A Practical Scheme to Compute Pessimistic Bilevel Optimization Problem

Bo Zeng
Department of Industrial Engineering
University of Pittsburgh, Pittsburgh, PA 15261

Abstract

In this paper, we present a new computation scheme for pessimistic bilevel optimization problem, which so far does not have any computational methods generally applicable yet. We first develop a tight relaxation and then design a simple scheme to ensure a feasible and optimal solution. Then, we discuss using this scheme to analyze and compute linear pessimistic bilevel problem and several extensions. We also provide demonstrations on illustrative examples, and a systematic numerical study on instances of two practical problems. Because of its simple structure and strong computational capacity, we believe that the developed scheme is of a critical value in studying and solving pessimistic bilevel optimization problems arising from practice.

1 Introduction

Bilevel optimization is a popular modeling and computing tool for non-centralized decision making problems where two decision makers (DMs), i.e., the upper level and the lower level DMs, interact sequentially. In this paper, we consider the *pessimistic formulation* (also known as the *weak formulation*) of bilevel optimization [32], which can be represented in the following mathematical form:

$$\mathbf{PBL}: \quad \Theta_p^* = \min_{\mathbf{x}} \max_{\mathbf{y}} \mathbf{F}(\mathbf{x},\mathbf{y}) \tag{1}$$

$$\text{s.t.} \quad \mathbf{x} \in \mathbf{X} \tag{2}$$

$$\mathbf{y} \in \mathbf{S}(\mathbf{x}) = \arg\min \Big\{ \mathbf{f}(\mathbf{x},\mathbf{y}) : \mathbf{y} \in \mathbf{Y}(\mathbf{x}) \Big\} \tag{3}$$

where $\mathbf{X} \subseteq \mathbb{R}^n$ and $\mathbf{Y}(\mathbf{x}) \subseteq \mathbb{R}^m$ for any $\mathbf{x}$. We mention that $\mathbf{x}$ or $\mathbf{y}$ variables are not necessary to be continuous and they can be discrete. The optimization problem defined in (1-2) is referred to as the upper level DM's (her) decision problem and the one appearing in (3) is called the lower level DM's (his) decision problem. Hence, $\mathbf{S}(\mathbf{x})$ represents the collection of optimal solutions of the lower level problem for a given $\mathbf{x}$.

Note that if we simply drop the *max* operation in (1), i.e.,

$$\min_{\mathbf{x}} \max_{\mathbf{y}} \mathbf{F}(\mathbf{x}, \mathbf{y}) \Rightarrow \min_{\mathbf{x},\, \mathbf{y}} \mathbf{F}(\mathbf{x}, \mathbf{y}),$$

the formulation in (1-3) becomes the *optimistic formulation* (also known as the *strong formulation*). We adopt **OBL** and Θ_o^* to denote the latter one and its optimal value, respectively. It is worth noting that although both **OBL** and **PBL** are introduced and investigated in the bilevel optimization literature, **OBL** is with an actual bilevel structure and **PBL** is indeed a three-level formulation. Given that bilevel optimization is generally referred to **OBL** in the mainstream of research literature, we use **OBL** and *bilevel optimization* interchangeably in the remainder of this paper, and explicitly mention pessimistic bilevel or **PBL** when needed.

Clearly, if $\mathbf{S}(\mathbf{x})$ is a singleton for every $\mathbf{x}$, **PBL** reduces to **OBL**, i.e., its three-level structure reduces to that of **OBL**. Nevertheless, $\mathbf{S}(\mathbf{x})$ might be a set with multiple elements, which is often observed in a few decomposition algorithms for bilevel optimization problem [33] (where a mixed integer program is solved through computing its bilevel optimization reformulation) and [42]. Given that, if the lower level DM is believed to be cooperative by always selecting a solution from $\mathbf{S}(\mathbf{x})$ in favor of the upper level DM, we can adopt the optimistic formulation. Nevertheless, if the lower level DM behaves non-cooperatively by taking a solution against the upper level DM's interest, we should employ the pessimistic formulation for modeling and computing.

The origin of bilevel optimization can be linked to the development of Stackelberg leader-follower game [39] on the investigation of market equilibrium in 1930s. Its first explicit mathematical formulation can be found in [11] in 1970s. Over the last 40 years, there is a great amount of research developed on bilevel optimization, ([22, 7, 19]). Theoretically, it is recognized that computing a bilevel optimization problem is not easy. Even for (optimistic) bilevel linear programming (LP) problem where both upper and lower level problems are LPs, it is NP-hard ([25, 6]). To address such challenge, many computational strategies and algorithms have been proposed, developed and investigated (see [22, 7, 19, 42] and references therein), the vast majority of which, however, is devoted to solving **OBL** formulation. As a result, those advanced and capable computing tools strongly support the real applications of optimistic bilevel models in transportation planning and capacity expansion ([12, 35]), government policy making ([8, 17]), revenue management ([13, 20]), electricity market ([27, 26]) and computational biology ([37, 14]).

It is generally believed that **PBL** is much more difficult and requires deeper analysis and investigations. Hence, **PBL** model and its variants recently have received a substantial amount of attention. One way to attack **PBL** is to compute a sequence of optimistic bilevel models that include a penalty term on the upper level objective value in the lower level objective function ([32, 34, 9]). Similarly, when **PBL** is pure linear, a penalty method is developed in [4], which solves **PBL** by a sequence of bilinear programs that penalize the difference between the primal and dual of lower level problem. In a series of new papers, [29, 31, 30], more general inner regularization techniques have been proposed, which can be used to systematically compute viscosity solutions for **PBL** and its variants with quasi-variational or quasi-equilibrium constraints. Note also that [43] fixes a flaw in [4] and customizes the

Kth-Best algorithm ([10]) to compute linear **PBL**. Very recently, [23, 41] study linear **PBL** through its single level form obtained by using strong duality or the optimal value function. A global solution procedure by enumerating appropriate basic matrices of the lower level problem and a descent algorithm for local solutions are also presented in [23]. A heuristic and iterative procedure is developed in [18], which directly computes relaxed pessimistic solutions of a mathematical program with equilibrium constraints.

One type of bilevel problem with coupled pessimistic constraints, i.e., constraints must be satisfied by the selected $\mathbf{x}$ and all solutions in $\mathbf{S}(\mathbf{x})$, is studied in [38, 40]. Authors argue that it is a generalization of the standard **PBL**. For the case where the feasible set of the lower level problem is not affected by the upper level variable $\mathbf{x}$, which is referred to as the independent **PBL**, an iterative procedure to compute ϵ-approximation solutions is developed. Another modeling extension is the strong-weak bilevel problem, which is a weighted sum of optimistic and pessimistic formulations [3, 16, 44, 2]. Clearly, by assigning different weights to them, we can have a trade-off between optimistic and pessimistic models, which in turn can be treated as two extreme situations of the strong-weak bilevel problem. In the aforementioned papers, various analytical properties, theoretical insights, and regularized computing methods have been developed for linear or nonlinear strong-weak bilevel problems in finite or infinite dimensional cases.

Overall, we note that almost all the aforementioned solution methods for general **PBL** are either iterative procedures that require a global solution at every update, or depend on non-trivial enumerative subroutines. Hence, they currently provide little support to practical **PBL** instances. It is worth mentioning that in [21, 24], various necessary optimality conditions for continuous **PBL** are derived, which could be helpful to develop fast algorithms to compute (local) optimal solutions.

In this paper, we focus on developing a new computing scheme for **PBL** formulation that has an optimal solution, and also demonstrate its usage to deal with some extensions. Different from existing research, our basic idea is to create and make use of a non-traditional bilevel optimization relaxation of **PBL**. Specifically, we introduce additional variables and constraints to directly describe the interaction between the upper and lower level DMs. Although it increases dimensionalities of both upper and lower level problems, which are in a relatively small scale, we obtain a tight bilevel relaxation to the three-level **PBL**, which then can be effectively solved by many existing solution methods. Hence, through computing such tight bilevel relaxation and a simple correction operation, an optimal solution to the original **PBL** can be recovered. Indeed, with this new computing scheme, we are able to solve instances of a couple of practical problems, where outputs from **OBL** and **PBL** could be significantly different.

The remainder of this paper is organized as follows. In Section 2, we develop a bilevel optimization formulation that is a tight relaxation to the general **PBL**, and present a two-step computing scheme to derive a feasible and optimal solution for **PBL**. In Section 3, we then discuss using this scheme to analyze and compute linear **PBL** and its several extensions. In Section 4, we first provide numerical illustrations using simple examples, and then present a systematic computational study on instances of two real applications. Section 5 concludes this paper.

2 A New Scheme for Level Reduction and Computation

In this paper, we make two rather standard assumptions: (a) functions $\mathbf{F}$ and $\mathbf{f}$ are continuous over their domains, and (b) $\mathbf{Y}(\mathbf{x})$ is compact for $\mathbf{x} \in \mathbf{X}$. With those two assumptions, it is clear that for a given $\mathbf{x}$, the lower level problem is either infeasible or has an optimal solution.

2.1 A Tight Relaxation for Level Reduction

As mentioned, pessimistic bilevel problem in (1-3) is a three-level problem. Apparently, it is more complicated than its bilevel optimistic counterpart **OBL**. Indeed, it is easy to see that **OBL** is a relaxation and its optimal value provides a lower bound to **PBL**. Because such relaxation is straightforward, we would say it is trivial. In the following, we first introduce a set of new variables and constraints that replicate those in the lower level, and then use them to directly capture the interaction between upper and lower level DMs, which leads to a non-trivial bilevel relaxation of **PBL**.

Lemma 1. *The following formulation **R-PBL** is a relaxation to **PBL**.*

$$\textbf{\textit{R-PBL}}: \qquad \tilde{\Theta}_p^* = \min_{\mathbf{x},\overline{\mathbf{y}}} \max_{\mathbf{y}\in\tilde{\mathbf{Y}}(\mathbf{x},\overline{\mathbf{y}})} \mathbf{F}(\mathbf{x},\mathbf{y}) \tag{4}$$

$$s.t.\ \mathbf{x} \in \mathbf{X} \tag{5}$$

$$\overline{\mathbf{y}} \in \mathbf{Y}(\mathbf{x}) \tag{6}$$

$$\tilde{\mathbf{Y}}(\mathbf{x},\overline{\mathbf{y}}) = \Big\{\mathbf{y} : \mathbf{y} \in \mathbf{Y}(\mathbf{x}), \mathbf{f}(\mathbf{x},\mathbf{y}) \le \mathbf{f}(\mathbf{x},\overline{\mathbf{y}})\Big\} \tag{7}$$

Proof. Consider a fixed $\mathbf{x}^* \in \mathbf{X}$ and $\mathbf{y}^* \in \arg\max\{\mathbf{F}(\mathbf{x}^*,\mathbf{y}) : \mathbf{y} \in \mathbf{S}(\mathbf{x}^*)\}$, i.e., solution $(\mathbf{x}^*,\mathbf{y}^*)$ is feasible to **PBL**. We set $\overline{\mathbf{y}}^* = \mathbf{y}^*$ and extend $(\mathbf{x}^*,\mathbf{y}^*)$ to a 3-tuple $(\mathbf{x}^*,\overline{\mathbf{y}}^*,\mathbf{y}^*)$. It is easy to see that the latter one satisfies constraints in (5-6) and $\mathbf{y}^*$ belongs to $\tilde{\mathbf{Y}}(\mathbf{x}^*,\overline{\mathbf{y}}^*)$. Moreover, because $\overline{\mathbf{y}}^* = \mathbf{y}^* \in \mathbf{S}(\mathbf{x}^*)$, we have

$$\mathbf{f}(\mathbf{x}^*,\overline{\mathbf{y}}^*) = \mathbf{f}(\mathbf{x}^*,\mathbf{y}^*) = \min\Big\{\mathbf{f}(\mathbf{x}^*,\mathbf{y}) : \mathbf{y} \in \mathbf{Y}(\mathbf{x}^*)\Big\}.$$

Then, comparing the definitions of $\tilde{\mathbf{Y}}(\mathbf{x}^*,\overline{\mathbf{y}}^*)$ and $\mathbf{S}(\mathbf{x}^*)$, it is clear that those two sets are identical, i.e., $\tilde{\mathbf{Y}}(\mathbf{x}^*,\overline{\mathbf{y}}^*) = \mathbf{S}(\mathbf{x}^*)$. Hence, $\mathbf{y}^* \in \arg\max\{\mathbf{F}(\mathbf{x}^*,\mathbf{y}) : \mathbf{y} \in \tilde{\mathbf{Y}}(\mathbf{x}^*,\overline{\mathbf{y}}^*)\}$. Therefore, the 3-tuple $(\mathbf{x}^*,\overline{\mathbf{y}}^*,\mathbf{y}^*)$ is feasible to **R-PBL**.

Finally, because they share the same objective function, it follows that **R-PBL** is a relaxation to **PBL**. □

Indeed, it can be proven that this relaxation is tight.

Lemma 2. *For a fixed* $\mathbf{x}^* \in \mathbf{X}$ *such that* $\mathbf{Y}(\mathbf{x}^*)$ *is not empty,* ***R-PBL*** *has an optimal solution* $(\mathbf{x}^*,\overline{\mathbf{y}}^*,\mathbf{y}^*)$ *such that*

$$\mathbf{f}(\mathbf{x}^*,\overline{\mathbf{y}}^*) = \min\Big\{\mathbf{f}(\mathbf{x}^*,\mathbf{y}) : \mathbf{y} \in \mathbf{Y}(\mathbf{x}^*)\Big\}, \tag{8}$$

and $\tilde{\mathbf{Y}}(\mathbf{x}^*, \overline{\mathbf{y}}^*) = \mathbf{S}(\mathbf{x}^*)$.

Proof. We first define the following function to help reveal the interaction between the upper and lower level DMs.

$$\phi(\mathbf{x}, \overline{\mathbf{y}}) = \max \Big\{ \mathbf{F}(\mathbf{x}, \mathbf{y}) : \mathbf{y} \in \mathbf{Y}(\mathbf{x}),\ \mathbf{f}(\mathbf{x}, \mathbf{y}) \leq \mathbf{f}(\mathbf{x}, \overline{\mathbf{y}}) \Big\}. \tag{9}$$

Clearly, it can be seen that the larger value of $f(\mathbf{x}, \overline{\mathbf{y}})$, the larger feasible set for the optimization problem in the right-hand-side of (9). Noting that it is a maximization problem, we can conclude that its optimal value, i.e., $\phi(\mathbf{x}, \overline{\mathbf{y}})$, is non-decreasing with respect to $\mathbf{f}(\mathbf{x}, \overline{\mathbf{y}})$, for a fixed $\mathbf{x}$.

