Dimensionality Reduction in Bilevel Linear Programming

Eneko C. Clemente[*,✉] and Oleg A. Prokopyev[*]

Abstract. We consider bilevel programs that involve a leader, who first commits to a mixed-integer decision, and a follower, who observes this decision and then responds rationally by solving a linear program (LP). Standard approaches often reformulate these bilevel optimization problems as single-level mixed-integer programs by exploiting the follower's LP optimality conditions. These reformulations introduce either complementarity or quadratic constraints, which become increasingly numerous and burdensome as the number of constraints in the follower's problem grows. This growth is arguably the main computational bottleneck for solving these problems at scale. Accordingly, we propose a dimensionality-reduction approach that projects the follower's constraints into a compressed representation of smaller size. Building on this representation, we develop a surrogate duality theory, from which we derive a feasibility-based lower bound and a surrogate-dual upper bound on the leader's optimal objective function value. We also demonstrate the existence of a family of projections that ensures provable approximation guarantees. Rather than searching for those projections, we employ linear sketching techniques, offering probabilistic approximation guarantees and highlighting trade-offs between dimensionality reduction and approximation quality. Finally, our preliminary numerical experiments illustrate the computational promise of dimensionality reduction techniques in bilevel programming.

1 Introduction

Bilevel programming models hierarchical decision-making processes involving two decision makers: a *leader*, who acts first, and a *follower*, who observes and responds to the leader's decisions (Dempe et al. 2015). Both decision makers are assumed to be rational, each solving their own optimization problem with linear objective functions. The leader accounts for the follower's rationality and anticipates the follower's optimal response in its own decision-making process. In particular, we adopt the leader's perspective and focus on settings in which the leader's decision variables may be both integer and continuous, while the follower's decisions are restricted to continuous variables.

Bilevel programs arise in a broad range of practical applications, from defense and security planning, particularly through the use of interdiction models (Brown et al. 2006; Gutin et al. 2015), to network design (Tawfik and Limbourg 2019; Q. Li et al. 2023), energy market (Aravena et al. 2021; He et al. 2024), pricing and revenue management (Labbé et al. 1998; Kuiteing et al. 2017), and facil-

[*]Department of Business Administration, University of Zurich, Switzerland. ✉ eneko.clemente@business.uzh.ch

ity location (Dan and Marcotte 2019; Lin et al. 2024). For a recent overview of bilevel programming and its applications, we refer the reader to the surveys by Kleinert et al. (2021) and Beck et al. (2023).

In this paper, we study a broad class of bilevel optimization problems with linear programs (LPs) at the lower level. Formally, these problems, further referred to as simply BLPs, are given by:

$$\begin{aligned} z^* := \max_{\mathbf{x},\mathbf{y}^*(\mathbf{x})} \ & \mathbf{a}^\top \mathbf{x} + \mathbf{d}^\top \mathbf{y}^*(\mathbf{x}) \\ \text{s.t.} \ & \mathbf{G}\mathbf{x} + \mathbf{H}\mathbf{y}^*(\mathbf{x}) \leq \mathbf{h}, \\ & \mathbf{y}^*(\mathbf{x}) \in \underset{\mathbf{y}\in\mathcal{Y}(\mathbf{x})}{\operatorname{argmin}} \ \mathbf{c}^\top \mathbf{y}, \\ & \mathbf{x} \in \{0,1\}^{n_1} \times \mathbb{R}^{n_2}_+, \ \mathbf{y}^*(\mathbf{x}) \in \mathbb{R}^m_+, \end{aligned} \tag{1}$$

where, given $n := n_1 + n_2$, we assume $\mathbf{a} \in \mathbb{Q}^n$, $\mathbf{d} \in \mathbb{Q}^m$, $\mathbf{c} \in \mathbb{Q}^m$, $\mathbf{G} \in \mathbb{Q}^{p\times n}$, $\mathbf{H} \in \mathbb{Q}^{p\times m}$, and $\mathbf{h} \in \mathbb{Q}^p$. In this formulation, $\mathbf{x}$ corresponds to the *leader's decision*, while $\mathbf{y}^*(\mathbf{x})$ denotes the *follower's optimal decision*, or equivalently, the *follower's rational response*. We refer to (1) as the upper-level problem or the leader's problem, and use these terms interchangeably throughout the paper.

