\} \\ y_k + \sum_{l=k+1}^{r} 2^{l-k-1} y_l &\le \sum_{v=0}^{n} C_{v+1,k}^r \lambda_v, && k \in \{1, \dots, r\} \\ \lambda_i &\ge 0, && i \in \{0, \dots, n\}, \\ y_k &\in \{0, 1\}, && k \in \{1, \dots, r\}, \end{aligned} \tag{BinZigZag}$$

with $r = \lceil \log_2 (n-1) \rceil$.



(3.3A) Initial segments

Model 6. General Integer Zig-Zag Formulation

$$
\begin{aligned}
\sum_{i=0}^{n} \lambda_i \bar{x}_i &= x, \\
\sum_{i=0}^{n} \lambda_i f(\bar{x}_i) &= z, \\
\sum_{i=0}^{n} \lambda_i &= 1, \\
\sum_{v=0}^{n} C^r_{v,k} \lambda_v &\leq y_k, && k \in \{1, \dots, r\}, \\
y_k &\leq \sum_{v=0}^{n} C^r_{v+1,k} \lambda_v, && k \in \{1, \dots, r\}, \\
\lambda_i &\geq 0, && i \in \{0, \dots, n\}, \\
y_k &\in \mathbb{Z}, && k \in \{1, \dots, r\},
\end{aligned}
\tag{IntZigZag}
$$

with $r = \lceil \log_2 (n-1) \rceil$.

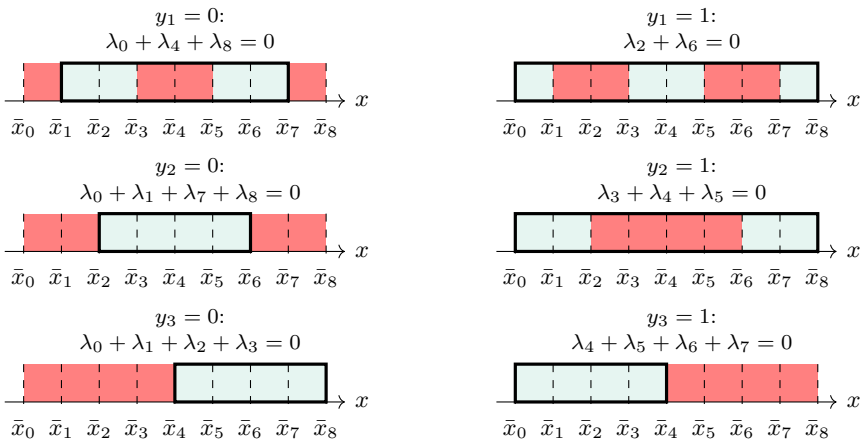
(LogAg) and the Zig-Zag methods (BinZigZag) and (IntZigZag) have different branching behaviors.

For an efficient solving of MILPs, it is important to quickly find tight dual bounds that are obtained by solving LP relaxations. As these models are all LP relaxations, they contains all feasible points, or, better said, the convex hull of all feasible points. If the LP relaxations are tight, we can find better dual bounds. In [20], the authors mention different metrics to evaluate the branching behavior. Besides that, we illustratively investigate this using an example with $n = 8$ uniformly sized segments in the following.

In Figure 3.3a, one can see the breakpoints $\bar{x}_0, \bar{x}_1, \dots, \bar{x}_8$. These breakpoints create eight segments that are, initially, all feasible, as there was no branching up to now. Feasible segments are displayed in light blue. Further, the convex hull of the feasible set is outlined by a thick rectangle. Here, the convex hull is equivalent to the full domain.

In Figure 3.3b, all possible branchings for (LogAg) are shown. As we have eight segments, there are 3 binary variables y_1, y_2 and y_3 that can be either zero or one, resulting in six different branches. All possibilities are displayed where the red parts mean that segments are excluded due to the chosen branch, blue segments are still feasible. There are three cases, where the convex hull also contains infeasible regions as they lie between feasible ones. Thus, a solution of the LP relaxation can also lie inside these infeasible regions. The ZigZag reformulations try to overcome this problem.

We visualize the branching for (IntZigZag) in Figure 3.3c, the behavior of (BinZigZag) is similar. In theory, there are more possible branching steps



(3.3B) Branching for (LogAg)

than the displayed ones, because in this reformulation the variables y_1, y_2 and y_3 are integral instead of binary. However, all other branchings would result in a combination of one infeasible branch and one branch that still contains the full domain, and, therefore, no progress at all. As one can see, all branchings lead to subdomains that are connected what means that their convex hulls are exactly the feasible set. Thus, it is not possible to have LP solutions that lie outside the feasible regions what should increase their quality and, thus, the dual bounds.

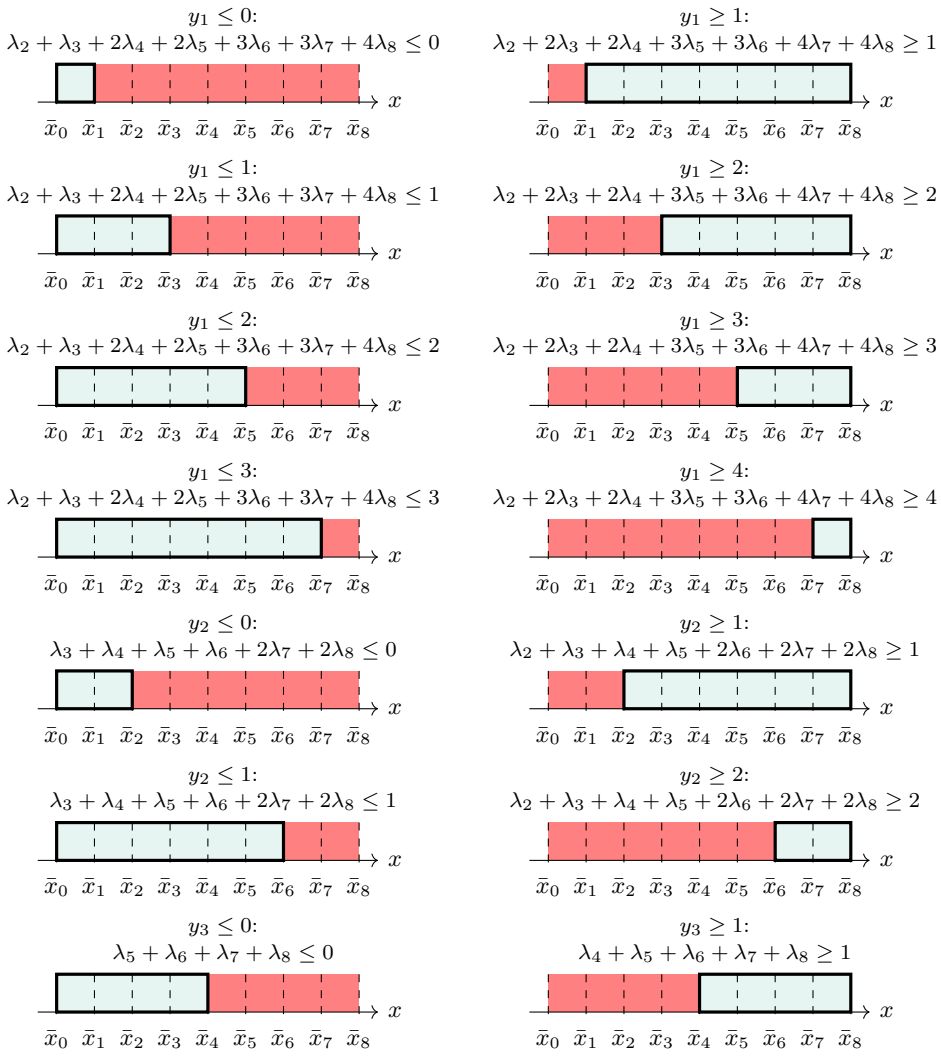
The following two equations show the inequalities we used to create the visualizations in Figures 3.3b and 3.3c. The continuous variables in (LogAg) are constrained by

$$\lambda_0 + \lambda_1 + \lambda_2 + \lambda_3 \leq y_3, \tag{3.18a}$$
$$\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 \leq 1 - y_3, \tag{3.18b}$$
$$\lambda_0 + \lambda_1 + \lambda_7 + \lambda_8 \leq y_2, \tag{3.18c}$$
$$\lambda_3 + \lambda_4 + \lambda_5 \leq 1 - y_2, \tag{3.18d}$$
$$\lambda_0 + \lambda_4 + \lambda_8 \leq y_1, \tag{3.18e}$$
$$\lambda_2 + \lambda_6 \leq 1 - y_1. \tag{3.18f}$$

When a binary variable is fixed, there is a subset of continuous variables that also need to be zero then. For example, when setting $y_1 = 0$, we obtain $\lambda_0 + \lambda_4 + \lambda_8 = 0$, i.e., $\lambda_0 = \lambda_4 = \lambda_8 = 0$ from (3.18e) and, thus, the feasible regions shown in the left upper plot of Figure 3.3b. For $y_1 = 1$, we, complementary, obtain $\lambda_2 = \lambda_6 = 0$ from (3.18f) what is visualized in the right upper plot.

In contrast, the inequalities in (IntZigZag) and (BinZigZag) have the form

$$\begin{aligned}&\lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5 + 3\lambda_6 + 3\lambda_7 + 4\lambda_8\\ &\leq y_1 + y_2 + 2y_3\\ &\leq \lambda_1 + \lambda_2 + 2\lambda_3 + 2\lambda_4 + 3\lambda_5 + 3\lambda_6 + 4\lambda_7 + 4\lambda_8,\end{aligned} \tag{3.19a}$$



(3.3c) Branching for (IntZigZag)

FIGURE 3.3. All possible branching steps for $n = 8$ segments. Branching results in segments being feasible or infeasible, what is represented by shading: The light blue areas represent feasible regions, while the red areas represent the infeasible segments. The bold rectangle outlines the convex hull.

$$\begin{aligned}
&\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + 2\lambda_7 + 2\lambda_8 \\
&\quad \le y_2 + y_3 \\
&\quad \le \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + 2\lambda_6 + 2\lambda_7 + 2\lambda_8, && \text{(3.19b)} \\
&\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 \\
&\quad \le y_3 \\
&\quad \le \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8, && \text{(3.19c)}
\end{aligned}$$

whereas the blue parts are the addends that are only present in (BinZigZag). It directly stands out that all inequalities only use adjacent variables.

Let us, for example, assume that we branch on y_1, using the branches $y_1 \leq 2$ and $y_1 \geq 3$, given in the third line of Figure 3.3c. From (3.19a), we obtain

$$\lambda_1 + \lambda_2 + 2\lambda_3 + 2\lambda_4 + 3\lambda_5 + 3\lambda_6 + 4\lambda_7 + 4\lambda_8 \leq 2 \tag{3.20}$$

for the left branch with $y_1 \leq 2$. Using one of the variables λ_6, λ_7, or λ_8 would result in a sum greater than two, as the sum of exactly two neighboring variables λ_i and λ_{i+1} has to be one. Thus, we cannot use the three rightmost segments. In contrast,

$$3 \leq \lambda_2 + \lambda_3 + 2\lambda_4 + 2\lambda_5 + 3\lambda_6 + 3\lambda_7 + 4\lambda_8 \tag{3.21}$$

emerges for $y_1 \geq 3$. Here, using any variable of $\lambda_0, \lambda_1, \ldots, \lambda_5$ would yield a sum less than three with the same argument as before. Thus, we can use exactly the three rightmost segments that were cut off in the other branch.

3.3. **Incremental model.** The incremental method was introduced by [27]. Similar to convex combination models, this formulation also employs $n-1$ binary variables $y_1, \ldots, y_{n-1}$ to represent the active segments. However, they are not one-hot encoded in this method, but all variables y_j with $j \leq i$ have a value of $y_j = 1$ if segment i is active. Further, if segment i is used, we have $y_j = 0$ for all $j > i$. The advantage of setting all leftmost binary variables to 1 is again to solve the branching issues described for (LogAg). In this formulation, there are n continuous variables $\delta_1, \ldots, \delta_n$. They are used similarly to the binary ones, i.e., $\delta_j = 1$ for all $j < i$ and $\delta_j = 0$ for all $j > i$. Moreover, the variable δ_i determines the exact position of x that is calculated by

$$x = \bar{x}_0 + \sum_{i=1}^{n} \delta_i (\bar{x}_i - \bar{x}_{i-1}). \tag{3.22}$$

The evaluation of $f(x)$ works equivalently. The following constraint is introduced to ensure that the continuous variables behave as expected:

$$0 \leq \delta_n \leq y_{n-1} \leq \delta_{n-1} \leq \cdots \leq y_1 \leq \delta_1 \leq 1. \tag{3.23}$$

Due to the variable names, the incremental method is also known as the δ-method. The convex combination formulations, on the other hand, are frequently referred to as the λ-methods. Again, one can find the model in (Inc) and a visual representation in Figure 3.4.

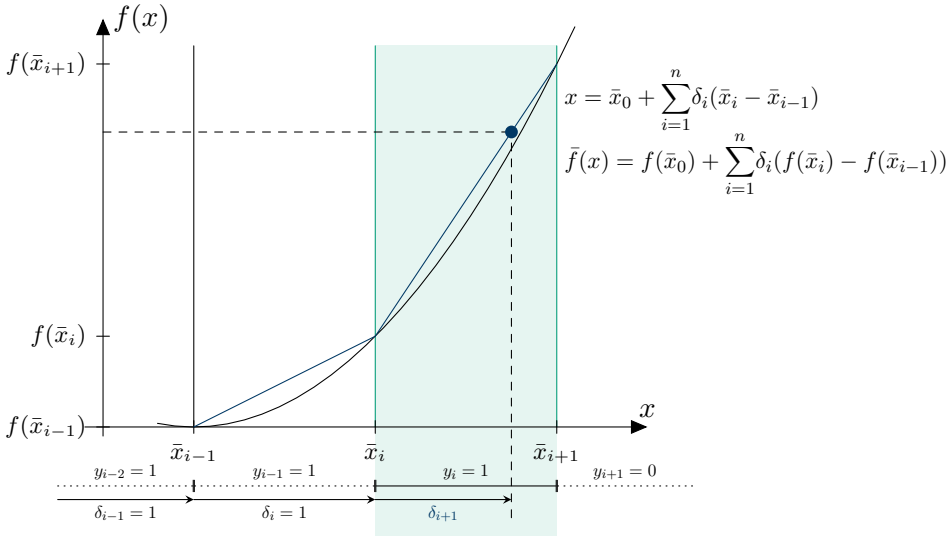


FIGURE 3.4. Representation of (Inc). The active segment is shaded in green, thus $y_j = 1$ for $j \leq i$ and $y_j = 0$ for $j > i$. The continuous variables δ_j represent the filled fraction of each segment, defining $(x, \bar{f}(x))$ as the cumulative sum of all segments up to the active one.

Model 7. Classical Incremental Method

$$\bar{x}_0 + \sum_{i=1}^{n} \delta_i(\bar{x}_i - \bar{x}_{i-1}) = x, \tag{Inc}$$

$$f(\bar{x}_0) + \sum_{i=1}^{n} \delta_i(f(\bar{x}_i) - f(\bar{x}_{i-1})) = z,$$

$$\begin{aligned} \delta_1 &\leq 1, \\ \delta_{i+1} &\leq y_i, && i \in \{1, \dots, n-1\}, \\ y_i &\leq \delta_i, && i \in \{1, \dots, n-1\}, \\ \delta_n &\geq 0, \\ y_i &\in \{0, 1\}, && i \in \{1, \dots, n-1\}. \end{aligned}$$

3.4. **Multiple choice model.** Another reformulation that employs a linear number of binary variables is the multiple choice model. As in the convex combination models, each variable y_i represents one segment i, and the variables are one-hot encoded. The main idea here is that there is also exactly one nonzero continuous variable. For each segment i, there is a variable x_i that represents the exact value on the x-axis. Every variable x_i yields one constraint of the form

$$y_i \bar{x}_{i-1} \leq x_i \leq y_i \bar{x}_i. \tag{3.24}$$

If we use segment i, we have $\bar{x}_{i-1} \leq x_i \leq \bar{x}_i$, and, thus, x_i is forced to lie exactly in the segment; otherwise, x_i is fixed to zero. Given the values of all variables x_i and parameters m_i and t_i that are chosen such that each linear

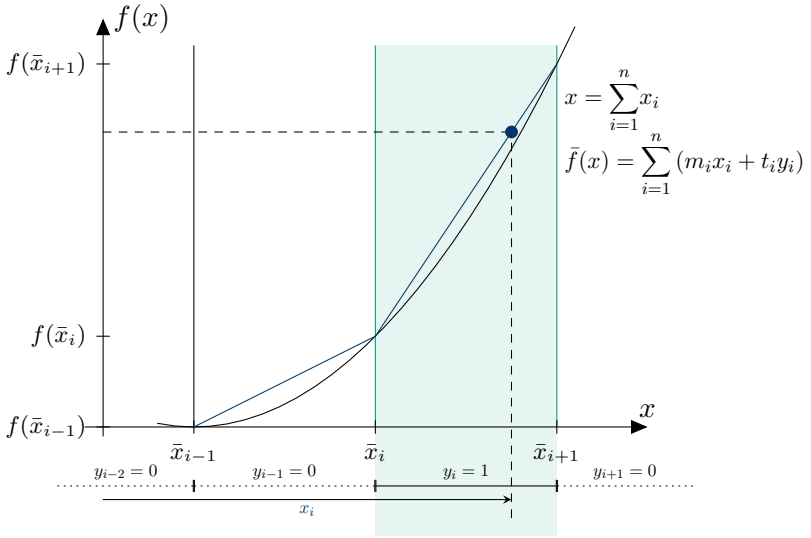


FIGURE 3.5. Visual representation of (MC). The active segment ($y_i = 1$) is shaded in green. Here, the continuous variables x_i represent the coordinate x restricted to the i-th segment, and the value $\bar{f}(x)$ is computed using the local slope m_i and intercept t_i of the active segment.

equation $m_i x + t_i$ connects $(\bar{x}_i, \bar{f}(\bar{x}_i))$ and $(\bar{x}_{i+1}, \bar{f}(\bar{x}_{i+1}))$, one can calculate the approximation value using the sum of the linear equations

$$\sum_{i=1}^{n} (m_i x_i + t_i y_i)\,. \tag{3.25}$$

Putting everything together, one obtains (MC) and the visual representation in Figure 3.5.

Model 8. Multiple Choice Model

$$\begin{aligned}
\sum_{i=1}^{n} x_i &= x, \\
\sum_{i=1}^{n} (m_i x_i + t_i y_i) &= z, \\
y_i \bar{x}_{i-1} \le x_i, &\qquad i \in \{1, \dots, n\}, \\
x_i \le y_i \bar{x}_i, &\qquad i \in \{1, \dots, n\}, \\
\sum_{i=1}^{n} y_i &= 1, \\
y_i \in \{0, 1\}, &\qquad i \in \{1, \dots, n\}.
\end{aligned} \tag{MC}$$

3.5. **Theoretical Comparison.** Finally, we provide an overview of the number of variables and constraints that are needed to model the MILP representations of this section in Table 3.1. For relaxations, another variable for each segment is introduced, i.e., n additional variables. This variable

TABLE 3.1. Sizes of all one-dimensional representations, assuming n segments and $n+1$ breakpoints.

Model	Constraints	Variables		
		Continuous	Binary	Integer
(Disag)	$n+3$	$2n$	n	0
(LogDisag)	$2\lceil \log_2 n \rceil + 3$	$2n$	$\lceil \log_2 n \rceil$	0
(Ag)	$n+5$	$n+1$	n	0
(LogAg)	$2\lceil \log_2 n \rceil + 3$	$n+1$	$\lceil \log_2 n \rceil$	0
(Inc)	$2n+1$	n	$n-1$	0
(MC)	$2n+3$	n	n	0
(BinZigZag)	$2\lceil \log_2(n-1) \rceil + 3$	$n+1$	$\lceil \log_2(n-1) \rceil$	0
(IntZigZag)	$2\lceil \log_2(n-1) \rceil + 3$	$n+1$	0	$\lceil \log_2(n-1) \rceil$

works as a bound for the maximum error ϵ and, thus, is the same for each representation.

We provide an overview of the number of variables and constraints that are needed to model the MILP representations of this section in Table 3.1. For relaxations, another variable for each segment is introduced, i.e., n additional variables. This variable works as a bound for the maximum error ϵ and, thus, is the same for each representation.

Apart from the size of the formulations, their computational performance is strongly related to their theoretical properties and branching behavior. According to [43], standard PWL encodings can have unbalanced branch-and-bound trees, which can lead to long runtimes. (Inc) tackles this issue by inherently functioning as a balanced branching scheme, which has a structural advantage. This advantage also extends to (BinZigZag) and (IntZigZag). Moreover, [42] investigates the models regarding local idealness and sharpness. Interestingly, while logarithmic formulations typically lack local ideality, their (and also our) findings show that they are still highly competitive in practice. The exponential reduction in the number of binary variables effectively compensates for this theoretical limitation, resulting in good computational performance.

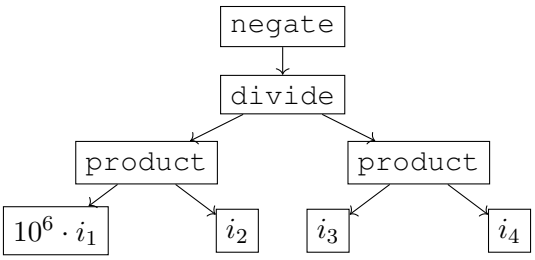
4. IMPLEMENTATION

After introducing the different possibilities to represent MILP relaxations of MINLPs, we go more into detail about how we implemented our framework and how we conducted the computational study. First, we describe the input instances before heading over to the reformulation and solving process.

4.1. **Input.** All input problems have to be stored in the Optimization Services Instance Language, short OSIL, format. The MINLPLIB contains all instances in this format. For other applications, in case, that a problem is not available in this format, our framework provides a conversion tool for `.nl` files. The OSIL format employs an XML vocabulary, which provides several advantages, the most notable of which is that nonlinearities are stored

```
1   <variables numberOfVariables="6">
2    <var name="i1" type="I" lb="12" ub="60"/>
3    [...]
4   </variables>
5   [...]
6   <nonlinearExpressions numberOfNonlinearExpressions="1">
7     <nl idx="0">
8      <negate>
9       <divide>
10       <product>
11        <variable idx="0" coef="1e6"/>
12        <variable idx="1"/>
13       </product>
14       <product>
15        <variable idx="2"/>
16        <variable idx="3"/>
17       </product>
18      </divide>
19     </negate>
20    </nl>
21  </nonlinearExpressions>
```

(4.1A) Parts of the OSIL code from instance `gear4.osil`.



(4.1B) Expression tree for the nonlinear term

FIGURE 4.1. Representations for the expression $-\frac{10^6 \cdot i_1 \cdot i_2}{i_3 \cdot i_4}$. This expression is a part of an instance from the MINLPLIB.

in a tree-based structure. In an OSIL file, each variable, constraint, etc. is saved along with various attributes such as coefficients, a name, or an index. These indices are then used as a reference, for instance to determine where a nonlinearity is used.

A nonlinearity is stored as an expression tree that includes tags for functions like power, products, and sums. An example for the nested expression $-\frac{10^6 \cdot i_1 \cdot i_2}{i_3 \cdot i_4}$ is given in Figure 4.1a, the corresponding expression tree is given in Figure 4.1b. For variable i_1 with index 0, there is an additional parameter, its coefficient of 10^6. The nonlinear term is used in the constraint with index 0, what is stated in line 7. In line 2, one can see how properties of the variables are stored: In the provided example, variable i_1 is an integer variable with $12 \leq i_1 \leq 60$. The index of 0 is implicitly given by the order in which the variables are defined. In line 11, variable i_1 is then used in the

nonlinear expression. It would have also been possible to store parts of this nonlinearity as quadratic equations: Therin, only two variable indices are stored together with a coefficient. More information about the OSIL format is provided by [14].

For our benchmarks, we use a subset of the problems provided in the MINLPLIB [9] that can be represented by the following nonlinear functions:

$$\begin{aligned} &f(x_1, x_2) = x_1 \cdot x_2,\ f(x_1, x_2) = \frac{x_1}{x_2},\ f(x) = \ln x,\ f(x) = \log_{10} x, \\ &f(x) = x^a \text{ (including } x^{-1} \text{ and } \sqrt{x}),\ f(x) = a^x \text{ (including } e^x), \\ &f(x) = \sin x,\ f(x) = \cos x,\ f(x) = \tanh x,\ f(x) = |x|. \end{aligned} \tag{2.1}$$

Furthermore, all variables must have both, lower and upper bounds; otherwise, a piecewise linear approximation or relaxation cannot be established. Using this selection criterion, only 306 instances (of 1615) remain in the benchmark set. To expand the set, we first apply presolving to the instances before selection. Presolving can establish missing bounds, allowing us to increase the benchmark set to 773 instances. Presolving is performed using SCIP [6]. In our numerical study in Section 5, we provide additional statistics about our benchmark instances.

4.2. **Reformulation.** Each input problem is now reformulated using a self-defined data format, first to a MINLP containing only one-dimensional nonlinearities and nonlinearities of the form $z = x_1 x_2$. Subsequently, the univariate transformation of bivariate products from Section 2 is performed. The resulting MINLP then serves as the basis for the various MILP relaxations.

Our self-defined data format for storing MINLPs is called `OSILData` and contains data structures for variables, objective(s), constraints, linear, quadratic, and nonlinear expressions. Each object of `OSILData` can later be solved in the same way, independently of the type of nonlinearities, what makes the results easily comparable. All nonlinear expressions are stored in a tree structure that can be easily modified. This structure is based on the expression trees of the original OSIL format. At the start of each solving process, an object of `OSILData` containing all given information is created from the initial MINLP.

4.2.1. *Creating one-dimensional nonlinearities.* The first reformulation step is to remove all higher-dimensional nonlinearities. Therefore, we investigate all nonlinearities recursively and reformulate each type one by one. We initially use two-dimensional multiplications, which we reformulate at the end. The first higher-dimensional nonlinearities that we want to remove are products with more than two factors, i.e., expressions of the form $\texttt{product}(x_1, x_2, \ldots, x_{n-1}, x_n)$. This is reformulated such that each product has two factors, resulting in the form $\texttt{product}(x_1, \texttt{product}(x_2, \texttt{product}(\ldots, \texttt{product}(x_{n-1}, x_n))))$. Next, each division of the form $z = \frac{x_1}{x_2}$ is reformulated by $z = x_1 \cdot \frac{1}{x_2}$, which is again a product of one-dimensional nonlinearities. Then, products of the form $x_1 x_2$ are reformulated using Equations (2.2) to (2.4).

After dealing with these special cases, we need to consider nested nonlinear equations with multiple variables. The OSIL expression tree format is

useful in this regard: We go through the tree recursively, replacing each nonlinearity $\mathrm{nl}_i(x)$ with a newly introduced variable z_i. In addition, we insert a new constraint $z_i = \mathrm{nl}_i(x)$, which creates a new expression tree. As we begin to replace the leafs, each newly created tree has a depth of at most one. Until now, the optimal solution for the resulting MINLP has not changed.

4.2.2. *Reformulation from one-dimensional MINLPs to MILP relaxations.* We now reformulate the MINLP based on the one-dimensional formulation using the MILP representations described in Section 3. Our implementation allows one to choose between relaxations and approximations; for the computational study conducted in this work we only consider piecewise linear relaxations. We restrict our study to relaxations as they provide valid bounds for the original MINLP. While approximations could also be used, we expect the performance of the MILP formulations to remain the same, as the underlying combinatorial structure is identical. Any relaxation has a larger set of feasible solutions than the initial MINLP; its size is determined by the error bound used.

The following approach is essentially the same for each PWL model. Since we reformulated all nonlinearities to one-dimensional nonlinear functions, each expression has the form $z = f(x)$ with a bounded variable x, i.e., $x^- \leq x \leq x^+$. Therefore, we can create a piecewise linear approximation function $\bar{f} : [x^-, x^+] \to \mathbb{R}$ for each nonlinear expression and extend it to a relaxation afterwards. We use an adjustable value ϵ that bounds the maximum error in each segment, i.e., we enforce

$$\left|\bar{f}(x) - f(x)\right| \leq \epsilon \tag{4.1}$$

for all x with $x^- \leq x \leq x^+$. The size of ϵ controls the number of breakpoints and, thus, segments we use. A smaller error bound leads to an increasing number of breakpoints and vice versa. While relative error tolerances could be implemented by scaling ϵ based on function values, we employ absolute errors to facilitate a more straightforward implementation and comparison of the different MILP models.

The approximation errors apply to each PWL function separately. When multiple approximations are used in the model, the cumulative effect can influence the final solution quality. In the worst case, the total error could scale with the number of PWL functions n, i.e., $\mathcal{O}(n\epsilon)$. However, optimization effects could reduce this. Since all compared formulations implement the same discretization, they are mathematically equivalent in terms of relaxation tightness and feasibility; thus, our comparison focuses strictly on their computational performance. Later in Section 5.2.4, we will look into the solution quality, i.e, the practical impact of the errors ϵ.

Algorithm 4.1 Breakpoint Generation

Input: Variable x with bounds $x^- \leq x \leq x^+$, function $f : [x^-, x^+] \to \mathbb{R}$ and error bound $\epsilon \in \mathbb{R}^+$
Output: Set of breakpoints $\bar{X}$

function GENERATEBREAKPOINTS(x, f, ϵ)
  Initialize $\bar{X} \leftarrow \{x^-\}$ ▷ Add leftmost breakpoint x^-
  $\bar{x}_{\text{cur}} \leftarrow x^-$
  while $x^+ - \bar{x}_{\text{cur}} > 10^{-6}$ **do**
    $l, u \leftarrow \bar{x}_{\text{cur}}, x^+$
    loop ▷ Binary Search
      $\epsilon_{\max} \leftarrow$ CREATEAPPROX($f, \bar{x}_{\text{cur}}, \frac{u+l}{2}$)
      if $u - l \leq 10^{-6}$ **then**
        $\bar{x}_{\text{cur}} \leftarrow \frac{u+l}{2}$
        $\bar{X} \leftarrow \bar{X} \cup \{\bar{x}_{\text{cur}}\}$ ▷ Next breakpoint is found
        break
      else if $\epsilon_{\max} \leq \epsilon$ **then**
        $l \leftarrow \frac{u+l}{2}$ ▷ Proceed on $[l, \frac{u+l}{2}]$
      else
        $u \leftarrow \frac{u+l}{2}$ ▷ Proceed on $[\frac{u+l}{2}, u]$
      end if
    end loop
  end while
  $\bar{X} \leftarrow \bar{X} \cup \{x^+\}$ ▷ Add rightmost breakpoint x^+
end function

function CREATEAPPROX(f, x_l, x_u) ▷ Connect endpoints linearly
  $m \leftarrow \frac{f(x_u) - f(x_l)}{x_u - x_l}$
  $t \leftarrow f(x_l) - m x_l$
  $\epsilon_{\max} \leftarrow \max\{|(mx + t) - f(x)| : x_l \leq x \leq x_u\}$
  return $\epsilon_{\max}$
end function

The set of breakpoints $\bar{X} = \{\bar{x}_0, \bar{x}_1, \ldots, \bar{x}_n\}$ is generated greedily as follows: Starting at the lower bound $\bar{x}_0 = x^-$, each subsequent breakpoint $\bar{x}_{i+1}$ is determined as the largest value in $(\bar{x}_i, x^+]$ such that it fulfills (4.1) for all $x \in [\bar{x}_i, \bar{x}_{i+1}]$. This point is found using a binary search, as shown in Algorithm 4.1. The process is repeated until the rightmost breakpoint x^+ is reached.

Finally, using all linear approximations

$$\bar{f}_i(x) = m_i x + t_i \text{ for } x \in [\bar{x}_{i-1}, \bar{x}_i], \tag{4.2}$$

we obtain the PWL approximation $\bar{f}(x)$. We enforce continuous piecewise linear approximations by interpolation, i.e., we have $m_i \bar{x}_{i-1} + t_i = f(\bar{x}_{i-1})$ and $m_i \bar{x}_i + t_i = f(\bar{x}_i)$ for $i = 1, \ldots, n$.

To create relaxations, we just add another variable, bounded by $-\epsilon \leq e \leq \epsilon$ that controls the maximum error and add it to the piecewise linear function $\bar{f}(x)$, i.e., we replace each nonlinearity by $\bar{f}(x) + e$.

In our implementation, one can choose which representation from Section 3 to use for the relaxation. Then, the respective constraints replace the occurring nonlinearities in the input problem.

4.3. **Solving.** Finally, we have a MILP relaxation for each different representation that can be solved using state-of-the-art MILP solution methods. To accomplish this, we first convert the MILPs from our own data structure to Pyomo models, cf. [10, 18]. We use Pyomo because it allows us to easily switch between different solvers. In our study, we use Gurobi Optimizer version 11.0.3 ([17]) to solve the MILPs using a time limit of four hours. Each problem is optimized on the NHR@FAU clusters using Intel Xeon Gold 6326 CPUs with four cores, a total of 32 GB RAM, and a base frequency of 2.9 GHz.

We store different information in each run. On the one hand, we examine the optimization result, which tells us whether a problem was solved to optimality, reached its time limit, or ran into an error. Infeasibility can only occur when we use approximations instead of relaxations. On the other hand, we are interested to understand more about how problems are solved: We store the primal and dual bounds on a regular basis and therefore can monitor the gap over time. Furthermore, we store the time it took until the first feasible solution was found. In the following Section 5, we present the numerical results and discuss them afterwards in Section 6.

5. Numerical results

In this section, we first present the benchmark set from the MINLPLIB and then numerically analyze how the different MILP models for PWL relaxations from Section 3 perform in practice.

5.1. **Benchmark instances.** As mentioned previously, our benchmark set is formed from a subset of the MINLPLIB and consists of 773 different instances. For 607 instances, at least one reformulation can build a model. These instances are listed in our benchmark set in Table E.1. In Figure 5.1, one can see the numbers of constraints and variables for each instance. On top, it shows the entire benchmark set and, on the bottom, all instances with less than 1000 constraints and variables. In both plots, each blue square represents the total number of constraints and variables in one benchmark instance. We have around 415 variables and 647 constraints on average per instance. The median numbers are 119 and 144, respectively.

On the left hand side, we further plotted red circles that consider only the nonlinear constraints. Similarly, on the right hand side, the green circles show only the number of binary and integer variables. There are, on average, around 68 binary/integer variables, as well as 147 nonlinear constraints per instance. The median number of binary/integer variables is 20, the median number of nonlinear constraints is 21.

Table 5.1 contains statistics about the reformulated benchmark instances. To make a conclusive comparison, we consider the set of instances for which a model could be created for all error bounds. The number of segments per instance and nonlinearity, as well as the segment length per nonlinearity, are calculated for each error bound. The mean and median number and length

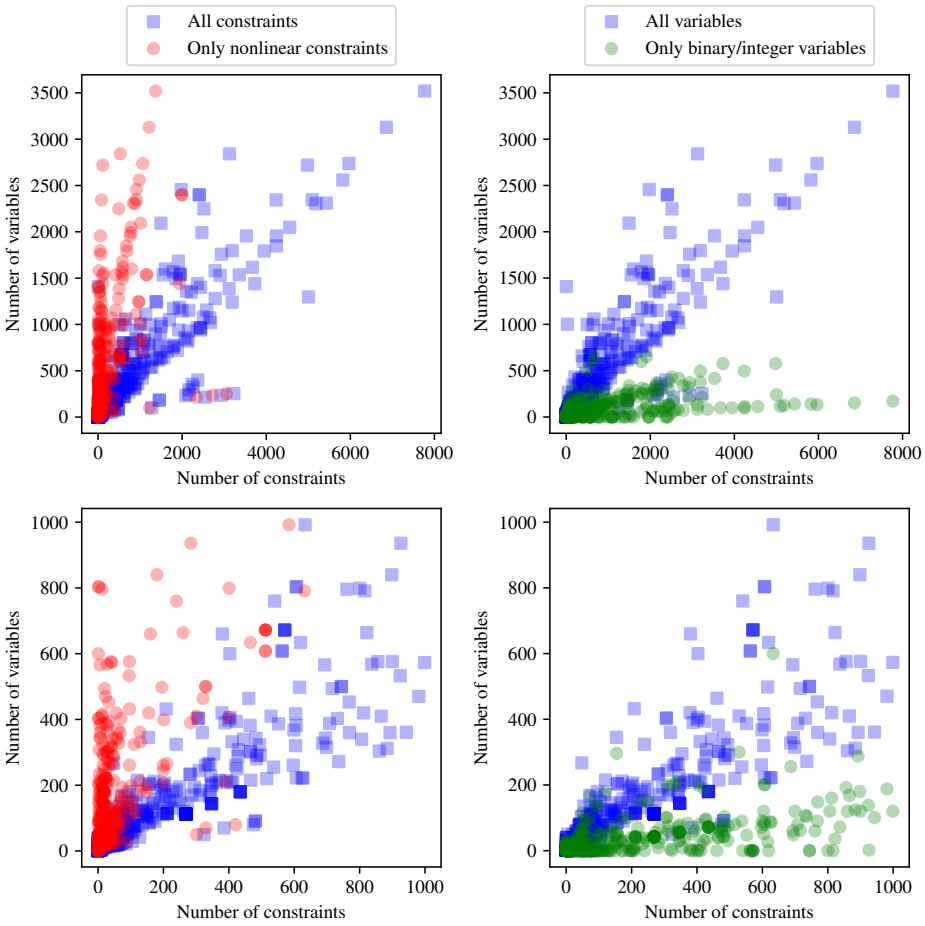


FIGURE 5.1. Total amount of variables and constraints for our benchmark instances. The blue squares represent - for comparison purposes - the total number of constraints and variables, whereas the red and green circles represent nonlinear constraint and binary/integer variable subsets, respectively.

of segments per nonlinearity are calculated for the entire set of nonlinearities instead of averaging over each instance separately.

We separate our benchmark set in four different ways:

(1) Difficulty: We solve all presolved problems using SCIP and Gurobi. Depending on the outcome, we classify each instance:
   Very easy: Both solvers solve the instance to optimality within a time limit of 60 seconds.
   Easy: Exactly one solver solves the instance to optimality within a time limit of 60 seconds.
   Medium: Neither solver solves the instance to optimality, but at least one finds a primal feasible solution within 60 seconds.
   Hard: All remaining problems (no feasible solution found).

(2) Type of nonlinearities: We use the predefined classification of nonlinearities $h\colon \mathbb{R}^n \to \mathbb{R}$ given by the MINLPLIB. The fundamental classes

TABLE 5.1. Sizes of the MILP reformulations. The instances have between 1 and 2818 nonlinearities, with a mean of 142.9 and a median of 24.

Error Bound	$\epsilon = 10^2$			
	Min	**Max**	**Mean**	**Median**
Segments per instance	2	752	46.78	28
Segments per NL	2	14	2.06	2
Segment length per NL	$1.36 \cdot 10^{-6}$	50000.00	13.53	1.50
Error Bound	$\epsilon = 10^0$			
	Min	**Max**	**Mean**	**Median**
Segments per instance	2	752	79.67	32
Segments per NL	2	126	3.52	2
Segment length per NL	$1.36 \cdot 10^{-6}$	9090.91	4.11	1.25
Error Bound	$\epsilon = 10^{-2}$			
	Min	**Max**	**Mean**	**Median**
Segments per instance	4	3856	480.02	123
Segments per NL	2	1249	21.19	7
Segment length per NL	$1.36 \cdot 10^{-6}$	1052.63	0.81	0.19
Error Bound	$\epsilon = 10^{-4}$			
	Min	**Max**	**Mean**	**Median**
Segments per instance	4	37386	4529.54	1027
Segments per NL	2	12474	199.96	53
Segment length per NL	$1.36 \cdot 10^{-6}$	117.51	0.12	0.02
Error Bound	$\epsilon = 10^{-6}$			
	Min	**Max**	**Mean**	**Median**
Segments per instance	4	373536	44959.64	10155
Segments per NL	2	124758	1984.82	521
Segment length per NL	$1.36 \cdot 10^{-6}$	13.01	0.02	$2.24 \cdot 10^{-3}$

are defined as:

$$\begin{aligned} \text{quadratic}(h) := (&\exists Q \in \mathbb{R}^{n\times n}, q \in \mathbb{R}^n, q_0 \in \mathbb{R} : \\ &\forall x \in \mathbb{R}^n : h(x) = q_0 + q^\top x + x^\top Q x) \end{aligned} \tag{5.1}$$

$$\begin{aligned} \text{polynomial}(h) := (&\exists K, |K| < \infty : \exists a_k \in \mathbb{N}^n_{\geq 0}, b_k \in \mathbb{R} : \\ &\forall x \in \mathbb{R}^n : h(x) = \sum_{k\in K} b_k \prod_{i\in\mathcal{N}} x_i^{a_{k,i}}) \end{aligned} \tag{5.2}$$

$$\text{signomial}(h) := (\exists K, |K| < \infty : \exists a_k \in \mathbb{R}^n, b_k \in \mathbb{R} :$$

$$\forall x \in \mathbb{R}^n : h(x) = \sum_{k \in K} b_k \prod_{i \in \mathcal{N}} x_i^{a_{k,i}}) \tag{5.3}$$

where $\mathcal{N}$ is the set of variable indices. Based on these definitions, we categorize the instances by their most complex constraint:

Quadratic: All nonlinear constraints are quadratic.

Polynomial: At least one constraint is strictly polynomial, and no constraint is signomial or general.

Signomial: At least one constraint is strictly signomial, and no constraint is general.

General: All remaining instances (e.g., involving exponential or trigonometric functions).

(3) Convexity: We rely on the pre-given information provided by the MINLPLIB, which refers to the properties of the continuous relaxation:

Convex: The continuous relaxation of the instance is known to be a convex optimization problem.

Non-Convex: The continuous relaxation is known to be a non-convex optimization problem.

Unknown: The MINLPLIB does not provide a definitive classification regarding the convexity of the instance.

(4) Variables: Depending on the domain of the variables, we classify the instances as:

Continuous: All variables are continuous.

Mixed-Binary: The instance contains binary variables, but no general integer variables.

Mixed-Integer: The instance contains general integer variables.

A full list of the benchmark set with details and their classifications is provided in Appendix E.

5.2. **Comparison of the MILP models for PWL relaxations.** We now present the computational results where we test all PWL MILP models from Section 3 on the previously explained benchmark set. First, we evaluate how many problems can be solved to optimality, and how long the solving process takes. Afterwards, we evaluate the solution quality over time.