On the other hand, for $\mathbf{x}^*$, because $\mathbf{f}$ is continuous, $\mathbf{Y}(\mathbf{x}^*)$ is compact, and by Weierstrass theorem, we have that $\min\{\mathbf{f}(\mathbf{x}^*, \overline{\mathbf{y}}) : \overline{\mathbf{y}} \in \mathbf{Y}(\mathbf{x}^*)\}$ has a minimizer. Let it be denoted by $\overline{\mathbf{y}}^*$. Then, by the fact that $\mathbf{F}$ is continuous, Weierstrass theorem again, and the aforementioned non-decreasing property, $\max\{\mathbf{F}(\mathbf{x}^*, \mathbf{y}) : \mathbf{y} \in \mathbf{Y}(\mathbf{x}^*), \mathbf{f}(\mathbf{x}^*, \mathbf{y}) \leq \mathbf{f}(\mathbf{x}^*, \overline{\mathbf{y}}^*)\}$ can be achieved, which is the minimum of $\phi(\mathbf{x}^*, \overline{\mathbf{y}})$ for $\overline{\mathbf{y}} \in \mathbf{Y}(\mathbf{x}^*)$. It also follows that $\tilde{\mathbf{Y}}(\mathbf{x}^*, \overline{\mathbf{y}}^*) = \mathbf{S}(\mathbf{x}^*)$. □

Consequently, the next result follows.

Proposition 3. (*i*) *One optimal solution to* ***R-PBL*** *solves* ***PBL****, and* (*ii*) *among* ***OBL****,* ***PBL*** *and* ***R-PBL*** *formulations, we have* $\Theta^*_o \leq \tilde{\Theta}^*_p = \Theta^*_p$.

This result suggests that we probably can compute the three-level **PBL** model by investigating a bilevel **R-PBL** formulation. It is worth pointing out that such level reduction, i.e., constructing the bilevel **R-PBL** for the three-level **PBL**, simply uses the primal representation of the lower level problem, which does not rely on the convexity, strong duality, or Karush-Kuhn-Tucker (KKT) conditions of that problem. Given **R-PBL**'s bilevel structure, almost all existing solution methods for mainstream bilevel optimization problems, including those for discrete or non-convex bilevel problems, can readily be applied with minor modifications. Hence, if a conversion procedure between a solution to **R-PBL** and that of the original **PBL** can be developed, we will be able to first compute **R-PBL** and then to derive a solution for **PBL**.

2.2 Computing PBL by R-PBL: A Complete Computational Scheme

Before we proceed to compute **PBL**, it would be worth examining that an optimal solution $(\mathbf{x}^\star, \mathbf{y}^\star)$ obtained for the trivial relaxation **OBL** is not feasible to **PBL**. Otherwise, we can simply report it as an optimal solution to **PBL**. The examination can be done easily by the *lexicographic method* often used in multi-objective optimization. Specifically in the examination, given that $\mathbf{f}(\mathbf{x}^\star, \mathbf{y}^\star)$ is the optimal lower level value with respect to $\mathbf{x}^\star$, we compute

$$\begin{aligned} \max \quad & \mathbf{F}(\mathbf{x}^\star, \mathbf{y}) \\ \text{s.t.} \quad & \mathbf{y} \in \mathbf{Y}(\mathbf{x}^\star), \quad \mathbf{f}(\mathbf{x}^\star, \mathbf{y}) \leq \mathbf{f}(\mathbf{x}^\star, \mathbf{y}^\star). \end{aligned}$$

Clearly, if its optimal value is equal to $\mathbf{F}(\mathbf{x}^\star, \mathbf{y}^\star)$, $(\mathbf{x}^\star, \mathbf{y}^\star)$ is feasible and optimal to **PBL** and we do not need to implement any additional operation. Hence, in the remaining part of this paper, we assume without loss of generality that the aforementioned examination step is done and it is necessary to compute **PBL** explicitly.

Although **R-PBL** is a tight relaxation, we point out that computing **R-PBL** may not directly generate a feasible solution to **PBL**. It is possible that the computing procedure for **R-PBL** produces a solution $(\mathbf{x}^*, \overline{\mathbf{y}}^*, \mathbf{y}^*)$ that does not satisfy (8), which indicates that $\mathbf{y}^*$ might not belong to $\mathbf{S}(\mathbf{x}^*)$. Actually, a trivial example can be found when $\mathbf{F}(\mathbf{x}, \mathbf{y}) = 0$. One optimal solution to its **R-PBL** formulation can be easily found with an $\mathbf{x}^*$ and $\overline{\mathbf{y}}^* = \mathbf{y}^* \in \arg\max\{\mathbf{f}(\mathbf{x}^*, \mathbf{y}) : \mathbf{y} \in \mathbf{Y}(\mathbf{x}^*)\}$. Clearly, whenever $\max\{\mathbf{f}(\mathbf{x}^*, \mathbf{y}) : \mathbf{y} \in \mathbf{Y}(\mathbf{x}^*)\} \neq \min\{\mathbf{f}(\mathbf{x}^*, \mathbf{y}) : \mathbf{y} \in \mathbf{Y}(\mathbf{x}^*)\}$, that solution is infeasible to its **PBL**. Next, we give a concrete bilevel model that is less straightforward.

$$\min \Big\{ x - 3y_1 : 0 \leq x \leq 10, (y_1, y_2) \in \mathbf{S}(x) = \arg\min\{y_2 : y_1 \leq x, y_1 \geq 0, y_2 \geq 0\} \Big\} \tag{10}$$

By computing its **R-PBL** formulation, we obtain an optimal solution $(x^*, \overline{y}_1^*, \overline{y}_2^*, y_1^*, y_2^*) = (0, 0, 2, 0, 2)$, which violates (8) and does not render (x^*, y_1^*, y_2^*) feasible to its **PBL**.

Nevertheless, such issue can be easily addressed by the following simple two-step operation, which we refer to as the *Relaxation-and-Correction* computational scheme for **PBL**.

---

(1) Relaxation: Compute the tight relaxation **R-PBL** and derive an optimal solution $(\mathbf{x}^*, \overline{\mathbf{y}}^*, \mathbf{y}^*)$. If

$$\mathbf{f}(\mathbf{x}^*, \overline{\mathbf{y}}^*) > \theta(\mathbf{x}^*) = \min \Big\{ \mathbf{f}(\mathbf{x}^*, \mathbf{y}) : \mathbf{y} \in \mathbf{Y}(\mathbf{x}^*) \Big\}, \tag{11}$$

perform the correction step. Otherwise, report $(\mathbf{x}^*, \mathbf{y}^*)$ as an optimal solution to **PBL**.

(2) Correction: Compute

$$\begin{aligned} \max \quad & \mathbf{F}(\mathbf{x}^*, \mathbf{y}) \\ \text{s.t.} \quad & \mathbf{y} \in \mathbf{Y}(\mathbf{x}^*),\ \mathbf{f}(\mathbf{x}^*, \mathbf{y}) \leq \theta(\mathbf{x}^*) \end{aligned}$$

and derive an optimal solution $\mathbf{y}'$. Report $(\mathbf{x}^*, \mathbf{y}')$ as an optimal solution to **PBL**. □

According to Lemma 2 and Proposition 3, the relaxation step produces the optimal value and an optimal upper level decision. Note that equation (11) is to verify $\tilde{\mathbf{Y}}(\mathbf{x}^*, \overline{\mathbf{y}}^*) = \mathbf{S}(\mathbf{x}^*)$. If not, the lexicographic method is implemented in the correction step, which guarantees an optimal pessimistic response from the lower level DM. Hence, the correctness of the *Relaxation-and-Correction* scheme is clear. Indeed, we can reduce the occurrence of correction step through replacing $\mathbf{f}(\mathbf{x}^*, \overline{\mathbf{y}}^*)$ by $\mathbf{f}(\mathbf{x}^*, \mathbf{y}^*)$ in (11). Regarding the example listed in (10), we note that the correction step is needed to derive a complete solution of its **PBL**, which yields $(x^*, y_1', y_2') = (0, 0, 0)$.

Within the *Relaxation-and-Correction* scheme, it can be seen that the essential computation of the original three-level formulation **PBL** reduces to that of a bilevel problem **R-PBL**. Compared to the optimistic counterpart **OBL**, in particular, the size of **R-PBL** does not increase much. Its lower level problem just has one more constraint than that of its optimistic counterpart, and its upper level problem just has one replica of the lower level variables and constraints. Hence, we believe that solving **R-PBL** might not demand drastically more computational expense.

Another critical observation, as shown in the following, is that the lower level problem in **R-PBL** is often a well-structured optimization problem.

Corollary 4. *If* $\mathbf{F}$ *is concave,* $\mathbf{f}$ *is convex with respect to* $\mathbf{y}$*, and* $\mathbf{Y}(\mathbf{x})$ *is a convex set for* $\mathbf{x} \in \mathbf{X}$*, the lower level problem in* ***R-PBL****, i.e.,*

$$\max \Big\{ \mathbf{F}(\mathbf{x}, \mathbf{y}) : \mathbf{y} \in \mathbf{Y}(\mathbf{x}),\ \mathbf{f}(\mathbf{x}, \mathbf{y}) \leq \mathbf{f}(\mathbf{x}, \overline{\mathbf{y}}), \Big\},$$

is a convex optimization problem with respect to $\mathbf{y}$.

One great advantage of this property is that if $\mathbf{F}$ and $\mathbf{f}$ are continuously differentiable and certain constraint qualification with $\mathbf{Y}(\mathbf{x})$ is satisfied, we can adopt KKT reformulation method, i.e., replacing this lower level problem by its KKT conditions. Hence, one more reduction of level can be achieved, which ultimately leads to a single level equivalence of **R-PBL**. As such single level optimization problem can readily be solved by a typical optimization package or a specialized solver for mathematical programs with equilibrium constraints, this computational scheme provides a remarkable convenience for practitioners. It is very different from a traditional understanding that those methods might not be able to solve **PBL**.

For **PBL** problems whose optimal solution may not exist, it has been proposed to compute their ϵ-approximations [32, 40] where $\mathbf{y} \in \mathbf{S}(\mathbf{x})$ in (3) is relaxed to

$$\mathbf{y} \in \mathbf{S}_{\epsilon}(\mathbf{x}) = \Big\{ \mathbf{y} \in \mathbf{Y}(\mathbf{x}) : \mathbf{f}(\mathbf{x}, \mathbf{y}) \leq \theta^{*}(\mathbf{x}) + \epsilon \Big\}. \tag{12}$$

One benefit is that compared to the original **PBL** model, the ϵ-approximation could have an optimal solution, although not guaranteed. Another significant advantage is that the ϵ-approximation provides a modeling flexibility. It can be used to capture the upper level DM's safety margin consideration to hedge against deviations from the expected lower level DM's decision [40]. Also, we note it can describe the lower level DM's tolerance level when he makes a non-cooperative decision against the upper level DM. So, an interesting study is to evaluate the impact of ϵ on the upper level DM's decisions and her objective function value. Even for the case where $\mathbf{S}(\mathbf{x})$ is always a singleton for every $\mathbf{x}$, i.e., the pessimistic formulation coincides with the optimistic one, we can make use of this modified **PBL** model to perform the sensitivity analysis on the upper level decisions for different ϵ. As shown in a practical study on bilevel gene knockout problem [5], it can be used to derive or evaluate an upper level gene knockout solution with respect to ϵ, where ϵ is interpreted as the modeling error of the lower level problem.

When an optimal solution exists, the *Relaxation-and-Correction* can easily be employed to compute the ϵ-approximation with some minor modifications: (*i*) in (7), replace $\mathbf{f}(\mathbf{x}, \mathbf{y}) \leq \mathbf{f}(\mathbf{x}, \overline{\mathbf{y}})$ by $\mathbf{f}(\mathbf{x}, \mathbf{y}) \leq \mathbf{f}(\mathbf{x}, \overline{\mathbf{y}}) + \epsilon$; (*ii*) in (11), replace $\mathbf{f}(\mathbf{x}^{*}, \overline{\mathbf{y}}^{*}) > \theta(\mathbf{x}^{*})$ by $\mathbf{f}(\mathbf{x}^{*}, \overline{\mathbf{y}}^{*}) > \theta(\mathbf{x}^{*}) + \epsilon$; and (*iii*) in the optimization problem of the correction step, replace $\mathbf{f}(\mathbf{x}^{*}, \mathbf{y}) \leq \theta(\mathbf{x}^{*})$ by $\mathbf{f}(\mathbf{x}^{*}, \mathbf{y}) \leq \theta(\mathbf{x}^{*}) + \epsilon$.

In the following, we apply our new computational scheme on the basic linear pessimistic problem and its extensions. Because the correction step is rather standard and easy to implement, we mainly consider the relaxation step, i.e., their associated **R-PBL** formulations and computation.

3 Linear Pessimistic Bilevel Optimization Problems

In the remainder of this paper, we will focus on linear pessimistic problems where all constraints and objective functions are linear in $\mathbf{x}$ and $\mathbf{y}$. We first consider the basic linear **PBL** where the lower level problem is an LP and present its single level reformulation. We then extend our study to: linear mixed integer **PBL** where the lower level problem includes discrete variables, bilevel problem with coupled pessimistic constraints, and strong-weak bilevel problem.

3.1 Basic Linear PBL Problem

Consider the following linear pessimistic bilevel problem

$$
\begin{aligned}
\mathbf{Linear-PBL}: \ \min \quad & \mathbf{cx} + \max \ \mathbf{d}_1\mathbf{y} && (13)\\
\text{s.t.} \quad & \mathbf{A}_1\mathbf{x} \le \mathbf{b}_1, \ \mathbf{x} \in \mathbb{R}^{n_c}_+ \times \mathbb{Z}^{n_d}_+ && (14)\\
& \mathbf{y} \in \mathbf{S}(\mathbf{x}) = \arg\min \Big\{ \mathbf{d}_2\mathbf{y} : \mathbf{B}_2\mathbf{y} \le \mathbf{b}_2 - \mathbf{A}_2\mathbf{x}, \ \mathbf{y} \in \mathbb{R}^m_+ \Big\} && (15)
\end{aligned}
$$

where $n = n_c + n_d$, $\mathbf{A}_1, \mathbf{b}_1, \mathbf{A}_2, \mathbf{B}_2, \mathbf{b}_2, \mathbf{c}, \mathbf{d}_1$ and $\mathbf{d}_2$ are of appropriate dimensions. Let $\mathbf{X} = \{\mathbf{x} \in \mathbb{R}^{n_c}_+ \times \mathbb{Z}^{n_d}_+ : \mathbf{A}_1\mathbf{x} \le \mathbf{b}_1\}$ and

$$
\Omega = \Big\{ (\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{n_c}_+ \times \mathbb{Z}^{n_d}_+ \times \mathbb{R}^m_+ : \mathbf{A}_1\mathbf{x} \le \mathbf{b}_1, \ \mathbf{A}_2\mathbf{x} + \mathbf{B}_2\mathbf{y} \le \mathbf{b}_2 \Big\}.
$$

Unless explicitly stated, it is assumed in this paper that Ω is a non-empty compact set. Hence, for a given $\mathbf{x}$, set $\mathbf{Y}(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^m_+ : \mathbf{B}_2\mathbf{y} \le \mathbf{b}_2 - \mathbf{A}_2\mathbf{x}\}$ is closed, if is not empty.