Follower's problem. The leader's decision $\mathbf{x}$ affects the responses available to the follower. This dependency is captured by the *follower's feasible region* $\mathcal{Y}(\mathbf{x})$ in (1) defined by a polyhedron:

$$\mathcal{Y}(\mathbf{x}) := \left\{\mathbf{y} \in \mathbb{R}^m_+ : \mathbf{F}\mathbf{y} = \mathbf{f} - \mathbf{L}\mathbf{x}\right\},$$

in the standard form, where $\mathbf{F} \in \mathbb{Q}^{q\times m}$, $\mathbf{L} \in \mathbb{Q}^{q\times n}$, and $\mathbf{f} \in \mathbb{Q}^q$. After observing the leader's decision, the follower rationally responds by solving a linear program parameterized by $\mathbf{x}$, given by:

$$\varphi(\mathbf{x}) := \min_{\mathbf{y}} \left\{\mathbf{c}^\top \mathbf{y} : \mathbf{y} \in \mathcal{Y}(\mathbf{x})\right\}, \tag{2}$$

where we denote by $\mathcal{R}(\mathbf{x})$ the set of all optimal solutions to (2), and $\varphi(\mathbf{x})$ is referred to as the *follower's value function.* Furthermore, we refer to (2) as the *lower-level problem* or the *follower's problem*, and use these terms interchangeably throughout the paper.

For some leader's decision $\mathbf{x}$, the follower's rational response $\mathbf{y}^*(\mathbf{x})$ may violate the leader's constraints, making $\mathbf{x}$ infeasible. Thus, the leader's problem (1) includes *coupling constraints* that link the follower's rational response to the leader's own feasibility conditions. Accordingly, we define:

$$\mathcal{X} := \left\{\mathbf{x} \in \{0,1\}^{n_1} \times \mathbb{R}^{n_2}_+ \ : \exists\, \mathbf{y}^*(\mathbf{x}) \in \mathcal{R}(\mathbf{x}) \text{ such that } \mathbf{G}\mathbf{x} + \mathbf{H}\mathbf{y}^*(\mathbf{x}) \leq \mathbf{h}\right\},$$

as the *leader's feasible set*, or the *upper-level feasible region.* Any decision $\mathbf{x} \in \mathcal{X}$ is referred to as a *leader's feasible decision.* Also, a pair of leader's and follower's decisions $(\mathbf{x}, \mathbf{y})$ is said to be *bilevel feasible* if $\mathbf{x} \in \{0,1\}^{n_1} \times \mathbb{R}^{n_2}$, $\mathbf{y} \in \mathcal{R}(\mathbf{x})$, and the coupling constraints $\mathbf{G}\mathbf{x} + \mathbf{H}\mathbf{y} \leq \mathbf{h}$ are satisfied.

Technical assumptions. For a given leader's feasible decision $\mathbf{x} \in \mathcal{X}$, the follower's problem (2) is an LP and may admit infinitely many optimal solutions. Hence, the leader's problem is not inherently well-defined, as its outcome may depend on which particular optimal solution the follower selects. In this study, we adopt the *optimistic* paradigm, under which the follower, when faced with multiple optimal solutions, selects the one most favorable to the leader. Formally, we require that:

$$\mathbf{y}^*(\mathbf{x}) \in \operatorname{argmax}\left\{\mathbf{d}^\top \mathbf{y} \,:\, \mathbf{c}^\top \mathbf{y} \leq \varphi(\mathbf{x}) \,,\, \mathbf{y} \in \mathcal{Y}(\mathbf{x}) \,,\, \mathbf{G}\mathbf{x} + \mathbf{H}\mathbf{y} \leq \mathbf{h}\right\}. \tag{3}$$

The optimistic paradigm is the most extensively studied variant in the literature (Kleinert et al. 2021). By contrast, in the pessimistic paradigm, the follower selects the optimal solution that is least favorable to the leader (Wiesemann et al. 2013). There also exist intermediate approaches which assume that the follower's response lies between these two extremes (Lagos and Prokopyev 2023).

Finally, the following assumptions are made throughout our study:

- **A1 :** $\mathcal{Y}(\mathbf{x}) \neq \emptyset \,,\, \forall \mathbf{x} \in \mathcal{X}$.
- **A2 :** $\exists \theta > 0$ such that $\|\mathbf{x}\|_1 \leq \theta \,,\, \forall \mathbf{x} \in \mathcal{X}$, and $\|\mathbf{y}\|_1 \leq \theta \,,\, \forall \mathbf{y} \in \mathcal{R}(\mathbf{x}) \,,\, \forall \mathbf{x} \in \mathcal{X}$.
- **A3 :** $\min\left\{\mathbf{c}^\top \mathbf{y}^*(\mathbf{x}) \,:\, \mathbf{y}^*(\mathbf{x}) \in \mathcal{R}(\mathbf{x}) \,,\, \mathbf{x} \in \mathcal{X}\right\} > 0$.

