# On Coupling Constraints in Linear Bilevel Optimization

Dorothee Henke, Henri Lefebvre, Martin Schmidt, Johannes Thürauf

Abstract. It is well-known that coupling constraints in linear bilevel optimization can lead to disconnected feasible sets, which is not possible without coupling constraints. However, there is no difference between linear bilevel problems with and without coupling constraints w.r.t. their complexity-theoretical hardness. In this note, we prove that, although there is a clear difference between these two classes of problems in terms of their feasible sets, the classes are equivalent on the level of optimal solutions. To this end, given a general linear bilevel problem with coupling constraints, we derive a respective problem without coupling constraints and prove that it has the same optimal solutions (when projected back to the original variable space).

## 1. Introduction

The research interest in bilevel optimization problems increased significantly over the last years and decades; see, e.g., Dempe and Zemkoho (2020) for a recent overview. However, and although serious advances have been made both w.r.t. theoretical aspects and algorithmic developments, there are still open questions. This is even the case for linear bilevel optimization problems that we consider in this note and that are given by[1]

$$\min_{x \in X, y} \quad c^\top x + d^\top y \tag{1a}$$

$$\text{s.t.} \quad Ax + By \geq a, \tag{1b}$$

$$y \in S(x), \tag{1c}$$

where $S(x)$ is the set of optimal solutions to the $x$-parameterized lower-level problem

$$\min_{y} \quad f^\top y \quad \text{s.t.} \quad Cx + Dy \geq b \tag{2}$$

and all variables are assumed to be continuous. In particular, we consider the optimistic linear bilevel problem, i.e., the leader is able to choose the $y$ that is the best w.r.t. the upper-level objective function if there are multiple optimal solutions to the follower's problem. Moreover, Problem (1) contains coupling constraints in (1b), i.e., upper-level constraints that explicitly depend on the lower-level variables. These coupling constraints are the main topic of this note. Instead, a bilevel problem with $B = 0$ does not have any coupling constraints. Throughout the remainder of the paper, we assume that the bilevel problem (1) is solvable, that $X \subset \mathbb{R}^n$ is a given polyhedron, and that all vectors and matrices have rational entries. Note that the linear bilevel problem always has an optimal solution that is at a vertex of the polyhedron obtained by intersecting $X$, the constraints in (1b), and in (2); see, e.g., Bard (1998).

*Date*: June 17, 2024.
2020 *Mathematics Subject Classification.* 90Cxx.
*Key words and phrases.* Bilevel optimization, Coupling constraints.

[1]For the ease of presentation, we omit stating dimensions of matrices and vectors.

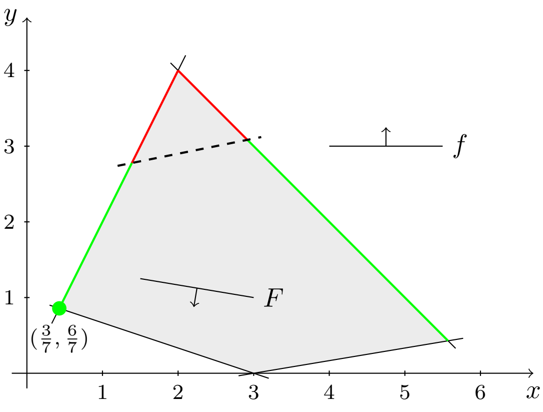


FIGURE 1. Visualization of Problem (3); mainly taken from Kleinert (2021).

For motivating our main research question, we briefly discuss coupling constraints in the following both w.r.t. their impact on the geometry of the bilevel feasible set and their impact on complexity.

1.1. **Geometry of the Feasible Set.** As an example, we consider the following linear bilevel problem taken from Kleinert (2021):

$$\min_{x,y} \quad F(x, y) = x + 6y \tag{3a}$$

$$\text{s.t.} \quad -x + 5y \leq 12.5, \tag{3b}$$

$$y \in S(x), \tag{3c}$$

with $S(x)$ being the set of optimal solutions to the lower-level problem

$$\min_{y} \quad f(x, y) = -y \tag{4a}$$

$$\text{s.t.} \quad 2x - y \geq 0, \tag{4b}$$

$$-x - y \geq -6, \tag{4c}$$

$$-x + 6y \geq -3, \tag{4d}$$

$$x + 3y \geq 3. \tag{4e}$$

Both levels are linear optimization problems and all variables are continuous.