Remark: It is shown in [4] that if $\mathbf{x}$ are continuous, i.e., $n_d = 0$, the aforementioned assumption ensures the existence of an optimal solution. Indeed, through Branch-and-Bound method, this result can easily extend to a more general case where $\mathbf{x}$ are mixed integer variables, i.e., $n_d \ge 1$.

Next, because of the convexity of linear functions and Corollary 4, we can easily develop the following result using KKT reformulation. Note that superscript t is used to indicate matrix transpose operation, and $(\mathbf{u}, \pi)$ are dual variables introduced for KKT reformulation.

Corollary 5. *For* **Linear−PBL** *problem, its tight **R-PBL** relaxation is*

$$
\begin{aligned}
\min \quad & \mathbf{cx} + \mathbf{d}_1\mathbf{y}\\
s.t. \quad & \mathbf{A}_1\mathbf{x} \le \mathbf{b}_1, \ \mathbf{A}_2\mathbf{x} + \mathbf{B}_2\overline{\mathbf{y}} \le \mathbf{b}_2\\
& \mathbf{x} \in \mathbb{R}^{n_c}_+ \times \mathbb{Z}^{n_d}_+, \overline{\mathbf{y}} \in \mathbb{R}^m_+\\
& \mathbf{y} \in \arg\max \Big\{ \mathbf{d}_1\mathbf{y} : \mathbf{B}_2\mathbf{y} \le \mathbf{b}_2 - \mathbf{A}_2\mathbf{x}, \ \mathbf{d}_2\mathbf{y} \le \mathbf{d}_2\overline{\mathbf{y}}, \ \mathbf{y} \in \mathbb{R}^m_+ \Big\}.
\end{aligned}
$$

Then, this tight relaxation is equivalent to the following single level formulation

$$\begin{aligned}
\min \quad & \mathbf{cx} + \mathbf{d}_1\mathbf{y} \\
s.t. \quad & \mathbf{A}_1\mathbf{x} \le \mathbf{b}_1,\ \mathbf{A}_2\mathbf{x} + \mathbf{B}_2\overline{\mathbf{y}} \le \mathbf{b}_2 \\
& 0 \le \mathbf{y} \perp (\mathbf{B}_2^t\mathbf{u} + \mathbf{d}_2^t\pi - \mathbf{d}_1^t) \ge 0, \quad 0 \le \mathbf{u} \perp (\mathbf{b}_2 - \mathbf{A}_2\mathbf{x} - \mathbf{B}_2\mathbf{y}) \ge 0 \\
& 0 \le \pi \perp (\mathbf{d}_2\overline{\mathbf{y}} - \mathbf{d}_2\mathbf{y}) \ge 0 \\
& \mathbf{x} \in \mathbb{R}_+^{n_c} \times \mathbb{Z}_+^{n_d}, \overline{\mathbf{y}} \in \mathbb{R}_+^m.
\end{aligned}$$

The last formulation is a mathematical program with linear complementarity constraints (MPCC). As previously mentioned, it can readily be solved by an optimization package specialized in this type of problems. Or, those linear complementarity constraints can be linearized using additional binary variables [6], which convert the whole formulation into a mixed integer program (MIP) that can be easily handled by widely adopted professional MIP solvers. Similar to using KKT reformulation strategy, another approach is by taking advantage of its minimax structure and the strong duality of the lower level problem, **R-PBL** can be converted into a single level optimization problem, whose objective function is bilinear and constraints are linear.

Sometimes, a bilevel problem involves multiple independent lower level DMs that have their own objective functions and constraints. It makes no difference to treat them as a single DM in the optimistic model as the resulting KKT reformulation still has a block structure where one block corresponds to one DM [15]. Nevertheless, when studying pessimistic bilevel problem, aggregating them into a single DM to build a single lower level problem may hide the nature of multiple lower level DMs. Indeed, we would rather treat them individually in **R-PBL** formulation as follows

$$\mathbf{A}_2^i\mathbf{x} + \mathbf{B}_2^i\overline{\mathbf{y}}^i \le \mathbf{b}_2^i,\ \text{and},\ \mathbf{y}^i \in \arg\max\left\{\mathbf{d}_1^i\mathbf{y}^i : \mathbf{B}_2^i\mathbf{y}^i \le \mathbf{b}_2^i - \mathbf{A}_2^i\mathbf{x},\ \mathbf{d}_2^i\mathbf{y}^i \le \mathbf{d}_2^i\overline{\mathbf{y}}^i,\ \mathbf{y}^i \in \mathbb{R}_+^m\right\},$$

where superscript i is used to indicate the i^{th} lower level DM's parameters and variables. It more accurately reflects decisions made by individual lower level DMs. Note that KKT reformulation can be applied to individual lower level problems, which preserves the block structure that is friendly to decomposition algorithms.

3.2 Mixed Integer PBL Problem

Consider the general mixed integer **PBL** where the lower level problem has discrete variables. Let $\mathbf{y} = (\mathbf{y}_c, \mathbf{y}_d)$ where $\mathbf{y}_c$ and $\mathbf{y}_d$ are continuous and discrete variables of $\mathbf{y}$ with m_c and m_d being their dimensions respectively. Then, we represent the corresponding expressions in (13) and (15) as

$$\min \mathbf{cx} + \max\ (\mathbf{d}_{1c}\mathbf{y}_c + \mathbf{d}_{1d}\mathbf{y}_d),$$

and

$$\mathbf{S}(\mathbf{x}) = \arg\min\left\{\mathbf{d}_{2c}\mathbf{y}_c + \mathbf{d}_{2d}\mathbf{y}_d : \mathbf{B}_{2c}\mathbf{y}_c + \mathbf{B}_{2d}\mathbf{y}_d \le \mathbf{b}_2 - \mathbf{A}_2\mathbf{x},\ \mathbf{y}_c \in \mathbb{R}_+^{m_c}, \mathbf{y}_d \in \mathbb{Z}_+^{m_d}\right\}. \quad (16)$$

We also denote the whole formulation as $\mathbf{MIP-PBL}$.

As mentioned in Section 2.1 that the construction of **R-PBL** relaxation does not rely on the convexity of the lower level problem, we next provide that tight relaxation of $\mathbf{MIP-PBL}$.

Corollary 6. *For* $\mathbf{MIP-PBL}$ *problem, its tight* ***R-PBL*** *relaxation is*

$$\min \quad \mathbf{c}\mathbf{x} + \mathbf{d}_{1c}\mathbf{y}_c + \mathbf{d}_{1d}\mathbf{y}_d \tag{17}$$

$$s.t. \quad \mathbf{A}_1\mathbf{x} \leq \mathbf{b}_1,\ \mathbf{A}_2\mathbf{x} + \mathbf{B}_{2c}\overline{\mathbf{y}}_c + \mathbf{B}_{2d}\overline{\mathbf{y}}_d \leq \mathbf{b}_2 \tag{18}$$

$$\mathbf{x} \in \mathbb{R}_+^{n_c} \times \mathbb{Z}_+^{n_d}, \overline{\mathbf{y}}_c \in \mathbb{R}_+^{m_c}, \overline{\mathbf{y}}_d \in \mathbb{Z}_+^{m_d} \tag{19}$$

$$(\mathbf{y}_c, \mathbf{y}_d) \in \arg\max \Big\{ \mathbf{d}_{1c}\mathbf{y}_c + \mathbf{d}_{1d}\mathbf{y}_d : \tag{20}$$

$$\mathbf{B}_{2c}\mathbf{y}_c + \mathbf{B}_{2d}\mathbf{y}_d \leq \mathbf{b}_2 - \mathbf{A}_2\mathbf{x}, \quad \mathbf{d}_{2c}\mathbf{y}_c + \mathbf{d}_{2d}\mathbf{y}_d \leq \mathbf{d}_{2c}\overline{\mathbf{y}}_c + \mathbf{d}_{2d}\overline{\mathbf{y}}_d \tag{21}$$

$$\mathbf{y}_c \in \mathbb{R}_+^{m_c}, \mathbf{y}_d \in \mathbb{Z}_+^{m_d} \Big\}. \tag{22}$$

For the aforementioned bilevel **R-PBL** formulation, because of the existence of discrete variables in the lower level problem, it is not feasible to apply the popular KKT reformulation method to derive a single level equivalent model. Indeed, even for the less complicated optimistic bilevel MIP, it may not have any optimal solution [28]. Nevertheless, by making a few modifications on this formulation, the resulting structural properties allow us to employ the recent *reformulation and decomposition* method to compute an exact or strong approximate solution [42]. In the following, we describe **R-PBL** and details of those modifications on **R-PBL**. For minimizing the distraction and also self-contained purpose, the tailored reformulation and decomposition procedure is presented in Appendix 1.

By taking advantage of the minimax structure in the objective function (17) with respect to (20), we first equivalently modify it as $\min \mathbf{c}\mathbf{x} + \eta$, and convert the set membership restriction in (20) into the following inequality

$$\eta \geq \max \Big\{ \mathbf{d}_{1c}\mathbf{y}_c + \mathbf{d}_{1d}\mathbf{y}_d : (21-22) \Big\}.$$

As a result, the whole **R-PBL** in Corollary 6 can be simplified as

$$\min \Big\{ \mathbf{c}\mathbf{x} + \eta : (18-19),\ \eta \geq \max\{\mathbf{d}_{1c}\mathbf{y}_c + \mathbf{d}_{1d}\mathbf{y}_d : (21-22)\} \Big\}. \tag{23}$$

We note that the aforementioned simplification reduces variables and constraints, which could provide a nontrivial computational advantage over the standard one in Corollary 6.

A sufficient condition ensuring the existence of an optimal solution to bilevel MIP problem is that the lower level problem has the *relatively complete response* property [42]. In the context of bilevel MIP in (23), this property is that for any possible $(\mathbf{x}, \overline{\mathbf{y}}_\mathbf{c}, \overline{\mathbf{y}}_\mathbf{d}, \mathbf{y}_\mathbf{d})$, the remaining lower level problem defined in (20-22), which is its LP portion, has a finite optimal value. However, in our case such property does not hold in general. Specifically, according to Lemma 2, for a fixed $\mathbf{x}^*$, **R-PBL** has an optimal solution with $(\overline{\mathbf{y}}_c^*, \overline{\mathbf{y}}_d^*)$ being an optimal solution to the original lower level problem in the $\mathbf{MIP-PBL}$ model. So, for such $(\mathbf{x}^*, \overline{\mathbf{y}}_c^*, \overline{\mathbf{y}}_d^*)$ combination, it is the most case that the lower level problem of (23), i.e.,

$$\max\Big\{\mathbf{d}_{1c}\mathbf{y}_c + \mathbf{d}_{1d}\mathbf{y}_d : (21-22)\Big\},$$

does not have the aforementioned property. The reason is that due to the second constraint in (21), the LP portion of the lower level problem could become infeasible when $\mathbf{y}_d \neq \overline{\mathbf{y}}_d^*$.

To address this issue, we need to introduce *artificial variables*, which are to be penalized with big-M coefficients, that allow constraint violations in (21) to avoid the infeasible situation. Explicitly, we modify the lower level problem in (23) as the following, where $\mathbf{1}$ is the l-dimensional all-ones row vector and $\mathbf{I}$ is the $l \times l$ identity matrix.

$$\eta \geq \max\Big\{\mathbf{d}_{1c}\mathbf{y}_c + \mathbf{d}_{1d}\mathbf{y}_d - M\mathbf{1}\mathbf{y}_a - My_f : \tag{24}$$

$$\mathbf{B}_{2c}\mathbf{y}_c + \mathbf{B}_{2d}\mathbf{y}_d - \mathbf{I}\mathbf{y}_a \leq \mathbf{b}_2 - \mathbf{A}_2\mathbf{x}, \quad \mathbf{d}_{2c}\mathbf{y}_c + \mathbf{d}_{2d}\mathbf{y}_d - y_f \leq \mathbf{d}_{2c}\overline{\mathbf{y}}_c + \mathbf{d}_{2d}\overline{\mathbf{y}}_d \tag{25}$$

$$\mathbf{y}_a \in \mathbb{R}^l_+,\ \mathbf{y}_c \in \mathbb{R}^{m_c}_+,\ \mathbf{y}_d \in \mathbb{Z}^{m_d}_+, y_f \in \mathbb{R}_+\Big\}. \tag{26}$$

Clearly, the new lower level problem in (24-26) has the relatively complete response property. As a result, the *extended reformulation* of the bilevel mixed integer **R-PBL** model in Corollary 6 is

$$\min\Big\{\mathbf{c}\mathbf{x} + \eta : (18-19),\ (24-26)\Big\}, \tag{27}$$

which can then be solved by the reformulation and decomposition method in [42], a procedure that is easily implementable with an optimization package for MPCC or an MIP solver.

Remark: It has been shown in [42] that if the bilevel MIP model has an optimal solution, it can be found by setting M to a large value in (27). Otherwise, an ϵ-optimal solution can be derived by adjusting the value of M. One example of the latter case can be found in Section 4.1.

3.3 Bilevel Problem with Coupled Pessimistic Constraints

One type bilevel problem with coupled constraints studied in [38, 40], under the linear settings, is formulated as

$$\mathbf{CP}: \quad \min \mathbf{c}\mathbf{x} \tag{28}$$

$$\text{s.t.} \quad \mathbf{A}_1\mathbf{x} \leq \mathbf{b}_1,\ \mathbf{x} \in \mathbb{R}^{n_c}_+ \times \mathbb{Z}^{n_d}_+ \tag{29}$$

$$\mathbf{f}^i\mathbf{x} + \mathbf{g}^i\mathbf{y} \leq t^i,\ \forall \mathbf{y} \in \arg\min_{\mathbf{y}}\Big\{\mathbf{d}^i\mathbf{y} : \mathbf{Q}^i\mathbf{y} \leq \mathbf{h}^i - \mathbf{P}^i\mathbf{x},\ \mathbf{y} \geq 0\Big\},\ i = 1, \ldots, p \tag{30}$$

where $\mathbf{f}^i$ and $\mathbf{g}^i$ are row vectors (like $\mathbf{c}$ in the objective function (28)) for all i. The optimization problem defining $\mathbf{y}$ in (30) is called the i^{th} lower level problem. Different from the conventional concept in optimistic models that the coupled constraints over $\mathbf{x}$ and $\mathbf{y}$, which appear in the upper level problem, should just be satisfied by *some* selected optimal $\mathbf{y}$, constraints in (30) are rather pessimistic as they must be satisfied by *all* optimal $\mathbf{y}$ of the associated lower level problems. In this regard, **CP**'s solution is just in $\mathbf{x}$ space, not in $(\mathbf{x}, \mathbf{y})$ space.