Assumptions **A1** and **A2** are relatively standard in bilevel programming (Kleinert et al. 2021); they ensure feasibility of the follower's problem for every leader's feasible decision and bound both feasible regions in terms of the ℓ_1-norm $\|\cdot\|_1$. We refer to the study by Rodrigues et al. (2025) for a discussion of the unbounded case. Assumption **A3** is also not particularly restrictive. Indeed, by Assumption **A2**, the follower's feasible set is bounded, which implies that the follower's objective function is bounded bellow by some constant $M \equiv M(\theta, \mathbf{c})$, where M may depend on the problem's dimension, as well as on θ and $\mathbf{c}$. Specifically, we have $\mathbf{c}^\top \mathbf{y} \geq M$ for all $\mathbf{y} \in \mathcal{Y}(\mathbf{x})$ and $\mathbf{x} \in \mathcal{X}$.

If $M \leq 0$, then one can introduce a follower's decision variable y_{m+1}, augment $\mathbf{y}$ to $(\mathbf{y}, y_{m+1})$, add $c_{m+1} := -M$ to the follower's cost vector $\mathbf{c}$, and, finally, enforce the constraint $y_{m+1} = 1$. These operations guarantee that the modified follower's problem satisfies Assumption **A3**. However, this transformation may influence the quality of the performance ratio when solving the leader's problem (1) approximately. Yet, the introduction of this auxiliary term M does not diminish the generality of our approach. Taken together, these three assumptions are mild and standard in the literature.

Value function reformulation. The leader's problem as described in (1) can be reformulated using the so-called *value-function* approach. In this formulation, the bilevel structure is ensured by explicitly incorporating the follower's value function as a constraint; see, for instance, the studies by Lozano and Smith (2017) and Tavaslıoğlu et al. (2019). The reformulated problem becomes:

$$
\begin{aligned}
z^* := \max_{\mathbf{x}, \mathbf{y}^*(\mathbf{x})} \ & \mathbf{a}^\top \mathbf{x} + \mathbf{d}^\top \mathbf{y}^*(\mathbf{x}) \\
\text{s.t.} \ & \mathbf{G}\mathbf{x} + \mathbf{H}\mathbf{y}^*(\mathbf{x}) \leq \mathbf{h}\,,\ \mathbf{y}^*(\mathbf{x}) \in \mathcal{Y}(\mathbf{x}), \\
& \mathbf{c}^\top \mathbf{y}^*(\mathbf{x}) \leq \varphi(\mathbf{x}), \\
& \mathbf{x} \in \{0,1\}^{n_1} \times \mathbb{R}_+^{n_2},\ \mathbf{y}^*(\mathbf{x}) \in \mathbb{R}_+^m.
\end{aligned}
$$

Broadly speaking, the main difficulty in solving bilevel programs typically arises from the value-function constraint $\mathbf{c}^\top \mathbf{y}^*(\mathbf{x}) \leq \varphi(\mathbf{x})$. Combined with the leader's and follower's feasibility constraints, this condition enforces optimality of the follower's decision with respect to its problem; see (2). Replacing the follower's optimality condition with these constraints converts the bilevel LP into a single-level reformulation that is generally non-convex due to the piecewise-linearity of $\varphi(\mathbf{x})$.

Main challenge. As discussed in more detail in Section 2, BLPs are strongly NP-hard in general (Hansen et al. 1992), but can often be tackled through single-level reformulations that rely on the necessary and sufficient optimality conditions of the follower's problem (Audet et al. 1997). These reformulations typically introduce complementarity constraints, most often in the form of quadratic or special ordered set (SOS) constraints.

However, as the number of constraints in the lower-level problem increases, so does the number of dual variables, which in turn increases the number or "complexity" of the constraints required to enforce optimality. In particular, the number of either complementarity constraints or the associated quadratic terms grows. As a result, one can argue that the number of constraints in $\mathcal{Y}(\mathbf{x})$ is a key determinant of the computational difficulty in solving BLPs at scale.

Contribution. To address the scalability limitations of solving bilevel programs, we propose a dimensionality reduction approach targeting the follower's problem. The core idea is conceptually simple: aggregate a subset of the follower's constraints to reduce the problem's complexity. We show that this idea can be rigorously formalized and may serve as a foundation for a broad class of solution techniques for bilevel programs. Our contributions are threefold, as summarized in Figure 1.

• **We apply constraints aggregation techniques to bilevel programming.** We exploit surrogate constraints to formalize a surrogate duality theory for BLPs, from which we derive a feasibility-based lower bound and a dual upper bound on the leader's optimal objective function value. These bounds are obtained by formulating two mathematical programs where the follower's constraints are *projected* (or aggregated) into a compressed representation of lower dimension, thereby reducing their number. Taken together, these bounds establish our *first Sandwich Theorem*.