The problem is visualized in Figure 1. The gray area is the set of points that satisfy the lower-level constraints. The points above the dashed line are infeasible w.r.t. the coupling constraint (3b). Due to the optimization direction of the follower, the green faces denote the bilevel feasible set. Note that the two red faces are not part of the bilevel feasible set since, for the respective $x$-values, the optimal replies $y$ by the follower violate the coupling constraint.

Hence, the example shows that the usage of coupling constraints makes it possible to model bilevel feasible sets that are disconnected. Theorem 3.3 in Benson (1989) states that the feasible set of a linear bilevel problem with $B = 0$ is always connected, which means that disconnected sets can only be modeled using coupling constraints—an aspect that gained some prominence since it allows to model mixed-binary linear problems using purely continuous linear bilevel models; see, e.g., Section 3 in Vicente et al. (1996) and Section 3.1 in Audet et al. (1997). Consequently, it seems to be the case that having coupling constraints introduces larger modeling capabilities.

1.2. **Complexity.** Already Jeroslow (1985) showed that linear bilevel problems are NP-hard. An alternative proof using a reduction from KNAPSACK can be found in Ben-Ayed and Blair (1990). Moreover, Hansen et al. (1992) proved that linear bilevel problems are strongly NP-hard by a reduction from KERNEL; see Problem GT57 in Garey and Johnson (1979). In particular, they even showed that the special case of min-max problems without coupling constraints is strongly NP-hard. The same holds true for the reduction from 3-SAT shown in Marcotte and Savard (2005). In Vicente et al. (1994) it is further shown that even checking local optimality of a given point is strongly NP-hard (again via 3-SAT). As before, the authors do not require coupling constraints to achieve this hardness result; see the linear bilevel problems in the proofs of Theorems 5.1 and 5.2 in Vicente et al. (1994). Hence, coupling constraints are not required to make linear bilevel problems strongly NP-hard. Moreover, Buchheim (2023) recently showed that the decision versions of linear bilevel problems with coupling constraints are in NP, which again implies that linear bilevel problems with coupling constraints are not harder than those without.

1.3. **Research Question.** Combining the two last discussions, we can summarize the following two conclusions:

(i) Using coupling constraints in linear bilevel problems allows for modeling a richer class of feasible sets.
(ii) These stronger modeling capabilities do not result in any change of the hardness of the resulting problem in terms of complexity theory.

Due to (ii), the question arises if we "really" increase modeling capabilities by using coupling constraints. While this is the case w.r.t. (dis-)connectedness of feasible sets, we prove that there is no difference on the level of optimal solutions. To do this, in the next section, given a linear bilevel problem with coupling constraints, we derive a linear bilevel problem without coupling constraints that has the same set of optimal solutions.

## 2. Exact Penalization of Coupling Constraints

In this section, we show that the bilevel problem (1) with coupling constraints can be reformulated as a bilevel problem without coupling constraints. To do this, we show that the violation of the coupling constraints (1b) can be exactly penalized in the objective function of the leader. This result is surprising for at least two reasons. First, we just saw that the feasible region of (1) may be disconnected and nonconvex. In this case, standard Lagrangian duality theory is usually limited and not as strong as in the convex case. Second, the resulting problem, i.e., after penalization, has a smooth objective function and no coupling constraint. Thus, it differs from classic exact penalty methods, which often require a nonsmooth penalty function.

Our key idea is to reformulate Problem (1) in a way so that the follower measures the violation of the coupling constraints directly in the lower-level problem, while the leader enforces that this violation is zero. Doing so leads to the following bilevel problem which contains a scalar and very simple coupling constraint:

$$\min_{x,y,\varepsilon} \quad c^\top x + d^\top y \tag{5a}$$

$$\text{s.t.} \quad x \in X, \tag{5b}$$

$$\varepsilon = 0, \tag{5c}$$

$$(y, \varepsilon) \in \tilde{S}(x). \tag{5d}$$

Here, $\tilde{S}(x)$ is the set of optimal solutions to the $x$-parameterized lower-level problem

$$\min_{y,\varepsilon} \quad f^\top y \tag{6a}$$
$$\text{s.t.} \quad Ax + By + \varepsilon e \geq a, \tag{6b}$$
$$Cx + Dy \geq b, \tag{6c}$$
$$\varepsilon \geq 0, \tag{6d}$$

where $e$ is the vector of ones. Essentially, $\varepsilon$ is an additional variable of the follower that measures the violation of the coupling constraints. The newly introduced coupling constraint (5c) enforces that it equals zero. Most importantly, we have $S(x) = \operatorname{proj}_y(\tilde{S}(x))$ for all leader's decisions $x$. In the next lemma, we show that Problem (5) is, indeed, equivalent to Problem (1).