As mentioned, it is argued in [38, 40] that this type of formulation generalizes the standard

PBL in (1-3). Also, for a special independent case where $\mathbf{P}^i$s are missing in those constraints, an iterative solution procedure using semi-infinite programming tools is developed [38, 40]. Next, by using the technique developed in Section 2, we show that the general dependent **CP** model can be reformulated into a format that is directly computable by many off-the-shelf packages or solvers.

Assume that $\mathbf{Y}^i(\mathbf{x}) = \{\mathbf{y} \in \mathbb{R}^{m_i}_+ : \mathbf{Q}^i\mathbf{y} \leq \mathbf{h}^i - \mathbf{P}^i\mathbf{x}\}$ is a non-empty compact set for $\mathbf{x} \in \mathbf{X}$ in this subsection. Instead of using "$\forall$" operation to describe the i^{th} coupled constraint in (30), we can represent it through pessimistic bilevel optimization as the following:

$$\mathbf{f}^i\mathbf{x} + \max_{\mathbf{y}\in\mathbf{S}^i(\mathbf{x})} \mathbf{g}^i\mathbf{y} \leq t^i, \tag{31}$$

where $\mathbf{S}^i(\mathbf{x}) = \arg\min\{\mathbf{d}^i\mathbf{y} : \mathbf{Q}^i\mathbf{y} \leq \mathbf{h}^i - \mathbf{P}^i\mathbf{x},\ \mathbf{y} \geq 0\}$. As a result, using arguments similar to those presented in Section 2.1, we have the next result.

Lemma 7. *For a fixed $\mathbf{x}^* \in \mathbf{X}$, it satisfies the i^{th} constraint in* (30) *if and only if there exist $\overline{\mathbf{y}}^{i*} \in \mathbb{R}^{m_i}_+$ and $\mathbf{y}^{i*} \in \arg\max\{\mathbf{g}^i\mathbf{y} : \mathbf{Q}^i\mathbf{y} \leq \mathbf{h}^i - \mathbf{P}^i\mathbf{x}^*,\ \mathbf{d}^i\mathbf{y} \leq \mathbf{d}^i\overline{\mathbf{y}}^{i*}, \mathbf{y} \in \mathbb{R}^{m_i}_+\}$, and they satisfy the following constraints*

$$\mathbf{Q}^i\overline{\mathbf{y}}^{i*} \leq \mathbf{h}^i - \mathbf{P}^i\mathbf{x}^*, \tag{32}$$

$$\mathbf{g}^i\mathbf{y}^{i*} \leq t^i - \mathbf{f}^i\mathbf{x}^*. \tag{33}$$

Proof. First, we prove the "$\Leftarrow$" direction. Consider a particular $\overline{\mathbf{y}}^{i*}$ such that

$$\overline{\mathbf{y}}^{i*} \in \Big\{\overline{\mathbf{y}} \in \mathbb{R}^{m_i}_+ : \mathbf{Q}^i\overline{\mathbf{y}}^i \leq \mathbf{h}^i - \mathbf{P}^i\mathbf{x}^*\Big\}.$$

Clearly, we have

$$\mathbf{S}^i(\mathbf{x}^*) = \arg\min\Big\{\mathbf{d}^i\mathbf{y} : \mathbf{Q}^i\mathbf{y} \leq \mathbf{h}^i - \mathbf{P}^i\mathbf{x}^*,\ \mathbf{y} \geq 0\Big\} \subseteq \Big\{\mathbf{y} : \mathbf{Q}^i\mathbf{y} \leq \mathbf{h}^i - \mathbf{P}^i\mathbf{x}^*,\ \mathbf{d}^i\mathbf{y} \leq \mathbf{d}^i\overline{\mathbf{y}}^{i*}, \mathbf{y} \geq 0\Big\}.$$

Then, it follows that

$$\max_{\mathbf{y}\in\mathbf{S}^i(\mathbf{x}^*)} \mathbf{g}^i\mathbf{y} \leq \max_{\mathbf{y}\in\{\mathbf{y}:\mathbf{Q}^i\mathbf{y}\leq\mathbf{h}^i-\mathbf{P}^i\mathbf{x}^*,\ \mathbf{d}^i\mathbf{y}\leq\mathbf{d}^i\overline{\mathbf{y}}^{i*},\ \mathbf{y}\geq 0\}} \mathbf{g}^i\mathbf{y}.$$

Let $\mathbf{y}^{i*}$ be an optimal solution to $\max\{\mathbf{g}^i\mathbf{y} : \mathbf{Q}^i\mathbf{y} \leq \mathbf{h}^i - \mathbf{P}^i\mathbf{x}^*,\ \mathbf{d}^i\mathbf{y} \leq \mathbf{d}^i\overline{\mathbf{y}}^{i*},\ \mathbf{y} \in \mathbb{R}^{m_i}_+\}$. So, if (33) is valid, we have

$$\mathbf{f}^i\mathbf{x}^* + \max_{\mathbf{y}\in\mathbf{S}^i(\mathbf{x}^*)} \mathbf{g}^i\mathbf{y} \leq \mathbf{f}^i\mathbf{x}^* + \mathbf{g}^i\mathbf{y}^{i*} \leq t^i,$$

which shows $\mathbf{x}^*$ satisfies the i^{th} constraint in (30).

Next, we prove the $\Rightarrow$ direction. Note that if (31) is satisfied by $\mathbf{x}^*$, the next inequality follows naturally.

$$\min\Big\{\mathbf{f}^i\mathbf{x}^* + \max_{\mathbf{y}\in\mathbf{S}^i(\mathbf{x}^*)} \mathbf{g}^i\mathbf{y}\Big\} \leq t^i. \tag{34}$$

Using the strategy in Section 2.1, we construct the following relaxation to the left-hand-side of (34).

$$\mathbf{f}^i\mathbf{x}^* + \min_{\overline{\mathbf{y}}} \max_{\mathbf{y}} \mathbf{g}^i\mathbf{y} \tag{35}$$

$$\text{s.t.} \quad \mathbf{Q}^i\overline{\mathbf{y}} \le \mathbf{h}^i - \mathbf{P}^i\mathbf{x}^*, \ \overline{\mathbf{y}} \ge 0 \tag{36}$$

$$\mathbf{y} \in \Big\{\mathbf{Q}^i\mathbf{y} \le \mathbf{h}^i - \mathbf{P}^i\mathbf{x}^*, \ \mathbf{d}^i\mathbf{y} \le \mathbf{d}^i\overline{\mathbf{y}}^i, \mathbf{y} \ge 0\Big\} \tag{37}$$

According to Lemma 2, the problem in (35-37) achieves its minimum when $\overline{\mathbf{y}}$ is an optimal solution of $\min\{\mathbf{d}^i\overline{\mathbf{y}} : \mathbf{Q}^i\overline{\mathbf{y}} \le \mathbf{h}^i - \mathbf{P}^i\mathbf{x}^*, \ \overline{\mathbf{y}} \ge 0\}$. Let $\overline{\mathbf{y}}^*$ be a such optimal solution. Then, according to Lemma 2 and (34), the next inequality is satisfied.

$$\mathbf{f}^i\mathbf{x}^* + \max\{\mathbf{g}^i\mathbf{y} : \mathbf{Q}^i\mathbf{y} \le \mathbf{h}^i - \mathbf{P}^i\mathbf{x}^*, \ \mathbf{d}^i\mathbf{y} \le \mathbf{d}^i\overline{\mathbf{y}}^{i*}, \mathbf{y} \ge 0\} \le t^i \tag{38}$$

Denote an optimal solution to the maximization problem in (38) by $\mathbf{y}^{i*}$. We have

$$\mathbf{f}^i\mathbf{x}^* + \mathbf{g}^i\mathbf{y}^{i*} \le t^i,$$

which, together with the selection criteria of $\overline{\mathbf{y}}^{i*}$, gives the desired results. □

We mention that superscript i is used in $\overline{\mathbf{y}}^i$ and $\mathbf{y}^i$ to indicate that those variables and associated constraints are introduced specifically for the i^{th} coupled constraint. So, if multiple coupled constraints exist, this equivalence result should be applied constraint-wise. Based on Lemma 7 and its proof, and using KKT reformulation, we can convert the **CP** model into a computationally friendly formulation.

Corollary 8. *The* **CP** *model defined in (28-30) is equivalent to the following single-level formulation.*

$$\begin{aligned}
&\min \mathbf{c}\mathbf{x} \\
&\ s.t.\ \mathbf{A}_1\mathbf{x} \le \mathbf{b}_1 \\
&\quad \mathbf{P}^i\mathbf{x} + \mathbf{Q}^i\overline{\mathbf{y}}^i \le \mathbf{h}^i, \ \mathbf{f}^i\mathbf{x} + \mathbf{g}^i\mathbf{y}^i \le t^i, \ i = 1, \dots, p \\
&\quad 0 \le \mathbf{y}^i \perp \Big((\mathbf{Q}^i)^t\mathbf{u}^i + (\mathbf{d}^i)^t\pi^i - (\mathbf{g}^i)^t\Big) \ge 0, \ i = 1, \dots, p \\
&\quad 0 \le \mathbf{u}^i \perp \big(\mathbf{h}^i - \mathbf{P}^i\mathbf{x} - \mathbf{Q}^i\mathbf{y}^i\big) \ge 0, \quad 0 \le \pi^i \perp \big(\mathbf{d}^i\overline{\mathbf{y}}^i - \mathbf{d}^i\mathbf{y}^i\big) \ge 0, \ i = 1, \dots, p \\
&\quad \mathbf{x} \in \mathbb{R}^{n_c}_+ \times \mathbb{Z}^{n_d}_+, \ \overline{\mathbf{y}}^i \in \mathbb{R}^{m_i}_+, \ i = 1, \dots, p.
\end{aligned}$$

Remark: (*i*) The reason we say that the aforementioned formulation is equivalent to **CP** is that the $\mathbf{x}$ portion of its optimal solution and its optimal value are optimal to **CP**. (*ii*) Note that we can deal with the nonlinear complementarity constraints in the aforementioned formulation by Branch-and-Bound or by enumerating the implied linear constraints, both of which are finite. Hence, it can be inferred that this formulation, as well as **CP**, either has an optimal solution, or is infeasible. Again, it can be directly computed by an MPCC optimization package or by any professional MIP solver, after linearization. (*iii*) If the non-empty assumption on $\mathbf{Y}^i(\mathbf{x})$ for $\mathbf{x} \in \mathbf{X}$ is not satisfied, similar to the modifications made in

(24-26), we can include artificial variables with big-M coefficients so that (33) is of a very negative left-hand-side value (through the help of artificial variables) whenever $\mathbf{x}$ causes set $\mathbf{Y}^i(\mathbf{x})$ empty.

3.4 Strong-Weak Bilevel Problem

Noting that the solution to the lower level problem might not be unique, the strong-weak bilevel optimization problem, also known as partially cooperative bilevel model, has been proposed and investigated in [3, 16, 44, 2]. It integrates the optimistic and pessimistic formulations through a weighted summation, where the weight coefficient can be interpreted as the cooperative probability of the lower level DM. Hence, knowing that the lower level DM just cooperates partially or stochastically, the upper level DM, instead of being completely optimistic or pessimistic, solves the strong-weak bilevel problem to derive an intermediate solution. Actually, it is interesting to observe that there exist situations where the lower level DM achieves his best interest by being partially or stochastically cooperative [16]. We believe that the strong-weak model has its significance in modeling and analyzing real situations.

Let β represent the weight parameter. The formulation of the strong-weak bilevel optimization problem is

$$\mathbf{SW-PBL}: \min \quad \mathbf{cx} + \beta f_1 + (1-\beta) f_2 \tag{39}$$

$$\text{s.t.} \quad \mathbf{A}_1\mathbf{x} \leq \mathbf{b}_1,\ \mathbf{x} \in \mathbb{R}_+^{n_c} \times \mathbb{Z}_+^{n_d} \tag{40}$$

$$f_1 = \min\Big\{\mathbf{d}_1\mathbf{y} : \mathbf{y} \in \mathbf{S}(\mathbf{x})\Big\},\ f_2 = \max\Big\{\mathbf{d}_1\mathbf{y} : \mathbf{y} \in \mathbf{S}(\mathbf{x})\Big\} \tag{41}$$

where $\mathbf{S}(\mathbf{x})$ is defined as in (15).
Remark: It is shown in [16, 44] that $\mathbf{SW-PBL}$ has an optimal solution if $\mathbf{x}$ are continuous, i.e., $n_d = 0$. Again, through Branch-and-Bound argument, this result extends to a more general case where $\mathbf{x}$ are mixed integer variables. Also, we would rather say its solution is simply in $\mathbf{x}$ space as the lower level DM's decision is not certain in general.

To solve $\mathbf{SW-PBL}$ problem, a couple of penalty methods have been developed in [16, 44], which iteratively adjust penalty coefficients and compute the resulting relaxations to ensure the convergence to an optimal solution. Next, we provide its tight **R-PBL** relaxation and the computationally friendly single level reformulation. Variables $\mathbf{y}^o$ and $\mathbf{y}^p$ are introduced to represent those in optimistic and pessimistic parts, respectively.