We extend the notion of surrogate duality to the bilevel setting and consider conditions under

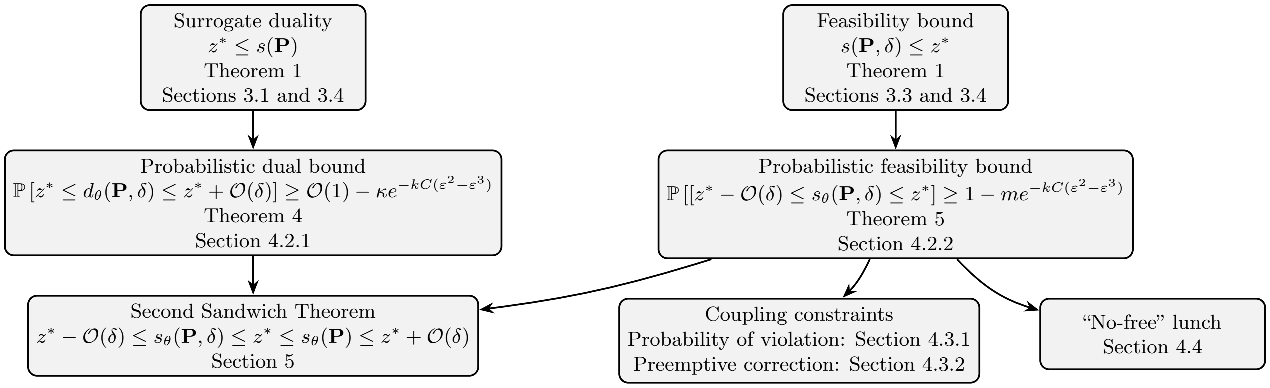


Figure 1: **Main contributions:** Given a projector $\mathbf{P} \in \mathbb{R}^{k\times q}$, we define $s(\mathbf{P})$ and $s_\theta(\mathbf{P})$ as the surrogate dual functions. Its randomized counterpart, denoted $d_\theta(\mathbf{P},\delta)$, accounts for projection uncertainty. The quantities $s(\mathbf{P},\delta)$ and $s_\theta(\mathbf{P},\delta)$ represent the feasibility-based bounds. The parameter $\delta > 0$ controls the approximation error incurred by projecting the follower's constraints to dimension k, and C is a constant; $\varepsilon = \mathcal{O}\left(\delta(1+\theta)^{-1}\right)$ and $\kappa \equiv \mathcal{O}\left(m(1+\theta\delta^{-1})^n\right)$.

which strong duality is guaranteed to hold. While the feasibility and dual programs may, in general, provide arbitrarily loose bounds, we characterize a class of projectors for which both bounds are provably within an $\mathcal{O}(\delta)$ absolute neighborhood of the true optimum. Specifically, this class consists of a subset of $\mathcal{O}(\delta)$-isometric mappings from $\mathbb{R}^q$ to some $\mathbb{R}^k$ where $k \equiv k(\delta)$ depends on the chosen distortion parameter δ. These mappings ensure that norms are approximately preserved in the lower-dimensional space. Accordingly, we obtain what we refer to as our *second Sandwich Theorem.*

• **We leverage random projection techniques to identify effective aggregation schemes and find solutions efficiently with probabilistic approximation guarantees.** While the two mathematical programs introduced in the Sandwich Theorems can provide meaningful bounds, identifying a "good" projector that narrows the gap between them remains challenging. To circumvent this difficulty, we leverage random projection techniques, also known as *linear sketching*, to derive probabilistic approximation guarantees for these bounds. We show that an $\mathcal{O}(\delta)$-approximate solution (in absolute terms) can be obtained with some probability $p \equiv p(\delta)$, provided the projected dimension $k \equiv k(\delta)$ is sufficiently large and the projector is sampled from a suitable distribution.

As a side result, we show that if both the leader's and follower's decision variables are continuous, then an $\mathcal{O}(\delta)$-approximate solution to the leader's problem can be computed in polynomial time with non-zero probability, assuming that the number of follower's constraints is sufficiently large. Furthermore, this probability asymptotically approaches one as the number of follower's constraints grows.

Also, we examine the impact of *coupling constraints*, which introduce additional challenges when applying random projection techniques. Accordingly, we propose a reformulated mathematical programs that, when feasible, still offer meaningful guarantees. In cases where the original formulations are used, we provide probabilistic bounds on the likelihood of violating the coupling constraints.

We close our theoretical contribution with a discussion on what we call the price of dimensionality reduction. Although random projection may appear to offer a “free lunch,” seemingly allowing a strongly NP-hard problem to be solved approximately via a randomized method, we show that this perspective is misleading. Indeed, the expectation over multiple randomized solutions may not be well-defined, providing no reliable guidance for derandomization. This fundamental limitation illustrates the uncertainty inherent to the method: while one can obtain good approximations with high probability, the lack of reliable averaging strategies is the price one pays for dimensionality reduction.