For many results, we use the so-called shifted geometric mean (SGM). The SGM of n numbers $t_1, \ldots, t_n$ is determined using the formula

$$\sqrt[n]{\prod_{i=1}^{n} (t_i + s)} - s. \tag{5.4}$$

Here, s represents an arbitrary shift applied to each term. This shift factor introduces a level of flexibility into the calculation, allowing us to adjust the significance of each term in the dataset. In our case, we want to decrease the impact of small runtimes. Following standard benchmarking practices (e.g. [31]), this prevents non-solver noise, such as I/O operations and initialization overhead, from distorting the comparative results. To improve the numerical stability of the computation, we employ an alternative formulation, namely

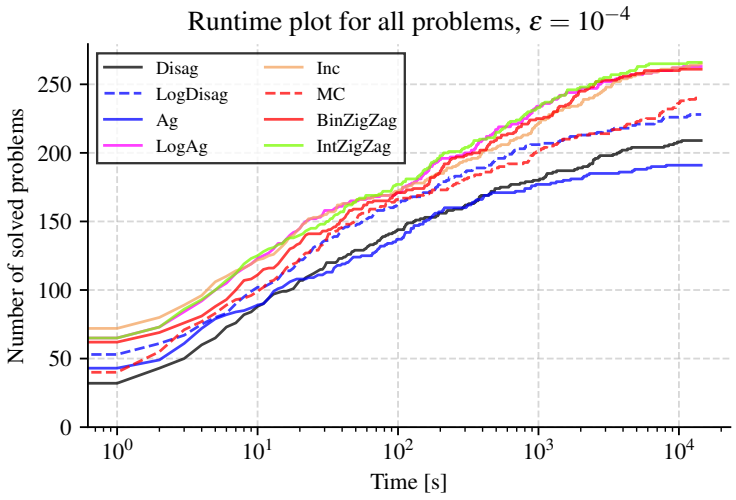
$$\sqrt[n]{\exp\left(\sum_{i=1}^{n} \ln(t_i + s)\right)} - s = \exp\left(\frac{\sum_{i=1}^{n} \ln(t_i + s)}{n}\right) - s. \tag{5.5}$$

In each MILP relaxation, we fix a maximal error bound ϵ. The break points are created depending on this error bound. To allow the model to use the maximal error bound, our framework creates two variables $\epsilon^+, \epsilon^- \in [0, \epsilon]$ every time, a non-linearity is replaced.

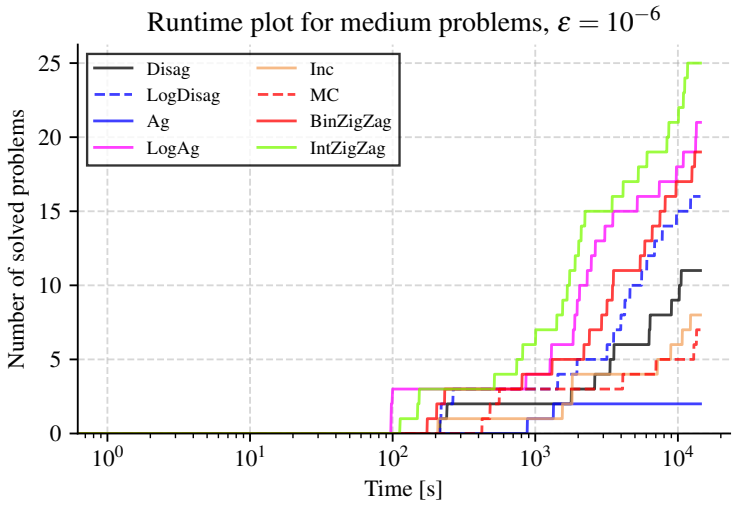
5.2.1. *Number of solved problems and runtimes.* First, we consider how many problems the various methods can solve. Figure 5.2a shows a plot that depicts the number of problems solved by each PWL model over time – here for $\epsilon = 10^{-4}$. We plot the number of optimally solved problems for each point in time between 0 and 14400 seconds (the time limit of 4 hours) and each method. The time is plotted on a logarithmic scale for better visibility. Each subfigure represents a different error bound, from now on, we will consider error bounds $\epsilon \in \{10^2, 10^0, 10^{-2}, 10^{-4}, 10^{-6}\}$ in all cases. We include large error bounds to evaluate the formulations under varying degrees of complexity. While these values may lead to solutions that are practically infeasible for the original MINLP, they help to study how the computational effort scales as the relaxation is refined toward tighter tolerances. All eight methods are plotted. The bottom plots show more differences as the problems become larger and thus more difficult with increasing error bounds.

In Table 5.2, we compare the runtimes until an optimal solution is found. If an instance is not solved during the time limit, the time limit is used as a runtime. Additional to the mean and median values, we use the shifted geometric mean (SGM) with $s = 10$, as proposed in (5.4).

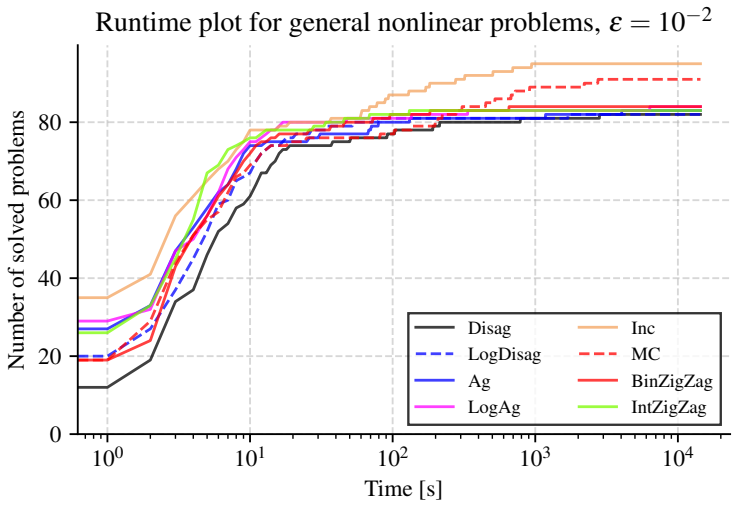
Table 5.3 distinguishes the instances according to whether they were solved within the time limit or not. As expected, the number of problems solved to optimality decreases as the error bounds become smaller. In some cases, we do not have information about the solution process for every method because of errors that occur during the steps of our framework as described in Section 4: On the one hand, the model creation time may be too long, and the solution process may not begin at all; on the other hand, when models become too large, some out-of-memory errors occur. These cases are



(A) Runtime plot for all problems, $\epsilon = 10^{-4}$



(B) Runtime plot for medium problems, $\epsilon = 10^{-6}$

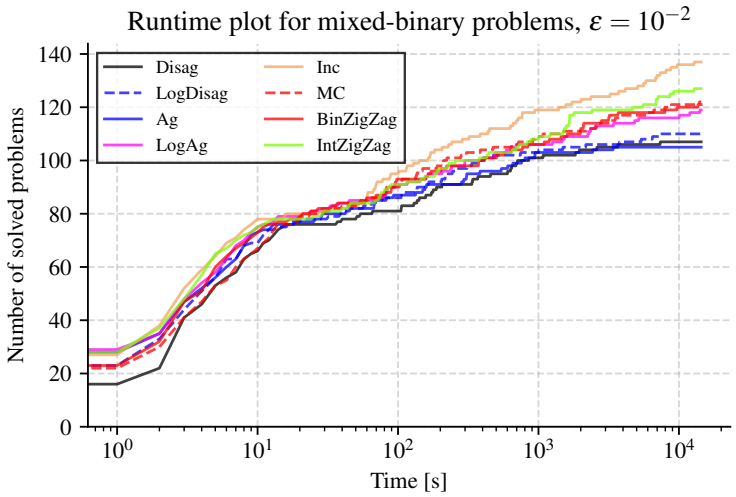


(C) Runtime plot for general nonlinear problems, $\epsilon = 10^{-2}$

omitted from our statistics for all methods. Figure 5.3 illustrates the number of instances reaching specific stages of the solution process for each MILP model. For every model, we plot a bar for each of the five error bounds. The different shades of blue represent these checkpoints: The whole process has started (STARTED), the MILP reformulation has been created (MIP CREATED), the MILP reformulation solution process has ended without errors (MIP SOLVED), and a primal or optimal solution of the relaxation has been found (PRIMAL/OPTIMAL SOLUTION FOUND), respectively.

Additionally, we present performance profiles in Appendix D.

5.2.2. *Subsets.* In Section 5.1, we explained how our benchmark set is subdivided. Now we want to make use of this to create reliable results. We created plots – similar to Figure 5.2a – but for the subsets of instances.



(D) Runtime plot for mixed-binary problems, $\epsilon = 10^{-2}$

FIGURE 5.2. Number of solved problems over time.

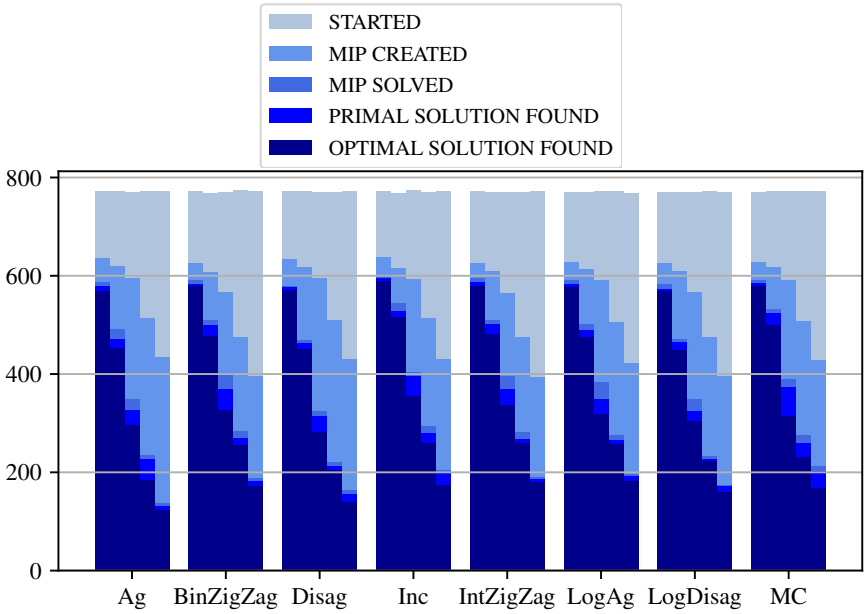


FIGURE 5.3. Solving progress of the different MILP methods. For each method, the bars represent the error bounds of $10^2, 10^0, 10^{-2}, 10^{-4}$, and 10^{-6}, from left to right.

For example, Figure 5.2b, shows the medium instance set for $\epsilon = 10^{-6}$ while Figures 5.2c and 5.2d distinguish depending on the type of nonlinear functions or the type of variables.

In total, we created 70 different plots, thus, the full results are moved to the appendix. Therein, they are grouped by subset. All results are given in Appendix C.

TABLE 5.2. Mean, median, and shifted geometric mean of the runtimes, depending on the error bound ϵ.

Error bound	$\epsilon = 10^2$			$\epsilon = 10^0$		
Runtime [s]	**Mean**	**Median**	**SGM**	**Mean**	**Median**	**SGM**
(Disag)	847.67	0.13	16.56	2852.49	5.35	75.68
(LogDisag)	899.86	0.13	19.25	2757.08	4.49	65.41
(Ag)	898.24	0.11	17.57	2838.93	4.32	71.18
(LogAg)	664.25	0.10	15.24	2116.79	3.60	52.28
(Inc)	**394.55**	**0.08**	**8.76**	**1091.61**	**3.11**	**29.14**
(MC)	640.23	0.11	13.25	1541.92	3.82	40.91
(BinZigZag)	631.83	0.10	13.17	2085.42	3.42	49.18
(IntZigZag)	547.76	0.09	12.62	1966.44	3.50	45.23
Error bound	$\epsilon = 10^{-2}$			$\epsilon = 10^{-4}$		
Runtime [s]	**Mean**	**Median**	**SGM**	**Mean**	**Median**	**SGM**
(Disag)	5015.56	172.01	317.78	5062.35	207.74	371.20
(LogDisag)	4198.54	77.26	213.28	4201.27	68.10	231.36
(Ag)	4491.85	74.85	236.16	5785.87	200.44	429.07
(LogAg)	3724.75	**42.52**	163.06	2563.17	**29.91**	126.61
(Inc)	**2646.75**	44.21	**121.33**	2648.00	31.60	138.80
(MC)	4087.79	73.77	218.22	4055.82	60.13	238.43
(BinZigZag)	3551.82	43.28	158.89	2625.05	44.21	144.72
(IntZigZag)	3241.05	49.92	143.68	**2401.44**	37.79	**121.02**
Error bound	$\epsilon = 10^{-6}$					
Runtime [s]	**Mean**	**Median**	**SGM**			
(Disag)	6081.46	841.82	694.25			
(LogDisag)	4687.49	157.04	358.92			
(Ag)	6687.15	974.21	767.21			
(LogAg)	**3242.70**	**73.48**	**174.72**			
(Inc)	3898.44	191.25	286.14			
(MC)	4468.54	328.05	439.52			
(BinZigZag)	4000.26	126.88	258.30			
(IntZigZag)	3375.64	87.50	198.68			

5.2.3. *Performance Across Subsets.* To ensure a reliable analysis of solver performance, we evaluate the benchmark results across the specific categories defined in Section 5.1. This differentiation allows us to identify how various model characteristics impact the efficiency of the PWL formulations.

TABLE 5.3. Solver results for the full benchmark set.

Error bound	$\epsilon = 10^2$		$\epsilon = 10^0$		$\epsilon = 10^{-2}$		$\epsilon = 10^{-4}$		$\epsilon = 10^{-6}$	
Solver result	**opt.**	**tl.**	**opt.**	**tl.**	**opt.**	**tl.**	**opt.**	**tl.**	**opt.**	**tl.**
(Disag)	570	29	450	100	282	137	204	98	139	86
(LogDisag)	569	30	449	101	305	114	222	80	159	66
(Ag)	569	30	452	98	296	123	185	117	123	102
(LogAg)	577	22	475	75	319	100	257	45	**183**	**42**
(Inc)	**589**	**10**	**516**	**34**	**356**	**63**	259	43	175	50
(MC)	579	20	499	51	315	104	232	70	168	57
(BinZigZag)	578	21	478	72	326	93	255	47	171	54
(IntZigZag)	580	19	482	68	336	83	**260**	**42**	181	44

While Figure 5.2a provides a global overview, the subset-specific plots reveal more nuanced trends. For instance, Figure 5.2b illustrates performance on medium-sized instances at a tight tolerance ($\epsilon = 10^{-6}$), whereas Figures 5.2c and 5.2d isolate the effects of general nonlinear functions and mixed-binary variables, respectively.

Due to a total of 70 distinct performance plots, the complete graphical results are provided in Appendix C.

5.2.4. *Solution qualities.* As we only solve relaxations, we further evaluate the gap between the optimal solution of the MINLP and the MILP relaxations' solution. Table 5.4 presents the corresponding results. Therein, additionally to the number of solved problems, we provide the median for the gaps between the optimal MINLP and MILP solutions x^* and x^{relax}, given by

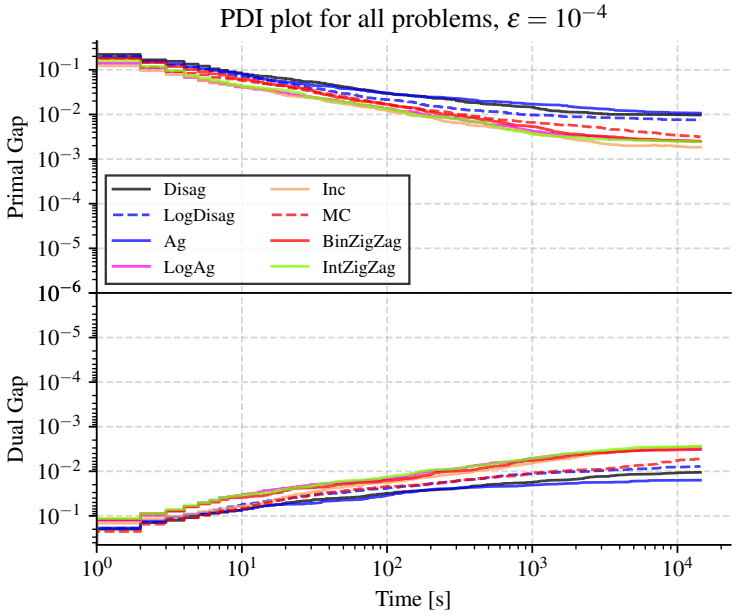
$$\frac{\left|c^\top x^* - c^\top x^{\text{relax}}\right|}{|c^\top x^*| + 10^{-10}}. \tag{5.6}$$

Adding 10^{-10} to the denominator prevents problems with instances that have an objective value of 0. As one can see, the gap is already very small for $\epsilon = 10^{-4}$, which means that using MILP relaxations, one can find fairly good dual bounds.

TABLE 5.4. Solution qualities of MILP relaxations, depending on the error bound ϵ, calculated by (5.6).

Error bound	$\epsilon = 10^2$	$\epsilon = 10^0$	$\epsilon = 10^{-2}$	$\epsilon = 10^{-4}$	$\epsilon = 10^{-6}$
Solved instances	556	428	267	156	103
Median gap	61.05%	30.56%	0.50%	0.01%	0.00%

Further, in Figure 5.4, we plot how the MILP solution improves over time. Therefore, we take the shifted geometric mean of all primal and dual gaps, i.e., the gaps between the primal/dual bounds and the optimal MILP



(A) $\epsilon = 10^{-4}$

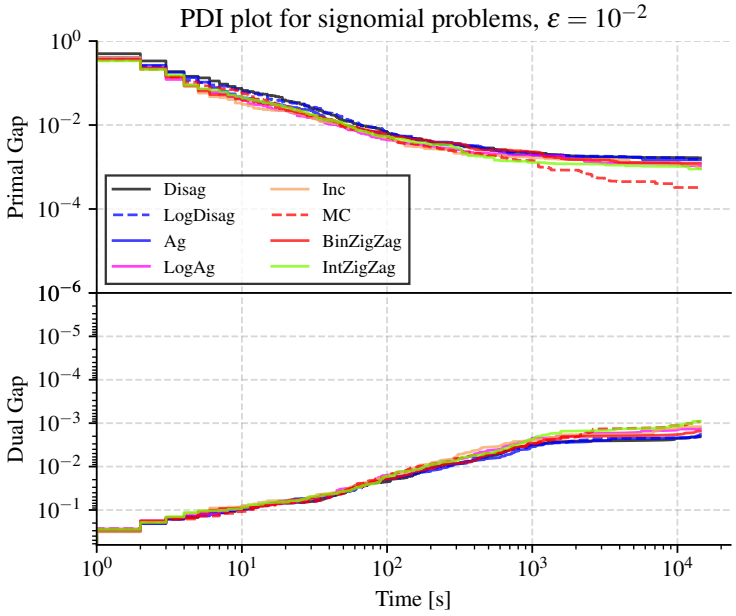
solution. The primal gap is calculated using

$$p := \begin{cases} 1 & \text{if } \overline{z} = \infty \text{ or } \overline{z} \cdot z^* < 0, \\ 0 & \text{if } \overline{z} = z^*, \\ \frac{|\overline{z} - z^*|}{\max\{|\overline{z}|, |z^*|\}} & \text{else,} \end{cases} \tag{5.7}$$

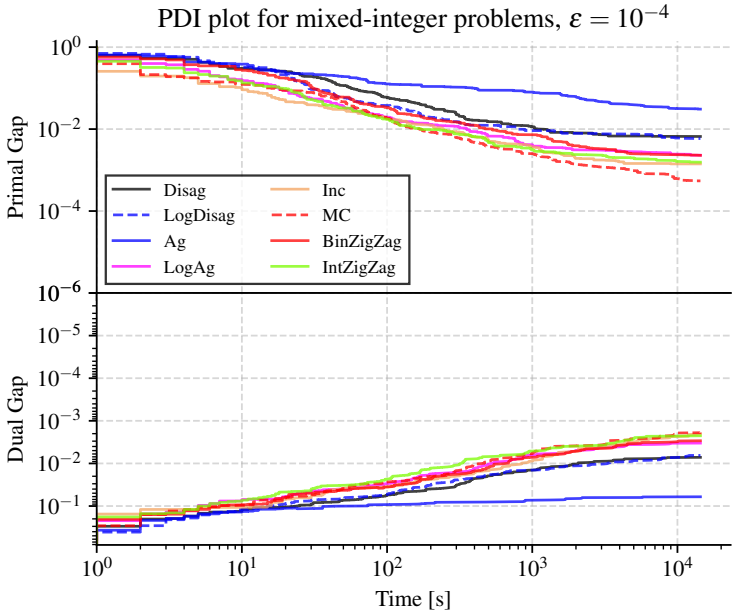
and, equivalently, the dual gap is calculated by

$$d := \begin{cases} 1 & \text{if } \underline{z} = -\infty \text{ or } \underline{z} \cdot z^* < 0, \\ 0 & \text{if } \underline{z} = z^*, \\ \frac{|\underline{z} - z^*|}{\max\{|\underline{z}|, |z^*|\}} & \text{else,} \end{cases} \tag{5.8}$$

where $\overline{z}$ and $\underline{z}$ are the current primal and dual bounds and $z^* = c^\top x^{\text{relax}}$ is the optimal solution of the MILP relaxation. The upper graphs in each subplot show how the primal gap shrinks, while the lower graphs show how the dual gap shrinks. Figure 5.4a shows the corresponding plot to Figure 5.2a. The other subfigures, Figures 5.4b and 5.4c, show the gap analysis for some subsets as described previously. Also for this type of analysis, all results are given in Appendix C.



(B) $\epsilon = 10^{-2}$



(C) $\epsilon = 10^{-4}$

FIGURE 5.4. SGM of primal and dual gap to optimal MILP solution over time.

6. Discussion

We now examine the presented results to draw meaningful conclusions from our computational study. For larger error bounds, the number of solved problems is quite similar for all methods, while already for $\epsilon = 10^0$ (Figure C.2), we can see some differences. For smaller error bounds, which correspond to larger models, the differences become more visible: A good example is $\epsilon = 10^{-4}$ (Figure 5.2a), where (Inc), (BinZigZag), (IntZigZag), and (LogAg) clearly outperform the remaining approaches. The runtimes in Table 5.2 also strengthen these findings.

Regarding the easy and very easy subsets, the methods behave quite similarly, see Figures C.6 to C.15. If a method stands out, it is typically one of the aforementioned models (Inc), (BinZigZag), (IntZigZag), and (LogAg). There are occasional outliers for (MC), but we will mention this again later on. For the medium and hard subset, (IntZigZag) and (Inc) are superior, see Figures C.16 to C.25. Specifically, (Inc) demonstrates an advantage for error bounds up to $\epsilon = 10^{-4}$, whereas (IntZigZag) dominates at $\epsilon = 10^{-6}$, as seen, for example, in Figure 5.2b.

In the case of general nonlinear instances, we can see that (Inc) has a very good performance, see Figures C.41 to C.45 – particularly highlighted in Figure 5.2c.

While the leading methods perform comparably well on convex problems, (Inc) stands out again for nonconvex problems, see Figures C.51 to C.55.

When considering the type of variables, we have the following: For mixed-binary (Figures C.61 to C.65) problems, again (Inc) is superior while for continuous (Figures C.56 to C.60) and mixed-integer (Figures C.66 to C.70) problems the methods behave more similarly. A good example is given in Figure 5.2d.

We further investigated smaller time limits in Table B.1, but the results coincide with our previous findings. This means, (Inc) solves the instances fastest, while the other methods catch up only for the easier subsets and very small error bounds.

Complementary to the runtimes, we also analyzed the solution process itself, i.e., how the optimality gaps decrease over time. However, before doing so, we investigated the extent to which our solution framework as described in Section 4 worked (Figure 5.3), where it stands out that (MC) is able to find primal solutions, i.e., solve feasibility problems quite fast. For larger error bounds, the gaps evolve more similarly for all methods, while, for smaller error bounds the differences become greater, see Figures C.1 to C.5 . Only the non-logarithmic convex combinations again show worse results than the remaining methods. For all error bounds, the logarithmic versions of the aggregated convex combination, i.e., (LogAg), (BinZigZag), and (IntZigZag) show a similar behavior. This is not surprising, as these methods only differ in how the branching scheme is defined, see Section 3.2.

There are specific instances where (MC) yields highly competitive results, as seen in Figures C.10, C.17, C.30, C.42, C.43 and C.62. However, its overall performance exhibits substantial variance across different problem sets. As suggested by our gap analysis in Figures 5.4b and 5.4c, this may be linked to its ability in finding primal feasible solutions.

Overall, the evaluation of the different methods leads to the following main conclusions:

- The incremental method shows the best performance in terms of the number of solved problems.
- The non-logarithmic convex combinations, (Ag) and (Disag), solve the fewest instances.
- In general, the logarithmic methods outperform their non-logarithmic counterparts, with the (IntZigZag) model being slightly superior to the other logarithmic models.
- For very large problems, (Inc) reaches its computational limits, and (IntZigZag) becomes a good alternative.
- (MC) might be a good alternative for feasibility problems.

Our observations suggest that the practical performance is influenced not only by formulation size, but also by structural properties such as branching behavior and local ideality. Finally, to emphasize the relevance of PWL relaxations in the context of solving MINLPs, we discuss how good the MILP relaxation's solutions are in general. Since we know the optimal solution for each MINLP as we restricted ourselves to these instances, we can calculate the gap between the optimal MILP and MINLP solutions, as described in (5.6). The optimal solution of a MILP relaxation is a valid dual bound for the corresponding MINLP. Thus, the results in Table 5.4, which shows the median values of all gaps, can be considered as relative optimality gaps for the MINLP problems. As we can see, the median gap is negligible already for $\epsilon = 10^{-2}$, which means that more than half of the MILP relaxation' solutions are similar to the corresponding optimal MINLP solution in terms of objective values. For $\epsilon \leq 10^{-4}$, also the mean and SGM values are fairly small.

The achievement of such small gaps suggests that PWL relaxations are a viable alternative to spatial branch-and-bound in practice. This is particularly noteworthy given our decomposed expression tree structure: although each individual nonlinearity introduces an error of up to ϵ that can propagate through the model, the resulting global gaps remain remarkably tight.

Remark 6.1. It is noteworthy that when evaluating the initial MINLPLIB problems instead of the presolved instances, one obtains similar results. This supports the reliability of our comparison and the conclusions made in this computational study.

7. Conclusion

In this paper, we compared various commonly used MILP models for PWL relaxations of one-dimensional nonlinear functions that are known in the literature. Using over 750 instances created from the MINLPLIB data set, we conducted a comprehensive computational study to determine a general performance of these models. Our results demonstrate the advantages of the incremental method as presented by [27] and are accompanied by a recommendation of this method for practical applications. Further, for very high-accuracy relaxations, the Zig-Zag formulations appear to be a viable alternative. Additionally, the multiple choice method offers a promising

formulation for quickly finding feasible solutions. Practitioners can use our framework to conduct such a study for their type of problems since it is publicly available, see [7].

Since our primary focus in this study was to evaluate different PWL models rather than assessing the overall performance of the solver, there are some points where our implementation could have been improved. For instance, a more sophisticated method for determining breakpoints or avoiding duplicates in variables and constraints would be desirable. It is important to note, however, that the inefficiencies in our framework remain consistent across all MILP models and do not affect the conclusions drawn from our research.

Future work may focus on different topics. First, this research only considers the maximum error for each segment, while overlooking factors such as the number of breakpoints and the specific characteristics of nonlinearities. Moreover, approaches that use PWL relaxations to solve MINLPs can benefit significantly from adaptivity by refining the PWL relaxations only locally in an iterative manner. Depending on the stages of the solution process, in this setting, PWL relaxations must thus be solved with both small and large numbers of segments. Consequently, the combination of different PWL models might be the best starting point for adaptive approaches, which is also supported by our results. One promising idea is to combine methods that are capable of finding primal solutions quickly, like the multiple choice method, with methods that are better suited for closing the optimality gap with high-accuracy PWL relaxations, like the incremental method or the integer Zig-Zag formulations. Finally, future research could extend this benchmark by evaluating the ideal formulations proposed by [26]. Comparing their theoretically stronger bounds against the empirical baseline of standard methods established here would be a valuable next step.

References

[1] Kevin-Martin Aigner, Robert Burlacu, Frauke Liers, and Alexander Martin. “Solving AC Optimal Power Flow with Discrete Decisions to Global Optimality”. In: *INFORMS Journal on Computing* 35.2 (2023), pp. 458–474. DOI: 10.1287/ijoc.2023.1270.

[2] Anantharam Balakrishnan and Stephen C. Graves. “A composite algorithm for a concave-cost network flow problem”. In: *Networks* 19.2 (1989), pp. 175–202. DOI: 10.1002/net.3230190202.

[3] Andreas Bärmann, Robert Burlacu, Lukas Hager, and Thomas Kleinert. “On piecewise linear approximations of bilinear terms: structural comparison of univariate and bivariate mixed-integer programming formulations”. In: *Journal of Global Optimization* 85.4 (2022), pp. 789–819. DOI: 10.1007/S10898-022-01243-y.

[4] Benjamin Beach, Robert Hildebrand, and Joey Huchette. “Compact mixed-integer programming formulations in quadratic optimization”. In: *Journal of Global Optimization* 84.4 (2022), pp. 869–912. DOI: 10.1007/s10898-022-01184-6.

[5] Pietro Belotti, Sonia Cafieri, Jon Lee, and Leo Liberti. "Feasibility-Based Bounds Tightening via Fixed Points". In: *Combinatorial Optimization and Applications*. Ed. by Weili Wu and Ovidiu Daescu. Vol. 6508. Springer Berlin Heidelberg, 2010, pp. 65–76. DOI: 10.1007/978-3-642-17458-2_7.

[6] Suresh Bolusani, Mathieu Besançon, Ksenia Bestuzheva, Antonia Chmiela, João Dionísio, Tim Donkiewicz, Jasper van Doornmalen, Leon Eifler, Mohammed Ghannam, Ambros Gleixner, Christoph Graczyk, Katrin Halbig, Ivo Hedtke, Alexander Hoen, Christopher Hojny, Rolf van der Hulst, Dominik Kamp, Thorsten Koch, Kevin Kofler, Jurgen Lentz, Julian Manns, Gioni Mexi, Erik Mühmer, Marc E. Pfetsch, Franziska Schlösser, Felipe Serrano, Yuji Shinano, Mark Turner, Stefan Vigerske, Dieter Weninger, and Lixing Xu. *The SCIP Optimization Suite 9.0*. ZIB-Report 24-02-29. Zuse Institute Berlin, Feb. 2024. URL: https://nbn-resolving.org/urn:nbn:de:0297-zib-95528.

[7] Kristin Braun. *GitHub - kristinbraun/pwl-t-rex — github.com*. https://github.com/kristinbraun/pwl-t-rex. 2025.

[8] Robert Burlacu, Björn Geißler, and Lars Schewe. "Solving mixed-integer nonlinear programmes using adaptively refined mixed-integer linear programmes". In: *Optimization Methods and Software* 35 (2019), pp. 37–64. DOI: 10.1080/10556788.2018.1556661.

[9] Michael R. Bussieck, Arne Stolbjerg Drud, and Alexander Meeraus. "MINLPLib—A Collection of Test Models for Mixed-Integer Nonlinear Programming". In: *INFORMS Journal on Computing* 15.1 (2003), pp. 114–119. DOI: 10.1287/ijoc.15.1.114.15159.

[10] Michael L Bynum, Gabriel A Hackebeil, William E Hart, Carl D Laird, Bethany L Nicholson, John D Siirola, Jean-Paul Watson, David L Woodruff, et al. *Pyomo-optimization modeling in python*. Vol. 67. Springer, 2021. DOI: 10.1007/978-3-030-68928-5.

[11] Carlos M. Correa-Posada and Pedro Sánchez-Martín. "Gas Network Optimization: A comparison of Piecewise Linear Models". 2014. URL: http://www.optimization-online.org/DB_HTML/2014/10/4580.html.

[12] Keely L. Croxton, Bernard Gendron, and Thomas L. Magnanti. "A Comparison of Mixed-Integer Programming Models for Nonconvex Piecewise Linear Cost Minimization Problems". In: *Management Science* 49.9 (2003), pp. 1268–1273. DOI: 10.1287/mnsc.49.9.1268.16570.

[13] Elizabeth D Dolan and Jorge J Moré. "Benchmarking optimization software with performance profiles". In: *Mathematical Programming* 91.2 (2002), pp. 201–213. DOI: 10.1007/s101070100263.

[14] Robert Fourer, Jun Ma, and Kipp Martin. "OSiL: An instance language for optimization". In: *Computational optimization and applications* 45.1 (2010), pp. 181–203. DOI: 10.1007/s10589-008-9169-6.

[15] Björn Geißler, Alexander Martin, Antonio Morsi, and Lars Schewe. "Using Piecewise Linear Functions for Solving MINLPs". In: *Mixed Integer Nonlinear Programming*. Ed. by Jon Lee and Sven Leyffer. Vol. 154. The IMA Volumes in Mathematics and its Applications.

Springer New York, 2012, pp. 287–314. DOI: 10.1007/978-1-4614-1927-3_10.

[16] Martin Gugat, Günter Leugering, Alexander Martin, Martin Schmidt, Mathias Sirvent, and David Wintergerst. "Towards simulation based mixed-integer optimization with differential equations". In: *Networks* (2018). DOI: 10.1002/net.21812.

[17] Gurobi Optimization, LLC. *Gurobi Optimizer Reference Manual*. 2024. URL: https://www.gurobi.com.

[18] William E Hart, Jean-Paul Watson, and David L Woodruff. "Pyomo: modeling and solving mathematical programs in Python". In: *Mathematical Programming Computation* 3.3 (2011), pp. 219–260. DOI: 10.1007/s12532-011-0026-8.

[19] M. M. Faruque Hasan and I.A. Karimi. "Piecewise linear relaxation of bilinear programs using bivariate partitioning". In: *AIChE Journal* 56.7 (2010), pp. 1880–1893. DOI: 10.1002/aic.12109.

[20] Joey Huchette and Juan Pablo Vielma. "Nonconvex Piecewise Linear Functions: Advanced Formulations and Simple Modeling Tools". In: *Operations Research* 71.5 (2022), pp. 1835–1856. DOI: 10.1287/opre.2019.1973.

[21] Joseph Andrew Huchette. "Advanced mixed-integer programming formulations : methodology, computation, and application". 2018. URL: https://dspace.mit.edu/handle/1721.1/119282.

[22] Robert G Jeroslow and James K Lowe. *Modelling with integer variables*. Springer, 1984.

[23] Jon Lee and Dan Wilson. "Polyhedral methods for piecewise-linear functions. I. The lambda method". In: *Discrete Appl. Math.* 108.3 (2001), pp. 269–285. DOI: 10.1016/S0166-218x(00)00216-x.

[24] Moritz Link and Stefan Volkwein. "Adaptive piecewise linear relaxations for enclosure computations for nonconvex multiobjective mixed-integer quadratically constrained programs". In: *Journal of Global Optimization* 87.1 (2023), pp. 97–132. DOI: 10.1007/s10898-023-01309-5.

[25] Andreas Lundell, Anders Skjäl, and Tapio Westerlund. "A reformulation framework for global optimization". In: *Journal of Global Optimization* 57.1 (2013), pp. 115–141. DOI: 10.1007/s10898-012-9877-4.

[26] Bochuan Lyu, Illya V Hicks, and Joey Huchette. "Building formulations for piecewise linear relaxations of nonlinear functions". In: *Operations Research* 74.1 (2026), pp. 484–499.

[27] Harry M Markowitz and Alan S Manne. "On the Solution of Discrete Programming Problems". In: *Econometrica* 25.1 (1957), pp. 84–110. DOI: 10.2307/1907744.

[28] Alexander Martin, Markus Möller, and Susanne Moritz. "Mixed Integer Models for the Stationary Case of Gas Network Optimization". In: *Mathematical Programming* 105.2 (2006), pp. 563–582. DOI: 10.1007/s10107-005-0665-5.

[29] Garth P McCormick. "Computability of global solutions to factorable nonconvex programs: Part I—Convex underestimating problems". In: *Mathematical programming* 10.1 (1976), pp. 147–175. DOI: 10.1007/bf01580665.

[30] R. Misener and C. A. Floudas. “Piecewise-Linear Approximations of Multidimensional Functions”. In: *Journal of Optimization Theory and Applications* 145.1 (2010), pp. 120–147. DOI: 10.1007/s10957-009-9626-0.

[31] Hans Mittelmann. *Benchmark for optimization software*. http://plato.asu.edu/bench.html.

[32] Antonio Morsi. “Solving MINLPs on Loosely-Coupled Networks with Applications in Water and Gas Network Optimization”. PhD thesis. Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), 2013.

[33] Manfred Padberg. “Approximating separable nonlinear functions via mixed zero-one programs”. In: *Operations Research Letters* 27.1 (2000), pp. 1–5. DOI: 10.1016/s0167-6377(00)00028-6.

[34] Steffen Rebennack. “Computing tight bounds via piecewise linear functions through the example of circle cutting problems”. In: *Mathematical Methods of Operations Research* 84 (1 2016), pp. 3–57. DOI: 10.1007/S00186-016-0546-0/FIGURES/15.

[35] Steffen Rebennack and Josef Kallrath. “Continuous piecewise linear delta-approximations for bivariate and multivariate functions”. In: *Journal of Optimization Theory and Applications* 167.1 (2015), pp. 102–117. DOI: 10.1007/s10957-014-0688-2.

[36] Steffen Rebennack and Josef Kallrath. “Continuous piecewise linear delta-approximations for univariate functions: computing minimal breakpoint systems”. In: *Journal of Optimization Theory and Applications* 167.2 (2015), pp. 617–643. DOI: 10.1007/s10957-014-0687-3.

[37] Steffen Rebennack and Vitaliy Krasko. “Piecewise Linear Function Fitting via Mixed-Integer Linear Programming”. In: *INFORMS Journal on Computing* 32.2 (2019), pp. 507–530. DOI: 10.1287/ijoc.2019.0890.

[38] Ricardo Rovatti, Claudia D’Ambrosio, Andrea Lodi, and Silvano Martello. “Optimistic MILP modeling of non-linear optimization problems”. In: *European Journal of Operational Research* 239.1 (2014), pp. 32–45. DOI: 10.1016/j.ejor.2014.03.020.

[39] Hanif D. Sherali. “On mixed-integer zero-one representations for separable lower-semicontinuous piecewise-linear functions”. In: *Operations Research Letters* 28.4 (2001), pp. 155–160. DOI: 10.1016/s0167-6377(01)00063-3.

[40] Abel Soares Siqueira, Raniere Gaia Costa da Silva, and Luiz-Rafael Santos. “Perprof-py: A Python Package for Performance Profile of Mathematical Optimization Software”. In: *Journal of Open Research Software* 4.1 (2016), p. 12. DOI: 10.5334/jors.81.

[41] Juan Pablo Vielma. “Mixed Integer Linear Programming Formulation Techniques”. In: *SIAM Review* 57.1 (2015), pp. 3–57. DOI: 10.1137/130915303.

[42] Juan Pablo Vielma, Shabbir Ahmed, and George L Nemhauser. “Mixed-Integer Models for Nonseparable Piecewise-Linear Optimization: Unifying Framework and Extensions”. In: *Operations Research* 58.2 (2010), pp. 303–315. DOI: 10.1287/opre.1090.0721.

[43] Sercan Yildiz and Juan Pablo Vielma. "Incremental and encoding formulations for Mixed Integer Programming". In: *Operations Research Letters* 41 (6 Nov. 2013), pp. 654–658. DOI: 10.1016/J.ORL.2013.09.004.

Declarations

Funding. This work has been done within the joint project "TrinkXtrem" funded by the Federal Ministry of Education and Research (BMBF) under the project number 02WEE1625B in the funding "Wasser-Extremereignisse" (WaX) of the Federal Program "Wasser:N" and as part of the announcement "Artificial Intelligence in Civil Security Research II" of the BMBF within the program "Research for Civil Security" of the Federal Government.

Competing Interests. There are no competing interests.

Author contributions. All authors contributed to the study's conception and design. Robert Burlacu came up with the original idea. Kristin Braun was responsible for the implementation and the computational results. The proof of Lemma 3.2 was conducted by Kristin Braun. All authors contributed to the first draft of the manuscript, and all authors provided feedback on previous drafts. All authors read and approved the final manuscript.

Data availability. All data used during the current study is available in the MINLPLIB: https://www.minlplib.org/

Appendix A. Branching scheme for the logarithmic branching convex combination model

The following Table A.1 gives a visualization of our branching scheme from Section 3.2.

Table A.1. Example for the iterative creation of L_s, R_s for $n = 16$.

	$\bar{x}_0$	$\bar{x}_1$	$\bar{x}_2$	$\bar{x}_3$	$\bar{x}_4$	$\bar{x}_5$	$\bar{x}_6$	$\bar{x}_7$	$\bar{x}_8$	$\bar{x}_9$	$\bar{x}_{10}$	$\bar{x}_{11}$	$\bar{x}_{12}$	$\bar{x}_{13}$	$\bar{x}_{14}$	$\bar{x}_{15}$	$\bar{x}_{16}$
T_4	L_4	L_4	L_4	L_4	L_4	L_4	L_4	L_4	-	R_4	R_4	R_4	R_4	R_4	R_4	R_4	R_4
T_3	L_3	L_3	L_3	L_3	-	R_3	R_3	R_3	R_3	R_3	R_3	R_3	-	L_3	L_3	L_3	L_3
T_2	L_2	L_2	-	R_2	R_2	R_2	-	L_2	L_2	L_2	-	R_2	R_2	R_2	-	L_2	L_2
T_1	L_1	-	R_1	-	L_1	-	R_1	-	L_1	-	R_1	-	L_1	-	R_1	-	L_1

Appendix B. Solving time for smaller time limits

For considering shorter time limits, we also provide statistics about the runtime. Table B.1 show the shifted geometric mean if the runtimes are limited to 100, 1000, and 10000 seconds, respectively. Every runtime that exceeds this time limit is then limited to the respective time limit.

TABLE B.1. Shifted geometric mean of the runtimes when the time limit is smaller.