Corollary 9. *For* $\mathbf{SW-PBL}$ *problem, its* ***R-PBL*** *model is*

$$\min \quad \mathbf{cx} + \beta f_1 + (1-\beta) f_2 \tag{42}$$

$$s.t. \quad \mathbf{A}_1\mathbf{x} \leq \mathbf{b}_1,\ \mathbf{x} \in \mathbb{R}_+^{n_c} \times \mathbb{Z}_+^{n_d} \tag{43}$$

$$\mathbf{A}_2\mathbf{x} + \mathbf{B}_2\overline{\mathbf{y}} \leq \mathbf{b}_2,\ \overline{\mathbf{y}} \in \mathbb{R}_+^m \tag{44}$$

$$f_1 = \mathbf{d}_1\mathbf{y}^o,\ \mathbf{y}^o \in \mathbf{S}(\mathbf{x}) \tag{45}$$

$$f_2 = \mathbf{d}_1\mathbf{y}^p,\ \mathbf{y}^p \in \arg\max\Big\{\mathbf{d}_1\mathbf{y}^p : \mathbf{B}_2\mathbf{y}^p \leq \mathbf{b}_2 - \mathbf{A}_2\mathbf{x},\ \mathbf{d}_2\mathbf{y}^p \leq \mathbf{d}_2\overline{\mathbf{y}}, \mathbf{y}^p \in \mathbb{R}_+^m\Big\}, \tag{46}$$

which is equivalent to the following single level formulation

$$\begin{aligned}
\min \quad & \mathbf{cx} + \beta f_1 + (1-\beta) f_2 \\
s.t. \quad & \mathbf{A}_1\mathbf{x} \le \mathbf{b}_1,\ \mathbf{A}_2\mathbf{x} + \mathbf{B}_2\overline{\mathbf{y}} \le \mathbf{b}_2,\ f_1 = \mathbf{d}_1\mathbf{y}^o,\ f_2 = \mathbf{d}_1\mathbf{y}^p \\
& 0 \le \mathbf{y}^o \perp (\mathbf{d}_2^t - \mathbf{B}_2^t\mathbf{v}) \ge 0, \quad 0 \ge \mathbf{v} \perp (\mathbf{b}_2 - \mathbf{A}_2\mathbf{x} - \mathbf{B}_2\mathbf{y}^o) \ge 0 \\
& 0 \le \mathbf{y}^p \perp (\mathbf{B}_2^t\mathbf{u} + \mathbf{d}_2^t\pi - \mathbf{d}_1^t) \ge 0, \quad 0 \le \mathbf{u} \perp (\mathbf{b}_2 - \mathbf{A}_2\mathbf{x} - \mathbf{B}_2\mathbf{y}^p) \ge 0 \\
& 0 \le \pi \perp (\mathbf{d}_2\overline{\mathbf{y}} - \mathbf{d}_2\mathbf{y}^p) \ge 0 \\
& \mathbf{x} \in \mathbb{R}_+^{n_c} \times \mathbb{Z}_+^{n_d},\ \overline{\mathbf{y}} \in \mathbb{R}_+^m.
\end{aligned}$$

Proof. Replacing f_1 and f_2 by their corresponding expressions, and according to the optimistic nature regarding f_1, the original $\mathbf{SW-PBL}$ can be equivalently written as

$$\min \quad \mathbf{cx} + \beta\mathbf{d}_1\mathbf{y}^o + (1-\beta)\max_{\mathbf{y}^p \in \mathbf{S}(\mathbf{x})} \mathbf{d}_1\mathbf{y}^p \tag{47}$$

$$\text{s.t.} \quad \mathbf{A}_1\mathbf{x} \le \mathbf{b}_1,\ \mathbf{x} \in \mathbb{R}_+^{n_c} \times \mathbb{Z}_+^{n_d} \tag{48}$$

$$\mathbf{y}^o \in \mathbf{S}(\mathbf{x}). \tag{49}$$

By treating $(\mathbf{x}, \mathbf{y}^o)$ as the upper level decisions, the formulation in (47-49) is actually a pessimistic bilevel problem as **PBL** in (1-3). Next, by using the strategy in Section 2.1, we derive a relaxation as the following

$$\begin{aligned}
\min \quad & \mathbf{cx} + \beta\mathbf{d}_1\mathbf{y}^o + (1-\beta)\mathbf{d}_1\mathbf{y}^p \\
\text{s.t.} \quad & \mathbf{A}_1\mathbf{x} \le \mathbf{b}_1,\ \mathbf{x} \in \mathbb{R}_+^{n_c} \times \mathbb{Z}_+^{n_d} \\
& \mathbf{A}_2\mathbf{x} + \mathbf{B}_2\overline{\mathbf{y}} \le \mathbf{b}_2, \overline{\mathbf{y}} \in \mathbb{R}_+^m \\
& \mathbf{y}^o \in \mathbf{S}(\mathbf{x}) \\
& \mathbf{y}^p \in \arg\max\Big\{\mathbf{d}_1\mathbf{y}^p : \mathbf{B}_2\mathbf{y}^p \le \mathbf{b}_2 - \mathbf{A}_2\mathbf{x},\ \mathbf{d}_2\mathbf{y}^p \le \mathbf{d}_2\overline{\mathbf{y}},\ \mathbf{y}^p \in \mathbb{R}_+^m\Big\},
\end{aligned}$$

which, according to Lemma 2, is a tight relaxation to the formulation in (47-49).

Then, using f_1 and f_2 to replace associated terms in the aforementioned formulation, it leads to the desired **R-PBL** model. Finally, by applying the KKT reformulation method to describe $\mathbf{S}(\mathbf{x})$ and $\arg\max\Big\{\mathbf{d}_1\mathbf{y}^p : \mathbf{B}_2\mathbf{y}^p \le \mathbf{b}_2 - \mathbf{A}_2\mathbf{x}, \mathbf{d}_2\mathbf{y}^p \le \mathbf{d}_2\overline{\mathbf{y}}, \mathbf{y}^p \in \mathbb{R}_+^m\Big\}$ respectively, we have the equivalent single level formulation of **R-PBL**. □

Remark: (*i*) Again, the equivalent single level formulation can be directly computed by an MPCC optimization package or by an MIP solver. We observe in numerical experiments that, instead of imposing equality restrictions, it often has computational advantage to employ $f_1 \ge \mathbf{d}_1\mathbf{y}^o$ and $f_2 \ge \mathbf{d}_1\mathbf{y}^p$, which, when an optimal solution is derived, are actually satisfied at equality. (*ii*) Another equivalent single level formulation can also be obtained if we explicitly impose the optimality requirement with respect to $\mathbf{d}_2\overline{\mathbf{y}}$ on $\overline{\mathbf{y}}$, and link optimistic and pessimistic parts by $f_1 \ge \mathbf{d}_1\overline{\mathbf{y}}$, which enables us to eliminate variables $\mathbf{y}^o$ and related constraints. Nevertheless, we prefer the single level formulation presented in Corollary 9, whose structure allows optimistic and pessimistic parts to be computed independently for a given $\mathbf{x}$. Indeed, for the situation where the optimality requirement is not directly available, e.g., a strong-weak formulation with an MIP lower level problem, we still can make use of the primal information of the lower level problem represented by $\bar{\mathbf{y}}$ to effectively compute the pessimistic portion, which is independent of the solution quality of $\mathbf{y}^o$.

4 Numerical Study

In this section, we apply the developed computing scheme to compute instances of linear **PBL** and its extensions, and demonstrate its computational efficacy. Note that there are a large number of bilevel optimization algorithms and methods that can be used to solve their bilevel tight relaxations. To provide a rather standard computational platform, a commercial solver, i.e., Gurobi 6.50 with default settings, is adopted to solve (linearized) KKT conditions based single level formulations of tight relaxations. In addition, all experiments are performed on a PC with Intel Xeon E5-1620 3.60GHz CPU and 32GB memory.

Our computational study contains two parts. In the first part, simple examples of **PBL** (and its extensions) with detailed parameters are used to illustrate steps of our computing scheme. In the second part, a systematic computational study on random instances of two practical models are reported, which shows the strength of this computing scheme in addressing real problems.

4.1 Simple Illustrative Examples

In this subsection, we first consider one simple linear pessimistic bilevel example appearing in [16]. Then, we extend it to have more general structures as in Sections 3.2-3.4. For those simple instances, we do not report their computational times as they are negligible.

Example 1. *Consider the following linear pessimistic bilevel problem adopted from [16].*

$$\min \quad -8x_1 - 6x_2 + \max(-25y_1 - 30y_2 + 2y_3 + 16y_4) \tag{50}$$

$$s.t. \quad x_1 + x_2 \leq 10,\ x_1, x_2 \geq 0 \tag{51}$$

$$(y_1, y_2, y_3, y_4) \in \mathbf{S}(x_1, x_2) = \arg\min\Big\{ -10y_1 - 10y_2 - 10y_3 - 10y_4 : \tag{52}$$

$$y_1 + y_2 + y_3 + y_4 \leq 10 - x_1 - x_2,\ -y_1 + y_4 \leq 0.8x_1 + 0.8x_2 \tag{53}$$

$$y_2 + y_4 \leq 4x_2,\ y_1, y_2, y_3, y_4 \geq 0\Big\}. \tag{54}$$

*According to Corollary 5, the associated **R-PBL** model is*

$$\begin{aligned}
&\min\ -8x_1 - 6x_2 - 25y_1 - 30y_2 + 2y_3 + 16y_4\\
&\quad s.t.\ x_1 + x_2 \le 10,\ \overline{y}_1 + \overline{y}_2 + \overline{y}_3 + \overline{y}_4 \le 10 - x_1 - x_2\\
&\qquad -\overline{y}_1 + \overline{y}_4 \le 0.8x_1 + 0.8x_2,\ \overline{y}_2 + \overline{y}_4 \le 4x_2,\ x_1, x_2 \ge 0,\ \overline{y}_1, \overline{y}_2, \overline{y}_3, \overline{y}_4 \ge 0\\
&\qquad (y_1, y_2, y_3, y_4) \in \arg\max\Big\{ -25y_1 - 30y_2 + 2y_3 + 16y_4 :\\
&\qquad y_1 + y_2 + y_3 + y_4 \le 10 - x_1 - x_2,\ -y_1 + y_4 \le 0.8x_1 + 0.8x_2,\ y_2 + y_4 \le 4x_2\\
&\qquad -10y_1 - 10y_2 - 10y_3 - 10y_4 \le -10\overline{y}_1 - 10\overline{y}_2 - 10\overline{y}_3 - 10\overline{y}_4,\ y_1, y_2, y_3, y_4 \ge 0\Big\}.
\end{aligned}$$

By solving its single level KKT reformulation in the form of that in Corollary 5, we obtain an optimal solution with $x_1^* = 10, x_2^* = 0$, $\overline{y}_j^* = y_j^* = 0$ for $j = 1, \dots, 4$, and the optimal value equal to -80. Because we have

$$\begin{aligned}
&-10\overline{y}_1^* - 10\overline{y}_2^* - 10\overline{y}_3^* - 10\overline{y}_4^* = 0\\
&= \min\Big\{ -10y_1 - 10y_2 - 10y_3 - 10y_4 :\\
&y_1 + y_2 + y_3 + y_4 \le 0,\ -y_1 + y_4 \le 8, y_2 + y_4 \le 0,\ y_1, y_2, y_3, y_4 \ge 0\Big\},
\end{aligned}$$

we can conclude that the $\mathbf{y}^*$-portion is optimal to the lower level problem and the correction step can simply be ignored. Finally, we report $(x_1, x_2, y_1, y_2, y_3, y_4) = (10, 0, 0, 0, 0, 0)$ as an optimal solution to the original pessimistic problem.

As demonstrated next (as well the example in (10)), it is not always the case that we can neglect the correction step.

Example 2. *Consider a modification of (50-54) where the objective function in (50) is changed to $\min 0x_1 + 0x_2 + \max(-2y_1 - 3y_2 - 2y_3 - 16y_4)$. By repeating the same solution procedure, we obtain an optimal solution with $x_1^* = x_2^* = 0$ and $\overline{y}_j^* = 0$ for $j = 1, \dots, 4$, and $y_1^* = 10, y_j^* = 0$ for j=2,3,4. Note that*

$$\begin{aligned}
&-10\overline{y}_1^* - 10\overline{y}_2^* - 10\overline{y}_3^* - 10\overline{y}_4^* = 0\\
&> -100 = \min\Big\{ -10y_1 - 10y_2 - 10y_3 - 10y_4 :\\
&y_1 + y_2 + y_3 + y_4 \le 10,\ -y_1 + y_4 \le 0,\ y_2 + y_4 \le 0,\ y_1, y_2, y_3, y_4 \ge 0\Big\},
\end{aligned}$$

which suggests that $(y_1^*, y_2^*, y_3^*, y_4^*)$ might not be optimal to the lower level problem. Hence, according to (11), we perform the correction step and derive an optimal solution (which is identical to $(x_1, x_2, y_1, y_2, y_3, y_4) = (0, 0, 10, 0, 0, 0)$) for this pessimistic bilevel problem. Certainly, as discussed after the relaxation-and-correction scheme, we can compare the value of $-10y_1^* - 10y_2^* - 10y_3^* - 10y_4^*$ and the optimal value of the lower level problem, which may allow us to ignore the correction step for this instance.

In the following, we solve a modification of (50-54) that is a mixed integer **PBL** instance.

Example 3. *Consider an instance extended from (50-54) with a new constraint on* (x_1, x_2) *in the upper level and an integer variable restriction on* y_3 *in the lower level. So, the lower level problem is a mixed integer program.*

$$
\begin{aligned}
\min \quad & -8x_1 - 6x_2 + \max(-25y_1 - 30y_2 + 2y_3 + 16y_4) \\
s.t. \quad & x_1 + x_2 \le 10, \ 2x_1 + 5x_2 \le 13, \ x_1, x_2 \ge 0 \\
& (y_1, y_2, y_3, y_4) \in \mathbf{S}(x_1, x_2) = \arg\min \Big\{ -10y_1 - 10y_2 - 10y_3 - 10y_4 : \\
& y_1 + y_2 + y_3 + y_4 \le 10 - x_1 - x_2, \ -y_1 + y_4 \le 0.8x_1 + 0.8x_2 \\
& y_2 + y_4 \le 4x_2, \ y_1, y_2, y_4 \ge 0, \ y_3 \in \mathbb{Z}_+ \Big\}.
\end{aligned}
$$

*The extended formulation of its tight **R-PBL** relaxation with artificial variables is*

$$
\begin{aligned}
\min \quad & -8x_1 - 6x_2 + \eta \\
s.t. \quad & x_1 + x_2 \le 10, \ 2x_1 + 5x_2 \le 13, \ x_1 + x_2 + \overline{y}_1 + \overline{y}_2 + \overline{y}_3 + \overline{y}_4 \le 10 \\
& -0.8x_1 - 0.8x_2 - \overline{y}_1 + \overline{y}_4 \le 0, \ -4x_2 + \overline{y}_2 + \overline{y}_4 \le 0 \\
& x_1, x_2 \ge 0, \ \overline{y}_1, \overline{y}_2, \overline{y}_4 \ge 0, \ \overline{y}_3 \in \mathbb{Z}_+ \\
& \eta \ge \max \Big\{ -25y_1 - 30y_2 + 2y_3 + 16y_4 - My_{a1} - My_{a2} - My_{a3} - My_f : \\
& y_1 + y_2 + y_3 + y_4 - y_{a1} \le 10 - x_1 - x_2, \ -y_1 + y_4 - y_{a2} \le 0.8x_1 + 0.8x_2 \\
& y_2 + y_4 - y_{a3} \le 4x_2, \ -10y_1 - 10y_2 - 10y_3 - 10y_4 - y_f \le -10\overline{y}_1 - 10\overline{y}_2 - 10\overline{y}_3 - 10\overline{y}_4 \\
& y_1, y_2, y_4, y_{a1}, y_{a2}, y_{a3}, y_f \ge 0, \ y_3 \in \mathbb{Z}_+ \Big\},
\end{aligned}
$$

where M *is set to* $5,000$ *in our numerical study. Using the procedure described in Appendix* 1 *to compute the aforementioned bilevel mixed integer program, we obtain an optimal solution with* $x_1^* = 6.005$*,* $x_2^* = 0$*,* $\overline{y}_1^* = 3.995$*,* $\overline{y}_2^* = \overline{y}_3^* = \overline{y}_4^* = 0$*,* $y_1^* = 0.995$*,* $y_2^* = y_4^* = 0$ *and* $y_3^* = 3$*, and the corresponding optimal value is* -66.915*. Then, deriving the optimal value of the original lower level problem given* (x_1^*, x_2^*) *and comparing it with respect to* $-10\overline{y}_1^* - 10\overline{y}_2^* - 10\overline{y}_3^* - 10\overline{y}_4^*$*, it can be seen that the correction step is not needed.*

The close proximity of x_1^* to an integer value indicates that the solution is ϵ-optimal and this **PBL** instance does not have any exact solution. Indeed, it is the actual case, which is confirmed by an analytical study presented in Appendix 2. It is worth mentioning that, interestingly, the optimistic counterpart of this instance has an optimal solution with $x_1 = 0$, $x_2 = 2$, $y_1 = y_3 = y_4 = 0$ and $y_2 = 8$, and the corresponding optimal value is -252.