**Lemma 2.1.** *For every bilevel feasible point $(x, y)$ of Problem (1), the point $(x, y, 0)$ is bilevel feasible for Problem (5) with the same objective value. For every bilevel feasible point $(x, y, \varepsilon)$ of Problem (5), the point $(x, y)$ is bilevel feasible for Problem (1) with the same objective value.*

*Proof.* Let $(x, y)$ be a bilevel feasible point of Problem (1). Then, $(x, y, 0)$ satisfies the upper- and lower-level constraints of Problem (5). The point $(y, 0)$ is optimal for the lower-level problem (6) since $\varepsilon$ is not part of the lower-level objective function and $f^\top y$ is minimal regarding Constraints (6c) due to the bilevel feasibility of $(x, y)$ for Problem (1).

Let now $(x, y, \varepsilon)$ be a bilevel feasible point of Problem (5). Then, the point $(x, y)$ satisfies the upper- and lower-level constraints of Problem (1) because $\varepsilon = 0$ holds. The point $y$ is optimal for the lower-level problem (2) since every optimal solution $\tilde{y}$ to the $x$-parameterized lower-level problem (2) can be extended to the feasible point $(\tilde{y}, \max\{\|a - Ax - By\|_\infty, 0\})$ of the $x$-parameterized lower-level problem (6) with the same objective value.

We finally note that Problem (1) and Problem (5) have the same upper-level objective functions, which proves the claim. $\square$

We now penalize the single coupling constraint (5c) of Problem (5) to obtain a bilevel problem without coupling constraints. Moreover, we show that there is a polynomial-sized (in the bit-encoding length of the original problem's data) penalty parameter so that this formulation is equivalent in terms of optimal solutions.

**Theorem 2.2.** *There is a polynomial-sized parameter $\kappa > 0$ so that the bilevel problem (without coupling constraints)*

$$\min_{x,y,\varepsilon} \quad c^\top x + d^\top y + \kappa\varepsilon \tag{7a}$$
$$\text{s.t.} \quad x \in X, \ (y, \varepsilon) \in \tilde{S}(x), \tag{7b}$$

*has the same set of optimal solutions as Problem (5). Again, $\tilde{S}(x)$ is the set of optimal solutions to the $x$-parameterized lower-level problem (6).*

The idea of the proof is as follows. First, we derive a single-level reformulation of the bilevel problem (5), using the KKT conditions of the follower's problem (6). Second, we apply results from augmented Lagrangian duality theory for mixed-integer linear problems to show that a polynomial-sized exact penalization parameter exists. Finally, we show that the resulting mixed-integer linear program is nothing but the KKT reformulation of Problem (7).

*Proof.* Since the lower-level problem (6) of Problem (5) is a linear program, we can replace it with its KKT conditions (Dempe and Dutta 2012), leading to

$$\min_{x,y,\varepsilon} \quad c^\top x + d^\top y \tag{8a}$$
$$\text{s.t.} \quad x \in X,\ \varepsilon = 0, \tag{8b}$$
$$Ax + By + \varepsilon e \geq a,\ Cx + Dy \geq b,\ \varepsilon \geq 0, \tag{8c}$$
$$B^\top \lambda + D^\top \mu = f,\ e^\top \lambda + \eta = 0, \tag{8d}$$
$$\lambda, \mu, \eta \geq 0, \tag{8e}$$
$$\lambda^\top (Ax + By + \varepsilon e - a) = 0,\ \mu^\top (Cx + Dy - b) = 0,\ \eta\varepsilon = 0. \tag{8f}$$

Using additional binary variables $z^\lambda, z^\mu, z^\eta$, and a sufficiently large big-$M$ value, we can reformulate the complementarity constraints (8f) as the mixed-integer linear constraints

$$\lambda \leq (e - z^\lambda)M,\ \mu \leq (e - z^\mu)M,\ \eta \leq (1 - z^\eta)M, \tag{9a}$$
$$Ax + By + \varepsilon e - a \leq z^\lambda M,\ Cx + Dy - b \leq z^\mu M,\ \varepsilon \leq z^\eta M. \tag{9b}$$