Error bound	$\epsilon = 10^{2}$			$\epsilon = 10^{0}$		
Time limit	100s	1000s	10000s	100s	1000s	10000s
(Disag)	7.86	12.65	16.08	17.53	40.01	70.15
(LogDisag)	8.78	14.60	18.69	16.03	35.22	60.53
(Ag)	8.37	13.21	17.05	16.46	37.37	65.95
(LogAg)	8.15	12.32	14.87	15.29	31.98	49.26
(Inc)	**5.26**	**7.32**	**8.63**	**12.87**	**21.72**	**28.27**
(MC)	6.82	10.55	12.96	15.14	27.90	39.19
(BinZigZag)	7.21	10.59	12.85	14.70	30.15	46.38
(IntZigZag)	7.51	10.44	12.34	14.26	28.30	42.76
Error bound	$\epsilon = 10^{-2}$			$\epsilon = 10^{-4}$		
Time limit	100s	1000s	10000s	100s	1000s	10000s
(Disag)	36.30	118.89	280.62	41.04	136.06	328.65
(LogDisag)	32.76	92.40	191.96	34.55	99.03	208.93
(Ag)	32.26	97.01	210.84	41.01	139.41	371.27
(LogAg)	29.20	76.51	148.44	28.76	73.66	119.37
(Inc)	**27.18**	**68.52**	**114.25**	**28.47**	78.00	**130.99**
(MC)	32.48	94.64	197.82	34.48	102.78	217.96
(BinZigZag)	29.39	76.75	145.38	31.54	83.15	136.19
(IntZigZag)	28.69	73.80	132.75	28.84	**72.35**	114.49
Error bound	$\epsilon = 10^{-6}$					
Time limit	100s	1000s	10000s			
(Disag)	54.42	210.54	601.99			
(LogDisag)	44.60	140.61	321.23			
(Ag)	52.09	217.11	648.88			
(LogAg)	**34.47**	**89.41**	**161.90**			
(Inc)	40.71	128.61	262.78			
(MC)	52.71	179.96	397.81			
(BinZigZag)	39.17	114.88	235.30			
(IntZigZag)	37.09	99.16	184.02			

Appendix C. Runtime plots

C.1. **Aggregate overview across categories.** The following figures show the runtime plots for all problems (aggregate overview across categories).

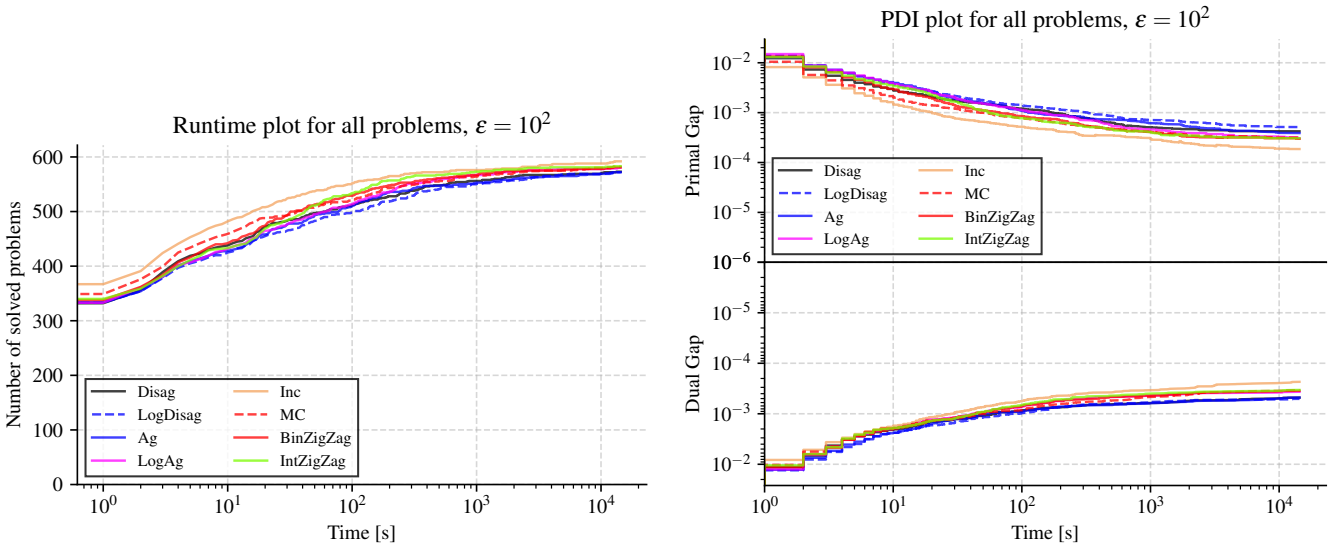


Figure C.1. Runtime and Primal-Dual Integral plots for all problems, $\epsilon = 10^2$

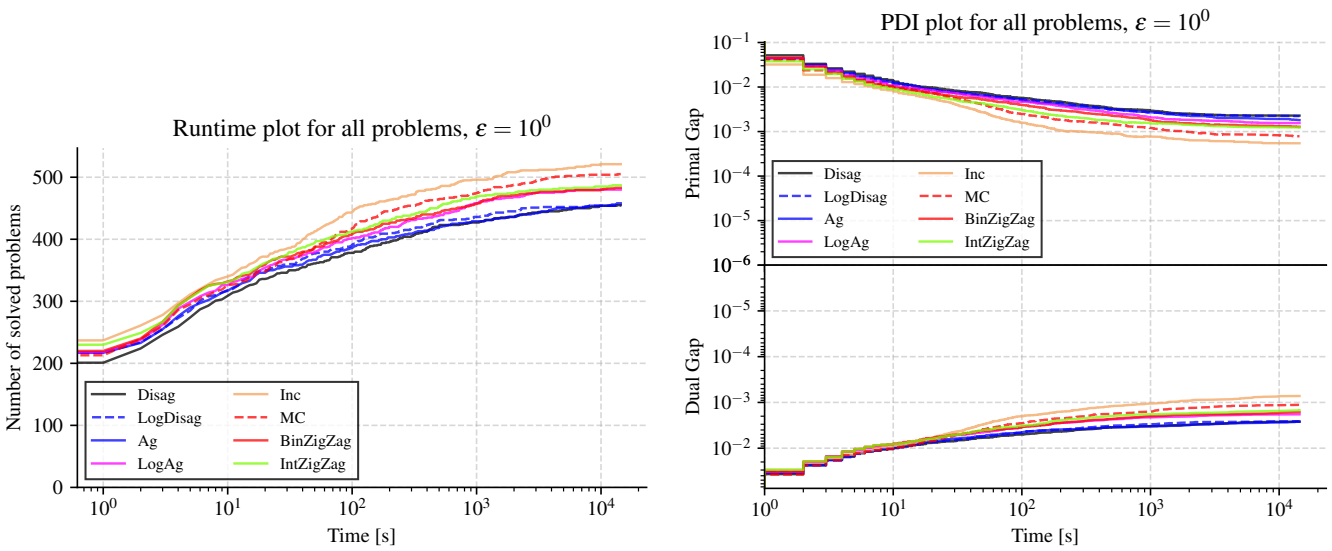


FIGURE C.2. Runtime and Primal-Dual Integral plots for all problems, $\epsilon = 10^0$

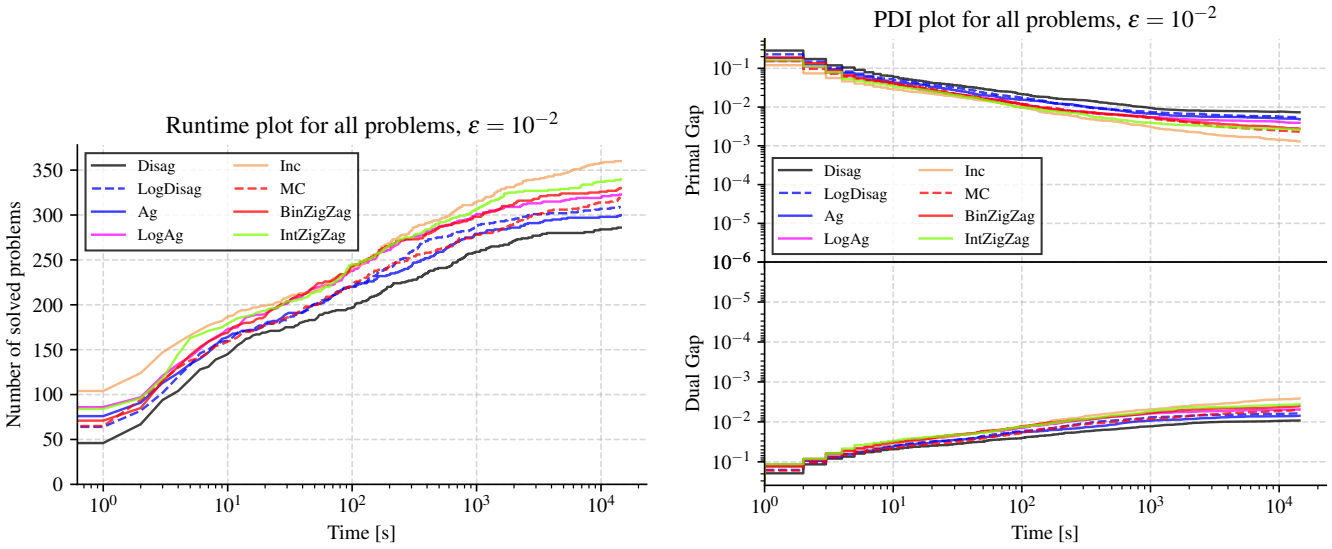


FIGURE C.3. Runtime and Primal-Dual Integral plots for all problems, $\epsilon = 10^{-2}$

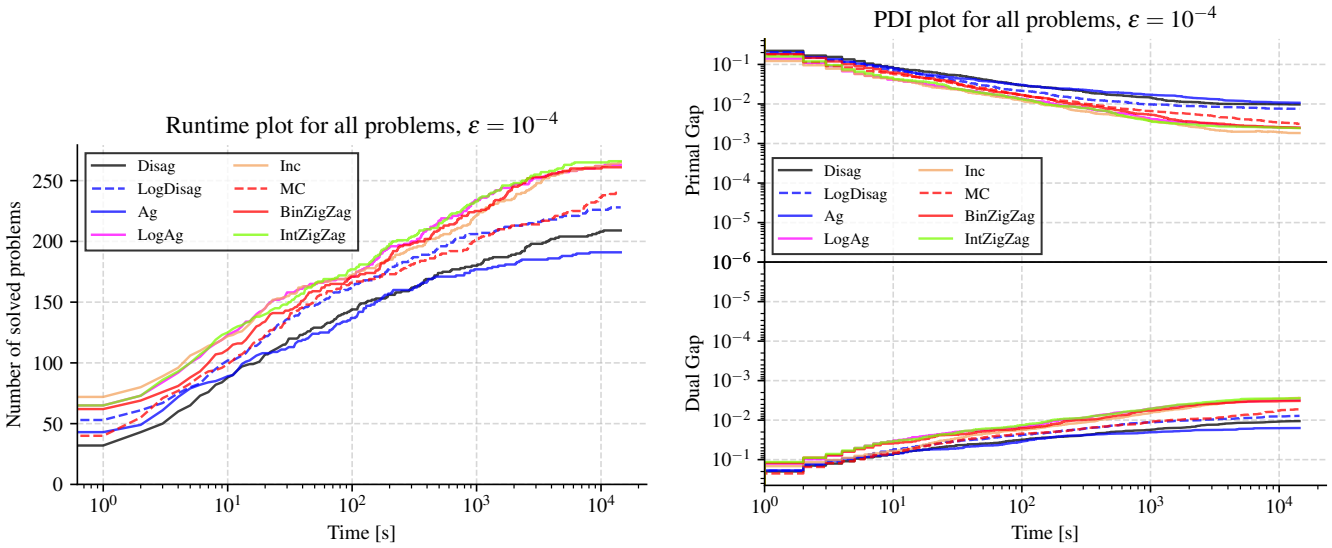


FIGURE C.4. Runtime and Primal-Dual Integral plots for all problems, $\epsilon = 10^{-4}$

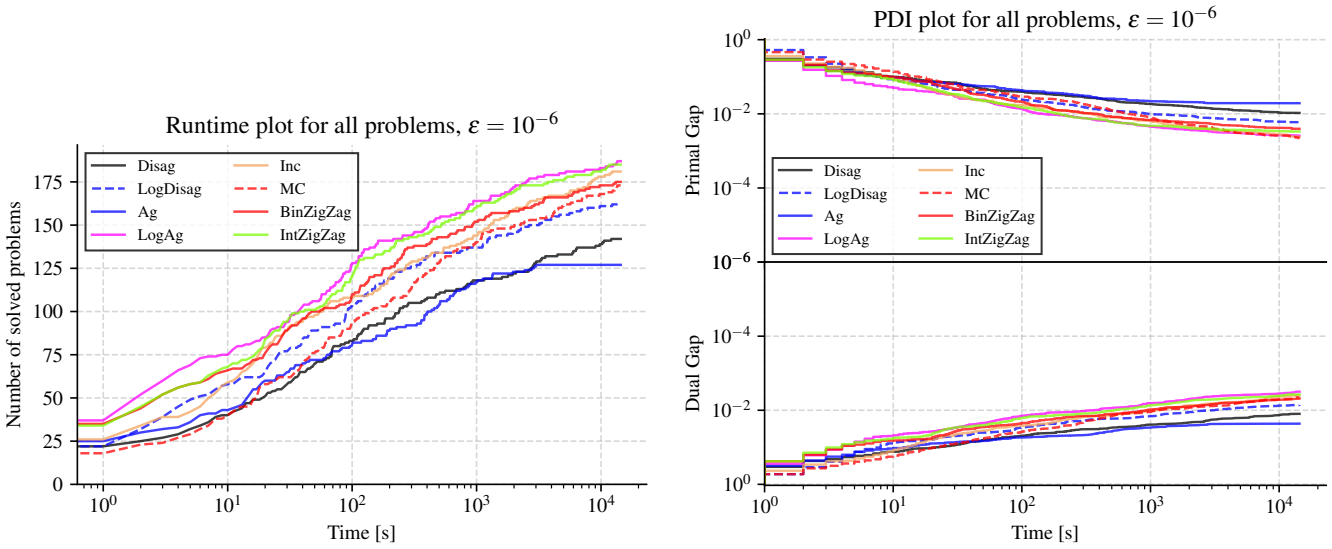


FIGURE C.5. Runtime and Primal-Dual Integral plots for all problems, $\epsilon = 10^{-6}$

C.2. **Runtime plots for very easy problems.** The following figures show the runtime plots for very easy problems.

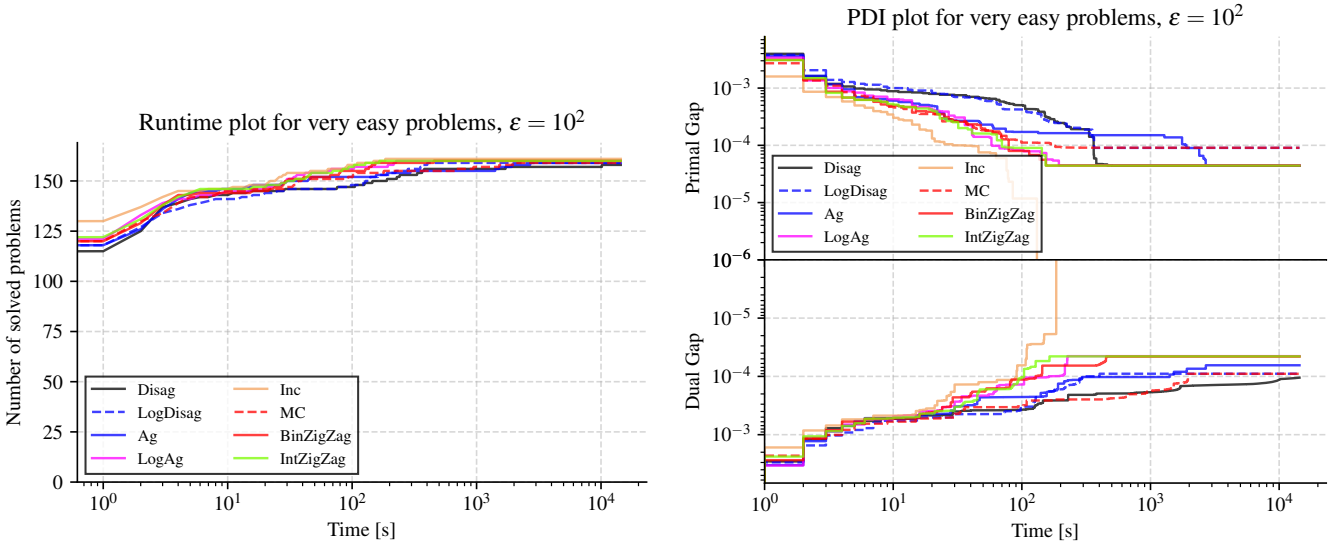


FIGURE C.6. Runtime and Primal-Dual Integral plots for very easy problems, $\epsilon = 10^2$

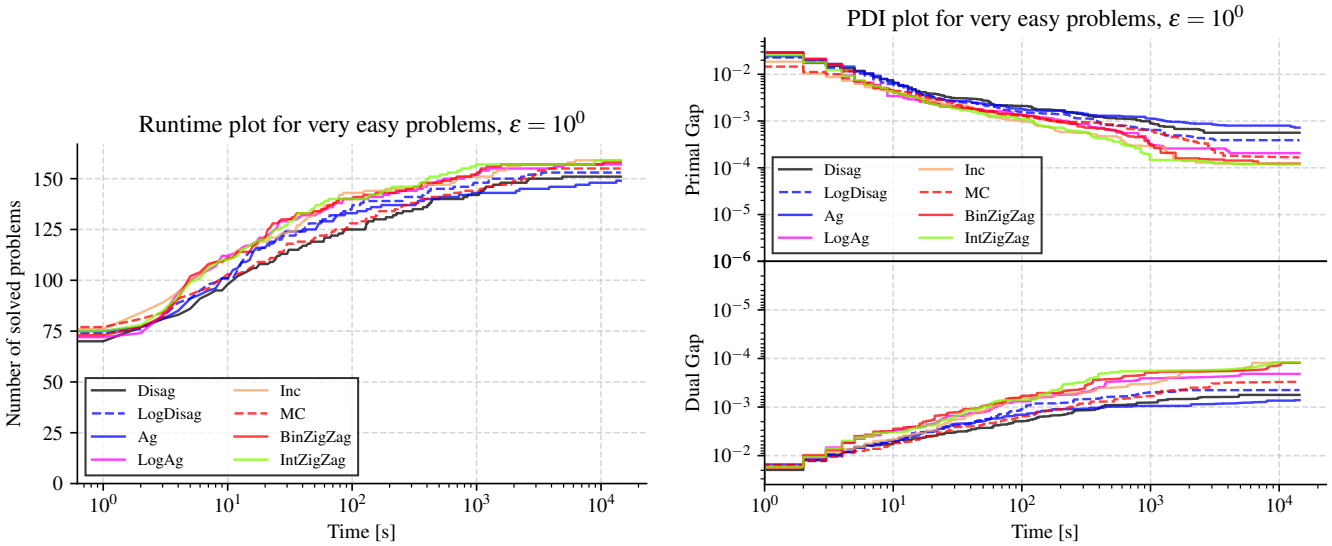


FIGURE C.7. Runtime and Primal-Dual Integral plots for very easy problems, $\epsilon = 10^0$

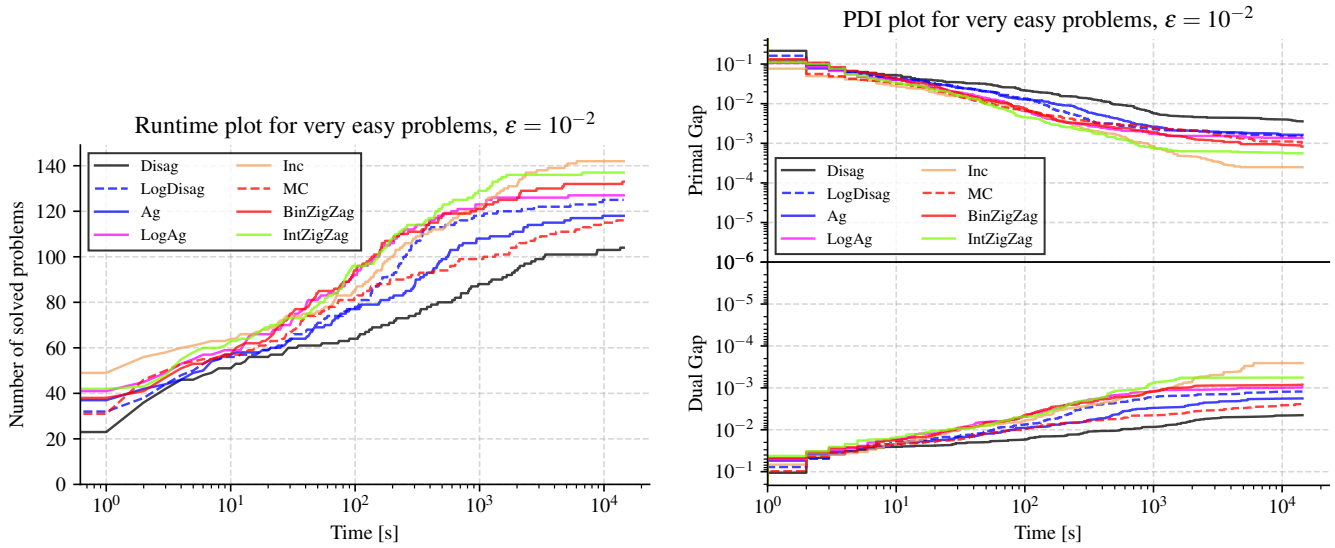


FIGURE C.8. Runtime and Primal-Dual Integral plots for very easy problems, $\epsilon = 10^{-2}$

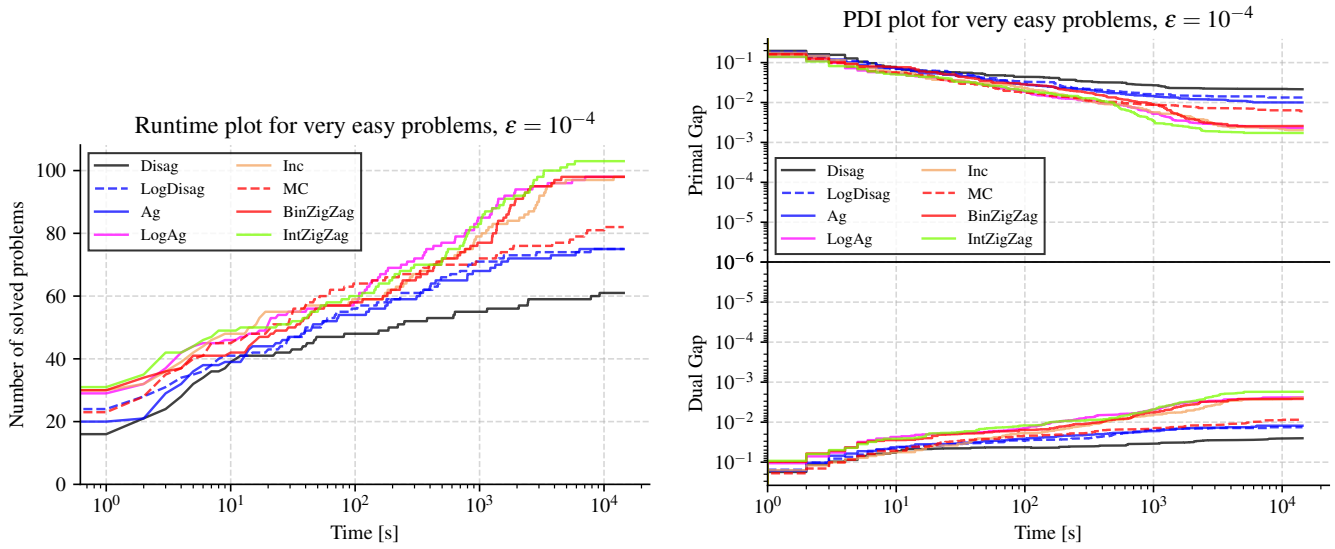


FIGURE C.9. Runtime and Primal-Dual Integral plots for very easy problems, $\epsilon = 10^{-4}$

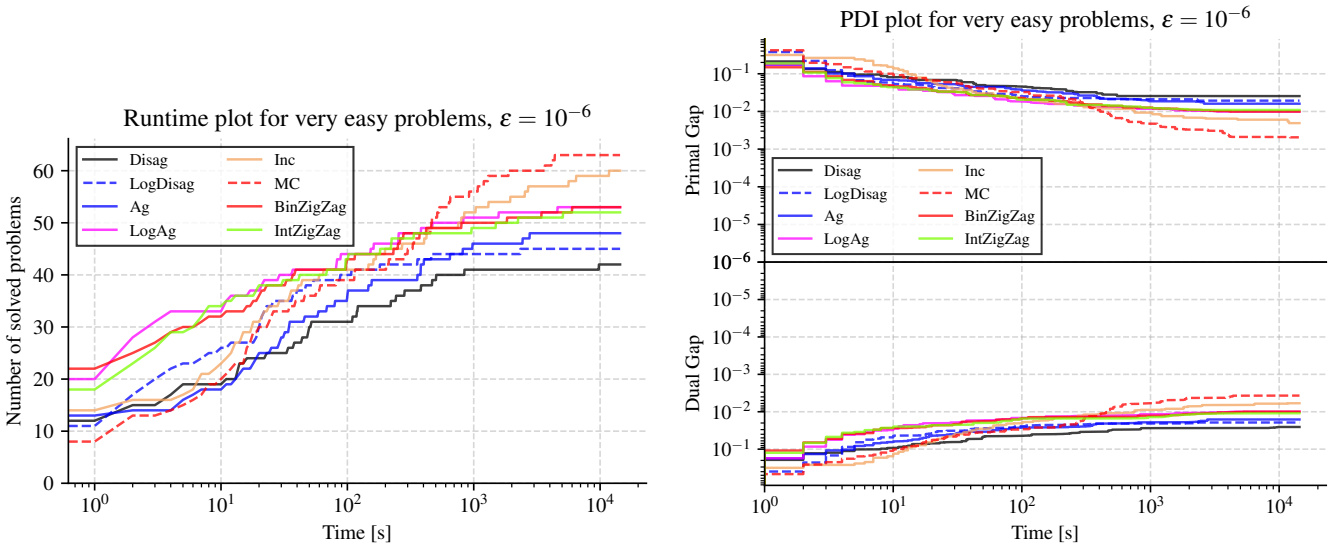


FIGURE C.10. Runtime and Primal-Dual Integral plots for very easy problems, $\epsilon = 10^{-6}$

C.3. **Runtime plots for easy problems.** The following figures show the runtime plots for easy problems.

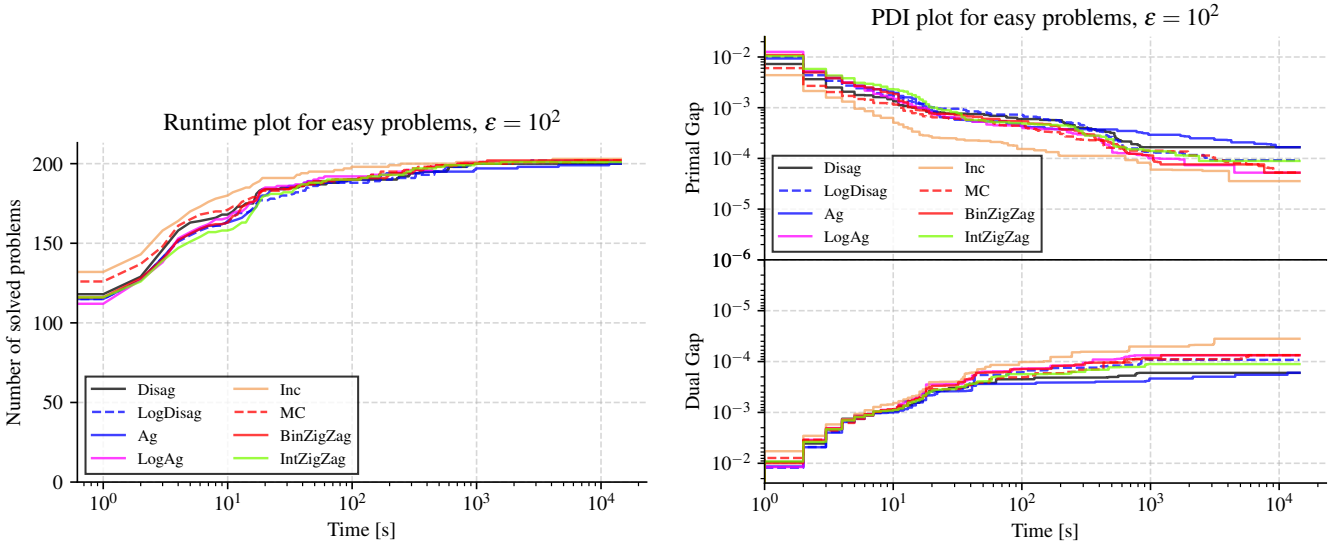


FIGURE C.11. Runtime and Primal-Dual Integral plots for easy problems, $\epsilon = 10^2$

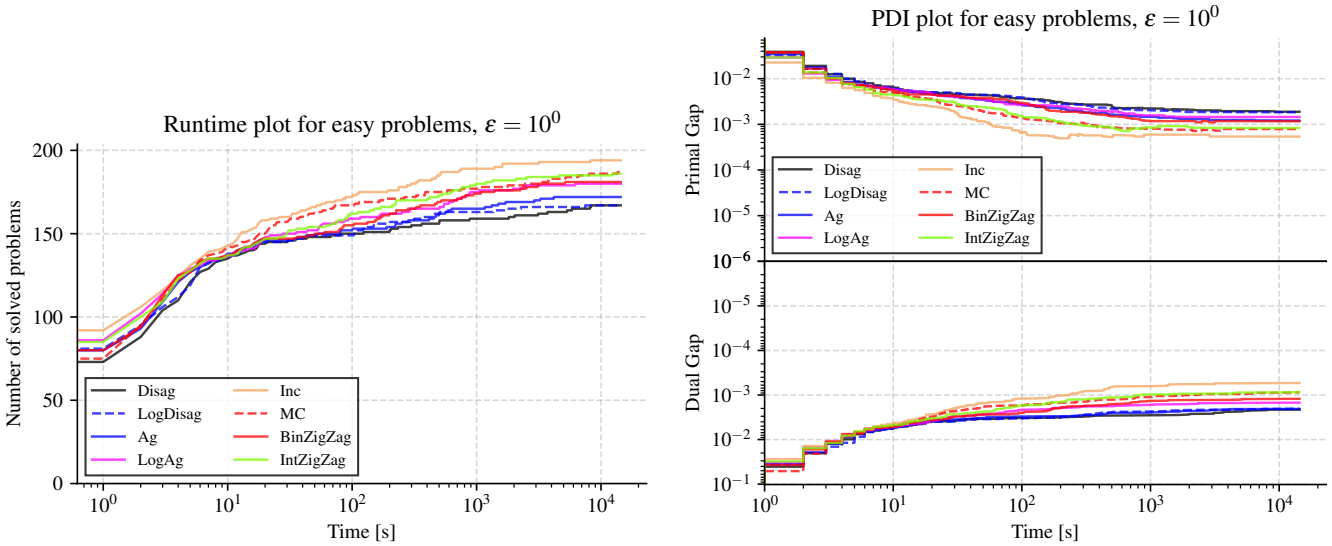


FIGURE C.12. Runtime and Primal-Dual Integral plots for easy problems, $\epsilon = 10^0$

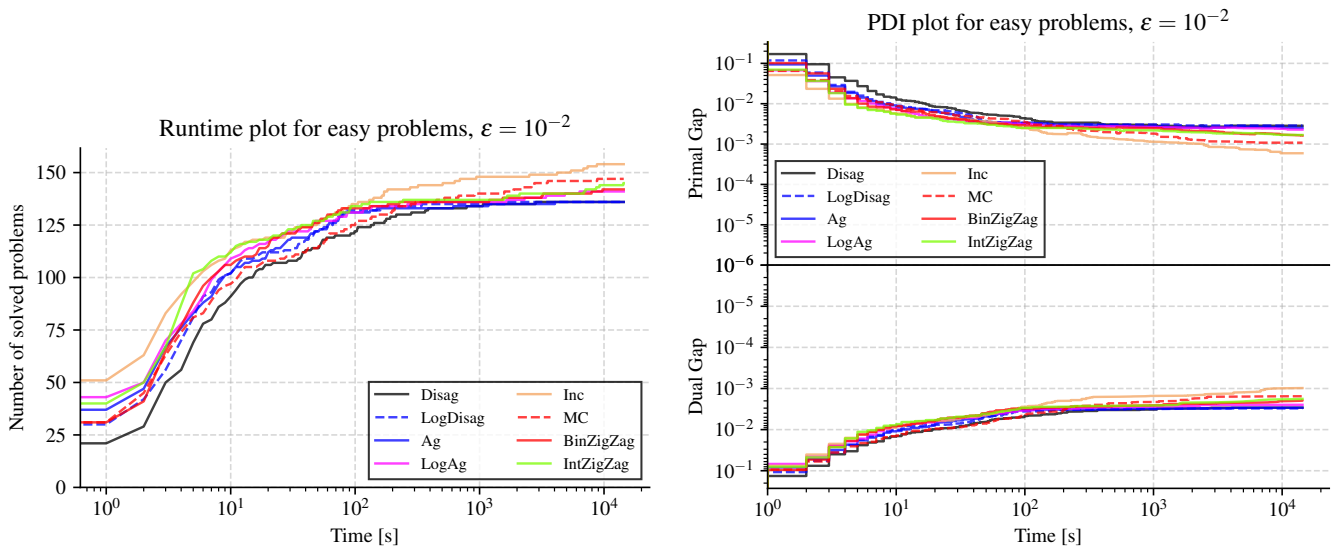


FIGURE C.13. Runtime and Primal-Dual Integral plots for easy problems, $\epsilon = 10^{-2}$

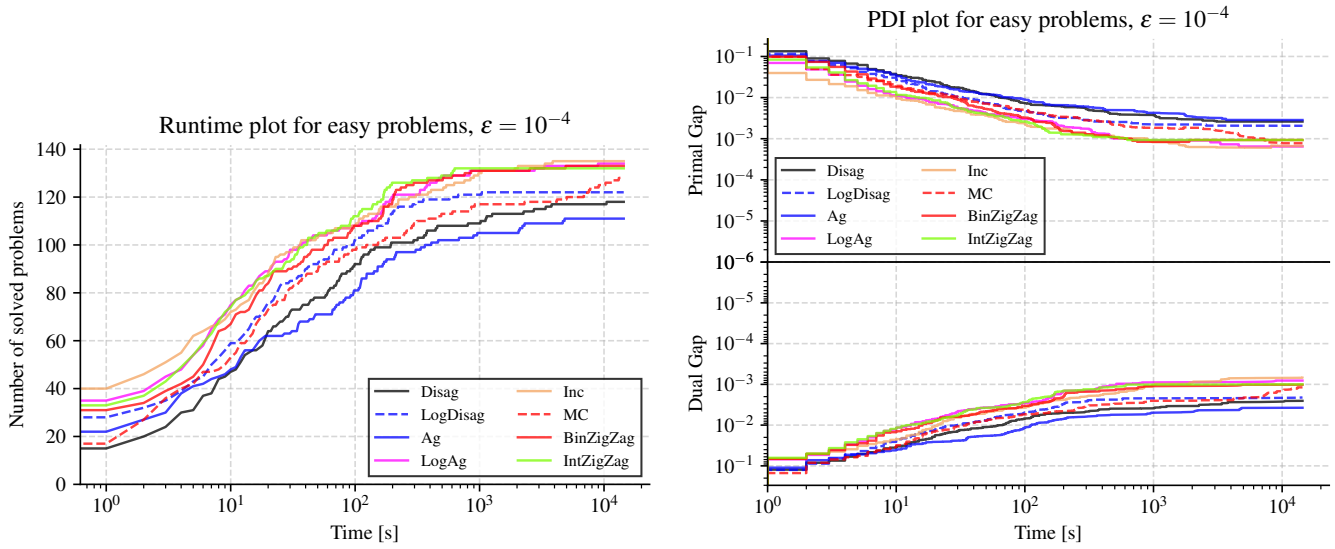


FIGURE C.14. Runtime and Primal-Dual Integral plots for easy problems, $\epsilon = 10^{-4}$

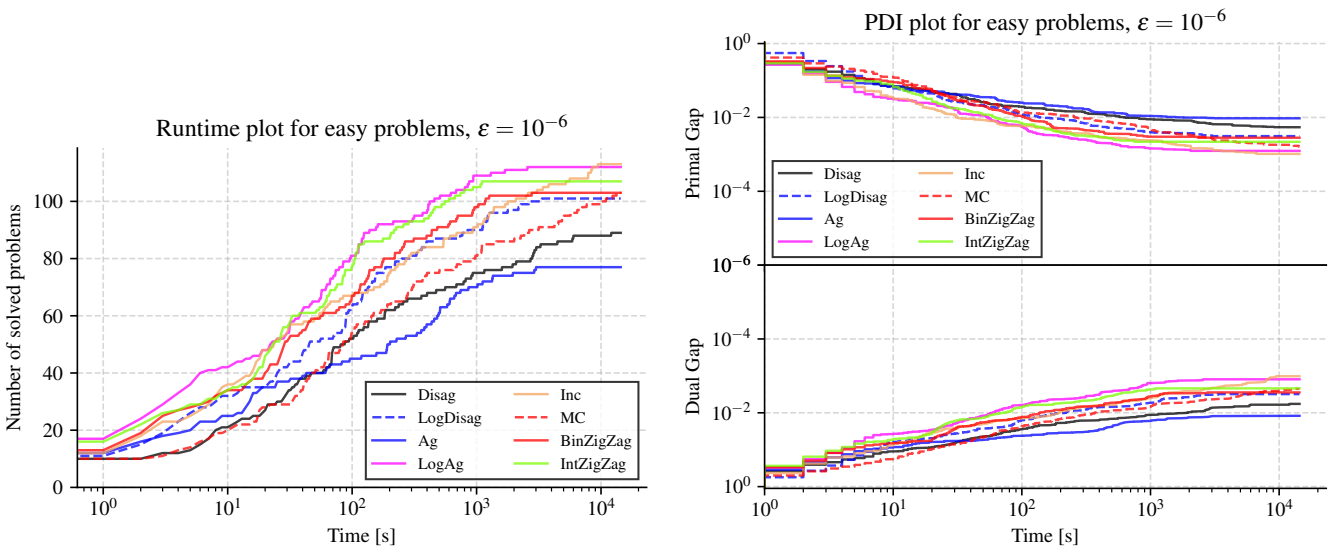


FIGURE C.15. Runtime and Primal-Dual Integral plots for easy problems, $\epsilon = 10^{-6}$

C.4. **Runtime plots for medium problems.** The following figures show the runtime plots for medium problems.