Next, we construct an instance based on (50-54) that has a coupled pessimistic constraint.

Example 4. *Consider the following problem*

$$
\begin{aligned}
\min \quad & -8x_1 - 6x_2 \\
s.t. \quad & x_1 + x_2 \le 10, \ x_1, x_2 \ge 0 \\
& 2x_1 + x_2 - y_1 - y_2 - y_3 - y_4 \le 0, \ \forall (y_1, y_2, y_3, y_4) \in \mathbf{S}(x_1, x_2)
\end{aligned}
$$

where $\mathbf{S}(x_1, x_2)$ *in the coupled constraint is defined as in (52-54). Note that* $\mathbf{Y}(x_1, x_2) =$

$\{(y_1,y_2,y_3,y_4)\in\mathbb{R}^4_+ : y_1+y_2+y_3+y_4\le 10-x_1-x_2,\ -y_1+y_4\le 0.8x_1+0.8x_2,\ y_2+y_4\le 4x_2\}$ *is a non-empty compact set for any feasible* (x_1,x_2). *Then, according to Lemma 7, we can reformulate the whole bilevel problem as*

$$
\begin{aligned}
\min \quad & -8x_1-6x_2\\
s.t. \quad & x_1+x_2\le 10,\ \overline{y}_1+\overline{y}_2+\overline{y}_3+\overline{y}_4+x_1+x_2\le 10\\
& -\overline{y}_1+\overline{y}_4-0.8x_1-0.8x_2\le 0,\ \overline{y}_2+\overline{y}_4-4x_2\le 0\\
& 2x_1+x_2+\max_{(y_1,y_2,y_3,y_4)\in\tilde{\mathbf{Y}}(x_1,x_2,\overline{y}_1,\overline{y}_2,\overline{y}_3,\overline{y}_4)}(-y_1-y_2-y_3-y_4)\le 0\\
& x_1,x_2,\overline{y}_1,\overline{y}_2,\overline{y}_3,\overline{y}_4\ge 0
\end{aligned}
$$

where $\tilde{\mathbf{Y}}(x_1,x_2,\overline{y}_1,\overline{y}_2,\overline{y}_3,\overline{y}_4)=\{(y_1,y_2,y_3,y_4)\in\mathbb{R}^4_+ : y_1+y_2+y_3+y_4\le 10-x_1-x_2,\ -y_1+y_4\le 0.8x_1+0.8x_2,\ y_2+y_4\le 4x_2, -10y_1-10y_2-10y_3-10y_4\le -10\overline{y}_1-10\overline{y}_2-10\overline{y}_3-10\overline{y}_4\}$. *By solving its single level reformulation in the form of that in Corollary 8, we obtain an optimal solution* $x_1^*=0$ *and* $x_2^*=5$.

Finally, we extend (50-54) into an instance of the strong-weak bilevel problem.

Example 5. *Consider the following instance of the strong-weak bilevel problem, which has been solved by iterative algorithms developed in [16, 44].*

$$
\begin{aligned}
\min \quad & -8x_1-6x_2+\beta f_1+(1-\beta)f_2\\
s.t. \quad & x_1+x_2\le 10,\ x_1,x_2\ge 0\\
& f_1=\min\{-25y_1-30y_2+2y_3+16y_4 : (y_1,y_2,y_3,y_4)\in\mathbf{S}(x_1,x_2)\}\\
& f_2=\max\{-25y_1-30y_2+2y_3+16y_4 : (y_1,y_2,y_3,y_4)\in\mathbf{S}(x_1,x_2)\}
\end{aligned}
$$

where $\mathbf{S}(x_1,x_2)$ *is defined as in (52-54). According to Corollary 9, its* ***R-PBL*** *model can be written as*

$$
\begin{aligned}
\min \quad & -8x_1-6x_2+\beta f_1+(1-\beta)f_2\\
s.t. \quad & x_1+x_2\le 10,\ \overline{y}_1+\overline{y}_2+\overline{y}_3+\overline{y}_4\le 10-x_1-x_2,\ -\overline{y}_1+\overline{y}_4\le 0.8x_1+0.8x_2\\
& \overline{y}_2+\overline{y}_4\le 4x_2,\ x_1,x_2,\overline{y}_1,\overline{y}_2,\overline{y}_3,\overline{y}_4\ge 0\\
& f_1=-25y_1^o-30y_2^o+2y_3^o+16y_4^o,\ (y_1^o,y_2^o,y_3^o,y_4^o)\in\mathbf{S}(x_1,x_2)\\
& f_2=-25y_1^p-30y_2^p+2y_3^p+16y_4^p\\
& (y_1^p,y_2^p,y_3^p,y_4^p)\in\arg\max\Big\{-25y_1-30y_2+2y_3+16y_4 :\\
& y_1+y_2+y_3+y_4\le 10-x_1-x_2,\ -y_1+y_4\le 0.8x_1+0.8x_2\\
& y_2+y_4\le 4x_2,\ -10y_1-10y_2-10y_3-10y_4\le -10\overline{y}_1-10\overline{y}_2-10\overline{y}_3-10\overline{y}_4,\\
& y_1,y_2,y_3,y_4\ge 0\Big\}.
\end{aligned}
$$

Again, we can compute its single level reformulation in the form of that in Corollary 9 to derive optimal solutions and values for different β. *For instance, if* β *is set to 0.2, an optimal solution is with* $(x_1^*,x_2^*)=(10,0)$ *and the optimal value is -80. If* β *is set to 0.5, an optimal* *solution is with* $(x_1^*, x_2^*) = (0, 0)$ *and the optimal value is -115. Results of both cases agree with those reported in [16, 44]. Moreover, by using the method of [44], we can easily consider a small number of values for* β *and depict the complete relationship between the optimal value and* β*.*

4.2 Instances of Practical Problems

In this subsection, we investigate the computational performance of our computing scheme on deriving pessimistic solutions of two practical bilevel optimization models, i.e., bilevel network design model for vehicle sharing program (VSP) [36] and bilevel model for gene knockout to achieve microbial strain optimization [14], which are referred to as *bilevel network design model* and *bilevel gene knockout model*, respectively. Note that those models are initially presented in optimistic forms, and optimistic solutions of instances with practical or random data, derived by KKT reformulation or strong duality based methods, have also been reported.

We mention that by using the computing scheme developed in this paper, pessimistic solutions for instances with actual data of the second model are provided in [5], which have revealed critical and interesting biological insights. Because our purpose here is to demonstrate the computational performance, rather than to provide economic or biological analysis, we generate instances of random parameters and of different sizes (available at [1]), and report computational results for their pessimistic bilevel forms.

Case 1: Computing Bilevel Network Design Model. We next present the pessimistic bilevel model for network design problem with a brief description. Detailed information regarding vehicle sharing program, modeling motivation, and the optimistic bilevel formulation can be found in [36]. We intentionally keep notations and representations same as those in [36], which minimizes the difficulties to appreciate this pessimistic model.

$$
\begin{aligned}
\max_{\mathbf{x},\mathbf{y},\mathbf{z}} \ \min_{\mathbf{v},\mathbf{w}\in S(\mathbf{x},\mathbf{y},\mathbf{z})} \ & \sum_{k\in K} \sum_{(i,j)\in A_s} r_{ij} v_{ijk} \\
s.t. \ & \sum_{i\in V_s} (C_s x_i + C_p y_i + C_v z_i) \leq C \\
& U x_i \geq y_i, \quad z_i \leq y_i, \quad y_i \leq y^{ub}, \qquad \forall i \in V_s \\
& x_i \in \{0,1\}, \quad y_i, z_i \in \mathbb{Z}_+, \qquad \forall i \in V_s \\
& S(\mathbf{x},\mathbf{y},\mathbf{z}) = \arg\min_{\mathbf{v},\mathbf{w}} \Big\{ \sum_{k\in K} \big(\sum_{(i,j)\in A} c_{ij} v_{ijk} + \sum_{i\in V} w_{ik} \big) : \\
& \sum_{j:(i,j)\in A} v_{ijk} - \sum_{j:(j,i)\in A} v_{jik} = D_{ik}, \qquad \forall i \in V, k \in K \\
& v_{ijk} \leq f_{ij} w_{ik}, \qquad \forall (i,j) \in A \backslash \underline{A}, k \in K \\
& U x_i \geq \sum_{k\in K} v_{ijk}, \quad U x_j \geq \sum_{k\in K} v_{ijk}, \qquad \forall (i,j) \in A_s \\
& \sum_{k\in K} \sum_{j:(i,j)\in A_s} v_{ijk} \leq z_i, \qquad \sum_{k\in K} \sum_{j:(j,i)\in A_s} v_{jik} \leq a(y_i - z_i), \qquad \forall i \in V_s \\
& w_{ik} \geq 0, \quad \forall i \in V, k \in K; \quad v_{ijk} \geq 0, \quad \forall (i,j) \in A, k \in K \Big\}.
\end{aligned}
$$

In this formulation, the upper level DM seeks to maximize her overall revenue that depends on aggregated traffic flows from the lower level DM. She makes decisions on the construction of vehicle sharing station at candidate site i (denoted by binary variable x_i), the capacity of that station (denoted by integer variable y_i), and the number of vehicles at i (denoted by integer variable z_i) in the transportation network, subject to the total budget (denoted by C), and capacity upper bound restrictions (represented by parameter U). The lower level DM, facing the system with vehicle sharing stations, selects an optimal route to minimize his traveling cost (for every origin-destination transportation pair $k \in K$), subject to flow balance constraints, capacity restrictions and waiting time estimations. Note that the pessimistic upper level objective function, i.e., $\max_{\mathbf{x},\mathbf{y},\mathbf{z}}(\min_{\mathbf{v},\mathbf{w}} \sum_{k\in K} \sum_{(i,j)\in A_s} r_{ij}v_{ijk})$, takes into consideration the non-cooperative behavior of travellers.

In Table 1, we present detailed performance data of our computing scheme for different instances, where columns "Nodes" and "Arcs" list the number of nodes and arcs in the transportation network; column "C" gives the overall budget; columns "DV" and "CV" provide the number of discrete variables (appearing in the upper level) and the number of continuous variables (appearing in the lower level); column "Const. (U/L)" displays numbers of constraints in upper and lower problems, respectively; column "Obj." reports the optimal value; and column "T(sec.)" gives the solution time of our computing scheme in seconds. We mention that instances of 5 nodes are based on the 5-node network given in [36] and other instances are randomly generated with GRID structure defined in [36].

Table 1: Computational Result of Pessimistic Bilevel Network Design Model

Nodes	Arcs	C	DV	CV	Const.(U/L)	Obj.	T(sec.)
5	14	10	12	38	13/30	0	0.02
		30	12	38	13/30	4	0.44
		50	12	38	13/30	10	0.23
		100	12	38	13/30	20	0.18
12	38	50	15	150	16/128	12.5	0.37
		100	15	150	16/128	25	1.68
		200	15	150	16/128	25	1.38
		400	15	150	16/128	25	4.84
17	56	50	21	219	22/155	15	0.57
		100	21	219	22/155	30	5.41
		200	21	219	22/155	30	3.19
		400	21	219	22/155	30	3.68
22	77	50	27	198	28/156	15	1.45
		100	27	198	28/156	27	234.19
		200	27	198	28/156	27	203.38
		400	27	198	28/156	27	118.54

To demonstrate differences between optimistic and pessimistic models, we compute optimal values of both models for a 17-node instance over budgets ranging from 0 to 400, which are presented as **OBL** and **PBL** curves in Figure 1.

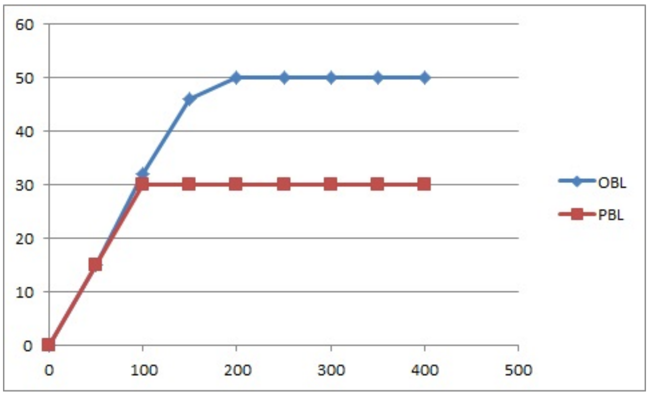

Figure 1: Optimal values of optimistic and pessimistic models

Case 2: Computing Bilevel Gene Knockout Model. In the following, we present the pessimistic bilevel gene knockout model with a brief description. Detailed information regarding gene knockout, modeling motivation, and the optimistic bilevel formulation can be found in [14]. Again, notations and representations are same as those in [14].

$$
\begin{aligned}
\max_{\mathbf{z}} \min_{\mathbf{v}\in S(\mathbf{z})} \quad & v_{chemical} \\
s.t. \sum_{j\in \mathsf{M}} & (1-z_j) \le K \\
& z_j \in \{0,1\}, \forall j \in \mathsf{M} \\
& S(\mathbf{z}) = \arg\max_{v} \Big\{ v_{biom} : \\
& \sum_{j\in \mathsf{M}} S_{ij} v_j = 0, \ \forall i \in \mathsf{N} \\
& v_{glc} = v_{glc\_uptake}, \quad v_{biom} \ge v_{biom}^{target} \\
& v_j^{min} z_j \le v_j \le v_j^{max} z_j, \ \forall j \in \mathsf{M} \Big\}
\end{aligned}
$$

In this formulation, the upper level DM seeks to maximize the cellular production of a particular chemical by removing up to K metabolic reactions (denoted by binary variable y_j) to modify the connectivity of metabolic network. Given that updated metabolic network, the lower level DM maximizes its cellular objective function, which is biomass formation, by determining reaction fluxes (denoted by continuous variable v_j) in the network, subject to fluxes balance requirements and bound restrictions. Again, the pessimistic upper level objective function, i.e., $\max_{\mathbf{z}} \min_{\mathbf{v}\in S(\mathbf{z})} \ v_{chemical}$, takes into consideration the impact of non-cooperative behavior of the cellular organism.