• **We numerically illustrate trade-offs induced by dimensionality reduction.** We provide preliminary insights into the practical implications of our techniques. In particular, we focus on the use of random projection, which we view as a natural first step toward exploring computationally in depth dimensionality reduction in bilevel programming. Using the min-cost flow interdiction problem as a testbed, we show that random projection can lead to meaningful runtime improvements on larger instances without significantly compromising bounds quality. We evaluate multiple projection strategies and find that their effectiveness depends on the structure of the follower’s constraints.

Organization. Section 2 reviews related work. Section 3 develops a surrogate duality theory. Section 4 presents an efficient random-projection procedure to construct lower and upper bounds with probabilistic approximation guarantees. Section 5 identifies a family of projectors for which the Sandwich Theorem can be strengthened with approximation guarantees. Section 6 reports preliminary computational experiments. Section 7 concludes. All proofs are deferred to the Appendix.

Notations. We denote the ℓ_1- and ℓ_2 norms by $\|\cdot\|_1$ and $\|\cdot\|_2$, respectively. Random variables are defined on a probability space $(\Omega, \mathcal{B}, \mathbb{P})$. A full summary of notations appears in Appendix A.

2 Literature review

Computational complexity. Although LPs can be solved in polynomial time (Khachiyan 1979; Karmarkar 1984), the introduction of a bilevel structure “increases” the computational complexity. Indeed, if decision variables at both levels are all continuous, then bilevel LPs are strongly NP-hard (Hansen et al. 1992). Consequently, no fully polynomial-time approximation algorithm is expected to exist for this class of problem unless P=NP. In fact, even finding a local optimum is known to be NP-hard (Prokopyev and Ralphs 2024) in that case. Moreover, bilevel LPs are closely related to disjunctive programs, as the two are, in a broad sense, equivalent (Basu et al. 2021).

Early developments and specialized algorithms. Bilevel programming is a distinct research area, due to its broad range of applications and exploitable structural properties. Early work focused on continuous leader–follower models, with methods such as complementary pivoting and early branch-and-bound (Júdice and Faustino 1992; Hansen et al. 1992). We refer to the survey by Kleinert et al. (2021) for historical context. For rare specialized classes, the bilevel structure enables approximation algorithms: a central example is interdiction attacker–defender models, where certain network variants admit constant-factor approximations and, in special cases, polynomial-time approximation schemes (Zenklusen 2010; Zenklusen 2015; Chestnut and Zenklusen 2017).

Reformulation-based approach. Bilevel programs with an LP at the lower level are typically solved using reformulation techniques that make them amenable to off-the-shelf solvers (IBM 2024; Gurobi 2024). These approaches first convert the bilevel program into a linear complementarity problem (Audet et al. 1997), which is then linearized to obtain a single-level mixed-integer linear program (MILP). The reformulation relies on optimality conditions, such as the Karush–Kuhn–Tucker conditions or strong duality (Fortuny-Amat and McCarl 1981; Bard and Falk 1982; Bialas and Karwan 1982), by replacing the follower's LP with its necessary and sufficient optimality constraints.

However, this process introduces quadratic terms that are in turn linearized using big-M constraints (Audet et al. 1997; Zare et al. 2019). This step inflates the single-level reformulation with additional binary variables and constraints and requires careful tuning of the big-M parameters, which may be difficult by itself (Kleinert et al. 2020). Also, the reformulation's "size" explodes with the number of constraints in the lower-level problem, which limits scalability, especially in the absence of exploitable structural properties in the follower's problem. Although widely used, these methods generally remain tractable only for small to moderately sized instances.

Constraint aggregation. Constraint aggregation, often applied through linear projection, serves as a technique to reduce the dimensionality of an optimization problem by combining multiple constraints into a smaller set. This idea lies at the core of *surrogate duality*, a well-established concept in mathematical programming. The foundational work of Glover (1965) introduces the principle of aggregating constraints into a single representative constraint in the context of single-level integer programming. Later contributions by Greenberg and Pierskalla (1970) and Glover (1975) further develop this approach and establish a formal theory to surrogate duality. However, despite their potential, none of these theoretical developments have been extended to bilevel programming.

In contrast to Lagrangian duality, which incorporates constraint violations as penalties in the objective function, surrogate duality replaces the original constraint set with a single surrogate constraint formed as a linear combination of the original ones. This transformation projects the feasible region into a compressed representation and results in stronger dual formulations (Karwan and Rardin 1979) compared to the Lagrangian dual. Surrogate duality also provides zero duality gaps for non-convex and integer programming problems (Glover 1975).