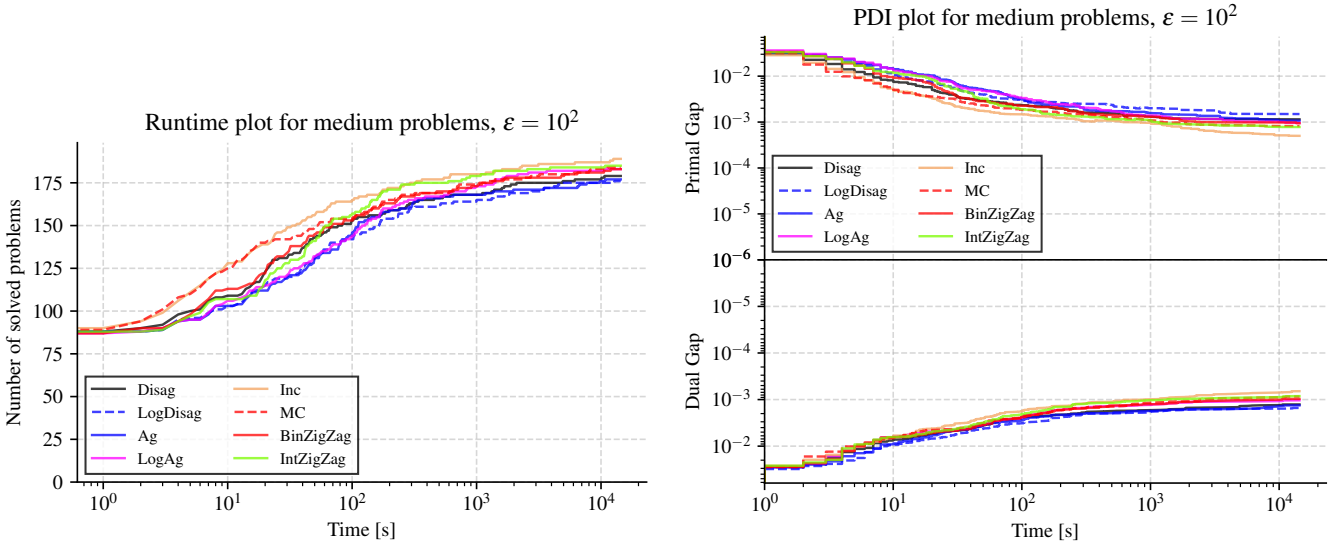


FIGURE C.16. Runtime and Primal-Dual Integral plots for medium problems, $\epsilon = 10^2$

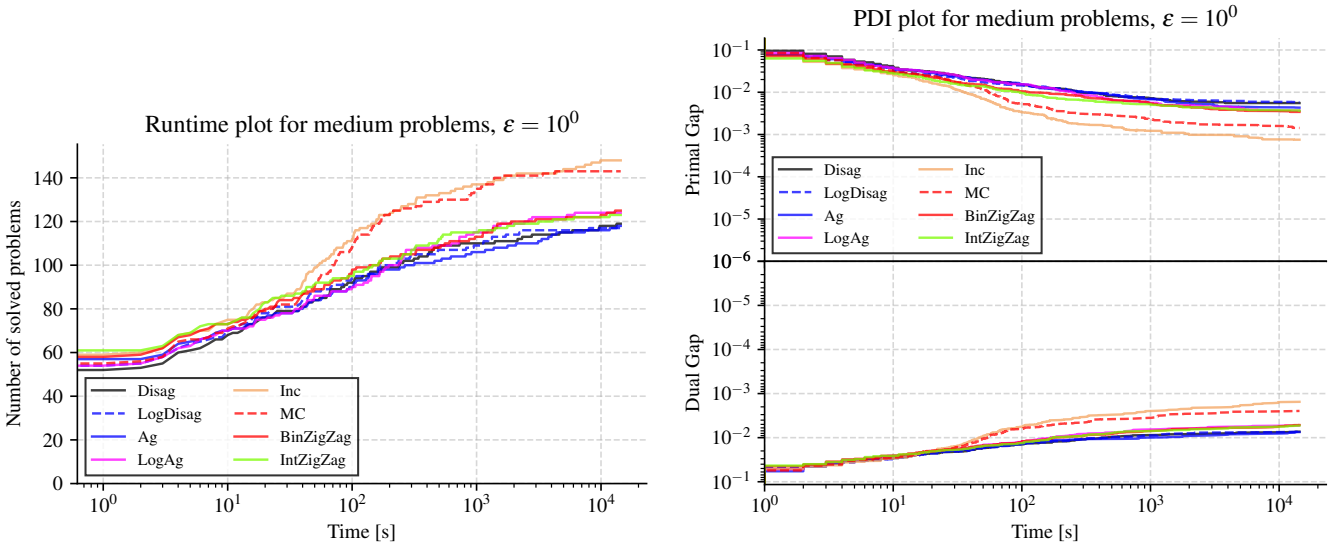


FIGURE C.17. Runtime and Primal-Dual Integral plots for medium problems, $\epsilon = 10^0$

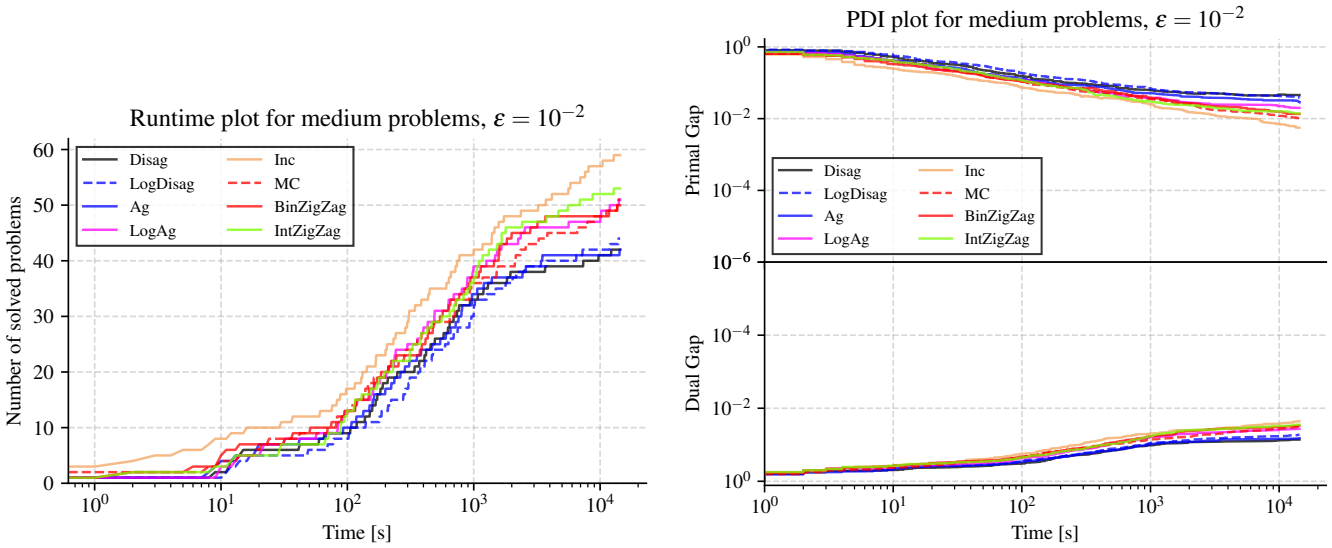


FIGURE C.18. Runtime and Primal-Dual Integral plots for medium problems, $\epsilon = 10^{-2}$

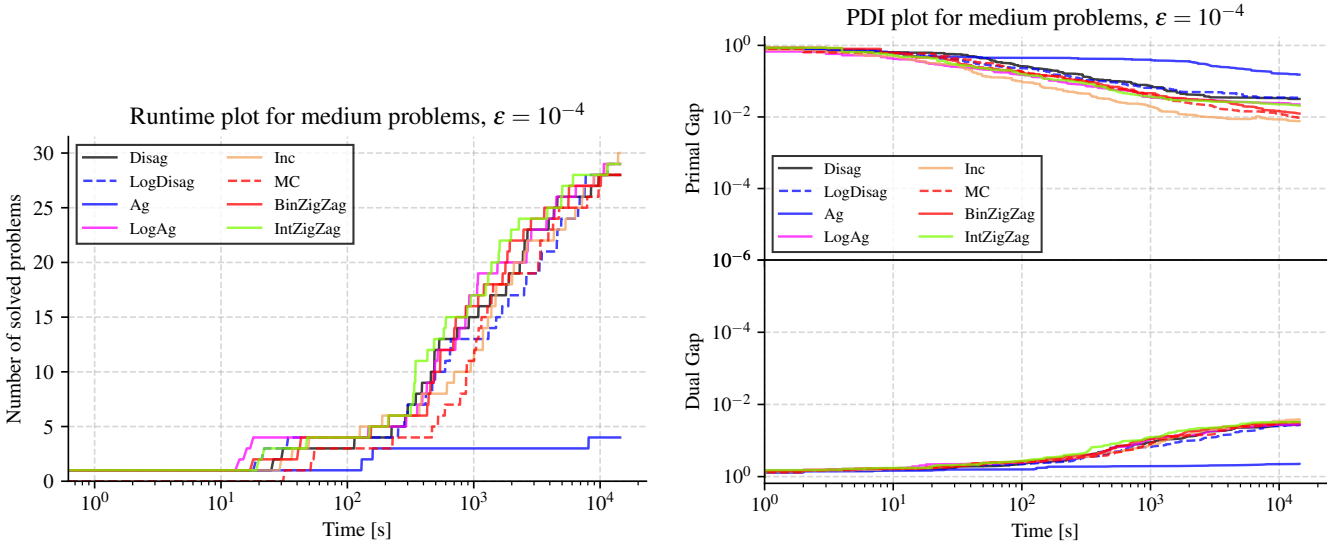


FIGURE C.19. Runtime and Primal-Dual Integral plots for medium problems, $\epsilon = 10^{-4}$

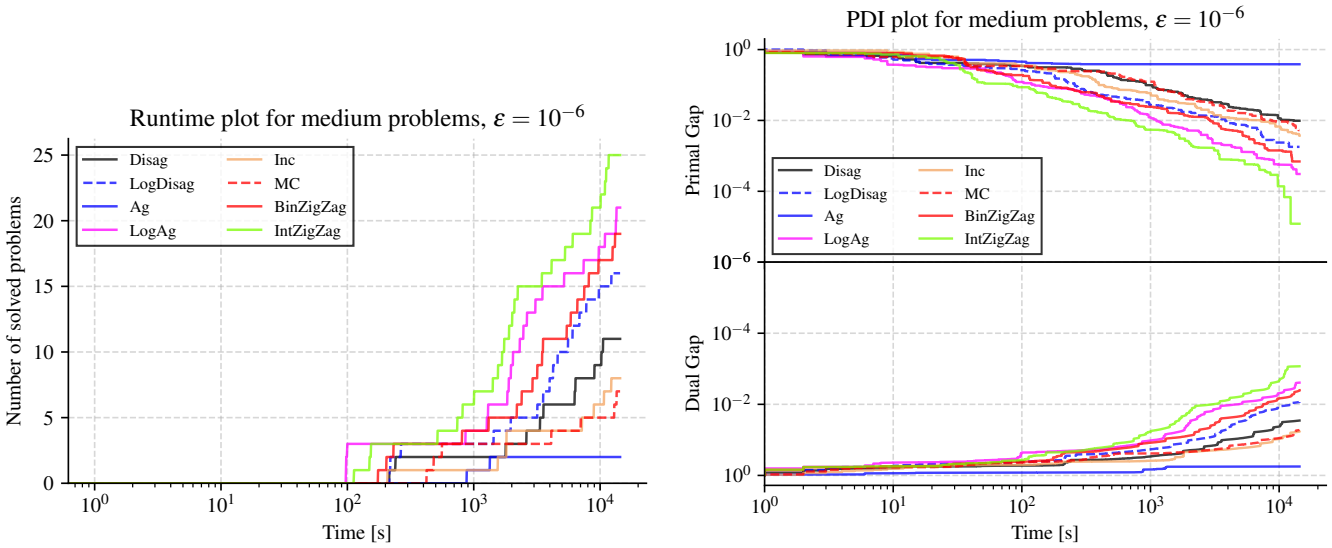


FIGURE C.20. Runtime and Primal-Dual Integral plots for medium problems, $\epsilon = 10^{-6}$

C.5. **Runtime plots for hard problems.** The following figures show the runtime plots for hard problems.

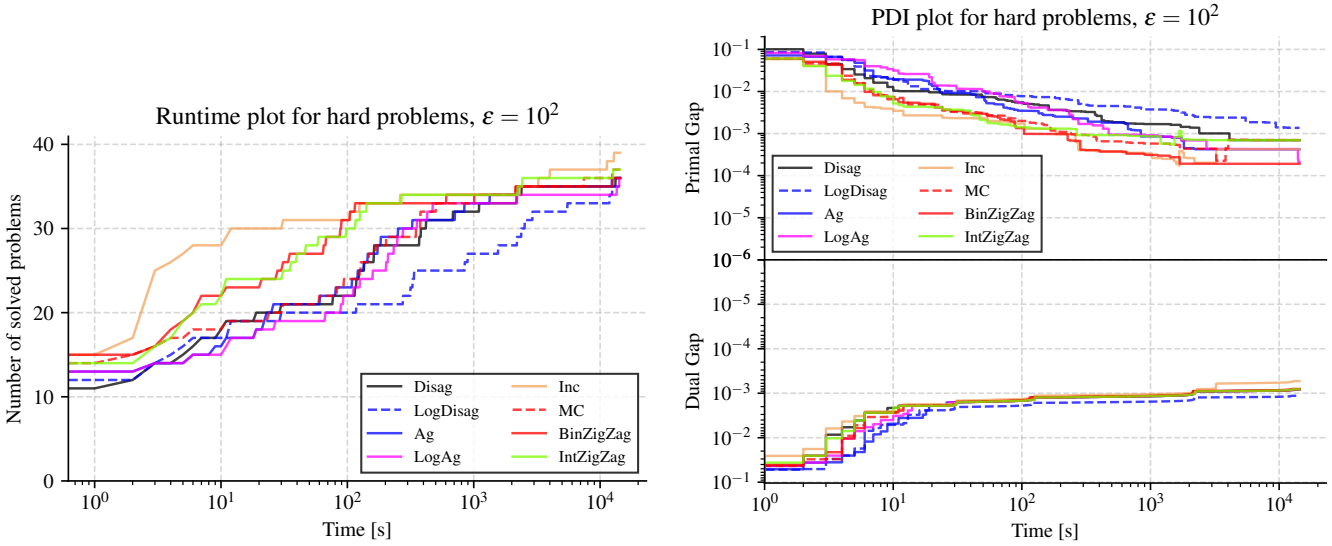


FIGURE C.21. Runtime and Primal-Dual Integral plots for hard problems, $\epsilon = 10^2$

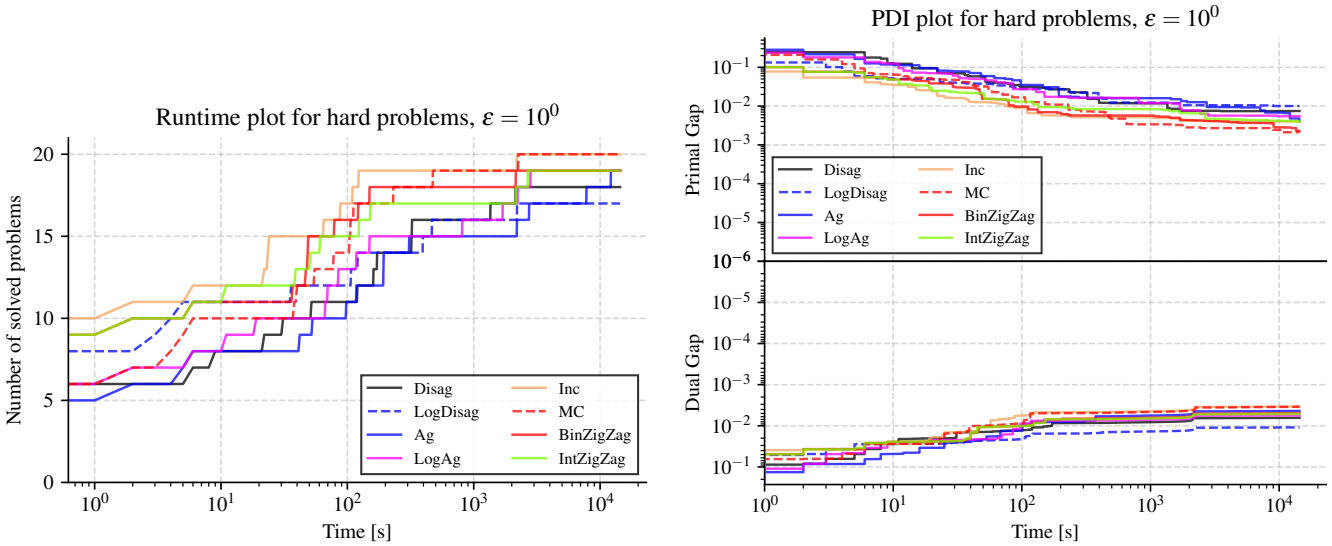


FIGURE C.22. Runtime and Primal-Dual Integral plots for hard problems, $\epsilon = 10^0$

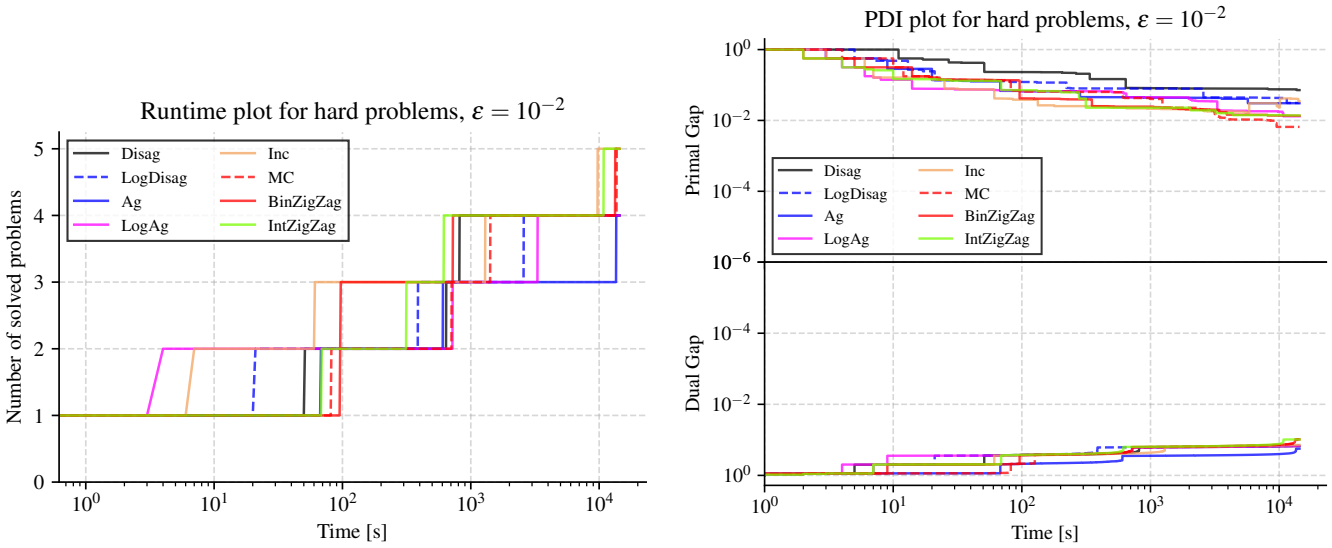


FIGURE C.23. Runtime and Primal-Dual Integral plots for hard problems, $\epsilon = 10^{-2}$

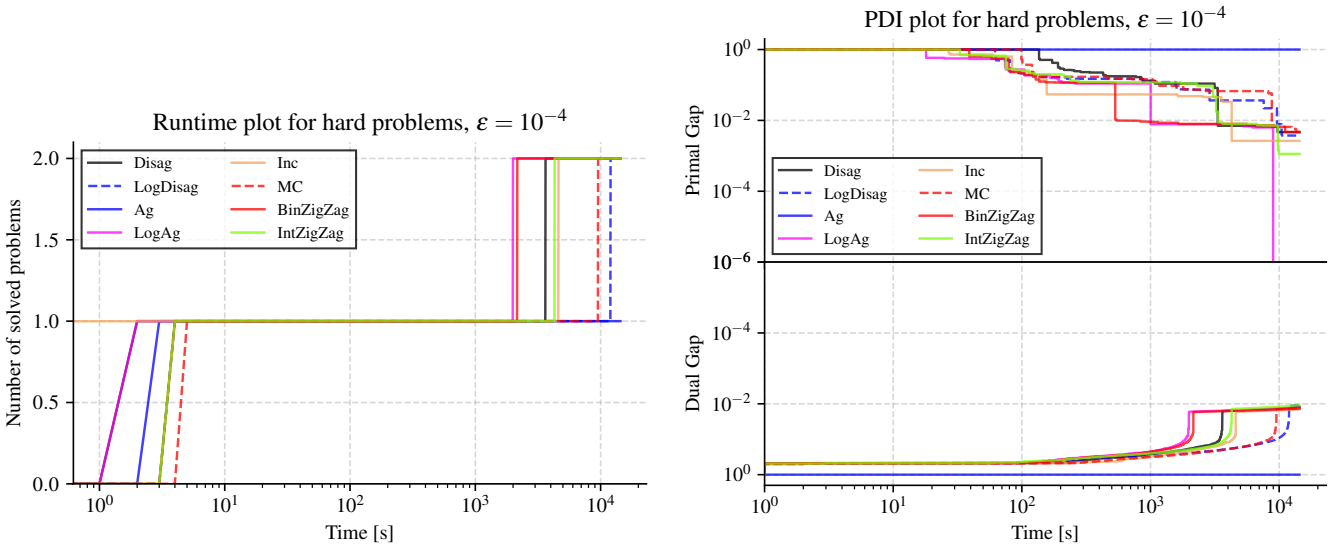


FIGURE C.24. Runtime and Primal-Dual Integral plots for hard problems, $\epsilon = 10^{-4}$

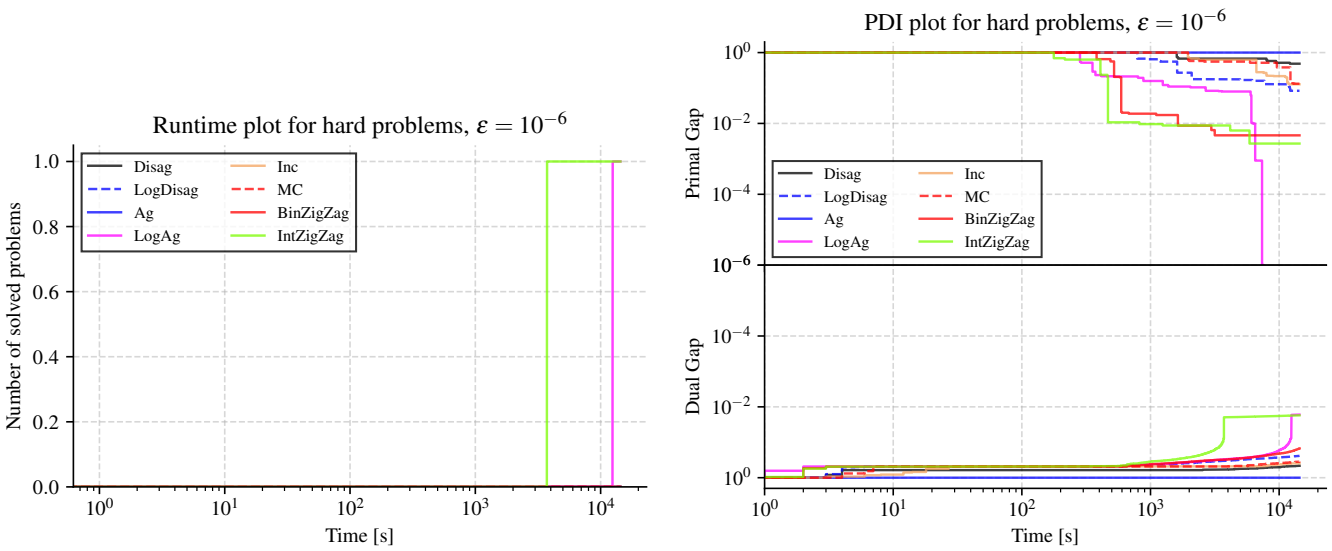


FIGURE C.25. Runtime and Primal-Dual Integral plots for hard problems, $\epsilon = 10^{-6}$

C.6. **Runtime plots for quadratic problems.** The following figures show the runtime plots for quadratic problems.

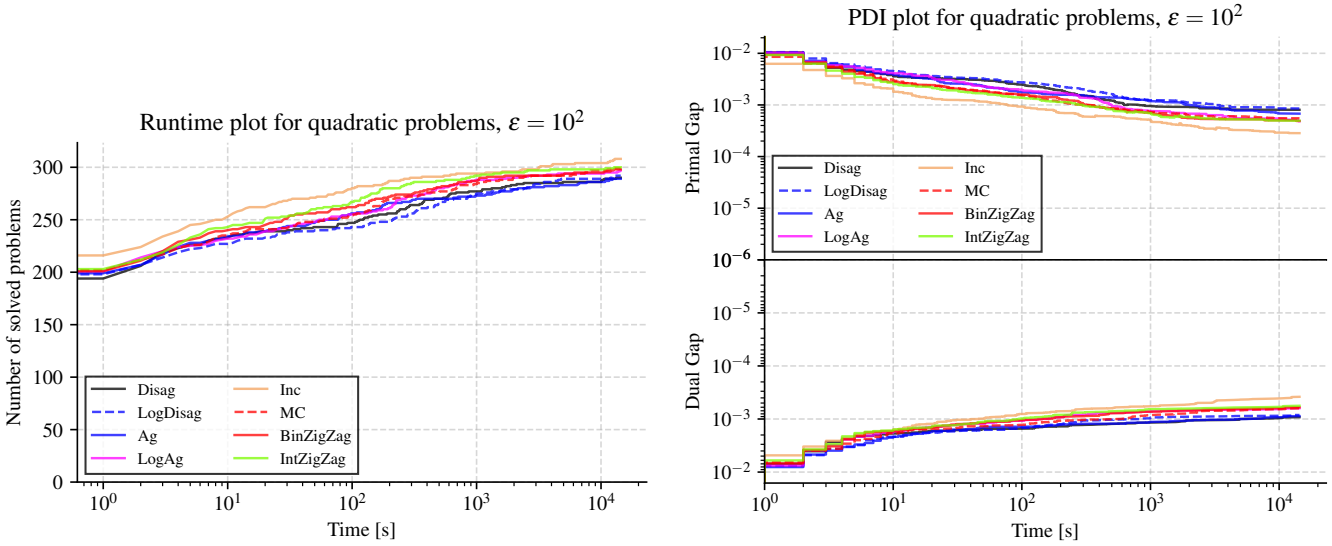


FIGURE C.26. Runtime and Primal-Dual Integral plots for quadratic problems, $\epsilon = 10^2$

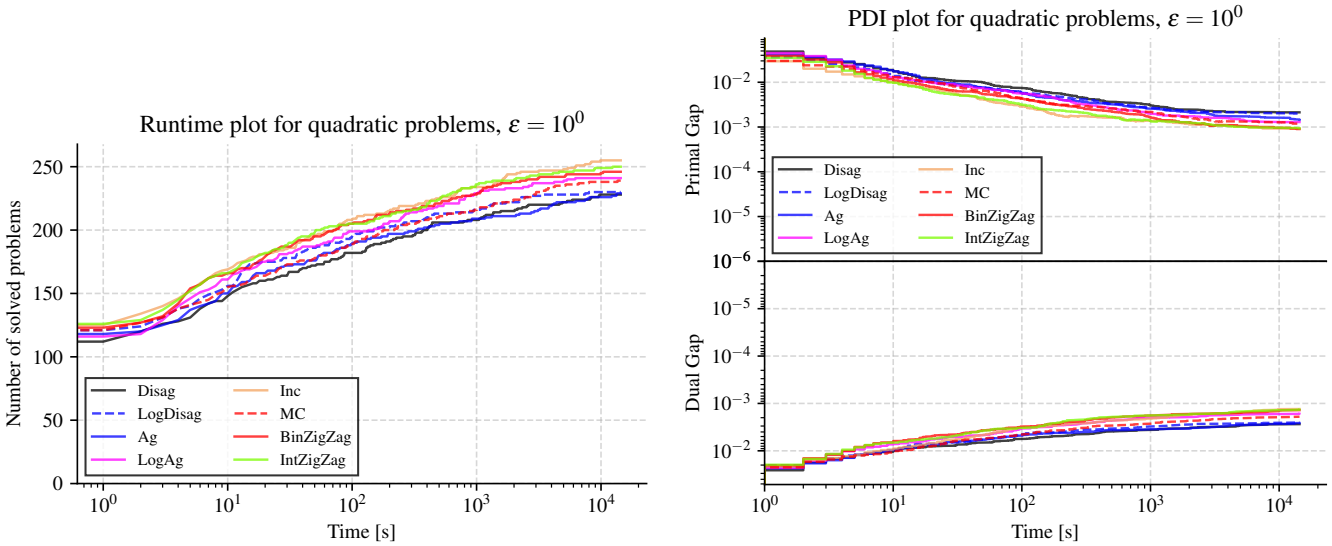


FIGURE C.27. Runtime and Primal-Dual Integral plots for quadratic problems, $\epsilon = 10^0$

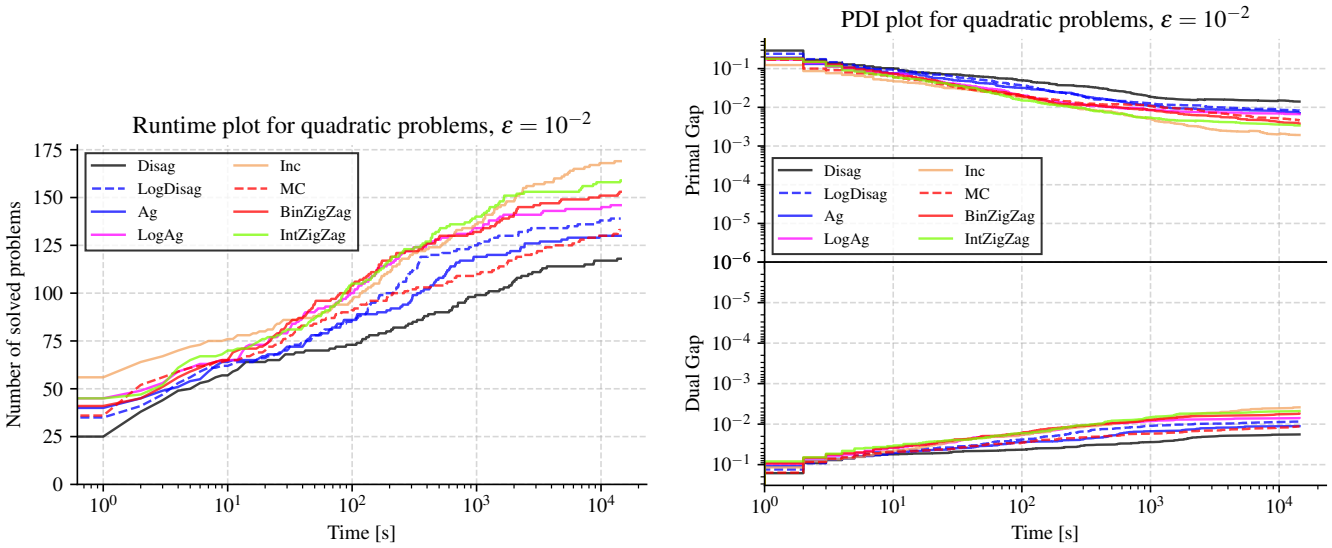


FIGURE C.28. Runtime and Primal-Dual Integral plots for quadratic problems, $\epsilon = 10^{-2}$

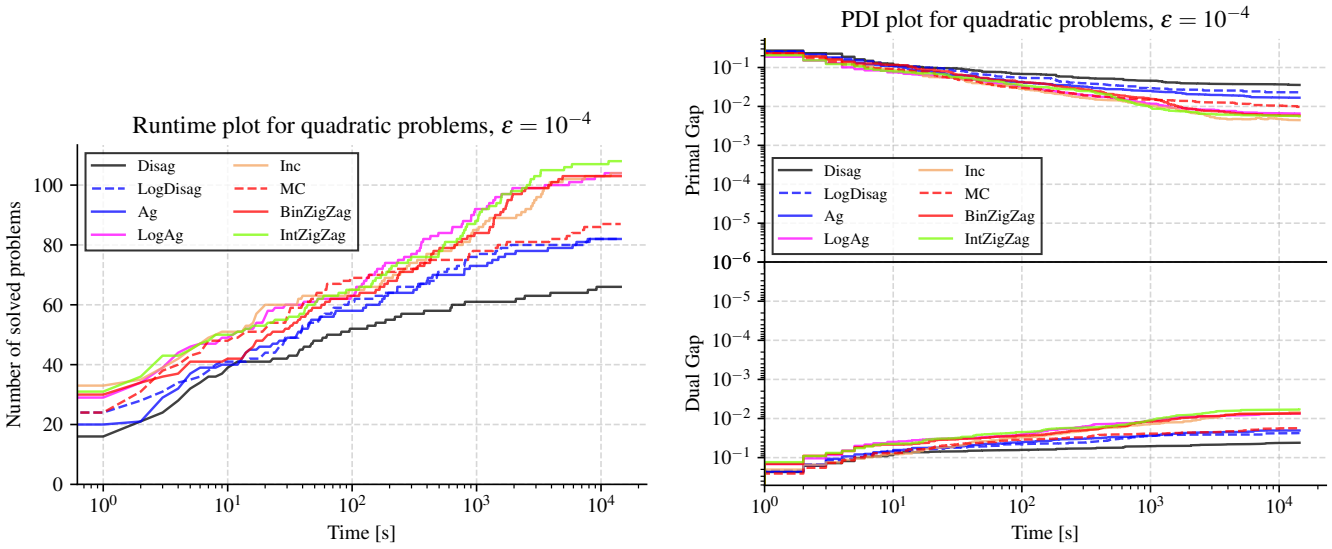


FIGURE C.29. Runtime and Primal-Dual Integral plots for quadratic problems, $\epsilon = 10^{-4}$

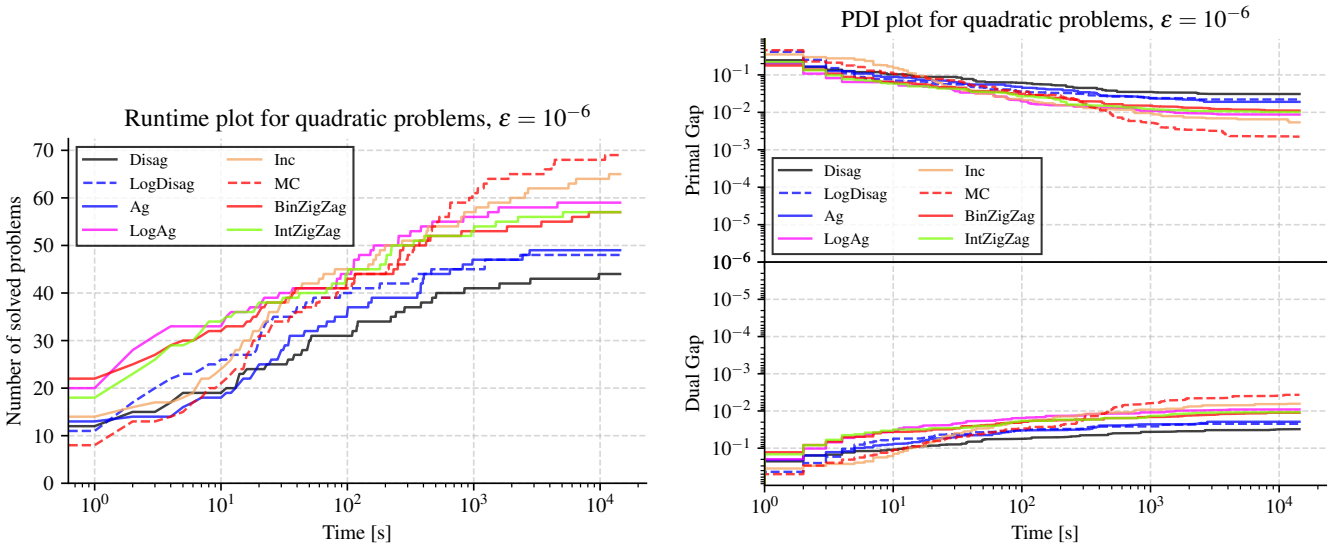


FIGURE C.30. Runtime and Primal-Dual Integral plots for quadratic problems, $\epsilon = 10^{-6}$

C.7. **Runtime plots for polynomial problems.** The following figures show the runtime plots for polynomial problems.

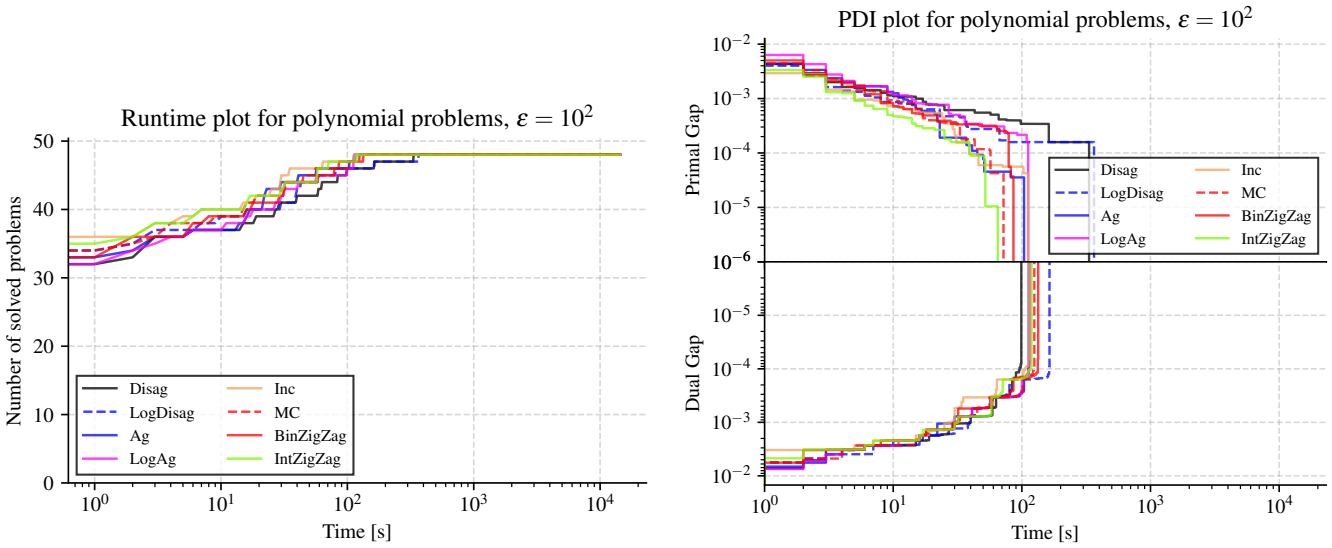


FIGURE C.31. Runtime and Primal-Dual Integral plots for polynomial problems, $\epsilon = 10^2$

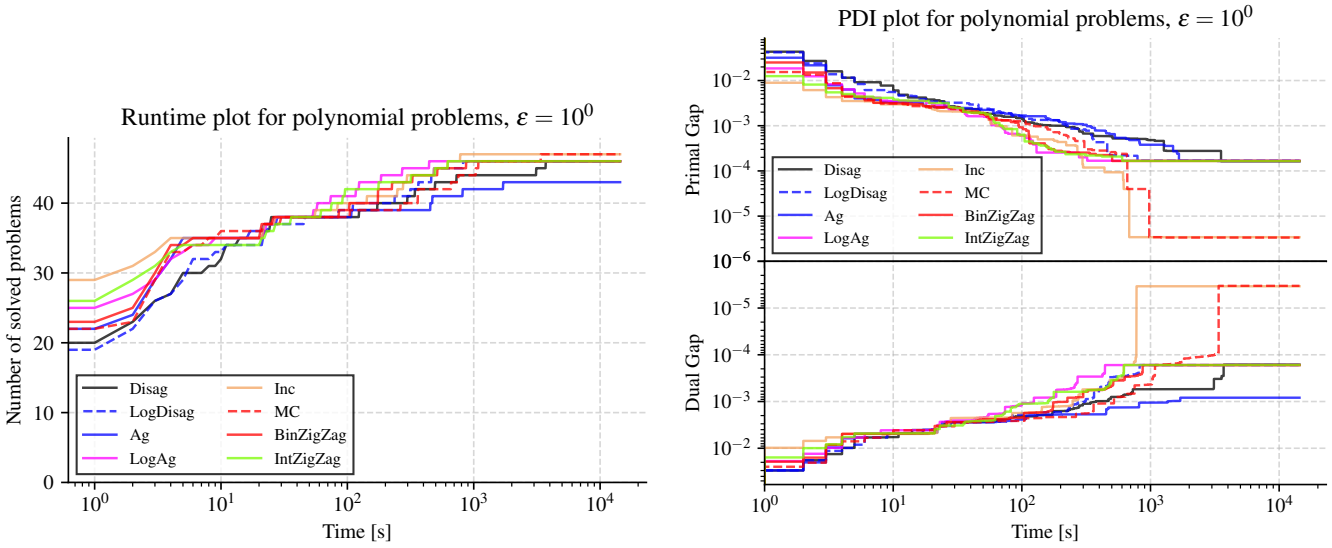


FIGURE C.32. Runtime and Primal-Dual Integral plots for polynomial problems, $\epsilon = 10^0$

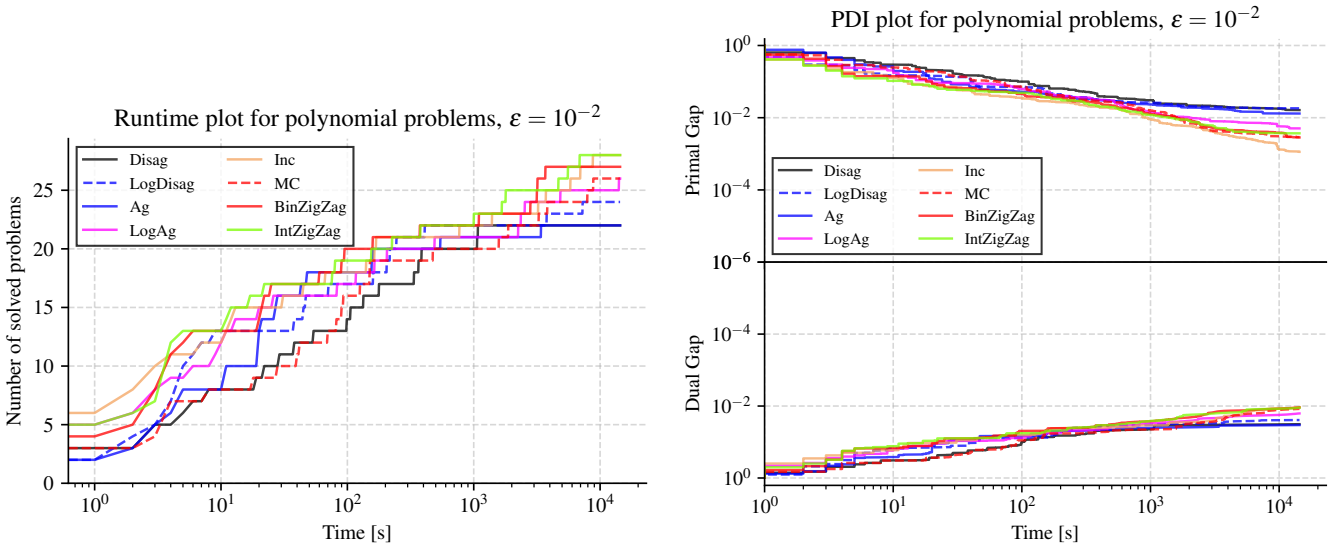


FIGURE C.33. Runtime and Primal-Dual Integral plots for polynomial problems, $\epsilon = 10^{-2}$

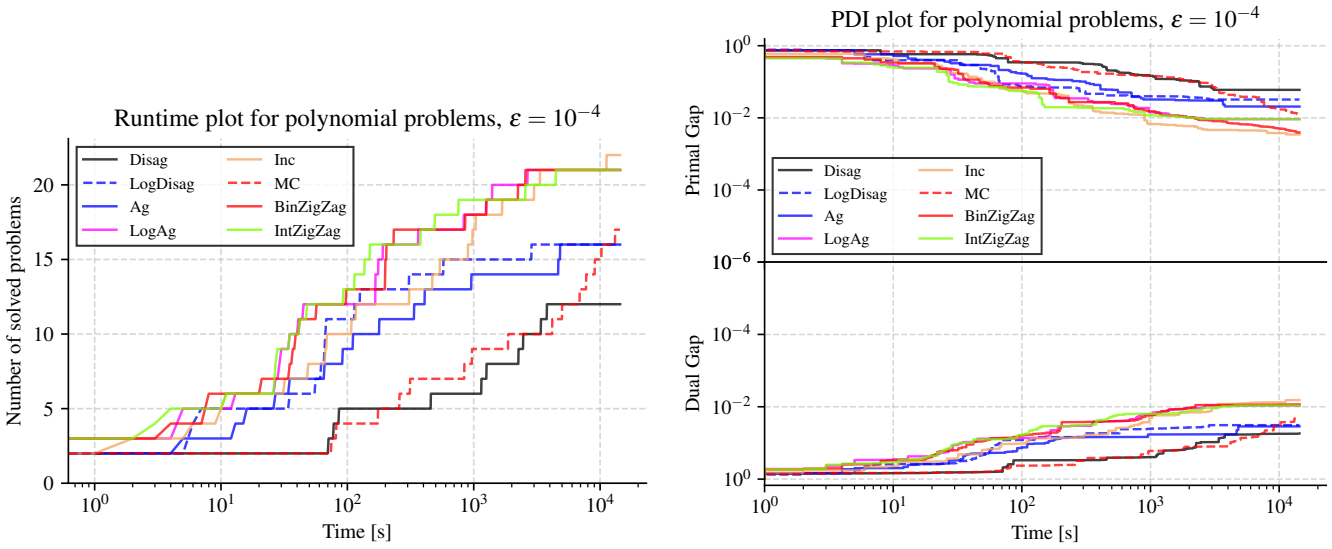


FIGURE C.34. Runtime and Primal-Dual Integral plots for polynomial problems, $\epsilon = 10^{-4}$

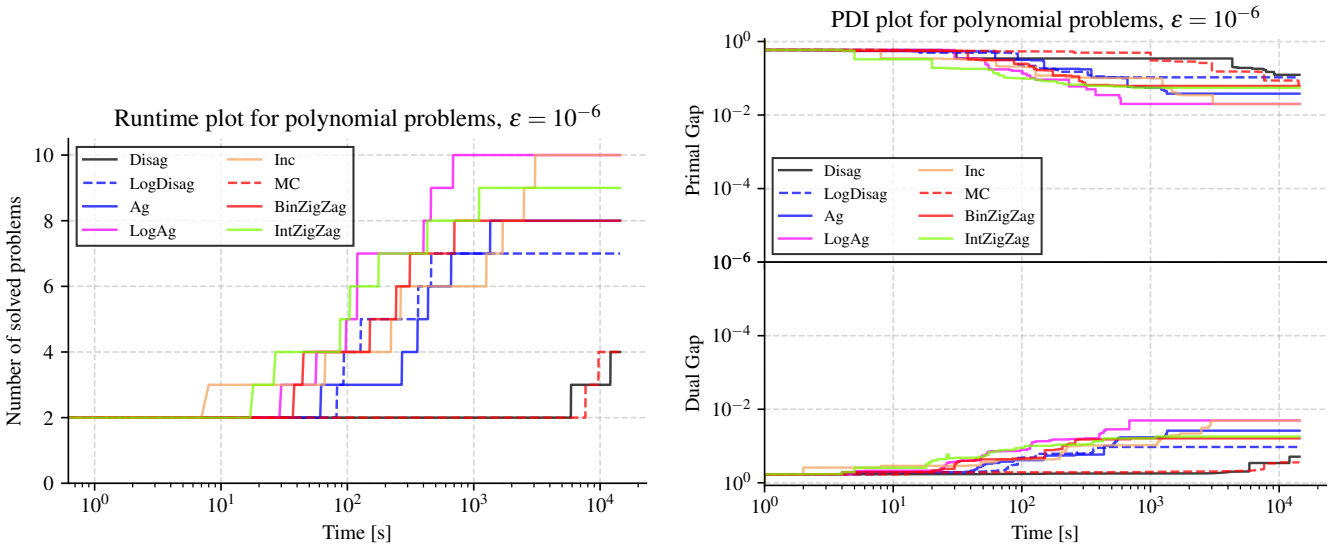


FIGURE C.35. Runtime and Primal-Dual Integral plots for polynomial problems, $\epsilon = 10^{-6}$

C.8. **Runtime plots for signomial problems.** The following figures show the runtime plots for signomial problems.