It is shown in Buchheim (2023) that a valid and polynomial-sized value for $M$ can be computed in polynomial time. Thus, the resulting problem is a mixed-integer linear program whose input data is polynomial-sized in the bit-encoding length of the input data of Problem (5). We now penalize the constraint $\varepsilon = 0$ in the $\ell_\infty$-sense and obtain

$$\min_{x,y,\varepsilon} \quad c^\top x + d^\top y + \kappa\varepsilon \tag{10a}$$
$$\text{s.t.} \quad x \in X, \tag{10b}$$
$$Ax + By + \varepsilon e \geq a,\ Cx + Dy \geq b,\ \varepsilon \geq 0, \tag{10c}$$
$$B^\top \lambda + D^\top \mu = f,\ e^\top \lambda + \eta = 0, \tag{10d}$$
$$\lambda, \mu, \eta \geq 0,\ (9). \tag{10e}$$

The existence of a finite and exact value for the penalty parameter $\kappa$ is guaranteed by Theorem 4 of Feizollahi et al. (2016), which states that the duality gap for the augmented Lagrangian dual of a solvable (mixed-integer) linear optimization problem can be closed by using a norm as the augmenting function and a sufficiently large but finite penalty parameter. Moreover, Proposition 1 of Feizollahi et al. (2016) ensures that the sets of optimal solutions of (10) and (8) are the same. Gu et al. (2020) show in Theorem 22 that the penalty parameter can be chosen to be of polynomial size in case of the $\ell_\infty$-norm. Finally, Problem (10) is the KKT reformulation of the bilevel problem (7). □

Note that the results from the literature that we use to conclude the proof are not constructive but pure existence results. Hence, we also do not state an explicit formula or big-$O$ expression for the value or the size of the parameter $\kappa$ here.

**Corollary 2.3.** *There is a polynomial-sized penalty parameter $\kappa > 0$ so that the following holds. For every bilevel optimal solution $(x, y)$ to Problem (1), the point $(x, y, 0)$ is bilevel optimal to Problem (7) with the same objective value. For every bilevel optimal point $(x, y, \varepsilon)$ of Problem (7), the point $(x, y)$ is bilevel optimal to Problem (1) with the same objective value.*

*Remark* 2.4. (a) Let us note that the reformulations of mixed-integer linear programs as linear bilevel problems presented in Vicente et al. (1996) and Audet et al. (1997) use two strategies for enforcing that continuous lower-level variables are binary. One is based on coupling constraints of the form $v = 0$, where $v$ is an auxiliary lower-level variable, and the other one corresponds to an exact penalization of these coupling constraints and therefore requires no coupling constraints in the final model. Our approach to penalize general coupling constraints in linear bilevel problems uses the same idea in the final step.

(b) The derivations in this section suggest that one could potentially extend the class of bilevel problems to which the results can be applied. First, the results in Gu et al. (2020) are valid for convex mixed-integer quadratic programs (MIQPs). Hence, this would allow for convex-quadratic objective functions of the leader and further integrality constraints in $X$. Moreover, we can also allow for convex-quadratic (but still continuous) problems in the lower level since their KKT conditions lead to polyhedral constraints in the KKT reformulation (8). However, the big-$M$s from Buchheim (2023) cannot be used directly anymore. Given that valid big-$M$s can also be found for quadratic programs in the lower level, the most general class of bilevel problems to which our results could be applied are those with convex MIQPs for the leader and convex QPs for the follower.

(c) Due to the finiteness of the penalty parameter in Corollary 2.3, the original bilevel problem with coupling constraints can also be solved by a finite sequence of bilevel problems without coupling constraints if we follow standard ideas of penalty methods.

## 3. Discussion

It has been known for at least 25 years that coupling constraints in linear bilevel problems can lead to disconnected feasible sets and that this is not possible without coupling constraints. However, we prove that there is no difference between these two types of linear bilevel problems on the level of optimal solutions. While, on the one hand, this closes a gap in the literature on linear bilevel optimization, it, on the other hand, also has some practical consequences. Many theoretical statements in the literature on bilevel optimization are made and shown for problems without coupling constraints—either simply for the ease of presentation or due to a lack of a proof for the case with coupling constraints. This note now allows for carrying over some of these results by transforming the given problem having coupling constraints into one without. We also point to the open question of efficient computation of the penalty parameter required in the last section. While we prove that it is polynomial-sized (in the bit-encoding length of the data of the given problem), the question on how to compute it in polynomial time is still open. Finally, the analogue question regarding the impact of coupling constraints is, to the best of our knowledge, still open for pessimistic bilevel problems. We are rather convinced that the approach applied in this paper cannot be directly transferred to the pessimistic case. Hence, the latter is an interesting and important research question.