Although surrogate duality offers strong theoretical foundations, it remains, to the best of our knowledge, rarely used in the context of single-level optimization. Nonetheless, a few notable implementations do exist. For instance, Trapp and Prokopyev (2015) apply constraints aggregation techniques in the context of stochastic optimization, Dinkel and Kochenberger (1978) apply surrogate duality in nonlinear programming, and more recently, Müller et al. (2022) extend it for nonlinear mixed-integer programs. Such developments remain scarce in the literature, primarily because identifying a "good" projection requires solving the surrogate dual problem, which is hard by itself.

Random projection in single-level optimization. Rather than relying on surrogate dual formulations to derive a suitable projector, a promising alternative is the use of random projection techniques. These techniques have shown early success in adjacent fields such as combinatorial optimization and information retrieval (Vempala 2005), as well as in nearest neighbor search (Indyk and Naor 2007). Recently, their application within optimization has expanded, beginning with convex optimization problems (Pilanci and Wainwright 2015), extending to linear programs with equality constraints (Vu et al. 2018), and further advancing to handle inequality and semidefinite constraints (Poirion et al. 2023), as well as conic and quadratic programs (Liberti et al. 2021).

Random projection techniques appears particularly useful for tackling very large-scale optimization problems, where the number of constraints exceeds the capacity of commercial solvers. These methods typically build upon the *Johnson–Lindenstrauss lemma*, originally proposed by Johnson and Lindenstrauss (1984). This lemma indicates that a set of vectors can be embedded into a lower-dimensional space while approximately preserving pairwise distances with high probability. This result, in turn, can be leveraged to provide probabilistic guarantees on the quality of the solution obtained when solving the optimization problem with projected constraints.

Other forms of randomization in single-level optimization. Randomization plays a growing role in mathematical programming and extends beyond the use of random projection ideas. In particular, randomization techniques have been studied in the context of sampling columns in large-scale linear programs. Recent work explores how column sampling can be leveraged to approximate solutions efficiently in high-dimensional settings, especially when full enumeration is computationally infeasible (X. Li and Ye 2022; Gao et al. 2023; Akchen and Mišić 2025).

This idea traces back to the use of constraint sampling in linear programming, introduced by De Farias and Van Roy (2004) in the context of approximate dynamic programming. It shares conceptual similarities with developments in sample-based convex optimization, where a series of contributions by Calafiore and Campi (2005), Campi and Garatti (2008), and Campi and Garatti (2018) establish rigorous probabilistic guarantees for solution quality. Another related direction is proposed by Bertsimas and Vempala (2004), who introduce a sampling-based method that iteratively reduces the solution space in convex programs.

Summary. Taken together, these various methodologies illustrate how randomization can serve as a principled and effective paradigm for improving the scalability of optimization algorithms. Random projection techniques, though relatively new in the field of mathematical programming, have already shown promise in single-level optimization, particularly when combined with aggregation methods. While the number of constraints remains a key bottleneck for solving bilevel linear programs at scale, such techniques are still largely absent from the bilevel optimization literature. In this work, we introduce a surrogate duality theory for bilevel programming and leverage random projection techniques to derive probabilistic approximation guarantees for otherwise intractable instances.

3 Formalizing dimensionality reduction via surrogate duality

To mitigate the computational burden of solving BLPs, we consider a dimensionality reduction approach that projects the follower's problem onto one with fewer constraints. We use the term *dimension* to refer to the number of constraints in the follower's problem. This section provides conceptual background for both Section 4 and Section 5, and is intentionally light on technical detail.

Throughout this study, we restrict our attention to dimensionality reduction schemes that are linear and uniform (applied identically to both the constraint matrix and right-hand side). Such mappings take the form $P : \mathbf{u} \mapsto \mathbf{P}\mathbf{u}$, where $\mathbf{u} \in \mathbb{R}^q$ and $\mathbf{P} \in \mathbb{R}^{k \times q}$ is a matrix defining a *linear projector*. Since all projection schemes considered in this work are linear, we refer directly to the matrix $\mathbf{P}$ as the projector, omitting further reference to the mapping P when clear from context.

3.1 Surrogate dual problem

Dimensionality reduction of constraints via linear projection essentially provides weighted combinations of the original constraints, *surrogate constraints*, which are the pillar of surrogate duality (Glover 1975). As shown below, this duality theory extends naturally to bilevel programs.