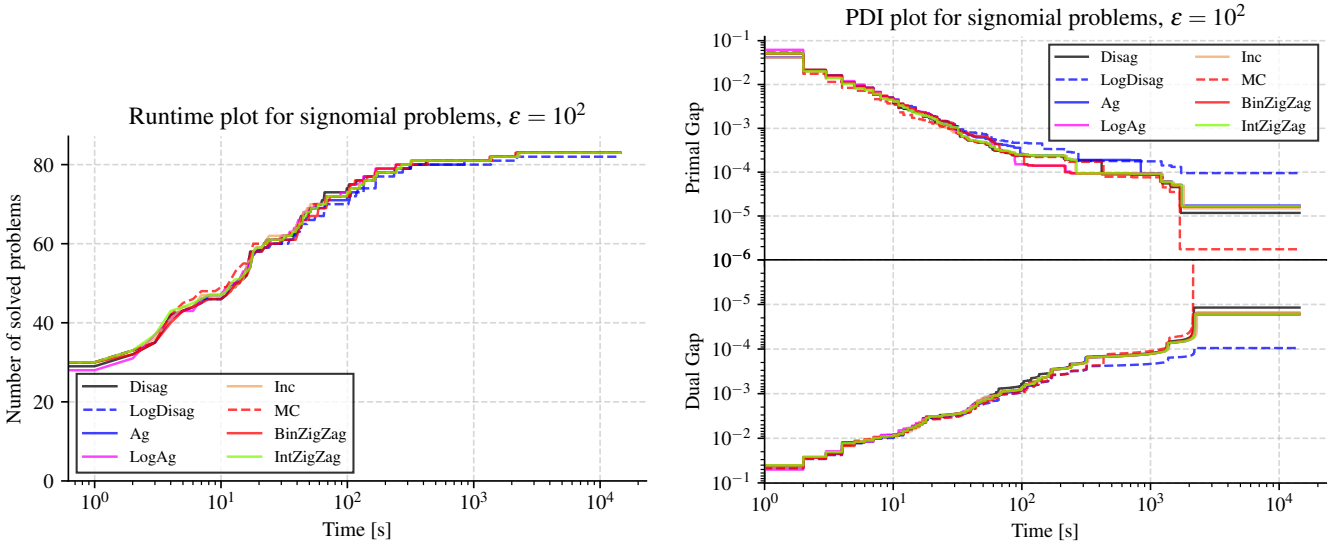


FIGURE C.36. Runtime and Primal-Dual Integral plots for signomial problems, $\epsilon = 10^2$

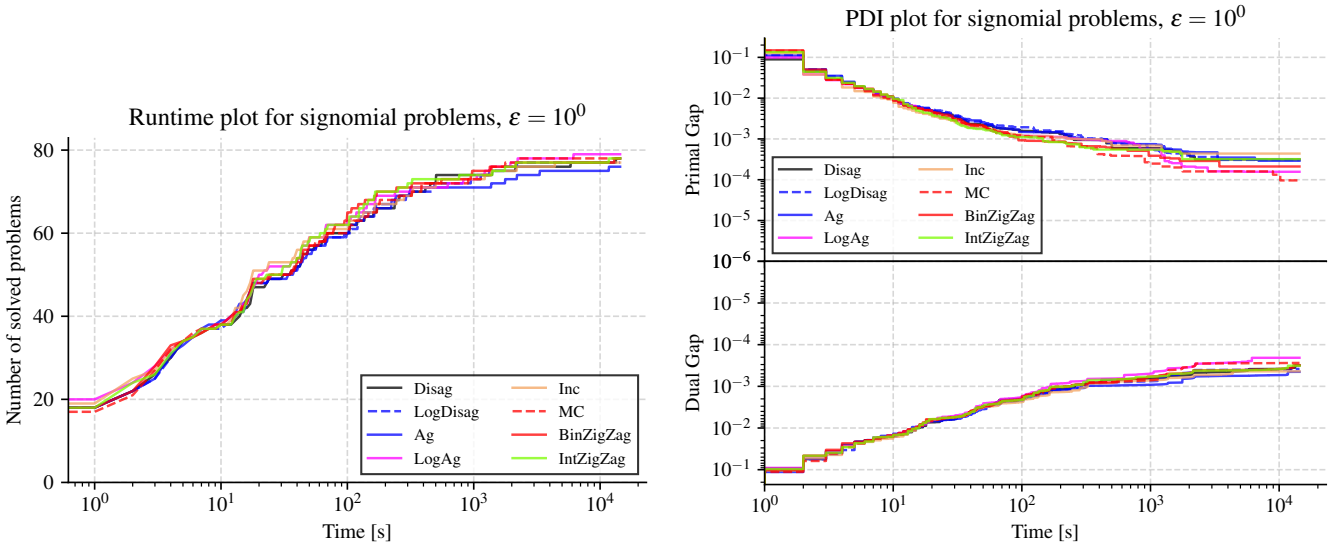


FIGURE C.37. Runtime and Primal-Dual Integral plots for signomial problems, $\epsilon = 10^0$

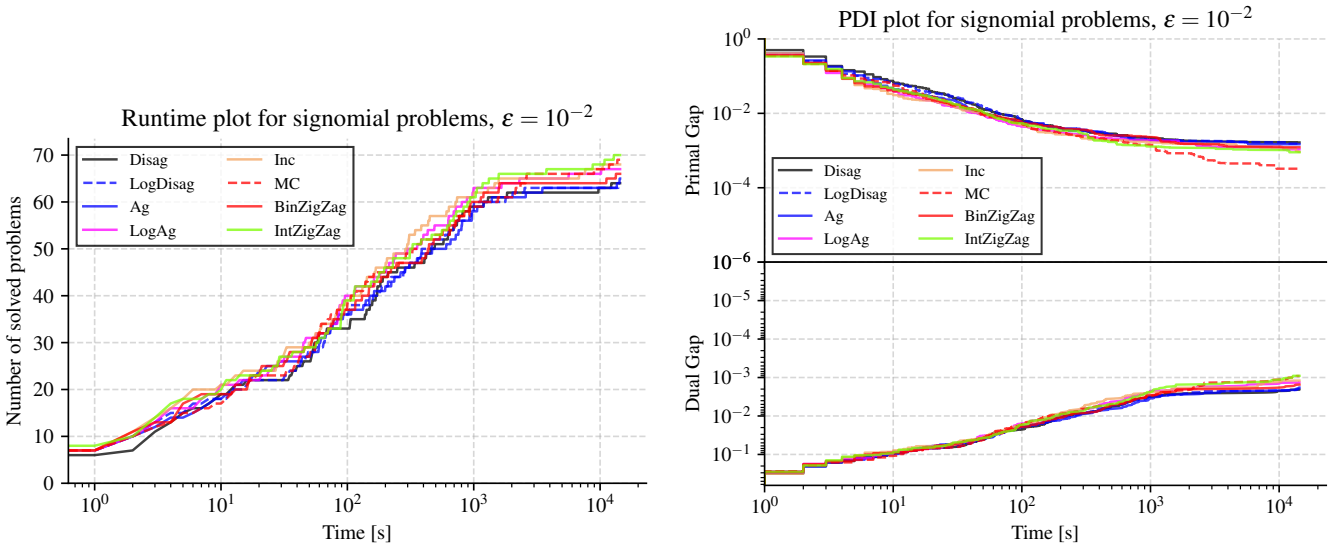


FIGURE C.38. Runtime and Primal-Dual Integral plots for signomial problems, $\epsilon = 10^{-2}$

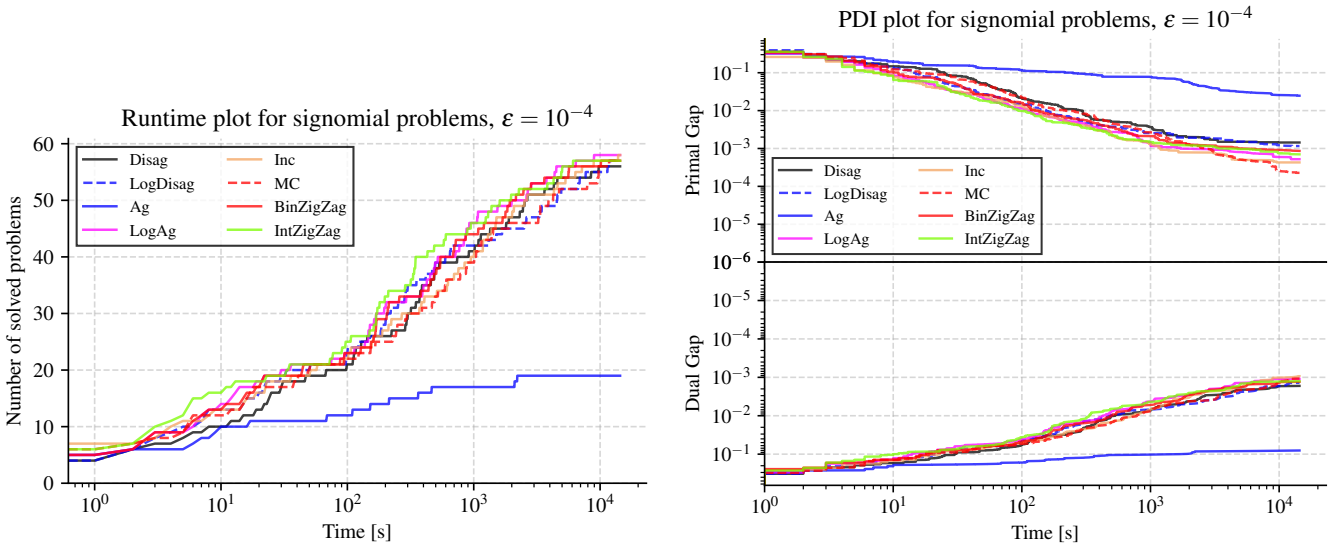


FIGURE C.39. Runtime and Primal-Dual Integral plots for signomial problems, $\epsilon = 10^{-4}$

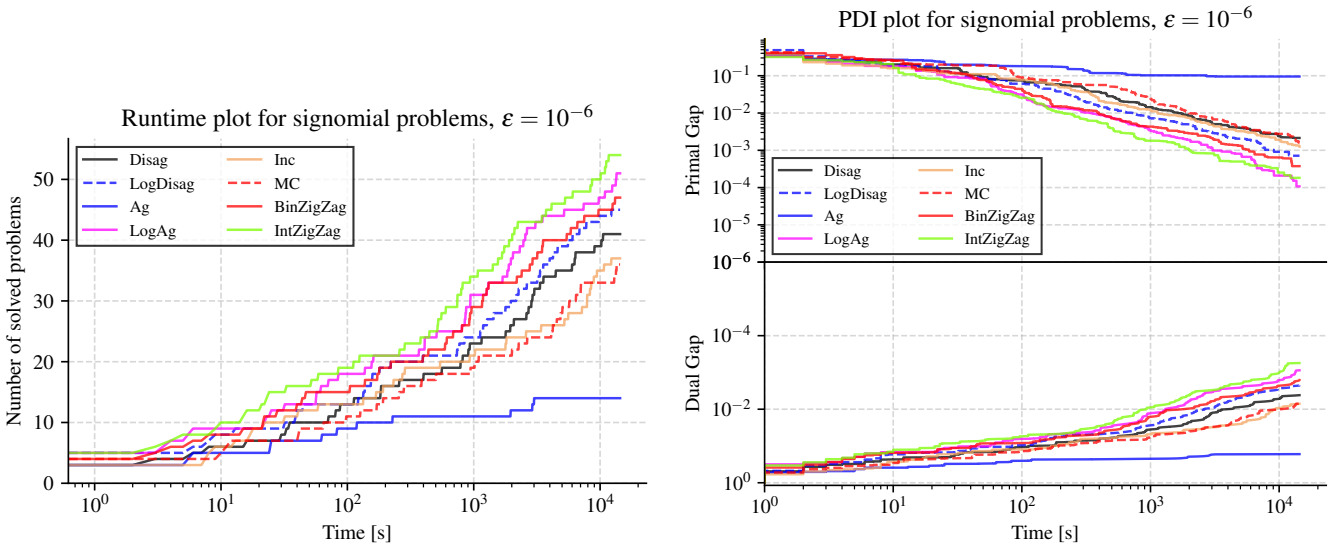


FIGURE C.40. Runtime and Primal-Dual Integral plots for signomial problems, $\epsilon = 10^{-6}$

C.9. **Runtime plots for general nonlinear problems.** The following figures show the runtime plots for general nonlinear problems.

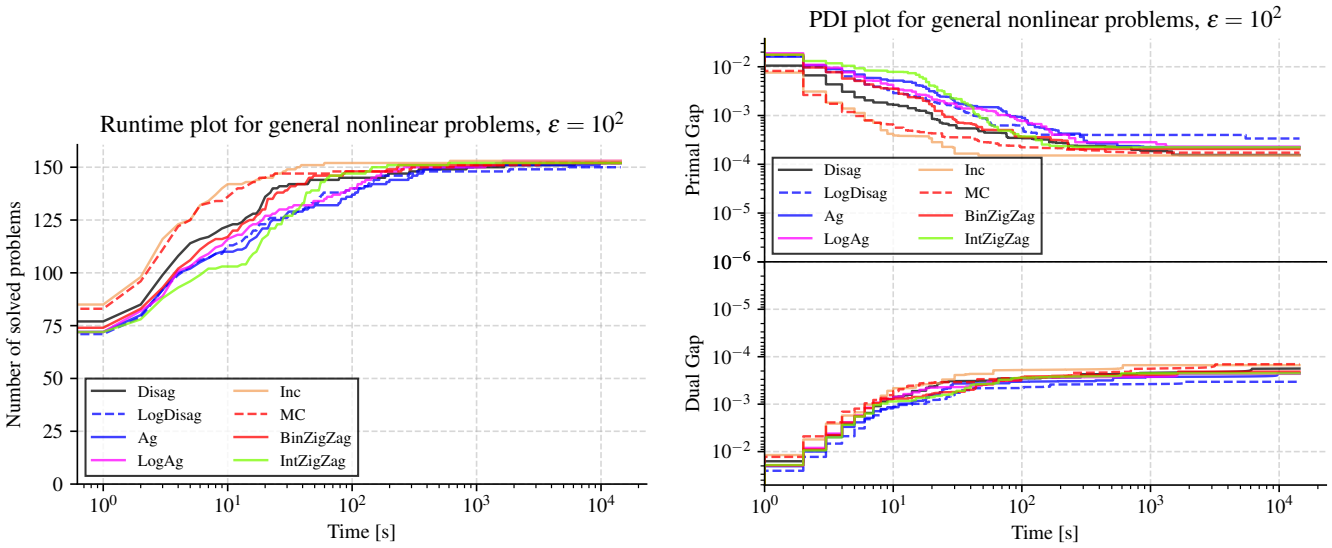


FIGURE C.41. Runtime and Primal-Dual Integral plots for general nonlinear problems, $\epsilon = 10^2$

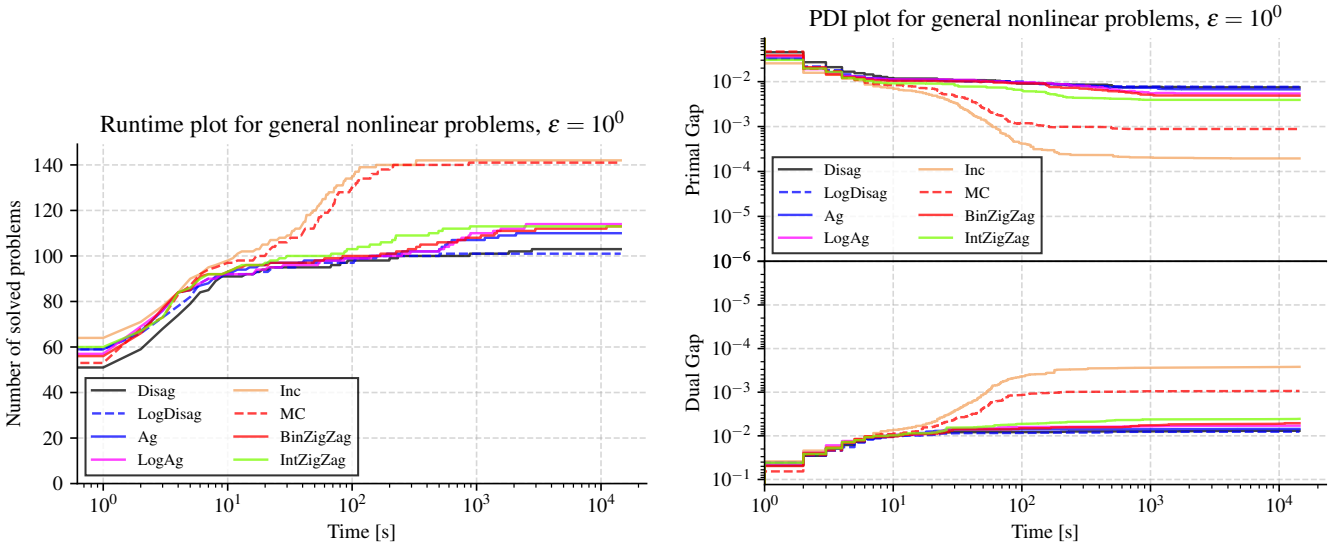


FIGURE C.42. Runtime and Primal-Dual Integral plots for general nonlinear problems, $\epsilon = 10^0$

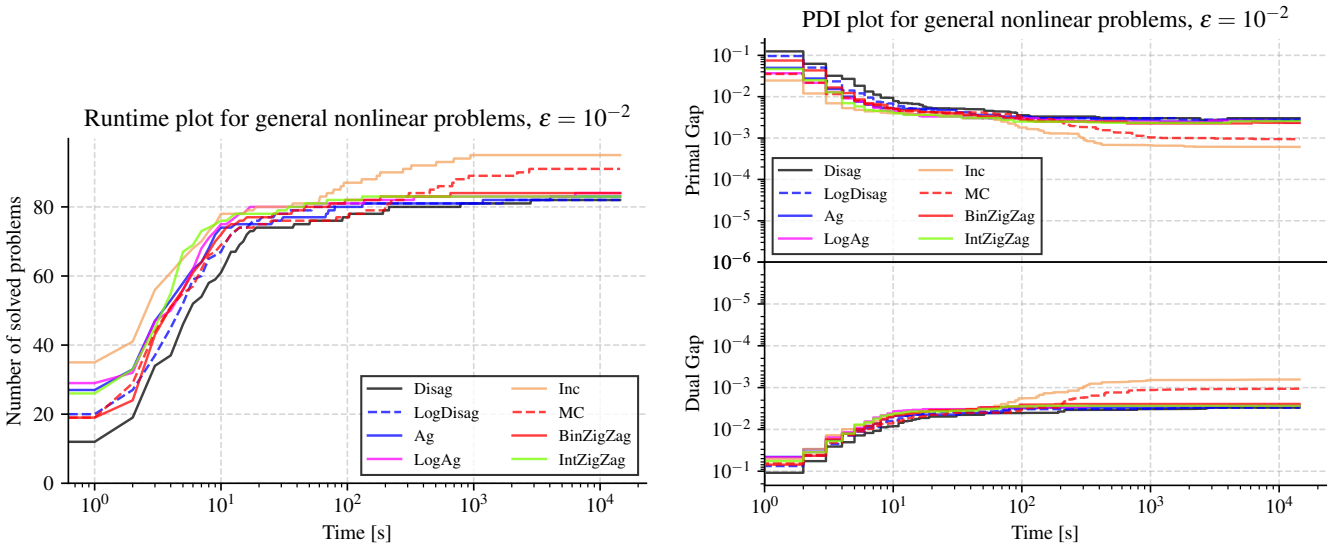


FIGURE C.43. Runtime and Primal-Dual Integral plots for general nonlinear problems, $\epsilon = 10^{-2}$

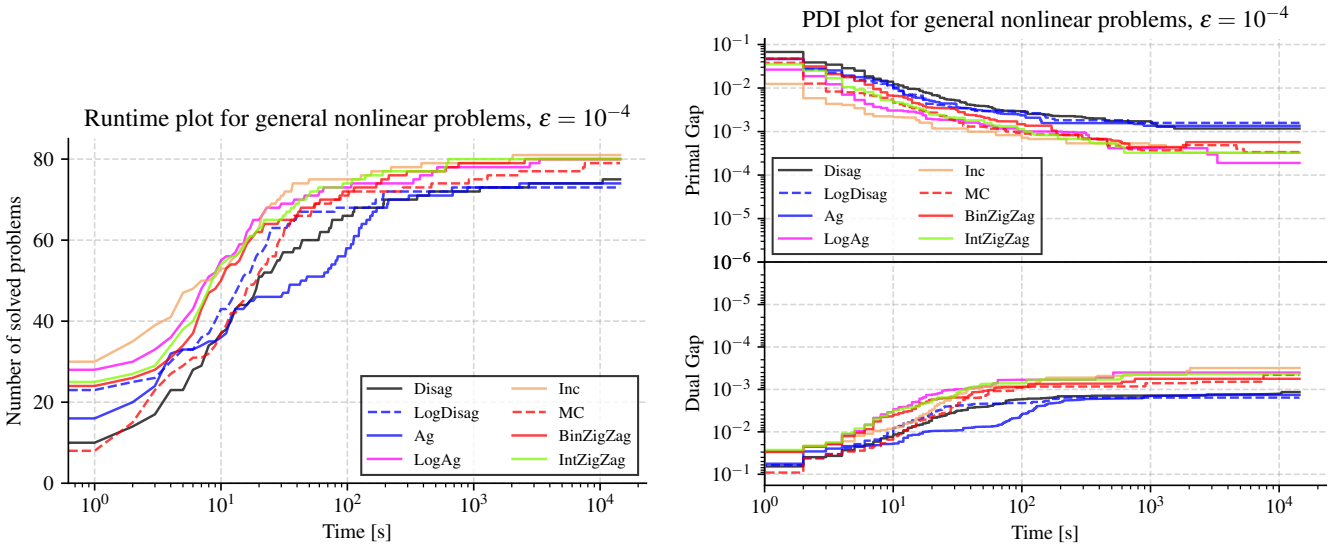


FIGURE C.44. Runtime and Primal-Dual Integral plots for general nonlinear problems, $\epsilon = 10^{-4}$

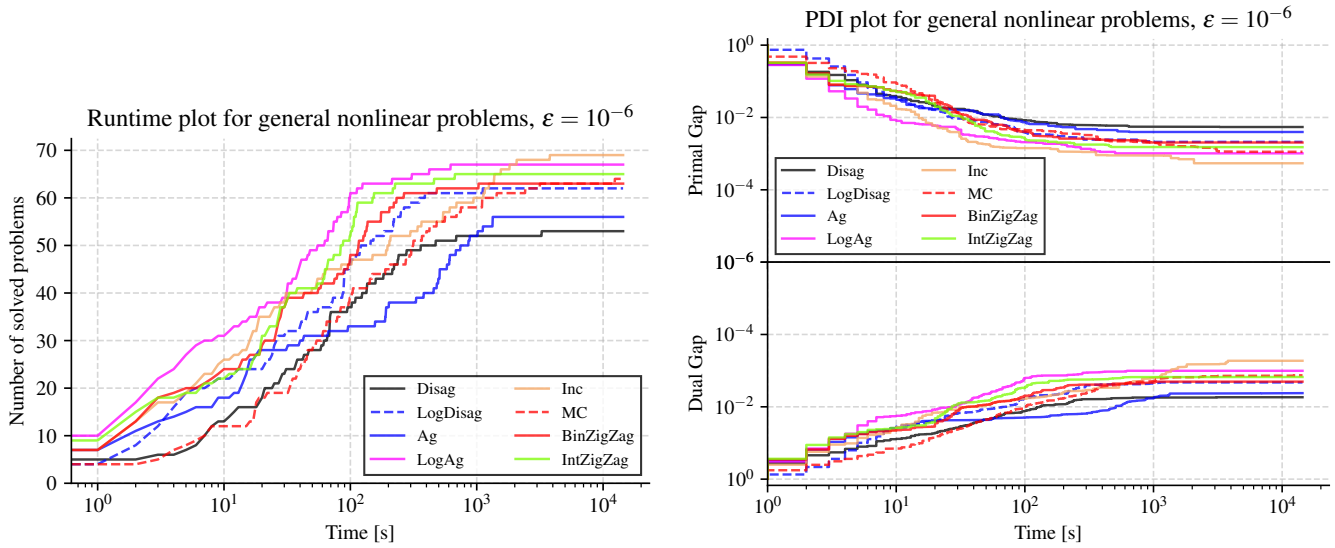


FIGURE C.45. Runtime and Primal-Dual Integral plots for general nonlinear problems, $\epsilon = 10^{-6}$

C.10. **Runtime plots for convex problems.** The following figures show the runtime plots for convex problems.

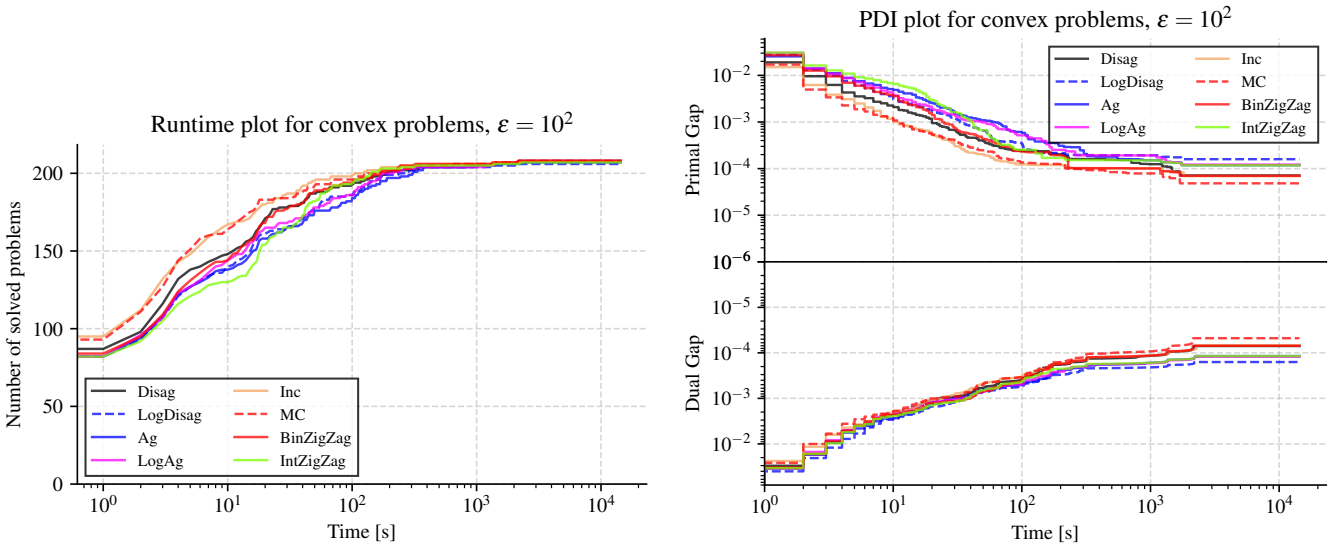


FIGURE C.46. Runtime and Primal-Dual Integral plots for convex problems, $\epsilon = 10^2$

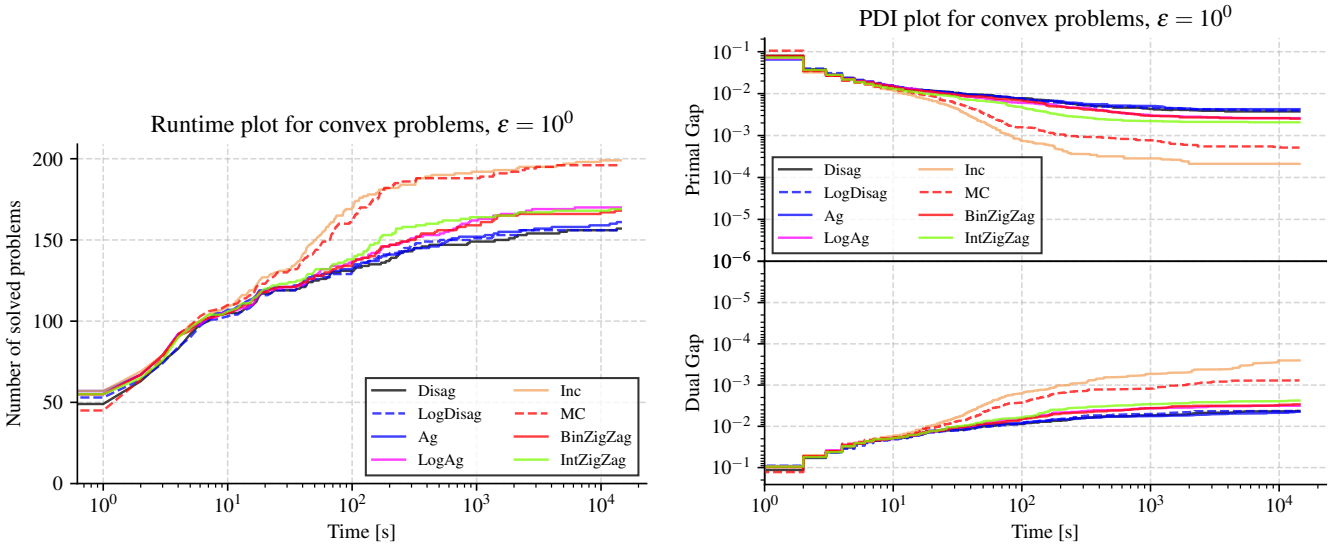


FIGURE C.47. Runtime and Primal-Dual Integral plots for convex problems, $\epsilon = 10^0$

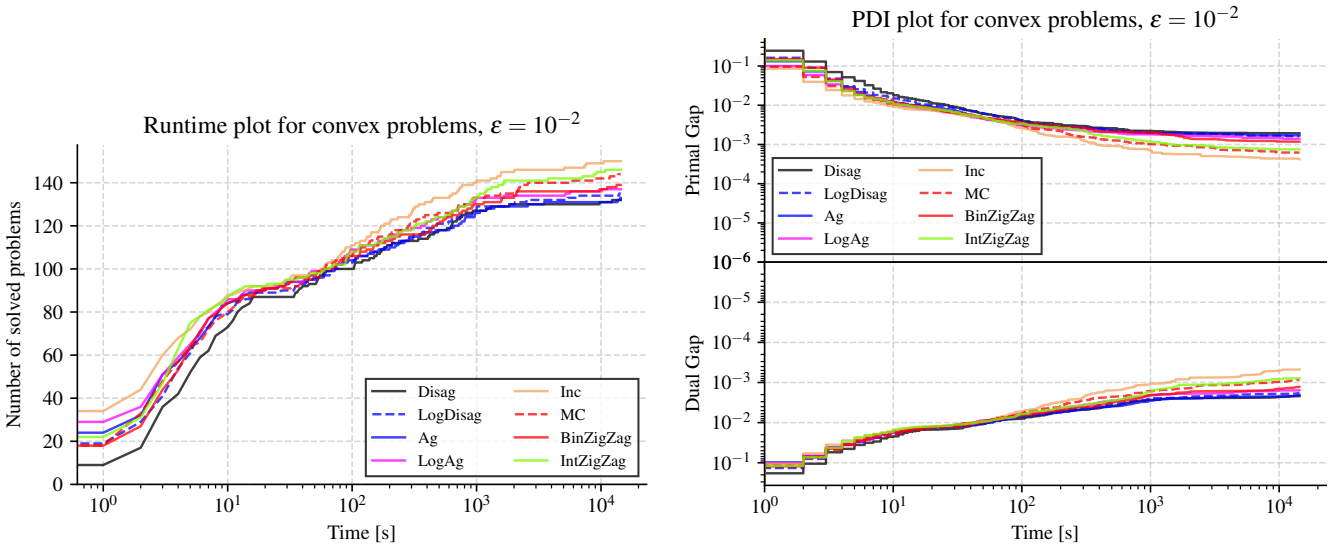


FIGURE C.48. Runtime and Primal-Dual Integral plots for convex problems, $\epsilon = 10^{-2}$

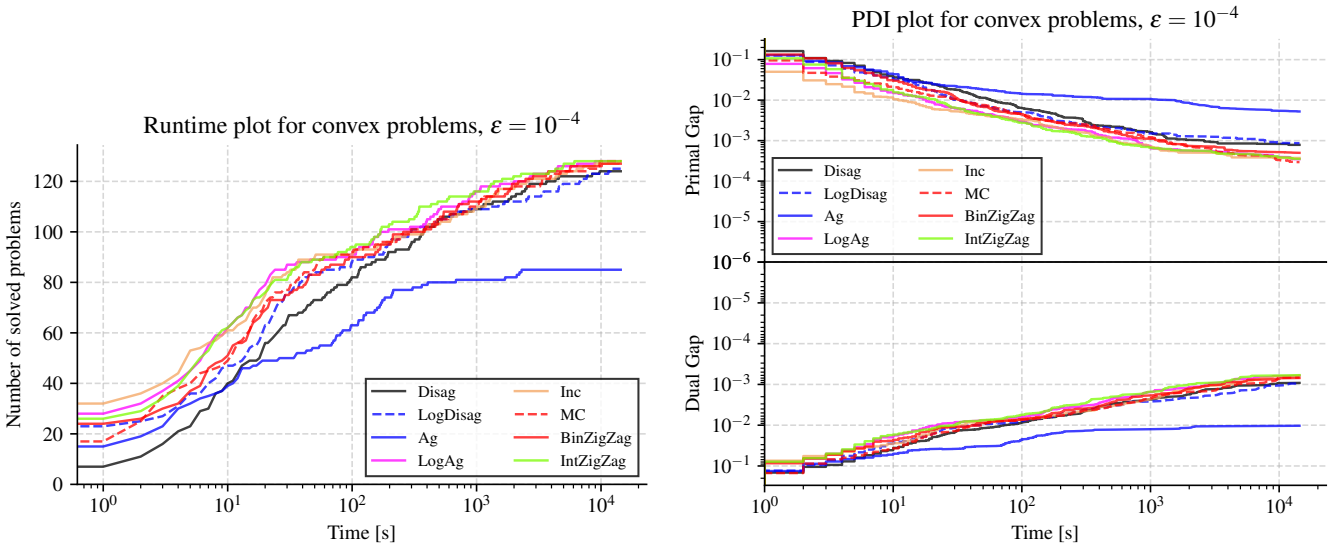


FIGURE C.49. Runtime and Primal-Dual Integral plots for convex problems, $\epsilon = 10^{-4}$

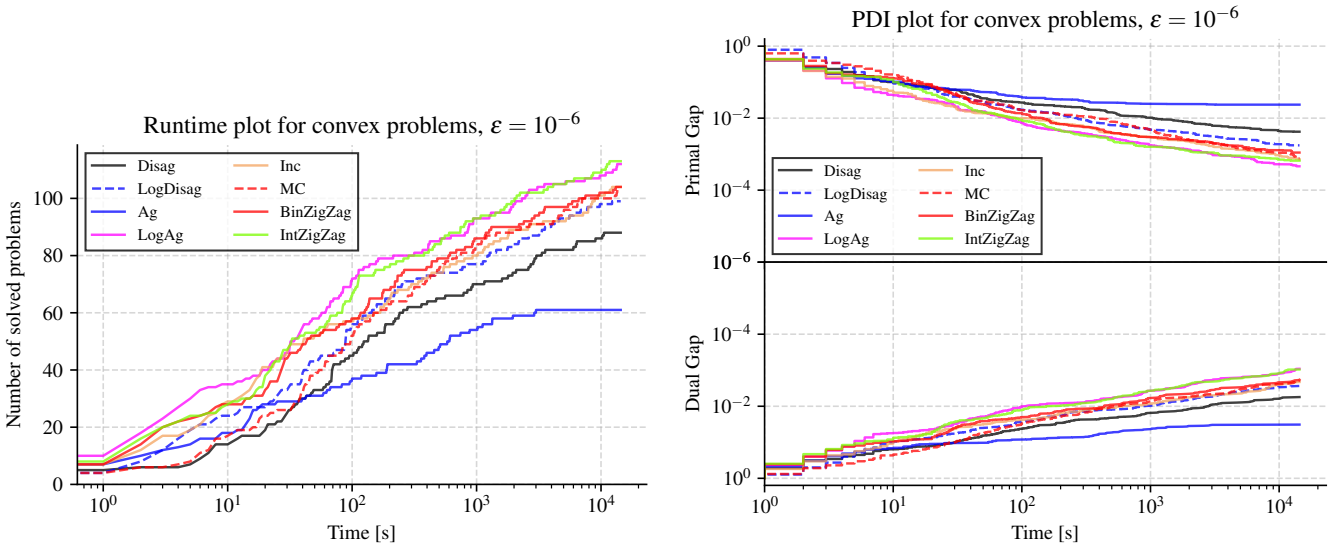


FIGURE C.50. Runtime and Primal-Dual Integral plots for convex problems, $\epsilon = 10^{-6}$

C.11. **Runtime plots for nonconvex problems.** The following figures show the runtime plots for nonconvex problems.