## Acknowledgements

The second and third author acknowledge the support by the German Bundesministerium für Bildung und Forschung within the project "RODES" (Förderkennzeichen 05M22UTB).

## Availability of data and materials

No data or code were generated or used during the study.

## References


Audet, C., P. Hansen, B. Jaumard, and G. Savard (1997). "Links between linear bilevel and mixed 0–1 programming problems." In: *Journal of Optimization Theory and Applications* 93.2, pp. 273–300. DOI: 10.1023/A:1022645805569.

Bard, J. F. (1998). *Practical bilevel optimization: algorithms and applications*. Vol. 30. Springer Science & Business Media. DOI: 10.1007/978-1-4757-2836-1.

Ben-Ayed, O. and C. E. Blair (1990). "Computational Difficulties of Bilevel Linear Programming." In: *Operations Research* 38.3, pp. 556–560. DOI: 10.1287/opre.38.3.556.

Benson, H. P. (1989). "On the structure and properties of a linear multilevel programming problem." In: *Journal of Optimization Theory and Applications* 60.3, pp. 353–373. DOI: 10.1007/BF00940342.

Buchheim, C. (2023). "Bilevel linear optimization belongs to NP and admits polynomial-size KKT-based reformulations." In: *Operations Research Letters* 51.6, pp. 618–622. DOI: 10.1016/j.orl.2023.10.006.

Dempe, S. and J. Dutta (2012). "Is bilevel programming a special case of a mathematical program with complementarity constraints?" In: *Mathematical Programming* 131.1-2, pp. 37–48. DOI: 10.1007/s10107-010-0342-1.

Dempe, S. and A. Zemkoho, eds. (2020). *Bilevel Optimization. Advances and Next Challenges*. Springer Optimization and Its Applications 161. Cham: Springer. DOI: 10.1007/978-3-030-52119-6.

Feizollahi, M. J., S. Ahmed, and A. Sun (2016). "Exact augmented Lagrangian duality for mixed integer linear programming." In: *Mathematical Programming* 161.1-2, pp. 365–387. DOI: 10.1007/s10107-016-1012-8.

Garey, M. R. and D. S. Johnson (1979). *Computers and Intractability*.

Gu, X., S. Ahmed, and S. S. Dey (2020). "Exact Augmented Lagrangian Duality for Mixed Integer Quadratic Programming." In: *SIAM Journal on Optimization* 30.1, pp. 781–797. DOI: 10.1137/19m1271695.

Hansen, P., B. Jaumard, and G. Savard (1992). "New branch-and-bound rules for linear bilevel programming." In: *SIAM Journal on scientific and Statistical Computing* 13.5, pp. 1194–1217. DOI: 10.1137/0913069.

Jeroslow, R. G. (1985). "The polynomial hierarchy and a simple model for competitive analysis." In: *Mathematical Programming* 32.2, pp. 146–164. DOI: 10.1007/BF01586088.

Kleinert, T. (2021). "Algorithms for Mixed-Integer Bilevel Problems with Convex Followers." PhD thesis. URL: https://opus4.kobv.de/opus4-trr154/frontdoor/index/index/docId/383.

Marcotte, P. and G. Savard (2005). "Bilevel Programming: A Combinatorial Perspective." In: *Graph Theory and Combinatorial Optimization*. Ed. by D. Avis, A. Hertz, and O. Marcotte. Boston, MA: Springer US, pp. 191–217. DOI: 10.1007/0-387-25592-3_7.

Vicente, L., G. Savard, and J. Júdice (1994). "Descent approaches for quadratic bilevel programming." In: *Journal of Optimization Theory and Applications* 81.2, pp. 379–399. DOI: 10.1007/BF02191670.

– (1996). "Discrete linear bilevel programming problem." In: *Journal of Optimization Theory and Applications* 89.3, pp. 597–614. DOI: 10.1007/BF02275351.