As mentioned, the aforementioned pessimistic bilevel model has been studied in [5] to generate biological significant solutions and insights. Different from that, we modify instances to have random data and different sizes to understand our solution scheme's computational capacity. Specifically, on top of the E.coli metabolic network, we randomly do: (i) perturb v_j^{min} and v_j^{max} by up to 25%, (ii) add 6 reactions (i.e., arcs in the network) to have 120 reactions in total, and (iii) define a subset of reactions as removable ones for knockout. Note that the whole network is rather sparse, given that it has 77 nodes and 120 arcs. In Table 2, which is in the same structure of Table 1, detailed computational results are presented, where column "Rmvbl. Rct." gives the cardinality of the subset of removable reactions in the network, which determines the number of binary variables in the upper level; column "K" defines the upper bound on the number of reactions to be removed.

Table 2: Computational Results of Pessimistic Bilevel Gene Knockout Model

Nodes	Rmvbl. Rct.	K	DV	CV	Const. (U/L)	Obj.	T(sec.)
77	60	3	60	120	1/319	0.31	14.79
		5	60	120	1/319	13.76	17.59
		10	60	120	1/319	18.91	15.07
		20	60	120	1/319	18.91	16.74
77	80	3	80	120	1/319	9.08	46.68
		5	80	120	1/319	79.27	16.68
		10	80	120	1/319	85.64	7.38
		20	80	120	1/319	85.64	6.97
77	100	3	100	120	1/319	110.19	10.46
		5	100	120	1/319	127.69	5.28
		10	100	120	1/319	129.00	11.62
		20	100	120	1/319	129.00	4.12
77	120	3	120	120	1/319	110.47	8.46
		5	120	120	1/319	129.50	15.34
		10	120	120	1/319	131.58	5.04
		20	120	120	1/319	131.58	11.59

Observations on Computational Performance. Based on results presented in Tables 1-2 and Figure 1, we note that (i) the computing scheme presented in this paper demonstrates a very strong solution capacity to deal with practical instances. This is more clear for bilevel gene knockout model. Although it has hundreds of variables and constraints, its sparse structure definitely reduces the computational burden. To the best of our knowledge, we have not seen such scale instances have been computed in the literature. Indeed, we are not aware of any substantial computational investigation on other methods or algorithms for pessimistic bilevel optimization problems, except for a systematic study presented in [40] on an iterative procedure to solve instances with a special independent property. Hence, our computing scheme is probably the most general and also capable tool to solve practical pessimistic bilevel instances. (ii) Nevertheless, we also observe that our computing scheme could be sensitive to the size of instances. In Table 1, when the number of nodes increases from 17 to 22, there is a non-trivial increase in the computational time. It indicates that if the problem is less sparse, its size could have a great impact on the algorithm performance. Hence, one future direction is to develop various enhancements so that it can scale gracefully with respect to the problem size. (*iii*) In terms of model's outputs, there could be a significant difference between optimistic and pessimistic ones. As can be seen in Figure 1, the optimal value of an optimistic model could be very much higher than that of its pessimistic counterpart. Within an optimistic environment, i.e., travellers are co-operative, one would anticipate that expanding a vehicle sharing program with a larger budget, up to 200, would always lead to a better revenue. However, under the pessimistic environment, the benefit from a larger budget might be much less significant or soon have a zero marginal value. With such an observation, either a more accurate model to describe travellers' behavior is needed, or a strong-weak bilevel optimization can be employed to derive a solution with the desired trade-off.

5 Conclusion

In this paper, we develop a tight relaxation of **PBL** and then design a simple computing scheme that helps us derive a solution being both feasible and optimal to **PBL**. We also discuss using this scheme to compute linear **PBL** and a few extensions. Then, we illustrate our computing scheme on several simple **PBL** examples, and perform a systematic numerical study on random instances of practical problems to demonstrate its effectiveness. Given that this scheme has convenient interfaces to existing research on bilevel problems and displays a strong computational capacity, we believe that it should be practically useful in solving pessimistic bilevel problems arising from real systems. One future research direction is to develop fast heuristics and problem-specific customizations to support this scheme so that it is scalable to large-scale instances.

Appendix 1: Reformulation and Decomposition Method for MIP – PBL

We first provide the complete extended reformulation of **R-PBL** for **MIP – PBL**.

$$
\begin{aligned}
\min \quad & \mathbf{c}\mathbf{x} + \eta : \\
\text{s.t.} \quad & \mathbf{A}_1\mathbf{x} \le \mathbf{b}_1,\ \mathbf{A}_2\mathbf{x} + \mathbf{B}_{2c}\overline{\mathbf{y}}_c + \mathbf{B}_{2d}\overline{\mathbf{y}}_d \le \mathbf{b}_2 \\
& \mathbf{x} \in \mathbb{R}_+^{n_c} \times \mathbb{Z}_+^{n_d}, \overline{\mathbf{y}}_c \in \mathbb{R}_+^{m_c},\ \overline{\mathbf{y}}_d \in \mathbb{Z}_+^{m_d} \\
& \eta \ge \max\Big\{\mathbf{d}_{1c}\mathbf{y}_c + \mathbf{d}_{1d}\mathbf{y}_d - M\mathbf{1}\mathbf{y}_a - My_f : \\
& \qquad \mathbf{B}_{2c}\mathbf{y}_c + \mathbf{B}_{2d}\mathbf{y}_d - \mathbf{I}\mathbf{y}_a \le \mathbf{b}_2 - \mathbf{A}_2\mathbf{x}, \quad \mathbf{d}_{2c}\mathbf{y}_c + \mathbf{d}_{2d}\mathbf{y}_d - y_f \le \mathbf{d}_{2c}\overline{\mathbf{y}}_c + \mathbf{d}_{2d}\overline{\mathbf{y}}_d \\
& \qquad \mathbf{y}_a \in \mathbb{R}_+^{l},\ \mathbf{y}_c \in \mathbb{R}_+^{m_c},\ \mathbf{y}_d \in \mathbb{Z}_+^{m_d}, y_f \in \mathbb{R}_+\Big\}.
\end{aligned}
$$

Let $\mathbb{Y} = \{\mathbf{y}_d^1, \ldots, \mathbf{y}_d^{|\mathbf{J}|}\}$ be the collection of all possible values of $\mathbf{y}_d$. Then, by enumerating $\mathbf{y}_d^j, j \in \mathbf{J}$, the last inequality can be equivalently written as

$$
\begin{aligned}
\eta \quad \ge \quad & \mathbf{d}_{1d}\mathbf{y}_d^j + \max\Big\{\mathbf{d}_{1c}\mathbf{y}_c - M\mathbf{1}\mathbf{y}_a - My_f : \\
& \qquad \mathbf{B}_{2c}\mathbf{y}_c - \mathbf{I}\mathbf{y}_a \le \mathbf{b}_2 - \mathbf{A}_2\mathbf{x} - \mathbf{B}_{2d}\mathbf{y}_d^j, \quad \mathbf{d}_{2c}\mathbf{y}_c - y_f \le \mathbf{d}_{2c}\overline{\mathbf{y}}_c + \mathbf{d}_{2d}\overline{\mathbf{y}}_d - \mathbf{d}_{2d}\mathbf{y}_d^j \\
& \qquad \mathbf{y}_a \in \mathbb{R}_+^{l},\ \mathbf{y}_c \in \mathbb{R}_+^{m_c}, y_f \in \mathbb{R}_+\Big\}\ \forall j \in \mathbf{J}.
\end{aligned}
$$

Note that every constraint for $j \in \mathbf{J}$ can be reformulated through KKT reformulation method, which eliminates max operation and leads to a single level representation of the extended reformulation of this **R-PBL**. Clearly, for a real problem, such single level representation could be extremely large. Nevertheless, a single level representation based on a subset of $\mathbf{J}$, e.g., the **MP** in the following decomposition algorithm, provides a relaxation and a lower bound. Moreover, by dynamically expanding that subset with a critical $\mathbf{y}_d^j$, which is determined by solving **SP** in the following decomposition algorithm, we can have a tighter relaxation and a stronger lower bound. Note also that a better feasible solution provides a stronger upper bound, which can be done by comparing solutions from **SP**s. Hence, through iteratively computing **MP** and **SP**, convergence to the optimal value can be achieved.

Based on the aforementioned reformulation strategy, we next describe the customized decomposition method, i.e., *the column-and-constraint generation algorithm* [42], to compute the extended reformulation of **R-PBL** for **MIP – PBL**. Let UB and LB be the upper and lower bounds respectively, k be the iteration index and ϵ be the optimality tolerance. Note that in the following **MP** formulation, $\mathbf{y}_d^j$ are constants for all j.

Column-and-Constraint Generation Algorithm for Bilevel MIP

Step 1: Set $LB = -\infty$, $UB = +\infty$, and $k = 0$.

Step 2: Solve the following master problem

$$
\begin{aligned}
\textbf{MP}: \quad & \underline{\Theta}_p^* = \mathbf{c}\mathbf{x} + \eta \\
\text{s.t.} \quad & \mathbf{A}_1\mathbf{x} \le \mathbf{b}_1, \qquad \mathbf{A}_2\mathbf{x} + \mathbf{B}_{2c}\overline{\mathbf{y}}_c + \mathbf{B}_{2d}\overline{\mathbf{y}}_d \le \mathbf{b}_2 \\
& \eta \ge \mathbf{d}_{1d}\mathbf{y}_d^j + \mathbf{d}_{1c}\mathbf{y}_c^j - M\mathbf{1}\mathbf{y}_a^j - My_f^j, \;\; 1 \le j \le k, \\
& 0 \le \mathbf{y}_c^j \perp (\mathbf{B}_{2c}^t\mathbf{u}^j + \mathbf{d}_{2c}^t\pi^j - \mathbf{d}_{1c}^t) \ge 0, \; 1 \le j \le k \\
& 0 \le \mathbf{y}_a^j \perp (M\mathbf{1}^t - \mathbf{I}\mathbf{u}^j) \ge 0, \; 1 \le j \le k \\
& 0 \le y_f^j \perp (M - \pi^j) \ge 0, \; 1 \le j \le k \\
& 0 \le \mathbf{u}^j \perp (\mathbf{b}_2 - \mathbf{A}_2\mathbf{x} - \mathbf{B}_{2d}\mathbf{y}_d^j - \mathbf{B}_{2c}\mathbf{y}_c^j + \mathbf{I}\mathbf{y}_a^j) \ge 0, \;\; 1 \le j \le k \\
& 0 \le \pi^j \perp (\mathbf{d}_{2c}\overline{\mathbf{y}}_c + \mathbf{d}_{2d}\overline{\mathbf{y}}_d - \mathbf{d}_{2d}\mathbf{y}_d^j - \mathbf{d}_{2c}\mathbf{y}_c^j + y_f) \ge 0, \;\; 1 \le j \le k \\
& \mathbf{x} \in \mathbb{R}_+^{n_c} \times \mathbb{Z}_+^{n_d}, \overline{\mathbf{y}}_c \in \mathbb{R}_+^{m_c}, \overline{\mathbf{y}}_d \in \mathbb{Z}_+^{m_d}
\end{aligned}
$$

Derive an optimal solution $(\mathbf{x}^*, \overline{\mathbf{y}}_c^*, \overline{\mathbf{y}}_d^*, \mathbf{y}_c^{1*}, \mathbf{y}_a^{1*}, y_f^{1*}, \mathbf{u}^{1*}, \pi^{1*}, \dots, \mathbf{y}_c^{k*}, \mathbf{y}_a^{k*}, y_f^{k*}, \mathbf{u}^{k*}, \pi^{k*})$, and update $LB = \underline{\Theta}_p^*$.

Step 3: If $UB - LB \le \epsilon$, return UB and the corresponding (incumbent) solution. Terminate. Otherwise, go to Step 4.

Step 4: Solve the following lower level problem for given $(\mathbf{x}^*, \overline{\mathbf{y}}_c^*, \overline{\mathbf{y}}_d^*)$, which serves as the subproblem.

$$
\begin{aligned}
\textbf{SP}: \; \varphi(\mathbf{x}^*, \overline{\mathbf{y}}_c^*, \overline{\mathbf{y}}_d^*) = \max \quad & \mathbf{d}_{1c}\mathbf{y}_c + \mathbf{d}_{1d}\mathbf{y}_d \\
\text{s.t.} \quad & \mathbf{B}_{2c}\mathbf{y}_c + \mathbf{B}_{2d}\mathbf{y}_d \le \mathbf{b}_2 - \mathbf{A}_2\mathbf{x}^* \\
& \mathbf{d}_{2c}\mathbf{y}_c + \mathbf{d}_{2d}\mathbf{y}_d \le \mathbf{d}_{2c}\overline{\mathbf{y}}_c^* + \mathbf{d}_{2d}\overline{\mathbf{y}}_d^* \\
& \mathbf{y}_c \in \mathbb{R}_+^{m_c}, \; \mathbf{y}_d \in \mathbb{Z}_+^{m_d}
\end{aligned}
$$

Derive an optimal solution $(\mathbf{y}_c^*, \mathbf{y}_d^*)$, and update $UB = \min\{UB, \mathbf{c}\mathbf{x}^* + \varphi(\mathbf{x}^*, \overline{\mathbf{y}}_c^*, \overline{\mathbf{y}}_d^*)\}$.