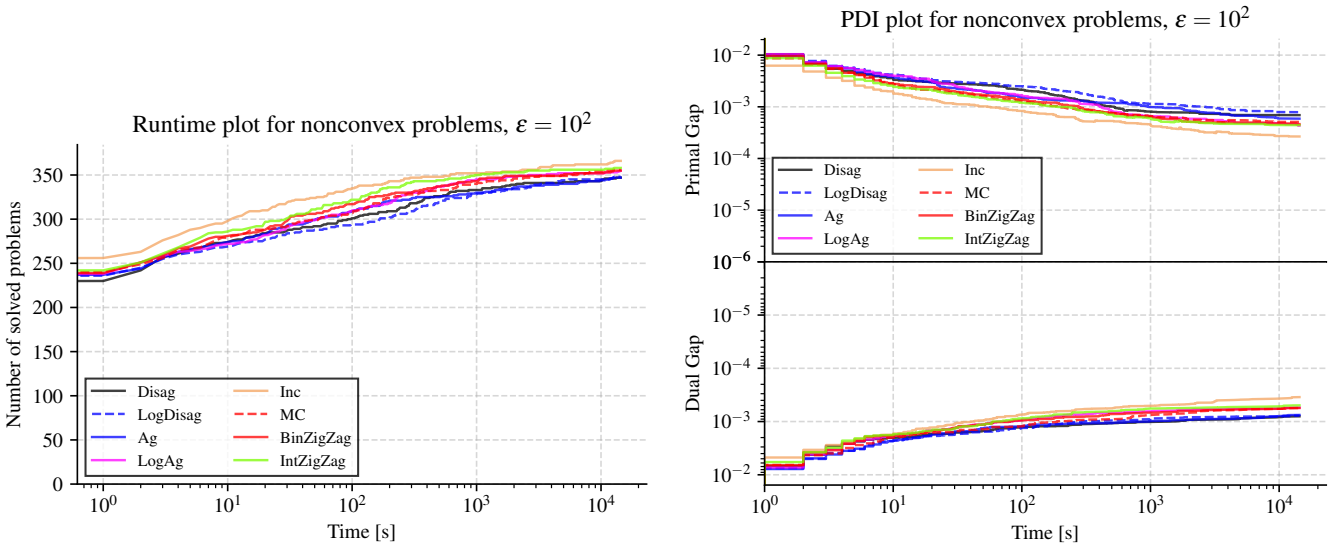


FIGURE C.51. Runtime and Primal-Dual Integral plots for nonconvex problems, $\epsilon = 10^2$

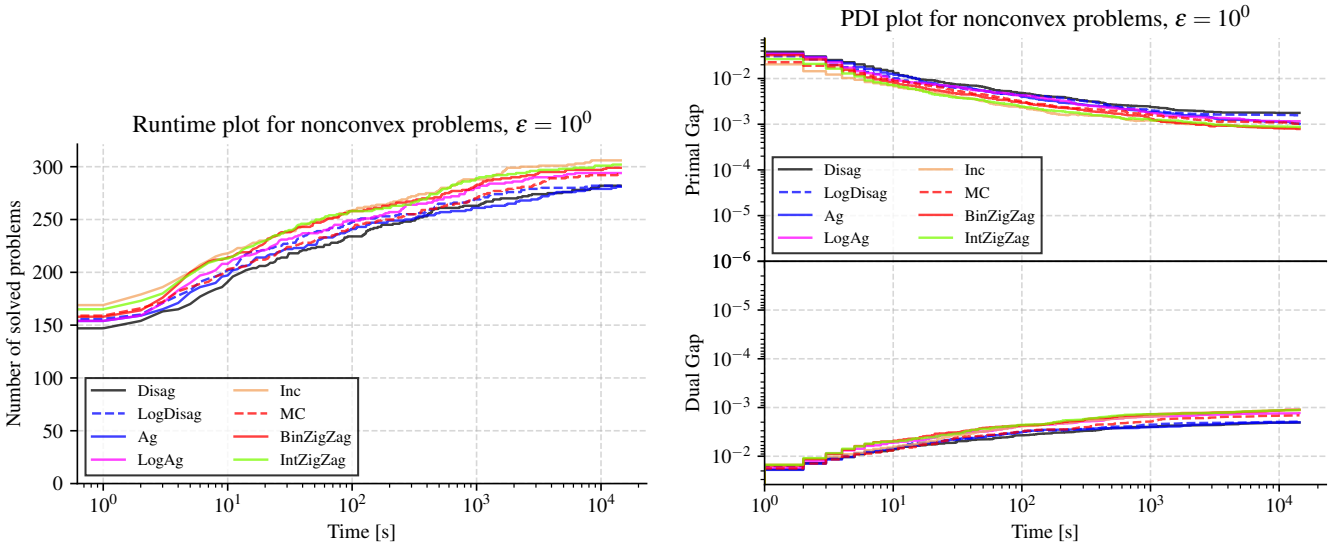


FIGURE C.52. Runtime and Primal-Dual Integral plots for nonconvex problems, $\epsilon = 10^0$

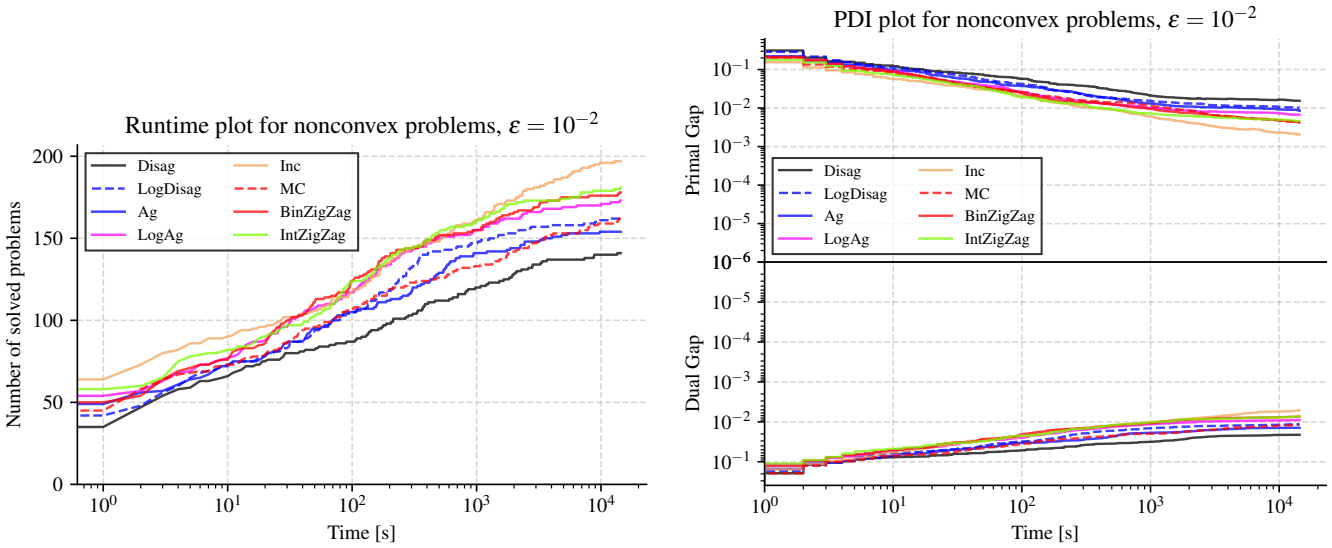


FIGURE C.53. Runtime and Primal-Dual Integral plots for nonconvex problems, $\epsilon = 10^{-2}$

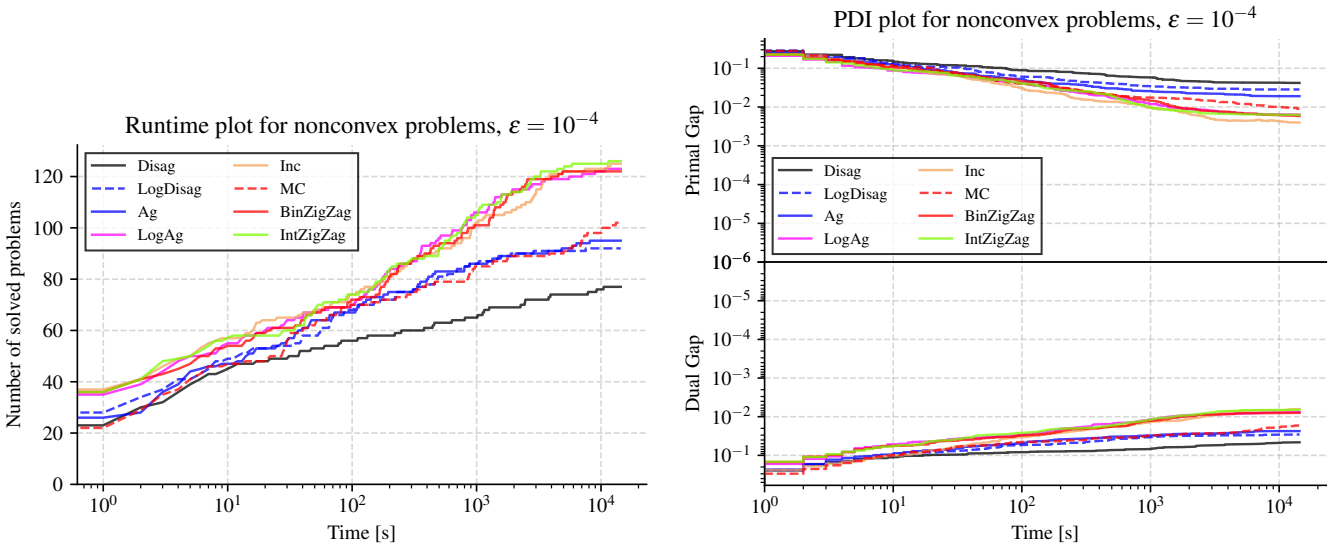


FIGURE C.54. Runtime and Primal-Dual Integral plots for nonconvex problems, $\epsilon = 10^{-4}$

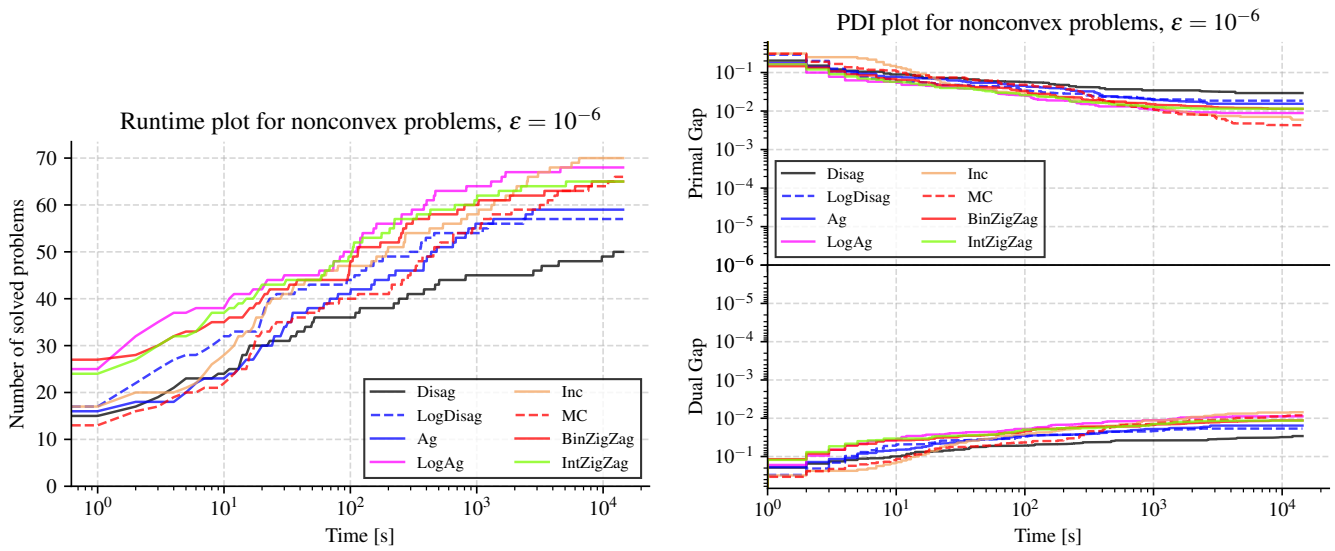


FIGURE C.55. Runtime and Primal-Dual Integral plots for nonconvex problems, $\epsilon = 10^{-6}$

C.12. **Runtime plots for continuous problems.** The following figures show the runtime plots for continuous problems.

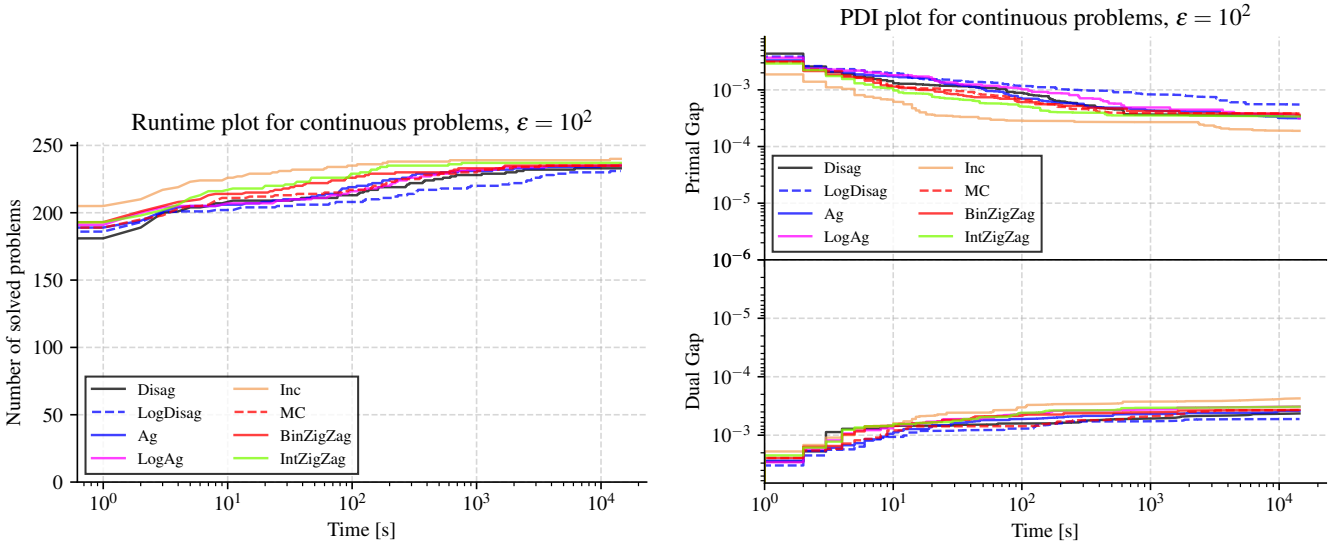


FIGURE C.56. Runtime and Primal-Dual Integral plots for continuous problems, $\epsilon = 10^2$

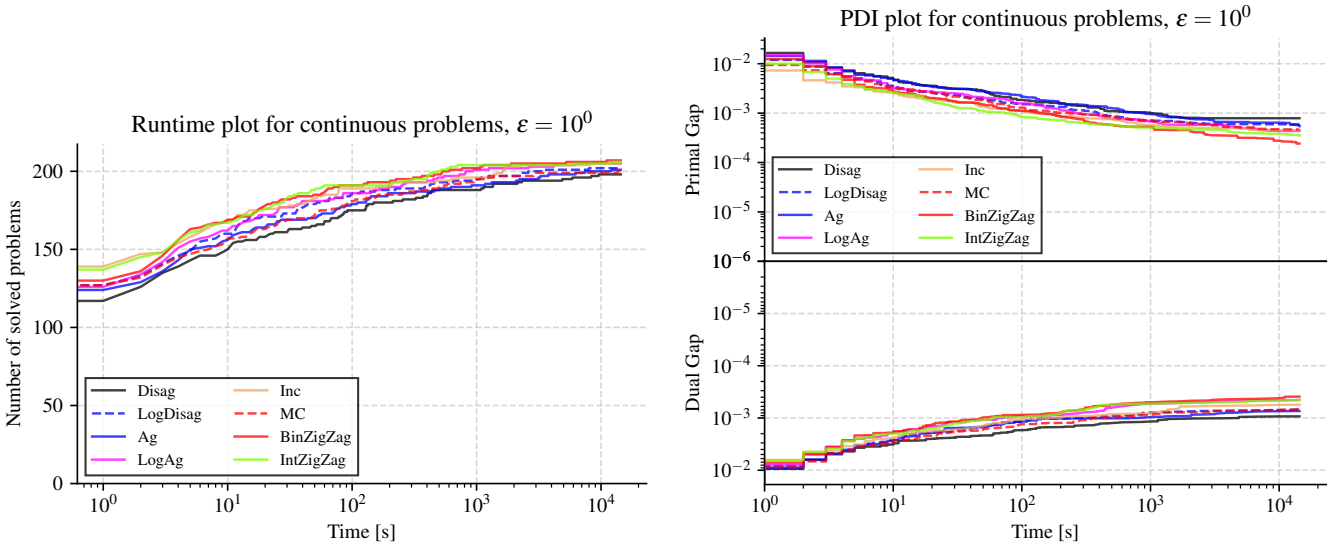


FIGURE C.57. Runtime and Primal-Dual Integral plots for continuous problems, $\epsilon = 10^0$

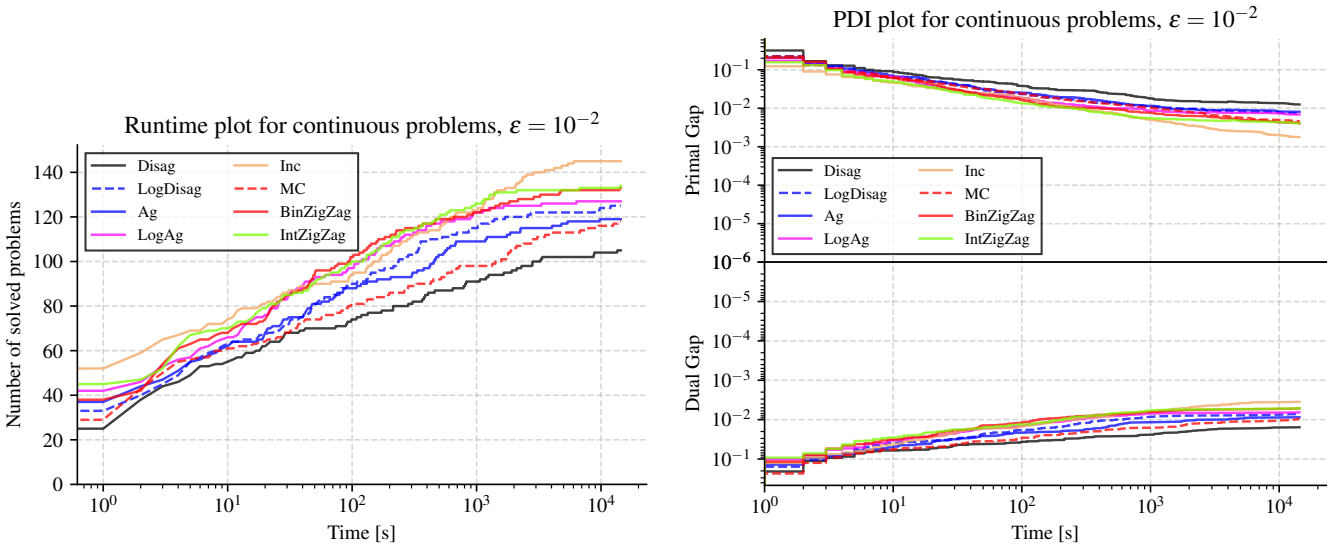


FIGURE C.58. Runtime and Primal-Dual Integral plots for continuous problems, $\epsilon = 10^{-2}$

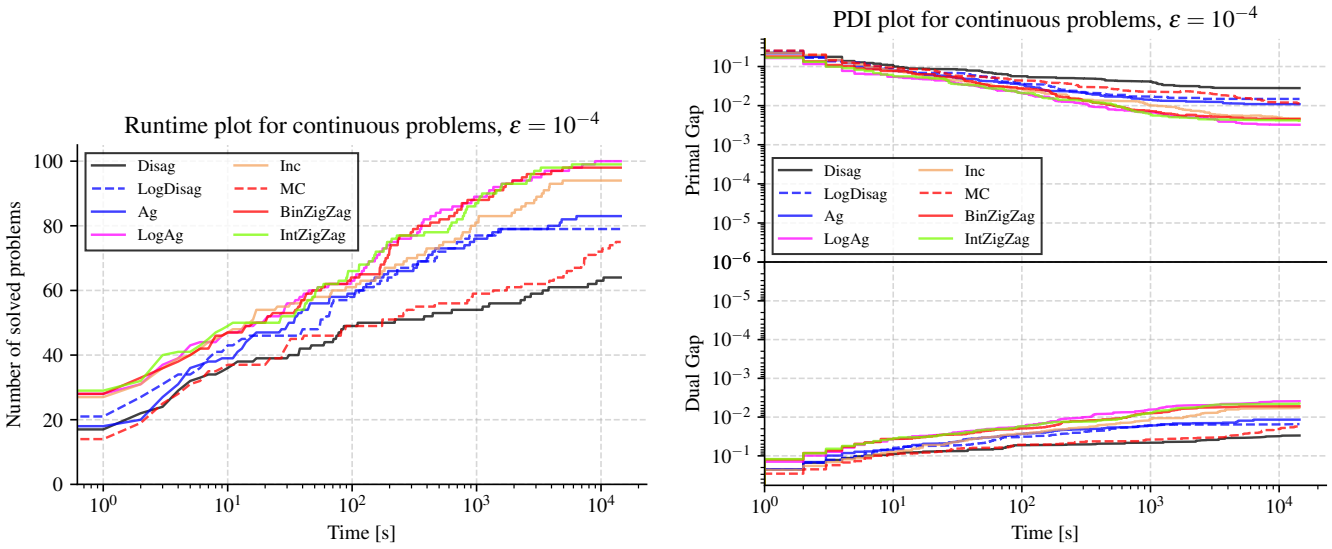


FIGURE C.59. Runtime and Primal-Dual Integral plots for continuous problems, $\epsilon = 10^{-4}$

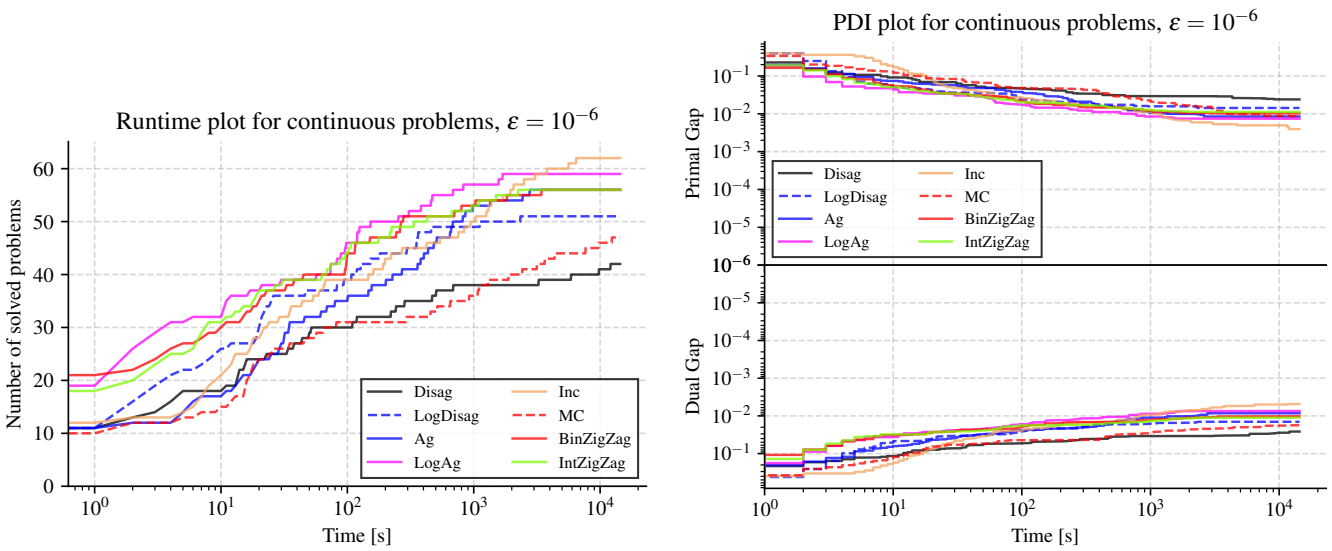


FIGURE C.60. Runtime and Primal-Dual Integral plots for continuous problems, $\epsilon = 10^{-6}$

C.13. **Runtime plots for mixed-binary problems.** The following figures show the runtime plots for mixed-binary problems.

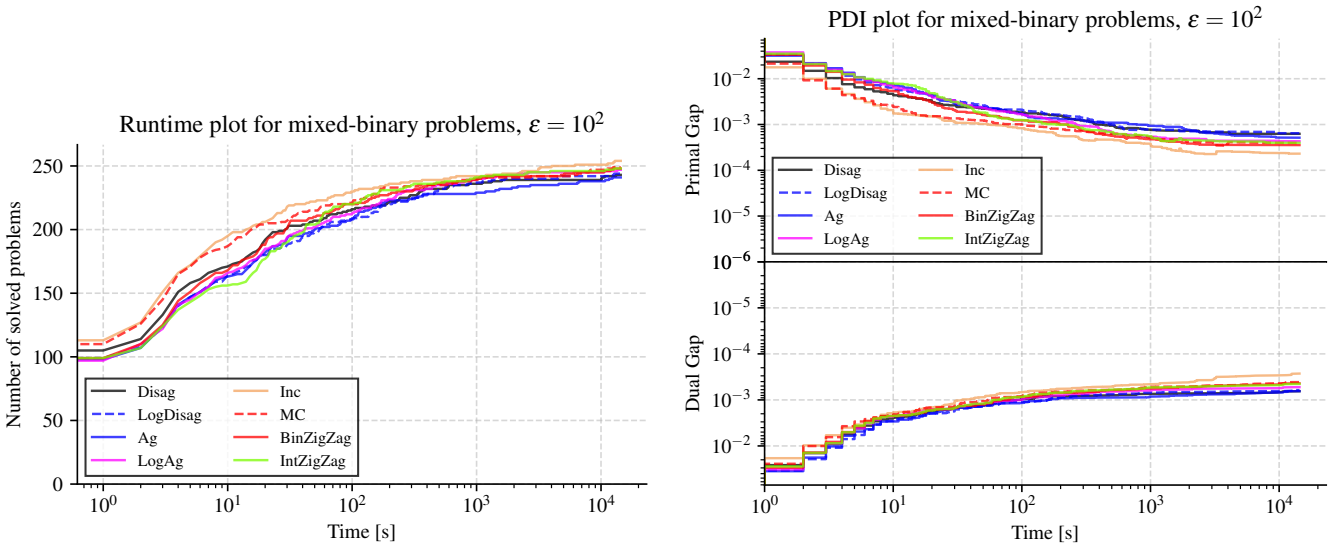


FIGURE C.61. Runtime and Primal-Dual Integral plots for mixed-binary problems, $\epsilon = 10^2$

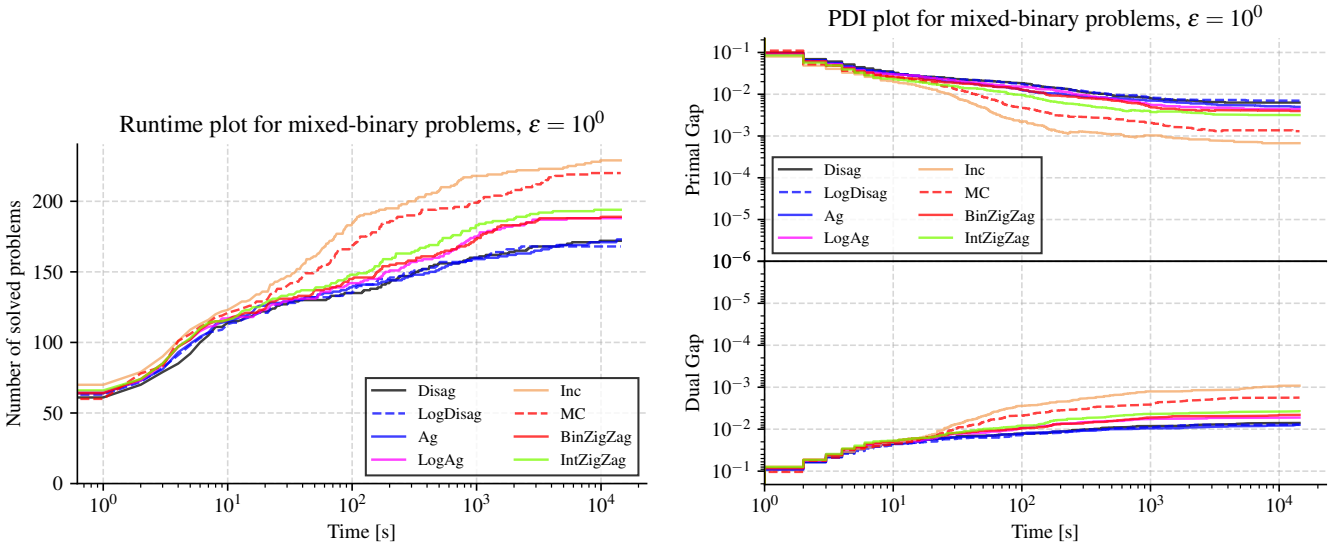


FIGURE C.62. Runtime and Primal-Dual Integral plots for mixed-binary problems, $\epsilon = 10^0$

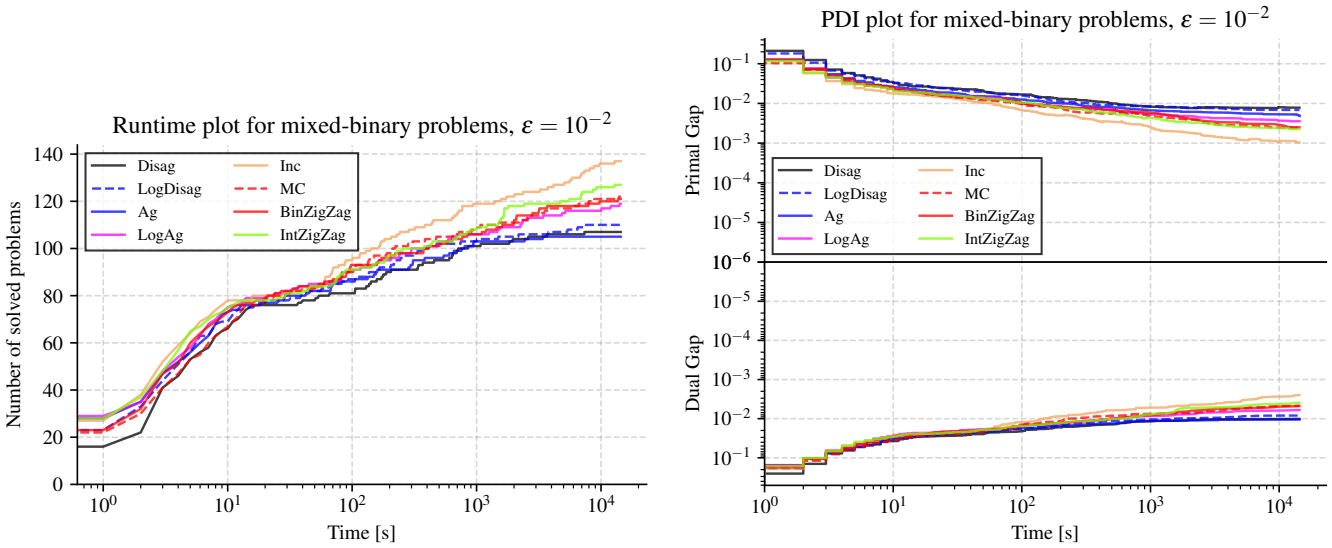


FIGURE C.63. Runtime and Primal-Dual Integral plots for mixed-binary problems, $\epsilon = 10^{-2}$

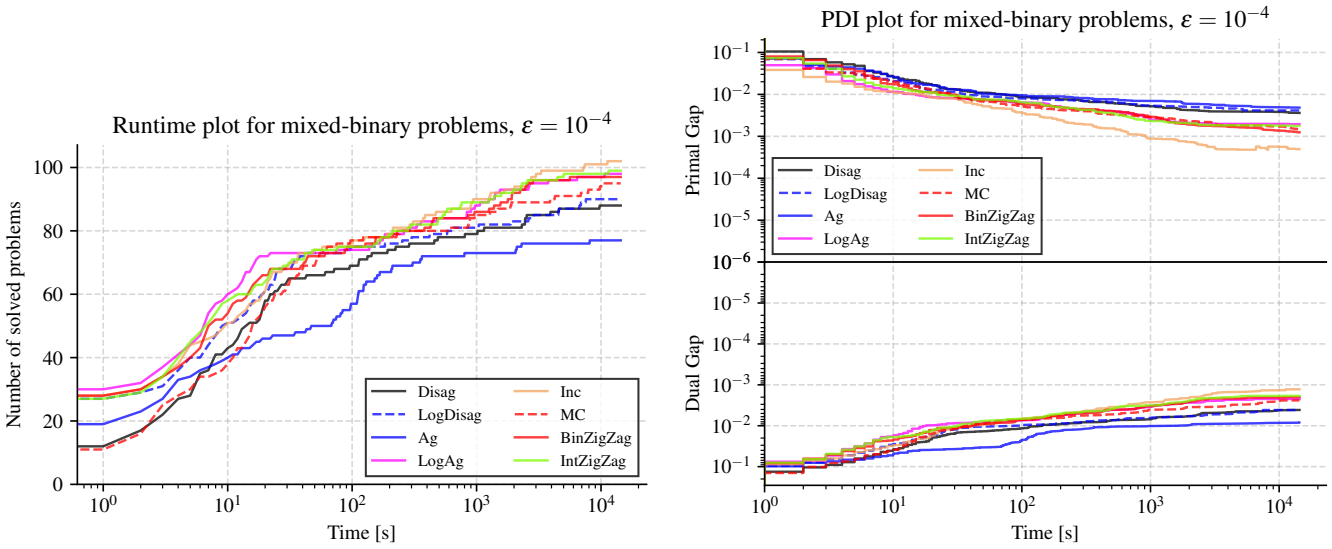


FIGURE C.64. Runtime and Primal-Dual Integral plots for mixed-binary problems, $\epsilon = 10^{-4}$

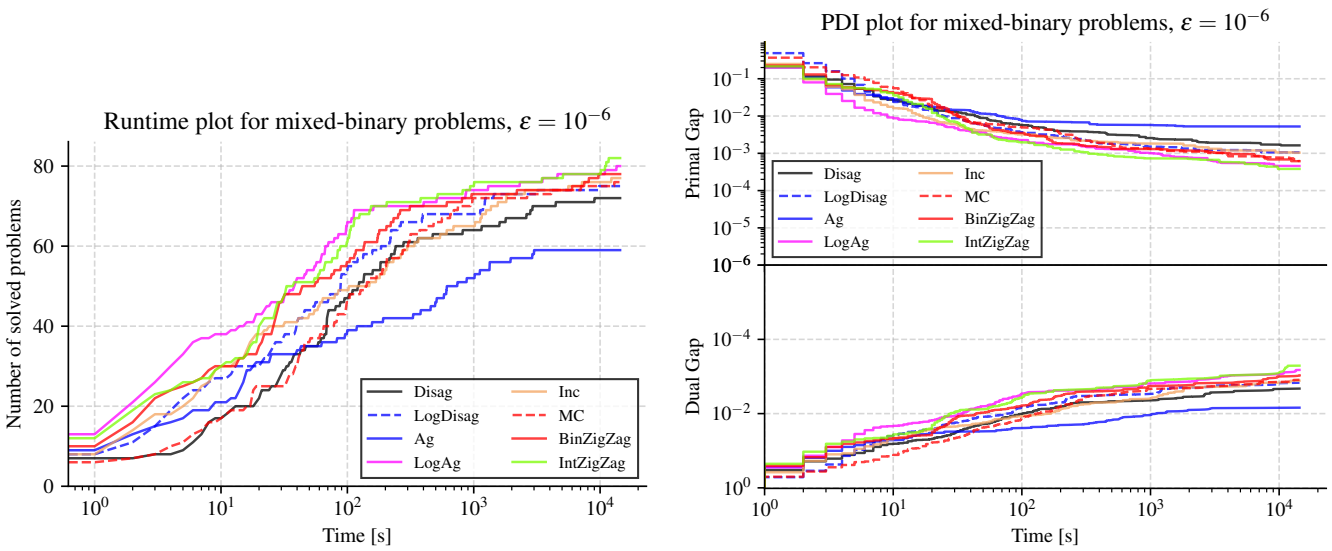


FIGURE C.65. Runtime and Primal-Dual Integral plots for mixed-binary problems, $\epsilon = 10^{-6}$

C.14. **Runtime plots for mixed-integer problems.** The following figures show the runtime plots for mixed-integer problems.

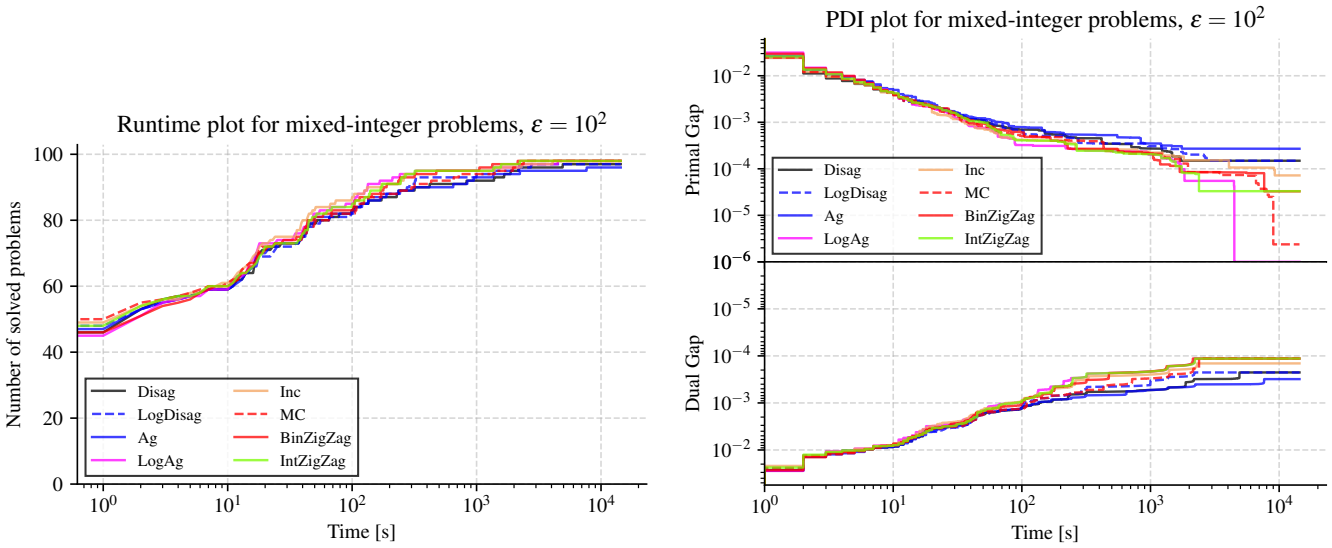


FIGURE C.66. Runtime and Primal-Dual Integral plots for mixed-integer problems, $\epsilon = 10^2$

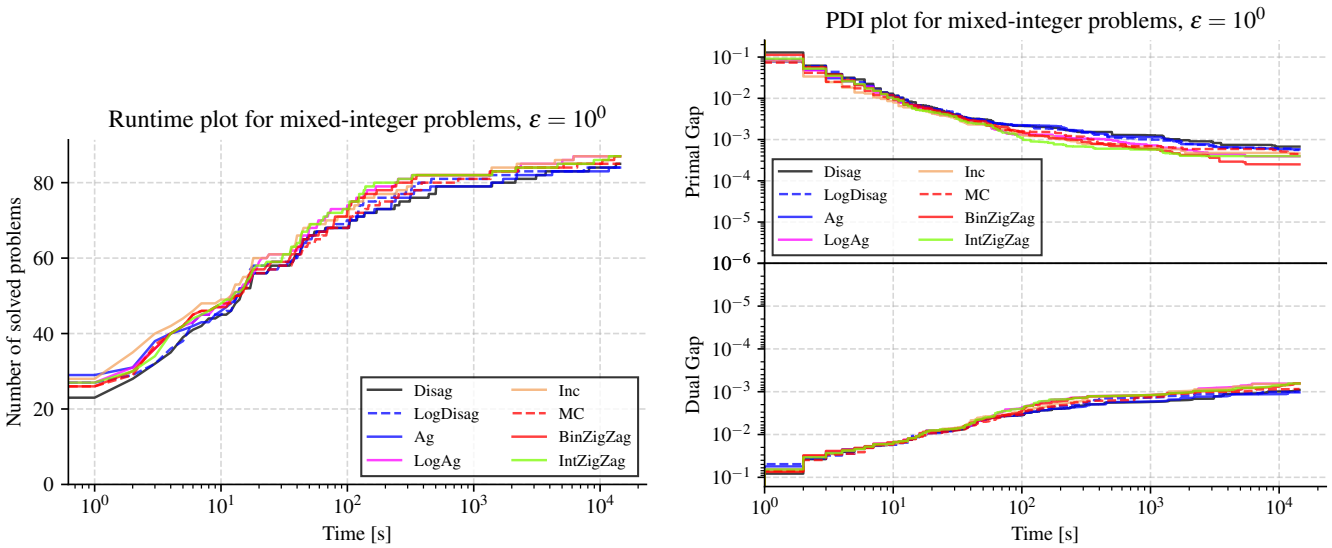


FIGURE C.67. Runtime and Primal-Dual Integral plots for mixed-integer problems, $\epsilon = 10^0$

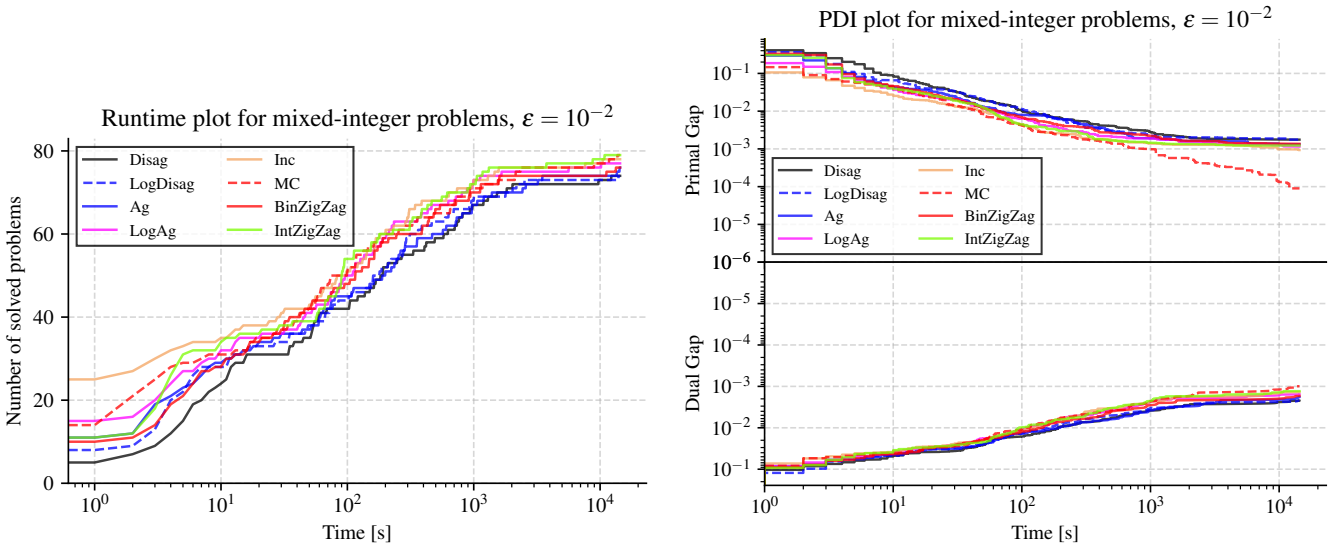


FIGURE C.68. Runtime and Primal-Dual Integral plots for mixed-integer problems, $\epsilon = 10^{-2}$