To proceed, we introduce a version of the follower's problem in which constraints are aggregated according to a linear projector $\mathbf{P}$. Specifically, we define the *projected follower's problem* as follows:

$$\varphi(\mathbf{x}, \mathbf{P}) := \min \left\{ \mathbf{c}^\top \mathbf{y} \, : \, \mathbf{P}\mathbf{F}\mathbf{y} = \mathbf{P}\big(\mathbf{f} - \mathbf{L}\mathbf{x}\big) \, , \, \mathbf{y} \in \mathbb{R}^m_+ \right\}, \tag{4}$$

where, throughout the paper, we refer to $\varphi(\mathbf{x}, \mathbf{P})$ as the *follower's projected value function.* In addition, the feasible region of the projected follower's problem is denoted by $\mathcal{Y}(\mathbf{x}, \mathbf{P})$.

By applying surrogate duality theory to the follower's problem, the projected follower's problem (4) and the original problem (2) satisfy weak duality (Glover 1975). Therefore, the follower's value function is bounded bellow by the projected one, i.e., $\varphi(\mathbf{x}, \mathbf{P}) \leq \varphi(\mathbf{x})$ for all leader's feasible decision $\mathbf{x} \in \mathcal{X}$. We now use the projected follower's problem to construct the so-called *surrogate problem* of the leader's problem (1). Specifically, we introduce the *surrogate dual function*:

$$\begin{aligned} s(\mathbf{P}) := \max\ & \mathbf{a}^\top \mathbf{x} + \mathbf{d}^\top \mathbf{y} \\ \text{s.t. } & \mathbf{G}\mathbf{x} + \mathbf{H}\mathbf{y} \leq \mathbf{h}\,,\ \mathbf{y} \in \mathcal{Y}(\mathbf{x}, \mathbf{P}) \\ & \varphi(\mathbf{x}, \mathbf{P}) \leq \mathbf{c}^\top \mathbf{y} \leq \varphi(\mathbf{x}) \\ & \mathbf{x} \in \{0,1\}^{n_1} \times \mathbb{R}_+^{n_2}\,,\ \mathbf{y} \in \mathbb{R}_+^m. \end{aligned} \tag{5}$$

The surrogate problem (5) mirrors the value-function reformulation of the leader's problem (1). The subtle difference is that the follower's feasible region is replaced by its projected counterpart with an additional constraint that involves the follower's projected value function. These constraints guarantee that any bilevel-feasible pair $\big(\mathbf{x}, \mathbf{y}^*(\mathbf{x})\big)$ for the leader's problem (1) is also feasible for the surrogate problem (5). Thus, we obtain a *weak duality* result; namely $z^* \leq s(\mathbf{P})$ for all projectors $\mathbf{P} \in \mathbb{R}^{k \times q}$, for all $k \in [q]$. Accordingly, $s(\mathbf{P})$ provides a surrogate dual-based upper bound (shortly referred to as *dual bound* throughout the study) for the leader's optimal objective function value.

The surrogate problem provides, in a sense, a basis for comparing different dimensionality reduction schemes. Consequently, it is natural to ask which projection leads to the strongest bound. To formalize this question, we introduce the *surrogate dual problem* as follows:

$$\inf_{\mathbf{P} \in \mathbb{R}^{k \times q}} s(\mathbf{P}),$$

where the infimum is taken over all linear projectors from $\mathbb{R}^q \to \mathbb{R}^k$.

As a side remark, the surrogate dual in the single-level context is originally defined in Glover (1975) over linear projectors from $\mathbb{R}^q$ to $\mathbb{R}$. However, the result also holds for any linear projector $\mathbf{P}^{(k)} : \mathbb{R}^q \to \mathbb{R}^k$, given $k \in [q]$. In particular, we obtain the following hierarchy of bounds:

$$\inf_{\mathbf{P}^{(1)} \in \mathbb{R}^{1 \times q}} s(\mathbf{P}^{(1)}) \geq \inf_{\mathbf{P}^{(2)} \in \mathbb{R}^{2 \times q}} s(\mathbf{P}^{(2)}) \geq \ldots \geq \inf_{\mathbf{P}^{(q)} \in \mathbb{R}^{q \times q}} s(\mathbf{P}^{(q)}).$$

Indeed, for $k \in [q]$, any projector $\mathbf{P}^{(k)} \in \mathbb{R}^{k \times q}$ can be used to construct a projection scheme into a higher dimension $\mathbb{R}^{(k+1) \times q}$ by adding a row with only zeros as components.

3.2 Strong duality

We identify a condition for which the surrogate dual attains the leader's optimal objective function value. Our argument proceeds in two steps: (i) we use results from surrogate duality for single-level convex optimization to show that, for the follower's problem, the surrogate problem provides a tighter bound than the Lagrangian dual; (ii) we show that the surrogate dual problem attains the leader's optimal objective function value for an optimal solution to the follower's dual problem.