Step 5: Set $\mathbf{y}_d^{k+1} = \mathbf{y}_d^*$, create variables $(\mathbf{y}_c^{k+1}, \mathbf{y}_a^{k+1}, y_f^{k+1}, \mathbf{u}^{k+1}, \pi^{k+1})$, and add the following constraints to **MP**. Set $k = k + 1$ and go to Step 2.

$$
\begin{aligned}
& \eta \ge \mathbf{d}_{1d}\mathbf{y}_d^{k+1} + \mathbf{d}_{1c}\mathbf{y}_c^{k+1} - M\mathbf{1}\mathbf{y}_a^{k+1} - My_f^{k+1} \\
& 0 \le \mathbf{y}_c^{k+1} \perp (\mathbf{B}_{2c}^t\mathbf{u}^{k+1} + \mathbf{d}_{2c}^t\pi^{k+1} - \mathbf{d}_{1c}^t) \ge 0 \\
& 0 \le \mathbf{y}_a^{k+1} \perp (M\mathbf{1}^t - \mathbf{I}\mathbf{u}^{k+1}) \ge 0, \qquad 0 \le y_f^{k+1} \perp (M - \pi^{k+1}) \ge 0 \\
& 0 \le \mathbf{u}^{k+1} \perp (\mathbf{b}_2 - \mathbf{A}_2\mathbf{x} - \mathbf{B}_{2d}\mathbf{y}_d^{k+1} - \mathbf{B}_{2c}\mathbf{y}_c^{k+1} + \mathbf{I}\mathbf{y}_a^{k+1}) \ge 0 \\
& 0 \le \pi^{k+1} \perp (\mathbf{d}_{2c}\overline{\mathbf{y}}_c + \mathbf{d}_{2d}\overline{\mathbf{y}}_d - \mathbf{d}_{2d}\mathbf{y}_d^{k+1} - \mathbf{d}_{2c}\mathbf{y}_c^{k+1} + y_f) \ge 0
\end{aligned}
$$

□

Appendix 2: Analysis of Mixed Integer PBL in Example 3

The following formulation is the mixed integer **PBL** instance in Example 3.

$$
\begin{aligned}
\min \quad & -8x_1 - 6x_2 + \max(-25y_1 - 30y_2 + 2y_3 + 16y_4) && (55)\\
\text{s.t.} \quad & x_1 + x_2 \le 10,\ 2x_1 + 5x_2 \le 13,\ x_1, x_2 \ge 0 && (56)\\
& (y_1, y_2, y_3, y_4) \in \mathbf{S}(x_1, x_2) = \arg\min \Big\{ -10y_1 - 10y_2 - 10y_3 - 10y_4 : && (57)\\
& y_1 + y_2 + y_3 + y_4 \le 10 - x_1 - x_2,\ -y_1 + y_4 \le 0.8x_1 + 0.8x_2 && (58)\\
& y_2 + y_4 \le 4x_2,\ y_1, y_2, y_4 \ge 0,\ y_3 \in \mathbb{Z}_+ \Big\}. && (59)
\end{aligned}
$$

We first consider the lower level problem defined in (57-59) for any feasible (x_1, x_2). Because of the connection between its objective function and its first constraint, we note that an optimal solution can always be obtained by setting $y_1 = 10 - x_1 - x_2$ and the associated optimal value is $-10(10 - x_1 - x_2)$. Hence, the three-level **PBL** in (55-59) can be simplified as the next bilevel problem

$$
\begin{aligned}
\min \quad & -8x_1 - 6x_2 + \eta\\
\text{s.t.} \quad & x_1 + x_2 \le 10,\ 2x_1 + 5x_2 \le 13,\ x_1, x_2 \ge 0\\
& \eta \ge \max \Big\{ -25y_1 - 30y_2 + 2y_3 + 16y_4 :\ y_1 + y_2 + y_3 + y_4 = 10 - x_1 - x_2\\
& -y_1 + y_4 \le 0.8x_1 + 0.8x_2,\ y_2 + y_4 \le 4x_2,\ y_1, y_2, y_4 \ge 0,\ y_3 \in \mathbb{Z}_+ \Big\}.
\end{aligned}
$$

Moreover, comparing y_1 and y_2's roles in the objective function and constraints, it follows that $y_2 = 0$ in any optimal solution, and constraints in the lower level problem can be further modified to

$$
\begin{aligned}
& \eta \ge \max \Big\{ -25y_1 + 2y_3 + 16y_4 :\ y_1 + y_3 + y_4 = 10 - x_1 - x_2\\
& -y_1 + y_4 \le 0.8x_1 + 0.8x_2,\ y_4 \le 4x_2,\ y_1, y_4 \ge 0,\ y_3 \in \mathbb{Z}_+ \Big\}\\
& = \max \Big\{ 2y_3 + \omega :\\
& y_3 \le 10 - x_1 - x_2,\ y_3 \in \mathbb{Z}_+\\
& \omega = \max\{-25y_1 + 16y_4 :\ y_1 + y_4 = 10 - x_1 - x_2 - y_3,\ -y_1 + y_4 \le 0.8x_1 + 0.8x_2\\
& y_4 \le 4x_2,\ y_1, y_4 \ge 0\} \Big\}.
\end{aligned}
$$

We separate y_3 from other variables as its possible values can be enumerated. In case that $y_3 \le 10 - x_1 - x_2$ is violated for a particular y'_3, we can set the corresponding objective function value to $-\infty$ by convention. Otherwise, the linear program problem defining ω is feasible. The feasible set and the optimal solution is displayed in Figure 2, where two critical points A and B are intersections of the first and second constraints, and the first and third constraints, respectively. Their analytical expressions are $(y_1, y_4)_A = (5 - 0.9x_1 - 0.9x_2 - 0.5y_3, 5 - 0.1x_1 - 0.1x_2 - 0.5y_3)$, and $(y_1, y_4)_B = (10 - x_1 - 5x_2 - y_3, 4x_2)$. The one with the smaller y_4 coordinate, along with the line segment $Q - P$, defines the feasible set, which is the bold line in Figure 2. Noting that the dot line represents an iso-profit line and the arrow represents the increasing direction to maximize ω, it is clear that either point A or B will be the optimal solution, which leads to

$$\omega = -45 + 20.9x_1 + 20.9x_2 + 4.5y_3, \text{ or } \omega = -250 + 25x_1 + 189x_2 + 25y_3.$$

In either case, we can, without changing y_3's feasibility (or infeasibility) status, always decrease

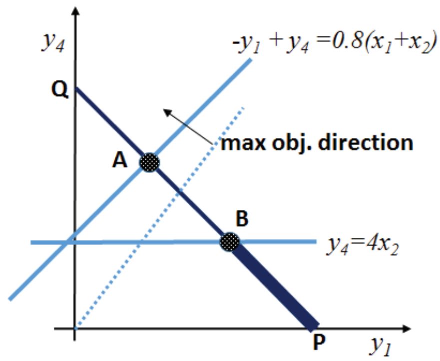

Figure 2: Solution of (y_1, y_4)

x_2 and add the decreased quantity to x_1 to achieve the same or a better (i.e., smaller, in terms of the upper level DM) value for ω. So, it is sufficient to fix $x_2 = 0$ for the upper level, which actually renders point B as the optimal solution and $\omega = -250 + 25x_1 + 25y_3$. As a result, the original **PBL** in (55-59) further reduces to the following problem

$$\begin{aligned} \min \quad & -8x_1 + \eta \\ \text{s.t.} \quad & x_1 \leq 10, \ 2x_1 \leq 13, \ x_1 \geq 0 \\ & \eta \geq \begin{cases} 27y_3 - 250 + 25x_1, & \text{if } x_1 + y_3 \leq 10 \\ -\infty, & \text{otherwise} \end{cases} \\ & \text{for } y_3 = 0, \dots, 10. \end{aligned}$$

With this formulation, the interaction between x_1 and y_3 and the resulting objective function value can be easily analyzed. Let $x_1 = k + \epsilon$ with $k \in 0, \dots, 6$ and $0 < \epsilon \leq 1$ (noting that $\epsilon \leq 0.5$ if $k = 6$). Then, feasible values for y_3 are $0, \dots, 10 - k - 1$ and the associated objective function value is $-250 + 17x_1 + 27\max\{0, \dots, 10 - k - 1\} = -7 + 17x_1 - 27k = -10k - 7 + 17\epsilon$. Clearly, the best possible objective function value is achieved when $k = 6$. Moreover, the smaller ϵ, the better (i.e., less) objective function value, which causes the infimum to be -67. However, this infimum cannot be achieved. When $\epsilon > 0$, feasible values for y_3 are $\{0, 1, 2, 3\}$, while when $\epsilon = 0$, feasible values for y_3 become $\{0, 1, 2, 3, 4\}$. For the latter case, the objective function value equals $-250 + 17 * 6 + 27\max\{0, 1, 2, 3, 4\} = -40$. With the aforementioned analysis, it can be confirmed that this mixed integer **PBL** does not have any exact solution.

Note that when $\epsilon = 0.005$, the associated ϵ-optimal value is -66.915, which matches the pessimistic result computed for Example 3 in Section 4.1.

Bibliography

[1] Random instances of – a practical scheme to compute pessimistic bilevel optimization problem. Available at `http://www.pitt.edu/~bzeng/`.

[2] A. Aboussoror, S. Adly, and F.E. Saissi. Strong-weak nonlinear bilevel problems: Existence of solutions in a sequential setting. *Set-Valued and Variational Analysis*, pages 1–20, 2016.

[3] A. Aboussoror and P. Loridan. Strong-weak stackelberg problems in finite dimensional spaces. *Serdica Mathematical Journal*, 21(2):151p–170p, 1995.

[4] A. Aboussoror and A. Mansouri. Weak linear bilevel programming problems: existence of solutions via a penalty method. *Journal of Mathematical Analysis and Applications*, 304(1):399–408, 2005.

[5] M. Apaydin, B. Zeng, and X. Qian. A reliable alternative of optknock for desirable mutant microbial strains. In *Biomedical and Health Informatics (BHI), 2016 IEEE-EMBS International Conference on*, pages 573–576. IEEE, 2016.

[6] C. Audet, P. Hansen, B. Jaumard, and G. Savard. Links between linear bilevel and mixed 0–1 programming problems. *Journal of Optimization Theory and Applications*, 93(2):273–300, 1997.

[7] J. Bard. *Practical bilevel optimization: algorithms and applications*. Springer, 1998.

[8] J. Bard, J. Plummer, and J. Claude Sourie. A bilevel programming approach to determining tax credits for biofuel production. *European Journal of Operational Research*, 120(1):30–46, 2000.

[9] M. Bergounioux and M. Haddou. A regularization method for ill-posed bilevel optimization problems. *RAIRO-Operations Research*, 40(1):19–35, 2006.

[10] W.F. Bialas and M.H. Karwan. Two-level linear programming. *Management Science*, 30(8):1004–1020, 1984.

[11] J. Bracken and J. McGill. Mathematical programs with optimization problems in the constraints. *Operations Research*, 21(1):37–44, 1973.

[12] L. Brotcorne, M. Labbé, P. Marcotte, and G. Savard. A bilevel model for toll optimization on a multicommodity transportation network. *Transportation Science*, 35(4):345–358, 2001.

[13] L. Brotcorne, M. Labbé, P. Marcotte, and G. Savard. Joint design and pricing on a network. *Operations Research*, 56(5):1104–1115, 2008.

[14] A. Burgard, P. Pharkya, and C. Maranas. Optknock: A bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnology and Bioengineering*, 84(6):647–657, 2003.

[15] H. Calvete and C. Galé. Linear bilevel multi-follower programming with independent followers. *Journal of Global Optimization*, 39(3):409–417, 2007.

[16] D. Cao and L. Leung. A partial cooperation model for non-unique linear two-level decision problems. *European Journal of Operational Research*, 140(1):134–141, 2002.

[17] R. Cassidy, M. Kirby, and W. Raike. Efficient distribution of resources through three levels of government. *Management Science*, 17(8):462–473, 1971.

[18] M Červinka, C Matonoha, and Jirí V Outrata. On the computation of relaxed pessimistic solutions to MPECs. *Optimization Methods and Software*, 28(1):186–206, 2013.

[19] B. Colson, P. Marcotte, and G. Savard. An overview of bilevel optimization. *Annals of Operations Research*, 153(1):235–256, 2007.

[20] J.-P. Côté, P. Marcotte, and G. Savard. A bilevel modelling approach to pricing and fare optimisation in the airline industry. *Journal of Revenue and Pricing Management*, 2(1):23–36, 2003.

[21] S. Dassanayaka. *Methods of variational analysis in pessimistic bilevel programming.* PhD thesis, Wayne State University, 2010.

[22] S. Dempe. *Foundations of bilevel programming.* Springer, 2002.

[23] S. Dempe, G. Luo, and S. Franke. Pessimistic bilevel linear optimization. Technical report, TU Bergakademie Freiberg, Fakultät für Mathematik und Informatik, 2016.

[24] S. Dempe, B. S. Mordukhovich, and A. B. Zemkoho. Necessary optimality conditions in pessimistic bilevel programming. *Optimization*, 63(4):505–533, 2014.

[25] A. Frangioni. On a new class of bilevel programming problems and its use for reformulating mixed integer problems. *European Journal of Operational Research*, 82(3):615–646, 1995.

[26] L. Garcés, A. Conejo, R. García-Bertrand, and R. Romero. A bilevel approach to transmission expansion planning within a market environment. *Power Systems, IEEE Transactions on*, 24(3):1513–1522, 2009.

[27] X. Hu and D. Ralph. Using EPECs to model bilevel games in restructured electricity markets with locational prices. *Operations Research*, 55(5):809–827, 2007.

[28] M. Köppe, M. Queyranne, and C. Ryan. Parametric integer programming algorithm for bilevel mixed integer programs. *Journal of Optimization Theory and Applications*, 146(1):137–150, 2010.

[29] M. B. Lignola and J. Morgan. Approximating security values of minsup problems with quasi-variational inequality constraints. *Pacific Journal of Optimization*, 10(4):749–765, 2014.

[30] M. B. Lignola and J. Morgan. Inner regularizations and viscosity solutions for pessimistic bilevel optimization problems. Technical report, Centre for Studies in Economics and Finance (CSEF), University of Naples, Italy, 2016.

[31] M. B. Lignola and J. Morgan. A method to bypass the lack of solutions in minsup problems under quasi-equilibrium constraints. *Optimization Letters*, 10(4):833–846, 2016.

[32] P. Loridan and J. Morgan. Weak via strong Stackelberg problem: New results. *Journal of Global Optimization*, 8(3):263–287, 1996.

[33] T. Magnanti and R. Wong. Accelerating Benders decomposition: Algorithmic enhancement and model selection criteria. *Operations Research*, 29(3):464–484, 1981.

[34] A. Malyshev and A. Strekalovsky. On global search for pessimistic solution in bilevel problems. In *Proceedings of the Toulouse Global Optimization Workshop*, pages 77–80, 2010.

[35] P. Marcotte. Network design problem with congestion effects: A case of bilevel programming. *Mathematical Programming*, 34(2):142–162, 1986.

[36] R. Nair and E. Miller-Hooks. Equilibrium network design of shared-vehicle systems. *European Journal of Operational Research*, 235(1):47–61, 2014.

[37] S. Ren, B. Zeng, and X. Qian. Adaptive bi-level programming for optimal gene knockouts for targeted overproduction under phenotypic constraints. *BMC Bioinformatics*, 14(Suppl 2):S17, 2013.

[38] A. Tsoukalas, W. Wiesemann, and B. Rustem. Global optimisation of pessimistic bi-level problems. *Lectures on global optimization*, 55:215–243, 2009.

[39] H. Von Stackelberg. *The theory of the market economy.* Oxford University Press, 1952.

[40] W. Wiesemann, A. Tsoukalas, P.-M. Kleniati, and B. Rustem. Pessimistic bilevel optimization. *SIAM Journal on Optimization*, 23(1):353–380, 2013.

[41] M Hosein Zare, Osman Y Özaltın, and Oleg A Prokopyev. On a class of bilevel linear mixed-integer programs in adversarial settings. *Journal of Global Optimization*, 71(1):91–113, 2018.

[42] B. Zeng and Y. An. Solving bilevel mixed integer program by reformulations and decomposition. Technical report, University of South Florida, available in *optimization-online.org*, 2014.

[43] Y. Zheng, D. Fang, and Z. Wan. A solution approach to the weak linear bilevel programming problems. *Optimization*, 65(7):1437–1449, 2016.

[44] Y. Zheng, Z. Wan, S. Jia, and G. Wang. A new method for strong-weak linear bilevel programming problem. *Journal Of Industrial and Management Optimization*, 11(2):529–547, 2015.