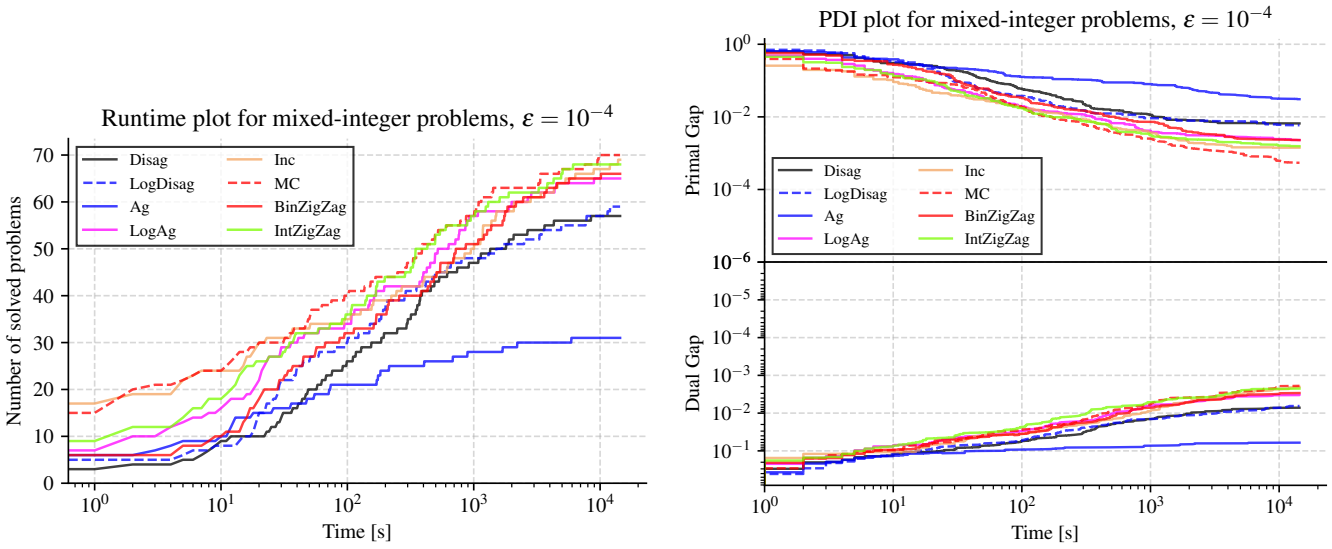


FIGURE C.69. Runtime and Primal-Dual Integral plots for mixed-integer problems, $\epsilon = 10^{-4}$

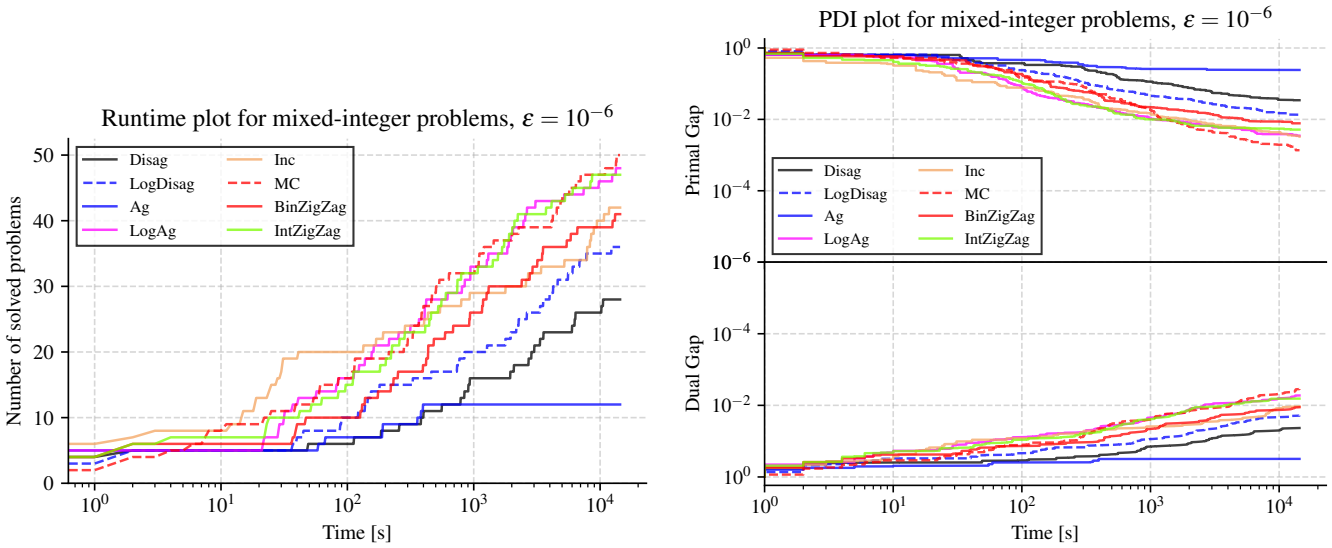
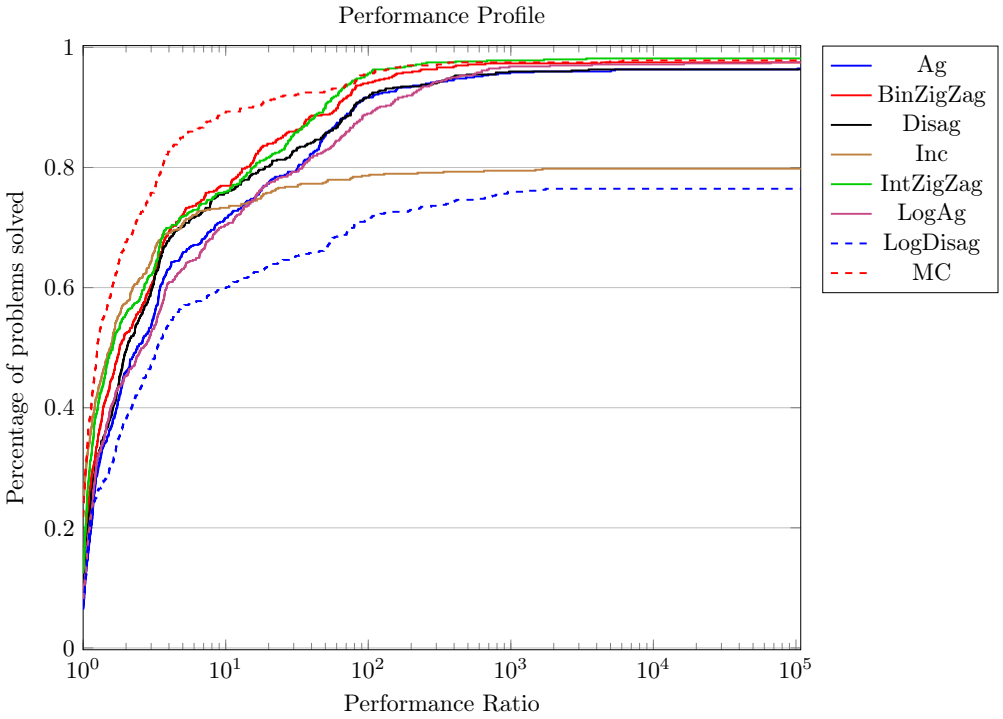


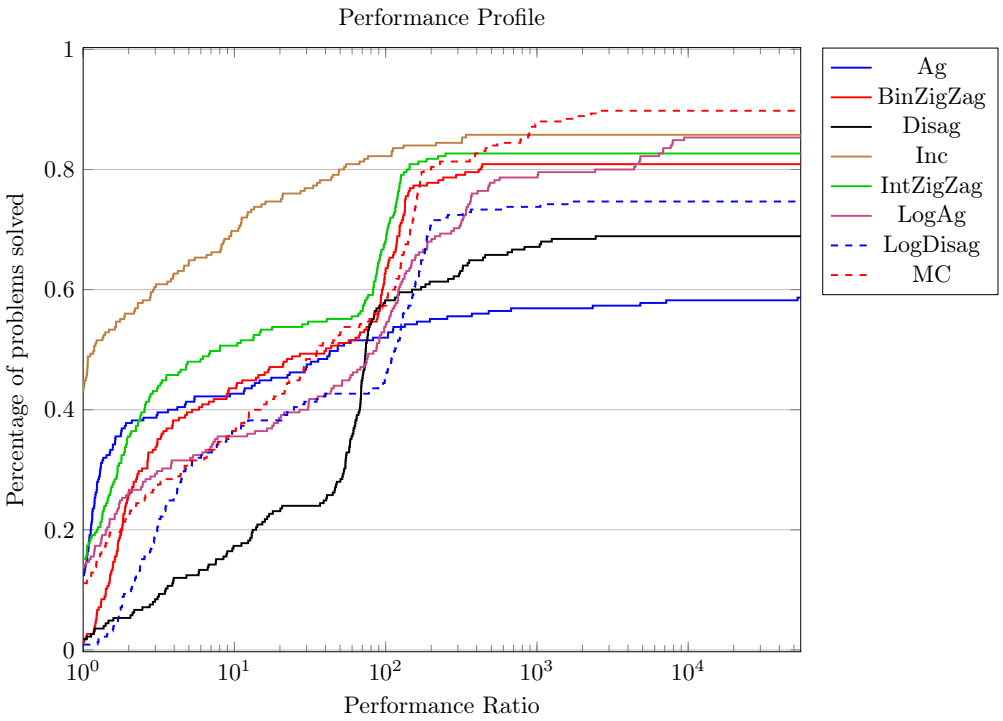
FIGURE C.70. Runtime and Primal-Dual Integral plots for mixed-integer problems, $\epsilon = 10^{-6}$



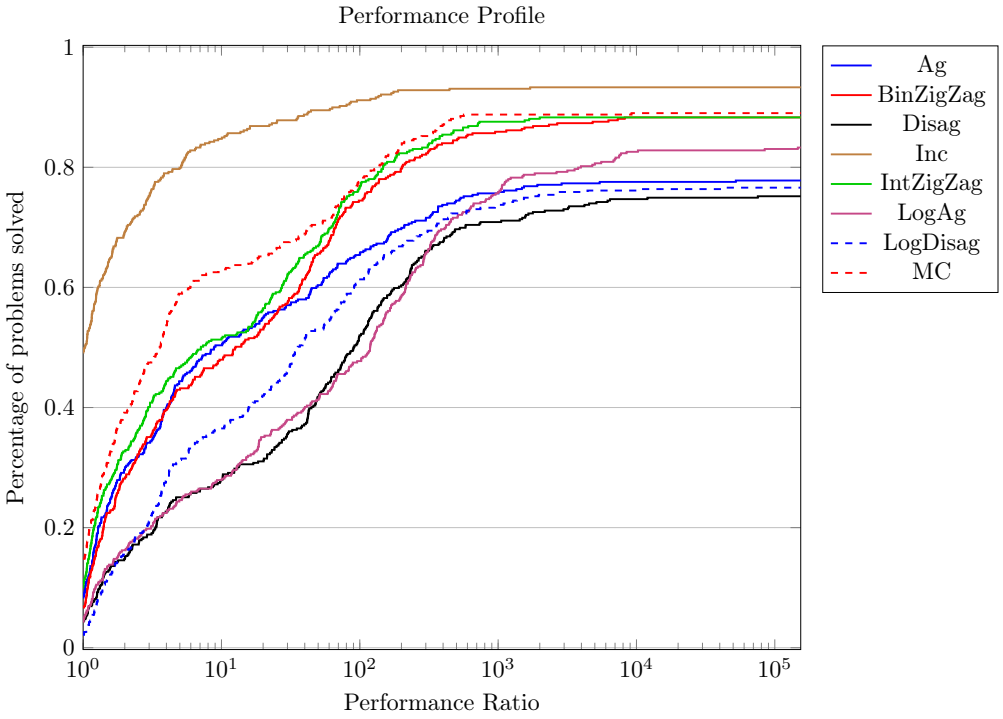
(D.1A) $\epsilon = 10^2$

Appendix D. Performance profiles

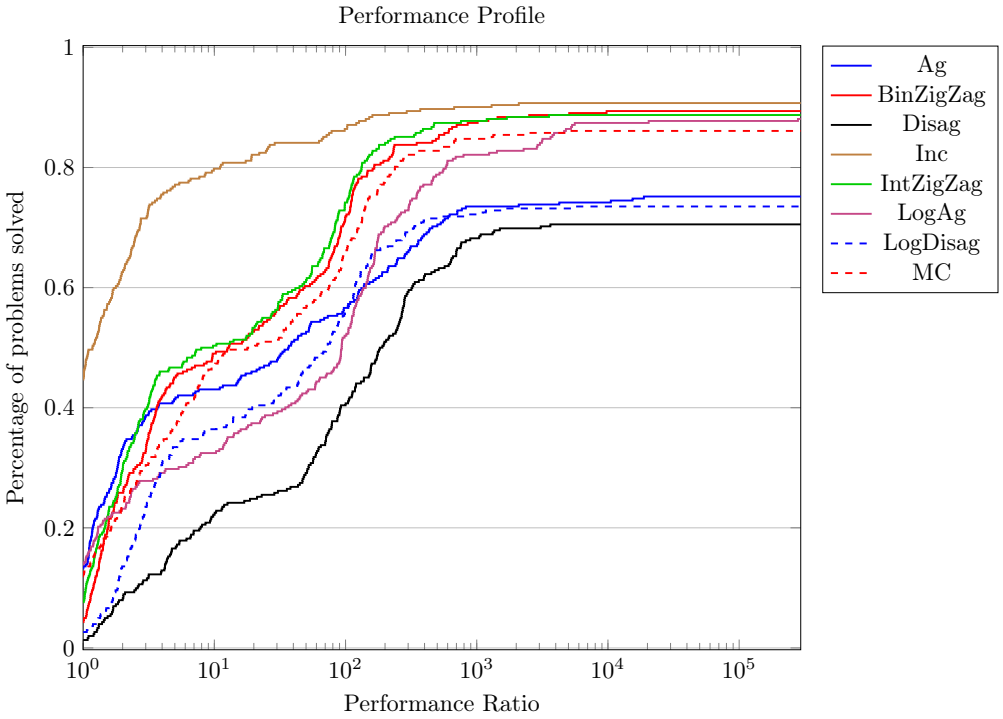
Additionally, we provide performance profile plots as proposed by [13] to illustrate the scaling of the runtimes, see Figures D.1 and D.2. The intention here is to obtain a more sophisticated picture of how the various methods perform if we allow the runtime to lie within a given factor of the best overall runtime. The performance profiles work as follows: Let $t_{p,s}$ be the runtime needed by MILP relaxation $\bar{s}$ to solve instance p. With the performance ratio $r_{p,\bar{s}} := t_{p,\bar{s}} / \min_s t_{p,s}$, the performance profile function value $P(\tau)$ is the percentage of problems solved by approach $\bar{s}$ such that the ratios $r_{p,\bar{s}}$ are within a factor $\tau \in \mathbb{R}$ of the best possible ratios. Figure D.1 considers the time until a first feasible solution was found while Figure D.2 considers the overall runtime until an instance was solved to optimality. All performance profiles are generated with the help of *perprof-py* ([40]).



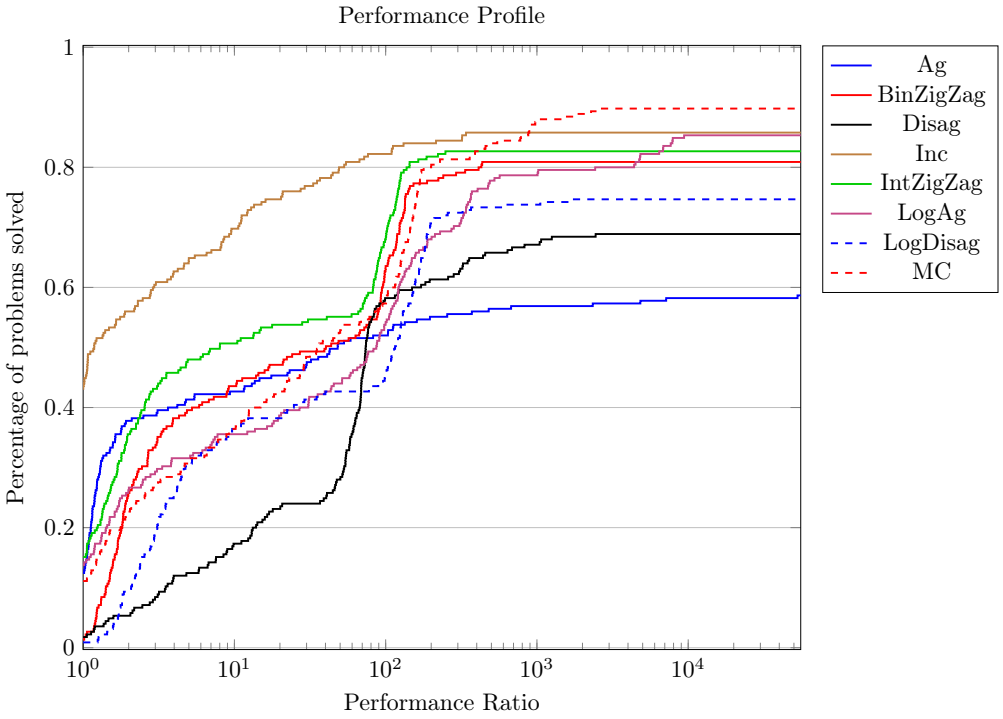
(D.1B) $\epsilon = 10^0$



(D.1C) $\epsilon = 10^{-2}$

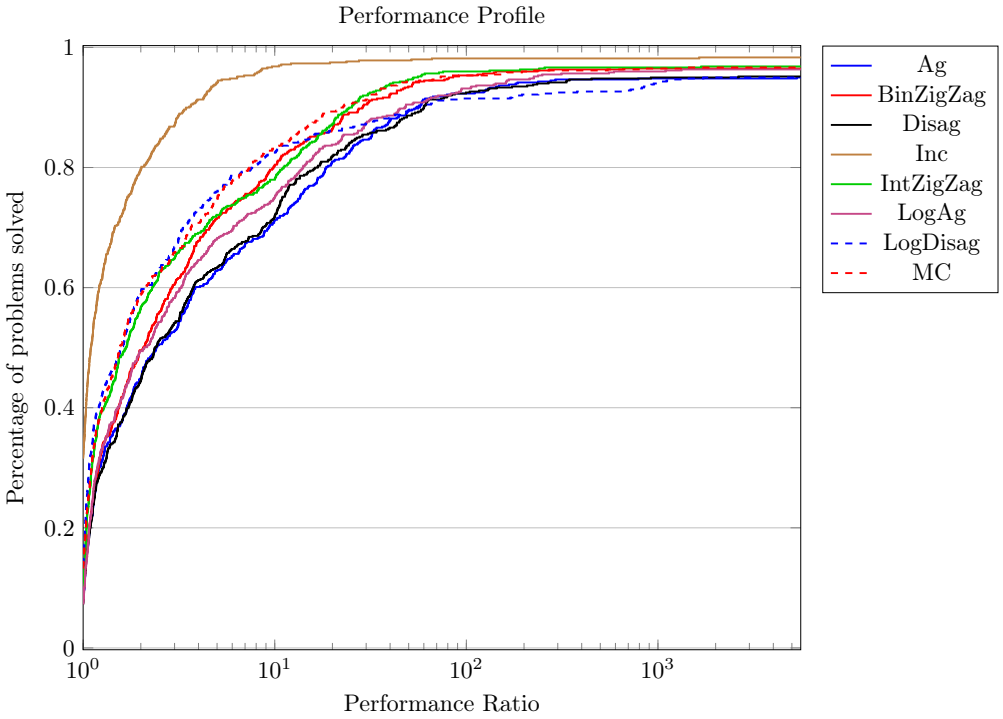


(D.1D) $\epsilon = 10^{-4}$

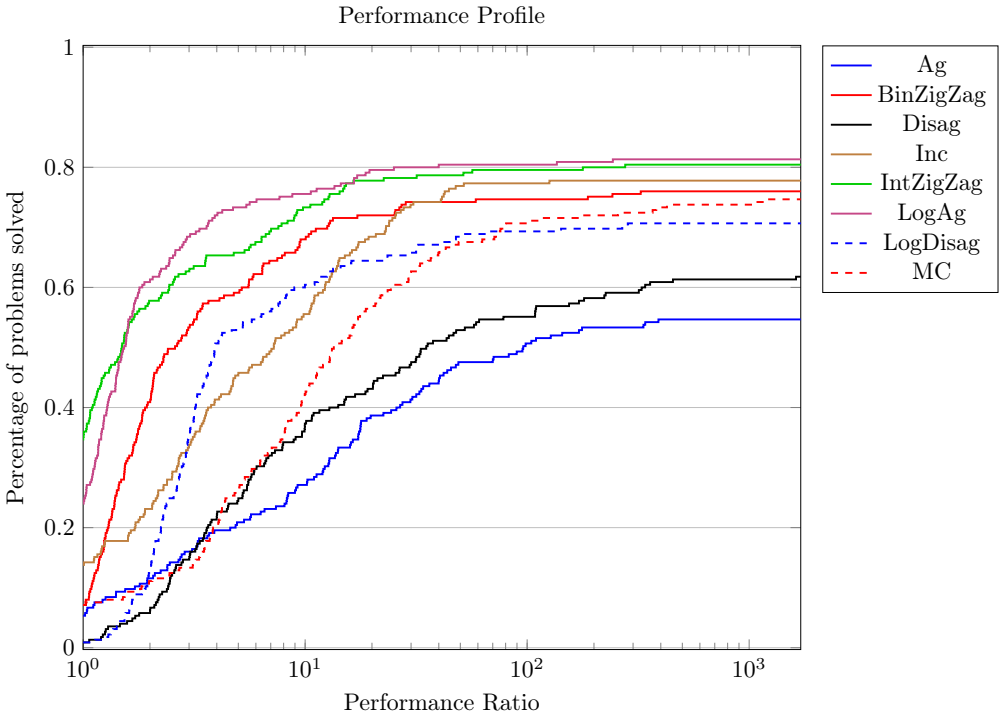


(D.1E) $\epsilon = 10^{-6}$

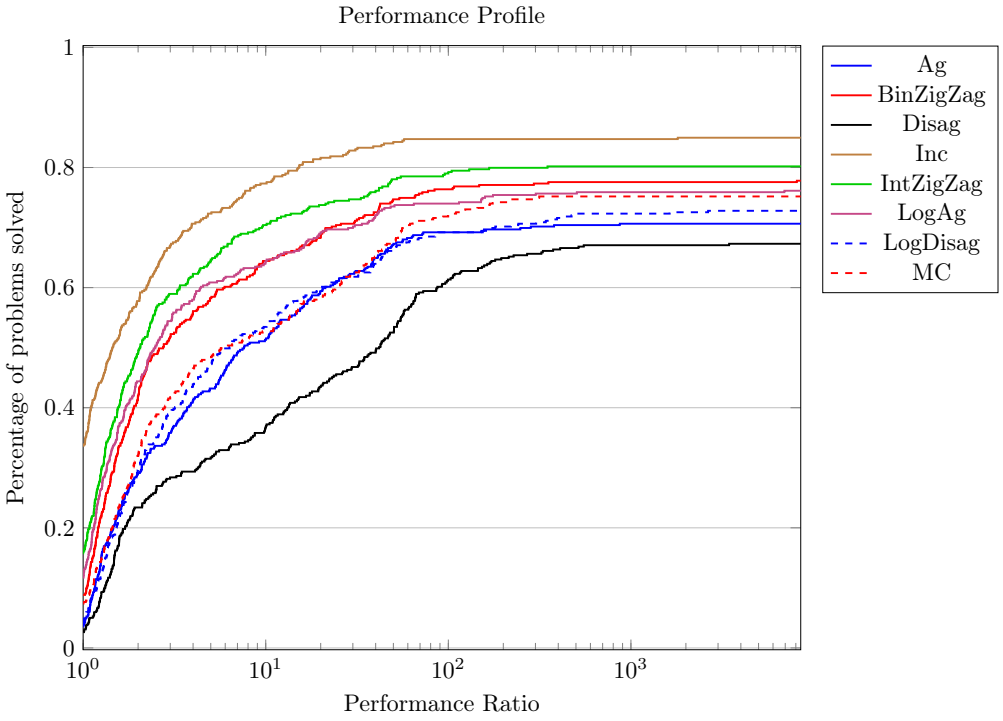
FIGURE D.1. Performance profiles considering the time until a primal solution is found.



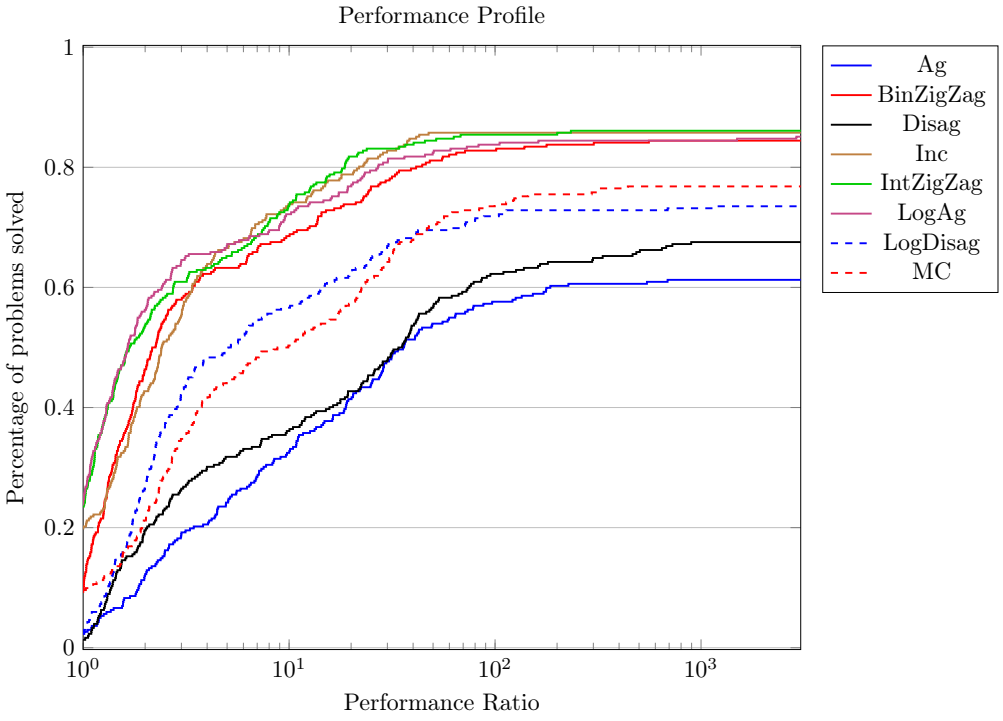
(D.2A) $\epsilon = 100$



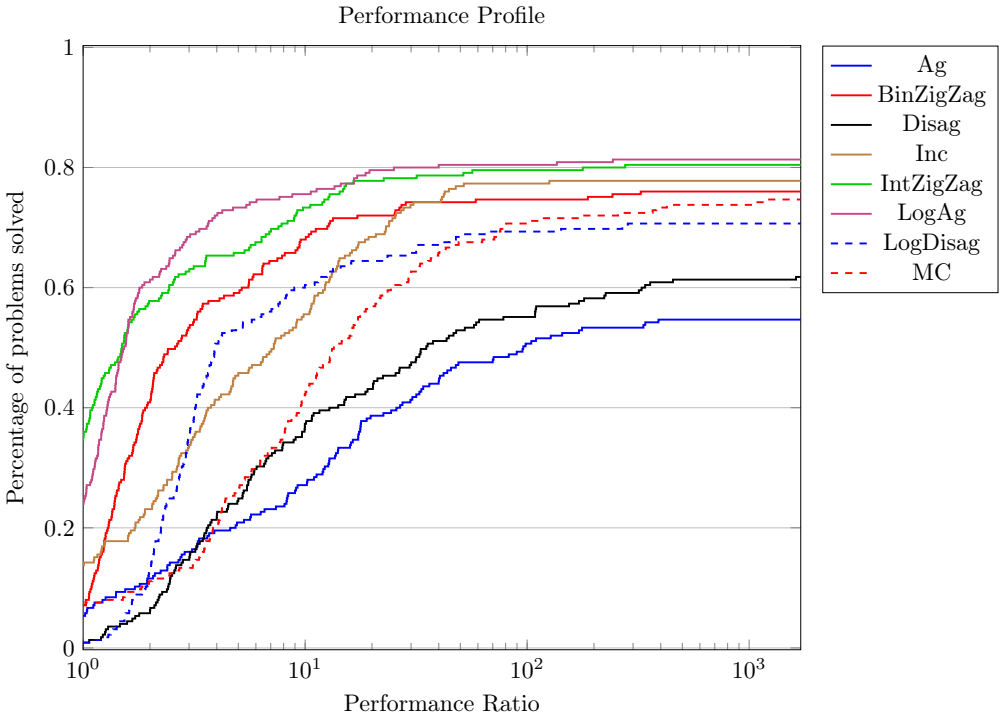
(D.2B) $\epsilon = 1$



(D.2C) $\epsilon = 10^{-2}$



(D.2D) $\epsilon = 10^{-4}$



(D.2E) $\epsilon = 10^{-6}$

FIGURE D.2. Performance profiles considering the time until an optimal solution is found.

Appendix E. Benchmark Set

Table E.1. Benchmark set

Instance	objective	Number of constr.	Number of nonlin. eqs.	Number of variables cont.	bin.	int.	Classification difficulty	type	convexity
arki0002	0.9755	1976	913	2456	0	0	medium	nonlin.	non-conv.
arki0015	-272.2998	1496	1012	2093	0	0	medium	nonlin.	non-conv.
ball_mk2_10	0.0	1	1	0	0	10	very easy	quad.	conv.
ball_mk2_30	0.0	1	1	0	0	30	easy	quad.	conv.
ball_mk3_10	infeasible	1	1	0	0	10	very easy	quad.	conv.
ball_mk3_20	infeasible	1	1	0	0	20	very easy	quad.	conv.
ball_mk3_30	infeasible	1	1	0	0	30	very easy	quad.	conv.
bayes2_30	0.0001	77	55	86	0	0	medium	quad.	non-conv.
bayes2_50	0.5202	77	55	86	0	0	medium	quad.	non-conv.
blend146	45.2966	624	24	135	87	0	medium	quad.	non-conv.
blend480	9.2266	884	32	188	124	0	medium	quad.	non-conv.
blend531	20.039	736	32	168	104	0	medium	quad.	non-conv.
blend718	7.3936	606	24	135	87	0	medium	quad.	non-conv.
blend721	13.5268	627	24	135	87	0	easy	quad.	non-conv.
blend852	53.9627	860	32	184	120	0	medium	quad.	non-conv.
btest14	-59.8174	93	86	135	0	0	medium	nonlin.	non-conv.
camshape100	-4.2841	200	101	199	0	0	medium	quad.	non-conv.
camshape200	-4.2785	400	201	399	0	0	medium	quad.	non-conv.
camshape400	-4.2757	800	401	799	0	0	medium	quad.	non-conv.
camshape800	-4.2743	1600	801	1599	0	0	hard	quad.	non-conv.
carton7	191.7295	687	64	72	200	56	easy	quad.	non-conv.
carton9	205.1371	893	68	72	216	72	easy	quad.	non-conv.
cecil_13	-115656.4997	898	180	660	180	0	medium	nonlin.	non-conv.
chakra	-179.1336	41	22	62	0	0	easy	sign.	non-conv.
chance	29.8944	3	1	4	0	0	easy	nonlin.	N/A
chp_partload	23.2981	2516	490	2203	45	0	easy	nonlin.	non-conv.
clay0203m	41573.2625	54	24	12	18	0	very easy	quad.	conv.
clay0204m	6545.0	90	32	20	32	0	very easy	quad.	conv.
clay0205m	8092.5	135	40	30	50	0	very easy	quad.	conv.
clay0303m	26669.1096	66	36	12	21	0	very easy	quad.	conv.
clay0304m	40262.3875	106	48	20	36	0	very easy	quad.	conv.
clay0305m	8092.5	155	60	30	55	0	very easy	quad.	conv.
crudeoil_lee1_05	79.75	1240	160	495	40	0	very easy	quad.	non-conv.
crudeoil_lee1_06	79.75	1503	192	594	48	0	very easy	quad.	non-conv.
crudeoil_lee1_07	79.75	1776	224	693	56	0	very easy	quad.	non-conv.
crudeoil_lee1_08	79.75	2059	256	792	64	0	very easy	quad.	non-conv.
crudeoil_lee1_09	79.75	2352	288	891	72	0	very easy	quad.	non-conv.
crudeoil_lee1_10	79.75	2655	320	990	80	0	very easy	quad.	non-conv.
crudeoil_lee2_05	96.1699	2581	420	1085	70	0	easy	quad.	non-conv.
crudeoil_lee2_06	101.1746	3117	504	1302	84	0	easy	quad.	non-conv.
crudeoil_lee2_07	101.1746	3670	588	1519	98	0	easy	quad.	non-conv.
crudeoil_lee2_08	101.1746	4240	672	1736	112	0	easy	quad.	non-conv.
crudeoil_lee2_10	101.1746	5431	840	2170	140	0	medium	quad.	non-conv.
crudeoil_lee3_05	85.4489	2786	490	1210	70	0	medium	quad.	non-conv.
crudeoil_lee3_06	85.4489	3359	588	1452	84	0	medium	quad.	non-conv.
crudeoil_lee3_07	85.4489	3949	686	1694	98	0	medium	quad.	non-conv.
crudeoil_lee3_08	85.4489	4556	784	1936	112	0	medium	quad.	non-conv.
crudeoil_lee3_09	85.4489	5180	882	2178	126	0	medium	quad.	non-conv.
crudeoil_lee3_10	85.4489	5821	980	2420	140	0	hard	quad.	non-conv.
crudeoil_lee4_05	132.5476	4241	760	1860	95	0	easy	quad.	non-conv.
crudeoil_lee4_06	132.5476	5093	912	2232	114	0	easy	quad.	non-conv.
crudeoil_lee4_07	132.5476	5965	1064	2604	133	0	easy	quad.	non-conv.
crudeoil_lee4_08	132.5476	6857	1216	2976	152	0	easy	quad.	non-conv.
crudeoil_lee4_09	132.5476	7769	1368	3348	171	0	easy	quad.	non-conv.
crudeoil_li01	5122.5645	695	56	296	48	0	medium	quad.	non-conv.

TABLE E.1. Benchmark set (cont.)

Instance	objective	Number of constr.	Number of nonlin. eqs.	Number of variables cont.	bin.	int.	Classification difficulty	type	convexity
crudeoil_li02	101567417.2	5004	54	1057	240	0	medium	quad.	non-conv.
crudeoil_li03	3487.1501	2442	192	832	132	0	hard	quad.	non-conv.
crudeoil_li05	3135.0625	1916	192	808	132	0	medium	quad.	non-conv.
crudeoil_li06	3355.0	2436	192	832	132	0	hard	quad.	non-conv.
crudeoil_pooling_ct2	10246.22	732	70	295	108	0	medium	quad.	non-conv.
crudeoil_pooling_ct4	13258.2597	924	95	395	138	0	medium	quad.	non-conv.
cvxnonsep_normcon20	-21.7491	1	1	10	0	10	easy	nonlin.	conv.
cvxnonsep_normcon20r	-21.7491	21	20	30	0	10	very easy	quad.	conv.
cvxnonsep_normcon30	-34.244	1	1	15	0	15	easy	nonlin.	conv.
cvxnonsep_normcon30r	-34.244	31	30	45	0	15	easy	quad.	conv.
cvxnonsep_normcon40	-32.6297	1	1	20	0	20	easy	nonlin.	conv.
cvxnonsep_normcon40r	-32.6297	41	40	60	0	20	easy	quad.	conv.
cvxnonsep_nsig20	80.9493	1	1	10	0	10	medium	sign.	conv.
cvxnonsep_nsig20r	80.9493	21	20	30	0	10	easy	nonlin.	conv.
cvxnonsep_nsig30	130.6287	1	1	15	0	15	medium	sign.	conv.
cvxnonsep_nsig30r	156.4267	31	30	45	0	15	easy	nonlin.	conv.
cvxnonsep_nsig40	133.9613	1	1	20	0	20	medium	sign.	conv.
cvxnonsep_nsig40r	133.9613	41	40	60	0	20	easy	nonlin.	conv.
cvxnonsep_pcon20	-21.5123	1	1	10	0	10	easy	nonlin.	conv.
cvxnonsep_pcon20r	-21.5123	20	19	29	0	10	easy	nonlin.	conv.
cvxnonsep_pcon30	-35.9868	1	1	15	0	15	easy	nonlin.	conv.
cvxnonsep_pcon30r	-35.9868	30	29	44	0	15	easy	nonlin.	conv.
cvxnonsep_pcon40	-46.5992	1	1	20	0	20	medium	nonlin.	conv.
cvxnonsep_pcon40r	-46.5992	40	39	59	0	20	easy	nonlin.	conv.
cvxnonsep_psig20	93.8114	0	1	10	0	10	medium	sign.	conv.
cvxnonsep_psig30	78.9989	0	1	15	0	15	medium	sign.	conv.
cvxnonsep_psig40	85.4958	0	1	20	0	20	medium	sign.	conv.
cvxnonsep_psig40r	86.5451	42	41	62	0	20	easy	nonlin.	conv.
edgecross10-060	459.0	481	1	47	44	0	medium	quad.	non-conv.
edgecross10-080	1037.0	481	1	17	74	0	very easy	quad.	non-conv.
edgecross14-078	725.0	1457	1	28	155	0	very easy	quad.	non-conv.
edgecross14-156	4310.0	1457	1	67	116	0	medium	quad.	non-conv.
ex1221	7.6672	5	2	2	3	0	easy	sign.	non-conv.
ex1223a	4.5796	9	5	3	4	0	very easy	quad.	conv.
ex1223b	4.5796	9	5	3	4	0	easy	nonlin.	conv.
ex1225	31.0	10	1	2	6	0	easy	sign.	non-conv.
ex1226	-17.0	5	1	2	3	0	easy	sign.	non-conv.
ex1263	19.6	55	4	20	72	0	very easy	quad.	non-conv.
ex1263a	19.6	35	4	0	4	20	very easy	quad.	non-conv.
ex1264	8.6	55	4	20	68	0	very easy	quad.	non-conv.
ex1264a	8.6	35	4	0	4	20	very easy	quad.	non-conv.
ex1265	10.3	74	5	30	100	0	very easy	quad.	non-conv.
ex1265a	10.3	44	5	0	5	30	very easy	quad.	non-conv.
ex1266	16.3	95	6	42	138	0	very easy	quad.	non-conv.
ex1266a	16.3	53	6	0	6	42	very easy	quad.	non-conv.
ex14_1_1	-0.0	4	4	3	0	0	easy	poly.	non-conv.
ex14_1_4	-0.0	4	4	3	0	0	easy	nonlin.	non-conv.
ex14_1_8	0.0	4	4	3	0	0	medium	nonlin.	non-conv.
ex14_1_9	-0.0	2	2	2	0	0	easy	nonlin.	N/A
ex14_2_2	0.0	5	4	4	0	0	easy	nonlin.	non-conv.
ex14_2_5	0.0	5	4	4	0	0	easy	nonlin.	non-conv.
ex14_2_8	0.0	5	4	4	0	0	easy	nonlin.	non-conv.
ex14_2_9	0.0	5	4	4	0	0	easy	nonlin.	non-conv.
ex2_1_1	-17.0	1	1	5	0	0	very easy	quad.	non-conv.
ex2_1_2	-213.0	2	1	6	0	0	very easy	quad.	non-conv.
ex2_1_4	-11.0	5	1	6	0	0	very easy	quad.	non-conv.
ex2_1_6	-39.0	5	1	10	0	0	very easy	quad.	non-conv.
ex2_1_9	-0.375	1	1	10	0	0	easy	quad.	non-conv.

TABLE E.1. Benchmark set (cont.)

Instance	objective	Number of constr.	Number of nonlin. eqs.	Number of variables cont.	bin.	int.	Classification difficulty	type	convexity
ex3_1_1	7049.248	6	3	8	0	0	very easy	quad.	non-conv.
ex3_1_2	-30665.5387	6	7	5	0	0	very easy	quad.	non-conv.
ex3_1_3	-310.0	6	3	6	0	0	very easy	quad.	non-conv.
ex3_1_4	-4.0	3	1	3	0	0	very easy	quad.	non-conv.
ex3pb	68.0097	31	5	24	8	0	easy	nonlin.	non-conv.
ex4	-8.0641	30	26	11	25	0	very easy	quad.	conv.
ex4_1_1	-7.4873	0	1	1	0	0	easy	poly.	non-conv.
ex4_1_2	-663.5001	0	1	1	0	0	easy	poly.	non-conv.
ex4_1_3	-443.6717	0	1	1	0	0	easy	poly.	non-conv.
ex4_1_4	0.0	0	1	1	0	0	easy	poly.	N/A
ex4_1_6	7.0	0	1	1	0	0	easy	poly.	N/A
ex4_1_7	-7.5	0	1	1	0	0	easy	poly.	non-conv.
ex4_1_9	-5.508	2	2	2	0	0	easy	poly.	non-conv.
ex5_2_2_case1	-400.0	6	3	9	0	0	very easy	quad.	non-conv.
ex5_2_2_case2	-600.0	6	3	9	0	0	very easy	quad.	non-conv.
ex5_2_2_case3	-750.0	6	3	9	0	0	very easy	quad.	non-conv.
ex5_3_2	1.8642	16	9	22	0	0	very easy	quad.	non-conv.
ex5_4_2	7512.2301	6	3	8	0	0	very easy	quad.	non-conv.
ex7_2_2	-0.3888	5	5	6	0	0	easy	sign.	non-conv.
ex7_3_6	infeasible	17	10	17	0	0	easy	poly.	non-conv.
ex8_1_1	-2.0218	0	1	2	0	0	easy	nonlin.	non-conv.
ex8_1_2	-1.0709	0	1	1	0	0	easy	nonlin.	non-conv.
ex8_3_13	-43.0895	72	54	115	0	0	hard	nonlin.	non-conv.
ex8_3_2	-0.4123	76	49	110	0	0	medium	quad.	non-conv.
ex8_3_3	-0.4166	76	49	110	0	0	medium	quad.	non-conv.
ex8_3_4	-3.58	76	49	110	0	0	medium	quad.	non-conv.
ex8_3_5	-0.0691	76	49	110	0	0	medium	quad.	non-conv.
ex8_3_8	-3.2561	93	65	126	0	0	medium	quad.	non-conv.
ex8_3_9	-0.763	45	27	78	0	0	medium	quad.	non-conv.
ex9_1_4	-37.0	9	4	10	0	0	very easy	quad.	non-conv.
ex9_2_3	0.0	15	6	16	0	0	very easy	quad.	non-conv.
flay02h	37.9473	51	2	42	4	0	easy	sign.	conv.
flay03h	48.9898	144	3	110	12	0	easy	sign.	conv.
flay03m	48.9898	24	3	14	12	0	easy	sign.	conv.
flay04h	54.4059	282	4	210	24	0	easy	sign.	conv.
flay04m	54.4059	42	4	18	24	0	easy	sign.	conv.
flay05h	64.4981	465	5	342	40	0	medium	sign.	conv.
flay05m	64.4981	65	5	22	40	0	easy	sign.	conv.
flay06h	66.9328	693	6	506	60	0	medium	sign.	conv.
flay06m	66.9328	93	6	26	60	0	medium	sign.	conv.
fo7	20.7298	211	14	72	42	0	medium	sign.	conv.
fo7_2	17.7493	211	14	72	42	0	easy	sign.	conv.
fo7_ar25_1	23.0936	269	14	70	0	42	easy	sign.	conv.
fo7_ar2_1	24.8398	269	14	70	0	42	easy	sign.	conv.
fo7_ar3_1	22.5175	269	14	70	0	42	easy	sign.	conv.
fo7_ar4_1	20.7298	269	14	70	0	42	easy	sign.	conv.
fo7_ar5_1	17.7493	269	14	70	0	42	easy	sign.	conv.
fo8	22.3819	273	16	90	56	0	medium	sign.	conv.
fo8_ar25_1	28.0452	347	16	88	0	56	medium	sign.	conv.
fo8_ar2_1	30.3406	347	16	88	0	56	medium	sign.	conv.
fo8_ar3_1	23.9101	347	16	88	0	56	easy	sign.	conv.
fo8_ar4_1	22.3819	347	16	88	0	56	easy	sign.	conv.
fo8_ar5_1	22.3819	347	16	88	0	56	medium	sign.	conv.
fo9	23.4643	343	18	110	72	0	medium	sign.	conv.
fo9_ar25_1	32.1864	435	18	108	0	72	medium	sign.	conv.
fo9_ar2_1	32.625	435	18	108	0	72	medium	sign.	conv.
fo9_ar3_1	24.8155	435	18	108	0	72	medium	sign.	conv.
fo9_ar4_1	23.4643	435	18	108	0	72	medium	sign.	conv.

TABLE E.1. Benchmark set (cont.)

Instance	objective	Number of constr.	Number of nonlin. eqs.	Number of variables cont.	bin.	int.	Classification difficulty	type	convexity
fo9_ar5_1	23.4643	435	18	108	0	72	hard	sign.	conv.
gabriel01	45.2444	467	48	143	72	0	medium	quad.	non-conv.
gabriel02	39.6097	597	96	190	71	0	hard	quad.	non-conv.
gabriel04	9.2266	943	128	260	101	0	hard	quad.	non-conv.
gabriel05	infeasible	1795	192	519	256	0	hard	quad.	non-conv.
gasprod_sarawak01	-32445.4049	212	34	93	38	0	very easy	quad.	non-conv.
gastrans	89.0858	149	24	85	21	0	very easy	poly.	non-conv.
gear	0.0	0	1	0	0	4	easy	sign.	N/A
gear3	0.0	4	1	4	0	4	easy	sign.	N/A
genpooling_lee1	-4640.0824	82	20	40	9	0	easy	quad.	non-conv.
genpooling_lee2	-3849.2654	92	30	44	9	0	easy	quad.	non-conv.
genpooling_meyer04	1086187.137	141	15	63	55	0	medium	quad.	non-conv.
genpooling_meyer10	1086187.137	423	33	207	187	0	medium	quad.	non-conv.
gkocis	-1.9231	8	2	8	3	0	easy	nonlin.	non-conv.
gsg_0001	2378.1605	112	1	78	0	0	easy	sign.	non-conv.
himmel16	-0.866	21	21	18	0	0	medium	quad.	non-conv.
house	-4500.0	8	3	8	0	0	very easy	quad.	non-conv.
hydroenergy1	209721.0066	428	48	192	96	0	medium	quad.	non-conv.
hydroenergy2	371811.847	856	96	384	192	0	medium	quad.	non-conv.
hydroenergy3	744963.7246	1498	168	672	336	0	medium	quad.	non-conv.
inscribedsquare01	0.9901	8	9	8	0	0	easy	nonlin.	non-conv.
inscribedsquare02	0.968	8	9	8	0	0	easy	nonlin.	non-conv.
kall_circles_c6a	2.1117	54	22	18	0	0	medium	quad.	non-conv.
kall_circles_c6b	1.9736	54	22	18	0	0	medium	quad.	non-conv.
kall_circles_c6c	2.7977	63	29	20	0	0	medium	quad.	non-conv.
kall_circles_c7a	2.6628	69	29	20	0	0	medium	quad.	non-conv.
kall_circles_c8a	2.5409	86	37	22	0	0	medium	quad.	non-conv.
kall_circlespolygons_c1p11	0.1996	48	21	43	0	0	very easy	quad.	non-conv.
kall_circlespolygons_c1p12	0.3396	48	21	43	0	0	medium	quad.	non-conv.
kall_circlespolygons_c1p13	0.3396	48	21	43	0	0	medium	quad.	non-conv.
kall_circlespolygons_c1p5a	2.8487	174	106	158	0	0	medium	quad.	non-conv.
kall_circlespolygons_c1p5b	3.7696	816	631	791	0	0	hard	quad.	non-conv.
kall_circlespolygons_c1p6a	3.744	1134	904	1110	0	0	hard	quad.	non-conv.
kall_circlesrectangles_c1r11	0.1996	52	23	49	0	0	very easy	quad.	non-conv.
kall_circlesrectangles_c1r12	0.3396	52	23	49	0	0	medium	quad.	non-conv.
kall_circlesrectangles_c1r13	0.2146	52	23	49	0	0	medium	quad.	non-conv.
kall_circlesrectangles_c6r1	7.1645	192	133	184	0	0	medium	quad.	non-conv.
kall_circlesrectangles_c6r29	6.2952	388	283	390	0	0	medium	quad.	non-conv.
kall_circlesrectangles_c6r39	6.1752	619	466	634	0	0	hard	quad.	non-conv.
kall_congruentcircles_c31	0.6438	16	4	10	0	0	very easy	quad.	non-conv.
kall_congruentcircles_c32	1.3759	16	4	10	0	0	very easy	quad.	non-conv.
kall_congruentcircles_c41	0.8584	24	7	12	0	0	very easy	quad.	non-conv.
kall_congruentcircles_c42	0.8584	24	7	12	0	0	very easy	quad.	non-conv.
kall_congruentcircles_c51	1.073	34	11	14	0	0	easy	quad.	non-conv.
kall_congruentcircles_c52	1.5371	34	11	14	0	0	very easy	quad.	non-conv.
kall_congruentcircles_c61	1.2876	46	16	16	0	0	medium	quad.	non-conv.
kall_congruentcircles_c62	1.2876	46	16	16	0	0	very easy	quad.	non-conv.
kall_congruentcircles_c63	1.2876	46	16	16	0	0	very easy	quad.	non-conv.
kall_congruentcircles_c71	1.5022	60	22	18	0	0	medium	quad.	non-conv.
kall_congruentcircles_c72	1.9663	60	22	18	0	0	easy	quad.	non-conv.
kall_diffcircles_10	11.9355	71	46	24	0	0	hard	quad.	non-conv.
kall_diffcircles_5a	5.1162	24	11	14	0	0	very easy	quad.	non-conv.
kall_diffcircles_5b	5.1162	24	11	14	0	0	easy	quad.	non-conv.
kall_diffcircles_6	7.7879	31	16	16	0	0	very easy	quad.	non-conv.
kall_diffcircles_7	7.1531	40	22	18	0	0	medium	quad.	non-conv.
kall_diffcircles_8	14.4813	49	29	20	0	0	medium	quad.	non-conv.
kall_diffcircles_9	13.3503	60	37	22	0	0	medium	quad.	non-conv.
kall_ellipsoids_tc02b	32.4	128	48	124	0	0	medium	nonlin.	non-conv.

TABLE E.1. Benchmark set (cont.)

Instance	objective	Number of constr.	Number of nonlin. eqs.	Number of variables cont.	bin.	int.	Classification difficulty	type	convexity
kall_ellipsoids_tc03c	36.4536	196	74	193	0	0	medium	nonlin.	non-conv.
kall_ellipsoids_tc05a	39.3979	461	321	464	0	0	hard	poly.	non-conv.
kport40	37.1758	48	38	153	3	111	hard	sign.	non-conv.
kriging_peaks-full010	0.2911	24	20	26	0	0	easy	nonlin.	non-conv.
kriging_peaks-full020	0.3724	44	40	46	0	0	medium	nonlin.	non-conv.
kriging_peaks-full030	-1.5866	64	60	66	0	0	medium	nonlin.	non-conv.
kriging_peaks-full050	-1.1566	104	100	106	0	0	medium	nonlin.	non-conv.
kriging_peaks-full100	-2.6375	204	200	206	0	0	medium	nonlin.	non-conv.
kriging_peaks-full200	-3.8902	404	400	406	0	0	medium	nonlin.	non-conv.
kriging_peaks-full500	-4.928	1004	1000	1006	0	0	hard	nonlin.	non-conv.
m3	37.8	43	6	20	6	0	easy	sign.	conv.
m6	82.2569	157	12	56	30	0	easy	sign.	conv.
m7	106.7569	211	14	72	42	0	easy	sign.	conv.
m7_ar25_1	143.585	269	14	70	0	42	easy	sign.	conv.
m7_ar2_1	190.235	269	14	70	0	42	easy	sign.	conv.
m7_ar3_1	143.585	269	14	70	0	42	easy	sign.	conv.
m7_ar4_1	106.7569	269	14	70	0	42	easy	sign.	conv.
m7_ar5_1	106.46	269	14	70	0	42	easy	sign.	conv.
mathopt4	0.0	2	2	2	0	0	easy	nonlin.	non-conv.
mathopt5_1	-0.9996	0	1	1	0	0	easy	nonlin.	N/A
mathopt5_2	-1.0	0	1	1	0	0	easy	nonlin.	N/A
mathopt5_3	-1.6164	0	1	1	0	0	easy	nonlin.	non-conv.
mathopt5_4	0.0	0	1	1	0	0	easy	poly.	N/A
mathopt5_5	-14.838	0	1	1	0	0	easy	nonlin.	N/A
mathopt5_7	-4.4367	0	1	1	0	0	easy	poly.	non-conv.
mathopt5_8	-0.6861	0	1	1	0	0	easy	poly.	N/A
mathopt6	-3.3069	0	1	2	0	0	easy	nonlin.	N/A
milinfract	2.6339	501	1	500	500	0	hard	nonlin.	non-conv.
no7_ar25_1	107.8153	269	14	70	0	42	easy	sign.	conv.
no7_ar2_1	107.8153	269	14	70	0	42	easy	sign.	conv.
no7_ar3_1	107.8153	269	14	70	0	42	medium	sign.	conv.
no7_ar4_1	98.5184	269	14	70	0	42	medium	sign.	conv.
no7_ar5_1	90.6227	269	14	70	0	42	medium	sign.	conv.
nuclear14a	-1.1296	633	584	392	600	0	hard	quad.	non-conv.
nuclear14b	-1.1276	1785	560	968	600	0	hard	quad.	non-conv.
nuclear25a	-1.1207	659	608	408	650	0	hard	quad.	non-conv.
nuclear25b	-1.1158	1909	583	1033	650	0	hard	quad.	non-conv.
nvs04	0.72	0	1	0	0	2	easy	poly.	N/A
nvs06	1.7703	0	1	0	0	2	easy	sign.	N/A
nvs10	-310.8	2	3	0	0	2	very easy	quad.	conv.
nvs11	-431.0	3	4	0	0	3	very easy	quad.	conv.
nvs12	-481.2	4	5	0	0	4	very easy	quad.	conv.
nvs13	-585.2	5	6	0	0	5	very easy	quad.	non-conv.
nvs15	1.0	1	1	0	0	3	very easy	quad.	conv.
nvs17	-1100.4	7	8	0	0	7	very easy	quad.	non-conv.
nvs18	-778.4	6	7	0	0	6	very easy	quad.	non-conv.
nvs23	-1125.2	9	10	0	0	9	medium	quad.	non-conv.
nvs24	-1033.2	10	11	0	0	10	medium	quad.	non-conv.
o7	131.6531	211	14	72	42	0	medium	sign.	conv.
o7_2	116.9459	211	14	72	42	0	medium	sign.	conv.
o7_ar25_1	140.412	269	14	70	0	42	medium	sign.	conv.
o7_ar2_1	140.412	269	14	70	0	42	medium	sign.	conv.
o7_ar3_1	137.9318	269	14	70	0	42	medium	sign.	conv.
o7_ar4_1	131.6531	269	14	70	0	42	medium	sign.	conv.
o7_ar5_1	116.9458	269	14	70	0	42	medium	sign.	conv.
o8_ar4_1	243.0707	347	16	88	0	56	medium	sign.	conv.
o9_ar4_1	236.1385	435	18	108	0	72	hard	sign.	conv.
oaer	-1.9231	7	2	6	3	0	easy	nonlin.	non-conv.

TABLE E.1. Benchmark set (cont.)

Instance	objective	Number of constr.	Number of nonlin. eqs.	Number of variables cont.	bin.	int.	Classification difficulty	type	convexity
oil	-0.9325	1546	418	1516	19	0	hard	nonlin.	non-conv.
oil2	-0.7333	926	284	934	2	0	medium	nonlin.	non-conv.
ortez	-9532.0391	74	27	69	18	0	easy	sign.	non-conv.
orth_d3m6	0.7071	62	52	25	0	0	medium	poly.	non-conv.
orth_d3m6_pl	0.7071	127	66	42	0	0	medium	poly.	non-conv.
orth_d4m6_pl	0.6495	86	41	42	0	0	medium	poly.	non-conv.
p_ball_10b_5p_2d_m	18.7186	109	50	30	50	0	very easy	quad.	conv.
p_ball_10b_5p_4d_m	71.3719	149	50	60	50	0	easy	quad.	conv.
p_ball_10b_7p_3d_m	109.8032	219	70	84	70	0	easy	quad.	conv.
p_ball_15b_5p_2d_m	6.5999	139	75	30	75	0	easy	quad.	conv.
p_ball_20b_5p_2d_m	2.4372	169	100	30	100	0	easy	quad.	conv.
p_ball_20b_5p_3d_m	19.7365	189	100	45	100	0	medium	quad.	conv.
p_ball_30b_10p_2d_m	41.934	529	300	110	300	0	medium	quad.	conv.
p_ball_30b_5p_2d_m	0.2916	229	150	30	150	0	medium	quad.	conv.
p_ball_30b_5p_3d_m	8.2183	249	150	45	150	0	medium	quad.	conv.
p_ball_30b_7p_2d_m	13.9338	337	210	56	210	0	medium	quad.	conv.
p_ball_40b_5p_3d_m	9.7772	309	200	45	200	0	medium	quad.	conv.
p_ball_40b_5p_4d_m	30.1327	329	200	60	200	0	medium	quad.	conv.
pindyck	-1170.4863	96	32	116	0	0	medium	nonlin.	non-conv.
pointpack02	2.0	3	1	5	0	0	very easy	quad.	non-conv.
pointpack04	1.0	10	6	9	0	0	very easy	quad.	non-conv.
pointpack06	0.3611	21	15	13	0	0	very easy	quad.	non-conv.
pointpack08	0.2679	36	28	17	0	0	medium	quad.	non-conv.
pointpack10	0.1775	55	45	21	0	0	medium	quad.	non-conv.
pointpack12	0.1511	78	66	25	0	0	medium	quad.	non-conv.
pointpack14	0.1217	105	91	29	0	0	medium	quad.	non-conv.
polygon25	-0.7798	324	301	50	0	0	hard	nonlin.	non-conv.
polygon50	-0.7839	1274	1226	100	0	0	hard	nonlin.	non-conv.
pooling_adhya1pq	-549.8031	49	20	33	0	0	very easy	quad.	non-conv.
pooling_adhya1stp	-549.8031	71	40	46	0	0	medium	quad.	non-conv.
pooling_adhya1tp	-549.8031	49	20	33	0	0	easy	quad.	non-conv.
pooling_adhya2pq	-549.8031	57	20	33	0	0	very easy	quad.	non-conv.
pooling_adhya2stp	-549.8031	79	40	46	0	0	medium	quad.	non-conv.
pooling_adhya2tp	-549.8031	57	20	33	0	0	easy	quad.	non-conv.
pooling_adhya3pq	-561.0447	74	32	52	0	0	easy	quad.	non-conv.
pooling_adhya3stp	-561.0447	109	64	72	0	0	medium	quad.	non-conv.
pooling_adhya3tp	-561.0447	74	32	52	0	0	medium	quad.	non-conv.
pooling_adhya4pq	-877.6457	77	40	58	0	0	very easy	quad.	non-conv.
pooling_adhya4stp	-877.6457	119	80	76	0	0	easy	quad.	non-conv.
pooling_adhya4tp	-877.6457	77	40	58	0	0	easy	quad.	non-conv.
pooling_bental4pq	-450.0	16	6	13	0	0	very easy	quad.	non-conv.
pooling_bental4tp	-450.0	16	6	13	0	0	very easy	quad.	non-conv.
pooling_bental5pq	-3500.0	86	60	92	0	0	very easy	quad.	non-conv.
pooling_bental5stp	-3500.0	149	120	119	0	0	medium	quad.	non-conv.
pooling_bental5tp	-3500.0	86	60	92	0	0	very easy	quad.	non-conv.
pooling_digabel16	-2410.6877	117	81	171	0	0	medium	quad.	non-conv.
pooling_digabel18	-689.1606	412	390	208	0	0	medium	quad.	non-conv.
pooling_digabel19	-4539.9122	171	128	212	0	0	medium	quad.	non-conv.
pooling_foulds2pq	-1100.0	34	16	36	0	0	very easy	quad.	non-conv.
pooling_foulds2stp	-1100.0	52	32	48	0	0	very easy	quad.	non-conv.
pooling_foulds2tp	-1100.0	34	16	36	0	0	very easy	quad.	non-conv.
pooling_foulds3pq	-8.0	571	512	672	0	0	very easy	quad.	non-conv.
pooling_foulds3stp	-8.0	1091	1024	832	0	0	medium	quad.	non-conv.
pooling_foulds3tp	-8.0	571	512	672	0	0	very easy	quad.	non-conv.
pooling_foulds4pq	-8.0	571	512	672	0	0	very easy	quad.	non-conv.
pooling_foulds4stp	-8.0	1091	1024	832	0	0	hard	quad.	non-conv.
pooling_foulds4tp	-8.0	571	512	672	0	0	very easy	quad.	non-conv.
pooling_foulds5pq	-8.0	563	512	608	0	0	very easy	quad.	non-conv.

TABLE E.1. Benchmark set (cont.)

Instance	objective	Number of constr.	Number of nonlin. eqs.	Number of variables cont.	bin.	int.	Classification difficulty	type	convexity
pooling_foulds5stp	-8.0	1079	1024	704	0	0	hard	quad.	non-conv.
pooling_foulds5tp	-8.0	563	512	608	0	0	easy	quad.	non-conv.
pooling_haverly1pq	-400.0	13	4	10	0	0	very easy	quad.	non-conv.
pooling_haverly1stp	-400.0	18	8	14	0	0	very easy	quad.	non-conv.
pooling_haverly1tp	-400.0	13	4	10	0	0	very easy	quad.	non-conv.
pooling_haverly2pq	-600.0	13	4	10	0	0	very easy	quad.	non-conv.
pooling_haverly2stp	-600.0	18	8	14	0	0	very easy	quad.	non-conv.
pooling_haverly2tp	-600.0	13	4	10	0	0	very easy	quad.	non-conv.
pooling_haverly3pq	-750.0	13	4	10	0	0	very easy	quad.	non-conv.
pooling_haverly3stp	-750.0	18	8	14	0	0	very easy	quad.	non-conv.
pooling_haverly3tp	-750.0	13	4	10	0	0	very easy	quad.	non-conv.
pooling_rt2pq	-4391.8259	52	18	34	0	0	very easy	quad.	non-conv.
pooling_rt2stp	-4391.8259	72	36	46	0	0	very easy	quad.	non-conv.
pooling_rt2tp	-4391.8259	52	18	34	0	0	very easy	quad.	non-conv.
pooling_sppa0pq	-35812.3336	744	329	500	0	0	hard	quad.	non-conv.
pooling_sppa0stp	-35812.3336	1083	658	616	0	0	hard	quad.	non-conv.
pooling_sppa0tp	-35812.3336	744	329	500	0	0	hard	quad.	non-conv.
pooling_sppa5pq	-27915.8392	1383	968	1245	0	0	hard	quad.	non-conv.
pooling_sppa5stp	-27829.022	2361	1936	1441	0	0	hard	quad.	non-conv.
pooling_sppa5tp	-27922.0748	1383	968	1245	0	0	hard	quad.	non-conv.
pooling_sppa9pq	-21933.994	2407	1992	2399	0	0	hard	quad.	non-conv.
pooling_sppa9tp	-21933.994	2407	1992	2399	0	0	hard	quad.	non-conv.
pooling_sppb0pq	-43412.4119	1957	1153	1537	0	0	hard	quad.	non-conv.
pooling_sppb0tp	-43372.8152	1957	1153	1537	0	0	hard	quad.	non-conv.
portfol_classical050_1	-0.0948	103	1	100	50	0	easy	quad.	conv.
portfol_classical200_2	-0.1101	403	1	400	200	0	medium	quad.	conv.
portfol_robust050_34	-0.0721	156	2	152	51	0	easy	quad.	non-conv.
portfol_robust100_09	-0.105	306	2	302	101	0	medium	quad.	non-conv.
portfol_robust200_03	-0.1291	606	2	602	201	0	medium	quad.	non-conv.
portfol_shortfall050_68	-1.0972	157	2	153	51	0	easy	quad.	non-conv.
portfol_shortfall100_04	-1.1179	307	2	303	101	0	medium	quad.	non-conv.
portfol_shortfall200_05	-1.1273	607	2	603	201	0	medium	quad.	non-conv.
pricing050	-1813.8291	5	5	50	0	0	medium	nonlin.	N/A
prob03	10.0	1	1	0	0	2	very easy	quad.	non-conv.
prob06	1.1771	2	2	2	0	0	very easy	quad.	non-conv.
prob09	-0.0	1	1	3	0	0	easy	poly.	non-conv.
prob10	3.4455	2	1	1	0	1	easy	nonlin.	N/A
procsel	-1.9231	7	2	7	3	0	easy	nonlin.	non-conv.
procurement2mot	212.0707	761	12	736	60	0	easy	sign.	conv.
product	-2142.9481	1925	132	1446	107	0	easy	quad.	non-conv.
product2	-2102.3893	3125	528	2714	128	0	medium	quad.	non-conv.
rbrock	0.0	0	1	2	0	0	easy	poly.	N/A
ringpack_10_1	-20.0665	385	330	20	50	0	medium	quad.	non-conv.
ringpack_10_2	-20.0665	475	420	20	60	0	medium	quad.	non-conv.
ringpack_20_1	-36.3387	2547	2337	40	175	0	medium	quad.	non-conv.
ringpack_20_2	-39.3413	2927	2717	40	195	0	medium	quad.	non-conv.
ringpack_20_3	-38.6317	3228	3055	40	213	0	medium	quad.	non-conv.
rsyn0805hfsg	1296.1206	429	3	271	37	0	easy	nonlin.	conv.
rsyn0805m	1296.1206	286	3	101	69	0	easy	nonlin.	conv.
rsyn0805m02hfsg	2238.3954	1045	6	552	148	0	easy	nonlin.	conv.
rsyn0805m02m	2238.3954	769	6	212	148	0	easy	nonlin.	conv.
rsyn0805m03hfsg	3068.9314	1698	9	828	222	0	easy	nonlin.	conv.
rsyn0805m03m	3068.9314	1284	9	318	222	0	easy	nonlin.	conv.
rsyn0805m04hfsg	7174.219	2438	12	1104	296	0	easy	nonlin.	conv.
rsyn0805m04m	7174.219	1886	12	424	296	0	easy	nonlin.	conv.
rsyn0810hfsg	1721.4477	483	6	301	42	0	easy	nonlin.	conv.
rsyn0810m	1721.4477	312	6	111	74	0	easy	nonlin.	conv.
rsyn0810m02hfsg	1741.3869	1188	12	622	168	0	medium	nonlin.	conv.

TABLE E.1. Benchmark set (cont.)

Instance	objective	Number of constr.	Number of nonlin. eqs.	Number of variables cont.	bin.	int.	Classification difficulty	type	convexity
rsyn0810m02m	1741.3869	866	12	242	168	0	easy	nonlin.	conv.
rsyn0810m03hfsg	2722.448	1935	18	933	252	0	medium	nonlin.	conv.
rsyn0810m03m	2722.448	1452	18	363	252	0	easy	nonlin.	conv.
rsyn0810m04hfsg	6581.9344	2784	24	1244	336	0	medium	nonlin.	conv.
rsyn0810m04m	6581.9344	2140	24	484	336	0	easy	nonlin.	conv.
rsyn0815hfsg	1269.9256	552	11	340	47	0	easy	nonlin.	conv.
rsyn0815m	1269.9256	347	11	126	79	0	easy	nonlin.	conv.
rsyn0815m02hfsg	1774.3973	1361	22	710	188	0	medium	nonlin.	conv.
rsyn0815m02m	1774.3973	981	22	282	188	0	easy	nonlin.	conv.
rsyn0815m03hfsg	2827.9259	2217	33	1065	282	0	medium	nonlin.	conv.
rsyn0815m03m	2827.9259	1647	33	423	282	0	easy	nonlin.	conv.
rsyn0815m04hfsg	3410.8543	3190	44	1420	376	0	medium	nonlin.	conv.
rsyn0815m04m	3410.8543	2430	44	564	376	0	medium	nonlin.	conv.
rsyn0820hfsg	1150.3005	604	14	365	52	0	easy	nonlin.	conv.
rsyn0820m	1150.3005	371	14	131	84	0	easy	nonlin.	conv.
rsyn0820m02hfsg	1092.0911	1500	28	770	208	0	medium	nonlin.	conv.
rsyn0820m02m	1092.0911	1074	28	302	208	0	easy	nonlin.	conv.
rsyn0820m03m	2028.8119	1809	42	453	312	0	easy	nonlin.	conv.
rsyn0820m04hfsg	2450.7722	3528	56	1540	416	0	medium	nonlin.	conv.
rsyn0820m04m	2450.7722	2676	56	604	416	0	medium	nonlin.	conv.
rsyn0830hfsg	510.072	716	20	432	62	0	easy	nonlin.	conv.
rsyn0830m	510.072	425	20	156	94	0	easy	nonlin.	conv.
rsyn0830m02hfsg	730.5072	1794	40	924	248	0	medium	nonlin.	conv.
rsyn0830m02m	730.5072	1272	40	372	248	0	easy	nonlin.	conv.
rsyn0830m03hfsg	1543.0593	2934	60	1386	372	0	medium	nonlin.	conv.
rsyn0830m03m	1543.0593	2151	60	558	372	0	easy	nonlin.	conv.
rsyn0830m04hfsg	2529.0734	4236	80	1848	496	0	medium	nonlin.	conv.
rsyn0830m04m	2529.0734	3192	80	744	496	0	easy	nonlin.	conv.
rsyn0840hfsg	325.5545	837	28	496	72	0	medium	nonlin.	conv.
rsyn0840m	325.5545	484	28	176	104	0	easy	nonlin.	conv.
rsyn0840m02hfsg	734.9835	2106	56	1072	288	0	medium	nonlin.	conv.
rsyn0840m02m	734.9835	1480	56	432	288	0	easy	nonlin.	conv.
rsyn0840m03m	2742.6457	2508	84	648	432	0	easy	nonlin.	conv.
rsyn0840m04hfsg	2564.4995	4980	112	2144	576	0	medium	nonlin.	conv.
rsyn0840m04m	2564.4995	3728	112	864	576	0	medium	nonlin.	conv.
sep1	-510.081	31	6	27	2	0	very easy	quad.	non-conv.
sfacloc1_2_80	12.7521	2088	15	169	62	0	medium	poly.	non-conv.
sfacloc1_2_90	17.8916	348	15	169	30	0	medium	poly.	non-conv.
sfacloc1_2_95	18.8501	208	15	162	9	0	medium	poly.	non-conv.
sfacloc1_3_80	8.5231	2161	15	231	62	0	medium	poly.	non-conv.
sfacloc1_3_90	11.622	421	15	231	30	0	medium	poly.	non-conv.
sfacloc1_3_95	12.3025	281	15	224	9	0	medium	poly.	non-conv.
sfacloc1_4_80	7.8791	2234	15	293	62	0	medium	poly.	non-conv.
sfacloc1_4_90	10.4575	494	15	293	30	0	medium	poly.	non-conv.
sfacloc1_4_95	11.1841	354	15	286	9	0	medium	poly.	non-conv.
sfacloc2_2_80	13.2795	2165	30	154	92	0	easy	poly.	non-conv.
sfacloc2_2_90	18.5941	393	30	154	60	0	easy	poly.	non-conv.
sfacloc2_2_95	19.5776	239	30	147	39	0	very easy	poly.	non-conv.
sfacloc2_3_80	11.0585	2268	45	216	107	0	medium	poly.	non-conv.
sfacloc2_3_90	15.0945	496	45	216	75	0	medium	poly.	non-conv.
sfacloc2_3_95	16.1511	342	45	209	54	0	very easy	poly.	non-conv.
sfacloc2_4_80	9.9531	2371	60	278	122	0	medium	poly.	non-conv.
sfacloc2_4_90	13.4115	599	60	278	90	0	medium	poly.	non-conv.
sfacloc2_4_95	14.2992	445	60	271	69	0	very easy	poly.	non-conv.
st_bpaf1a	-45.3797	10	1	10	0	0	very easy	quad.	non-conv.
st_bpaf1b	-42.9626	10	1	10	0	0	very easy	quad.	non-conv.
st_bpk1	-13.0	6	1	4	0	0	very easy	quad.	non-conv.
st_bsj3	-86768.55	1	1	6	0	0	very easy	quad.	non-conv.

TABLE E.1. Benchmark set (cont.)

Instance	objective	Number of constr.	Number of nonlin. eqs.	Number of variables cont.	bin.	int.	Classification difficulty	type	convexity
st_bsj4	-70262.05	4	1	6	0	0	very easy	quad.	non-conv.
st_e01	-6.6667	1	1	2	0	0	very easy	quad.	non-conv.
st_e02	201.1593	3	3	3	0	0	very easy	quad.	non-conv.
st_e05	7049.2493	3	2	5	0	0	very easy	quad.	non-conv.
st_e06	0.0	3	1	3	0	0	easy	poly.	non-conv.
st_e07	-400.0	7	3	10	0	0	very easy	quad.	non-conv.
st_e08	0.7418	2	2	2	0	0	very easy	quad.	non-conv.
st_e09	-0.5	1	2	2	0	0	very easy	quad.	non-conv.
st_e12	-4.5142	3	1	4	0	0	easy	sign.	non-conv.
st_e13	2.0	2	1	1	1	0	very easy	quad.	non-conv.
st_e15	7.6672	5	2	2	3	0	easy	sign.	non-conv.
st_e18	-2.8284	4	2	2	0	0	very easy	quad.	non-conv.
st_e19	-118.7049	2	2	2	0	0	easy	poly.	non-conv.
st_e21	-13.4019	6	1	6	0	0	easy	sign.	non-conv.
st_e22	-85.0	5	1	2	0	0	very easy	quad.	non-conv.
st_e23	-1.0833	2	1	2	0	0	very easy	quad.	non-conv.
st_e24	3.0	4	1	2	0	0	very easy	quad.	non-conv.
st_e26	-185.7792	4	1	2	0	0	very easy	quad.	non-conv.
st_e27	2.0	6	1	2	2	0	very easy	quad.	non-conv.
st_e30	-1.5811	15	5	14	0	0	very easy	quad.	non-conv.
st_e31	-2.0	135	5	88	24	0	very easy	quad.	non-conv.
st_e33	-400.0	6	3	9	0	0	very easy	quad.	non-conv.
st_e34	0.0156	4	4	6	0	0	very easy	quad.	non-conv.
st_e40	30.4142	8	4	1	0	3	easy	poly.	non-conv.
st_e41	641.8236	2	3	4	0	0	easy	sign.	non-conv.
st_glmp_fp3	-12.0	8	1	4	0	0	very easy	quad.	non-conv.
st_glmp_kk92	-12.0	8	1	4	0	0	very easy	quad.	non-conv.
st_ht	-1.6	3	1	2	0	0	very easy	quad.	non-conv.
st_iqpbk1	-621.4878	7	1	8	0	0	very easy	quad.	non-conv.
st_iqpbk2	-1195.2256	7	1	8	0	0	very easy	quad.	non-conv.
st_miqp3	-6.0	1	1	0	0	2	very easy	quad.	conv.
st_miqp4	-4574.0	4	1	3	0	3	very easy	quad.	conv.
st_miqp5	-333.8889	13	1	5	0	2	very easy	quad.	conv.
st_pan1	-5.2837	4	1	3	0	0	very easy	quad.	non-conv.
st_ph1	-230.1173	5	1	6	0	0	very easy	quad.	non-conv.
st_ph11	-11.2812	4	1	3	0	0	very easy	quad.	non-conv.
st_ph12	-22.625	4	1	3	0	0	very easy	quad.	non-conv.
st_ph13	-11.2812	10	1	3	0	0	very easy	quad.	non-conv.
st_ph14	-229.7222	10	1	3	0	0	very easy	quad.	non-conv.
st_ph15	-392.7037	4	1	4	0	0	very easy	quad.	non-conv.
st_ph2	-1028.1173	5	1	6	0	0	very easy	quad.	non-conv.
st_ph3	-420.2348	5	1	6	0	0	very easy	quad.	non-conv.
st_qpc-m3a	-382.695	10	1	10	0	0	very easy	quad.	non-conv.
st_qpc-m3b	0.0	10	1	10	0	0	very easy	quad.	non-conv.
st_qpk2	-12.25	12	1	6	0	0	very easy	quad.	non-conv.
st_qpk3	-36.0	22	1	11	0	0	very easy	quad.	non-conv.
st_robot	0.0	8	7	8	0	0	very easy	quad.	non-conv.
st_rv1	-59.9439	5	1	10	0	0	very easy	quad.	non-conv.
st_rv2	-64.4807	10	1	20	0	0	very easy	quad.	non-conv.
st_rv3	-35.7607	20	1	20	0	0	very easy	quad.	non-conv.
st_rv7	-138.1875	20	1	30	0	0	very easy	quad.	non-conv.
st_rv8	-132.6616	20	1	40	0	0	very easy	quad.	non-conv.
st_rv9	-120.1531	20	1	50	0	0	very easy	quad.	non-conv.
st_testgr1	-12.8116	5	1	0	0	10	very easy	quad.	conv.
st_testgr3	-20.59	20	1	0	0	20	very easy	quad.	conv.
st_testph4	-80.5	10	1	0	0	3	very easy	quad.	conv.
supplychain	2260.2566	30	6	24	3	0	very easy	quad.	non-conv.
syn05hfsg	837.7324	58	3	37	5	0	easy	nonlin.	conv.

TABLE E.1. Benchmark set (cont.)

Instance	objective	Number of constr.	Number of nonlin. eqs.	Number of variables cont.	bin.	int.	Classification difficulty	type	convexity
syn05m02hfsg	3032.7354	151	6	84	20	0	easy	nonlin.	conv.
syn05m02m	3032.7354	101	6	40	20	0	easy	nonlin.	conv.
syn05m03hfsg	4027.3718	249	9	126	30	0	easy	nonlin.	conv.
syn05m03m	4027.3718	174	9	60	30	0	easy	nonlin.	conv.
syn05m04hfsg	5510.3873	362	12	168	40	0	easy	nonlin.	conv.
syn05m04m	5510.3873	262	12	80	40	0	easy	nonlin.	conv.
syn10hfsg	1267.3536	112	6	67	10	0	easy	nonlin.	conv.
syn10m	1267.3536	54	6	25	10	0	easy	nonlin.	conv.
syn10m02hfsg	2310.3007	294	12	154	40	0	easy	nonlin.	conv.
syn10m02m	2310.3007	198	12	70	40	0	easy	nonlin.	conv.
syn10m03hfsg	3354.6828	486	18	231	60	0	easy	nonlin.	conv.
syn10m03m	3354.6828	342	18	105	60	0	easy	nonlin.	conv.
syn10m04hfsg	4557.0623	708	24	308	80	0	easy	nonlin.	conv.
syn10m04m	4557.0623	516	24	140	80	0	easy	nonlin.	conv.
syn15hfsg	853.2847	181	11	106	15	0	easy	nonlin.	conv.
syn15m	853.2847	89	11	40	15	0	easy	nonlin.	conv.
syn15m02hfsg	2832.7489	467	22	242	60	0	easy	nonlin.	conv.
syn15m02m	2832.7489	313	22	110	60	0	easy	nonlin.	conv.
syn15m03hfsg	3850.1818	768	33	363	90	0	easy	nonlin.	conv.
syn15m03m	3850.1818	537	33	165	90	0	easy	nonlin.	conv.
syn15m04hfsg	4937.4777	1114	44	484	120	0	medium	nonlin.	conv.
syn15m04m	4937.4777	806	44	220	120	0	easy	nonlin.	conv.
syn20hfsg	924.2633	233	14	131	20	0	easy	nonlin.	conv.
syn20m	924.2633	113	14	45	20	0	easy	nonlin.	conv.
syn20m02hfsg	1752.1332	606	28	302	80	0	medium	nonlin.	conv.
syn20m02m	1752.1332	406	28	130	80	0	easy	nonlin.	conv.
syn20m03hfsg	2646.9509	999	42	453	120	0	medium	nonlin.	conv.
syn20m03m	2646.9509	699	42	195	120	0	easy	nonlin.	conv.
syn20m04hfsg	3532.7439	1452	56	604	160	0	medium	nonlin.	conv.
syn20m04m	3532.7439	1052	56	260	160	0	easy	nonlin.	conv.
syn30hfsg	138.1596	345	20	198	30	0	medium	nonlin.	conv.
syn30m	138.1596	167	20	70	30	0	easy	nonlin.	conv.
syn30m02hfsg	399.6831	900	40	456	120	0	hard	nonlin.	conv.
syn30m02m	399.6831	604	40	200	120	0	easy	nonlin.	conv.
syn30m03hfsg	654.1542	1485	60	684	180	0	medium	nonlin.	conv.
syn30m03m	654.1542	1041	60	300	180	0	easy	nonlin.	conv.
syn30m04hfsg	865.722	2160	80	912	240	0	medium	nonlin.	conv.
syn30m04m	865.722	1568	80	400	240	0	easy	nonlin.	conv.
syn40hfsg	67.7133	466	28	262	40	0	medium	nonlin.	conv.
syn40m	67.7133	226	28	90	40	0	easy	nonlin.	conv.
syn40m02hfsg	388.7724	1212	56	604	160	0	medium	nonlin.	conv.
syn40m02m	388.7724	812	56	260	160	0	easy	nonlin.	conv.
syn40m03hfsg	395.148	1998	84	906	240	0	medium	nonlin.	conv.
syn40m03m	395.148	1398	84	390	240	0	easy	nonlin.	conv.
syn40m04hfsg	901.7511	2904	112	1208	320	0	medium	nonlin.	conv.
syn40m04m	901.7511	2104	112	520	320	0	easy	nonlin.	conv.
t1000	0.0	28	24	1002	0	0	easy	nonlin.	non-conv.
tanksize	1.2686	74	21	38	9	0	easy	sign.	non-conv.
tln12	90.5	72	12	0	12	156	medium	quad.	non-conv.
tln2	5.3	12	2	0	2	6	very easy	quad.	non-conv.
tln4	8.3	24	4	0	4	20	very easy	quad.	non-conv.
tln5	10.3	30	5	0	5	30	easy	quad.	non-conv.
tln6	15.3	36	6	0	6	42	easy	quad.	non-conv.
tln7	15.0	42	7	0	7	56	medium	quad.	non-conv.
tloss	16.3	53	6	0	6	42	very easy	quad.	non-conv.
tls4	8.3	64	4	16	85	4	medium	sign.	conv.
tls5	10.3	90	5	25	131	5	medium	sign.	conv.
tls6	15.3	120	6	36	173	6	medium	sign.	conv.

TABLE E.1. Benchmark set (cont.)

Instance	objective	Number of constr.	Number of nonlin. eqs.	Number of variables cont.	bin.	int.	Classification difficulty	type	convexity
tls7	15.0	154	7	49	289	7	hard	sign.	conv.
tltr	48.0667	54	3	0	12	36	very easy	quad.	non-conv.
torsion25	-0.4175	4	2	1408	0	0	medium	quad.	non-conv.
trig	-3.7625	1	2	1	0	0	easy	nonlin.	N/A
util	999.5788	167	4	117	28	0	very easy	quad.	non-conv.
wastepaper3	0.0189	30	16	25	27	0	easy	sign.	non-conv.
wastepaper4	0.0035	38	20	32	44	0	medium	sign.	non-conv.
wastepaper5	0.0008	46	24	39	65	0	medium	sign.	non-conv.
wastepaper6	0.0001	54	28	46	90	0	medium	sign.	non-conv.
waterno2_01	19.4567	204	65	157	9	0	easy	poly.	non-conv.
waterno2_02	39.5714	410	130	314	18	0	easy	poly.	non-conv.
waterno2_03	115.0045	616	195	471	27	0	medium	poly.	non-conv.
waterno2_04	145.4398	822	260	628	36	0	medium	poly.	non-conv.
waterno2_06	282.888	1234	390	942	54	0	medium	poly.	non-conv.
waterno2_09	922.5953	1852	585	1413	81	0	medium	poly.	non-conv.
waterno2_12	2263.3584	2470	780	1884	108	0	hard	poly.	non-conv.
watertreatnd_conc	348336.7614	319	29	355	5	0	easy	sign.	non-conv.
watertreatnd_flow	348337.0417	379	155	415	5	0	hard	sign.	non-conv.
waterund01	86.8333	38	14	40	0	0	medium	quad.	non-conv.
waterund08	164.4898	95	40	90	0	0	medium	quad.	non-conv.
waterund11	104.8861	64	28	64	0	0	medium	quad.	non-conv.
waterund14	329.5698	135	66	125	0	0	medium	quad.	non-conv.
waterund17	157.0944	66	27	74	0	0	medium	quad.	non-conv.
waterund18	238.7333	64	28	60	0	0	medium	quad.	non-conv.
waterund22	323.5051	135	66	146	0	0	medium	quad.	non-conv.
waterund25	410.6354	87	36	121	0	0	medium	quad.	non-conv.
waterund27	556.6752	208	96	432	0	0	medium	quad.	non-conv.
waterund28	1812.1703	540	240	760	0	0	medium	quad.	non-conv.
waterund32	638.7168	380	160	660	0	0	medium	quad.	non-conv.
waterund36	662.807	239	110	324	0	0	medium	quad.	non-conv.

APPENDIX F. OVERVIEW OVER FIGURES AND TABLES

LIST OF FIGURES

List of Tables