To begin, we define the Lagrangian dual (Bertsimas and Tsitsiklis 1997) of the follower's problem. Specifically, given a follower's feasible decision $\mathbf{y} \in \mathcal{Y}(\mathbf{x})$, for some $\mathbf{x} \in \mathcal{X}$, and $\boldsymbol{\lambda} \in \mathbb{R}^{1\times q}$, we define:

$$\mathcal{L}\big(\mathbf{y}, \boldsymbol{\lambda}\,;\, \mathbf{x}\big) := \mathbf{c}^\top \mathbf{y} + \boldsymbol{\lambda}\big(\mathbf{f} - \mathbf{F}\mathbf{y} - \mathbf{L}\mathbf{x}\big).$$

Next, to proceed, we introduce the *Lagrangian problem*:

$$\mathcal{L}\big(\boldsymbol{\lambda}\,;\, \mathbf{x}\big) := \min_{\mathbf{y}} \mathcal{L}(\mathbf{y}, \boldsymbol{\lambda}\,;\, \mathbf{x}) = \begin{cases} \boldsymbol{\lambda}(\mathbf{f} - \mathbf{L}\mathbf{x}), & \text{if } \mathbf{c}^\top - \boldsymbol{\lambda}\mathbf{F} \geq \mathbf{0}, \\ -\infty, & \text{otherwise.} \end{cases}$$

Although the Lagrangian dual is a well-established approach for deriving lower bounds in convex optimization, we show that, within the context of LP, the surrogate dual can be considered stronger than its Lagrangian counterpart. The original result is due to Glover (1975). Specifically, we have that $\mathcal{L}\big(\mathbf{P}\,;\, \mathbf{x}\big) \leq \varphi(\mathbf{x}, \mathbf{P}) \leq \varphi(\mathbf{x})$, for all projectors $\mathbf{P} \in \mathbb{R}^{1\times q}$. Indeed, let $\mathbf{y}$ be a feasible solution to the projected follower's problem (4). Then, we obtain the following inequality:

$$\mathbf{c}^\top \mathbf{y} = \mathbf{c}^\top \mathbf{y} + \mathbf{P}\big(\mathbf{f} - \mathbf{F}\mathbf{y} - \mathbf{L}\mathbf{x}\big) \geq \mathcal{L}\big(\mathbf{P}\,;\, \mathbf{x}\big),$$

by the definition of Lagrangian and the feasibility of $\mathbf{y}$. Since the follower's problem is feasible for all $\mathbf{x} \in \mathcal{X}$, we conclude by the LP strong duality that there exists some extreme points $\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(K)}$ of the polyhedron defined by the constraints $\mathcal{D} := \big\{\boldsymbol{\lambda} \in \mathbb{R}^{1\times q} \,:\, \mathbf{c}^\top - \boldsymbol{\lambda}\mathbf{F} \geq \mathbf{0}\big\}$ that satisfy:

$$\max_{\mathbf{P}\in\mathbb{R}^{1\times q}} \mathcal{L}\big(\mathbf{P}\,;\, \mathbf{x}\big) = \max_{i\in[K]} \mathcal{L}\big(\mathbf{e}^{(i)}\,;\, \mathbf{x}\big) = \varphi(\mathbf{x}).$$

Accordingly, there exists some $i \in [K]$ such that $\varphi(\mathbf{x}, \mathbf{e}^{(i)}) = \varphi(\mathbf{x})$. In particular, this relation holds for all leader's feasible decision $\mathbf{x} \in \mathcal{X}$ (with index i that may depend on $\mathbf{x}$). This result makes the follower's optimality constraint binding in (5). However, if the projected follower's problem admits multiple optimal solutions, then an optimal solution to (1) need not be optimal for (5) when using $\mathbf{P} = \mathbf{e}^{(i)}$. In that case, a duality gap may persist. Hence, an "optimal" projector has to balance a trade-off. It should not only approximate the projected follower's value function well, but also, loosely speaking, shrink $\mathcal{Y}(\mathbf{x}, \mathbf{P})$ to exclude competing optimal solutions. If the projected follower's problem has a unique optimal solution for every $\mathbf{x} \in \mathcal{X}$, then strong duality follows. Specifically, under this assumption, there exists $i \in [K]$ with $s\left(\mathbf{e}^{(i)}\right) = z^*$ and the hierarchy of bounds collapses.

As shown above, the surrogate dual has limited algorithmic value from a solution-methods perspective in its current form. Indeed, solving it appears to require enumerating the extreme points of the follower's dual feasible region. That enumeration seems necessary but not sufficient; determining additional conditions that close this gap remains an open question.

3.3 Surrogate constraints & feasibility problem

Dimensionality reduction ideas can also be used to provide a feasibility bound (i.e., a lower bound) on the leader's optimal objective function value. To simplify our discussion in this section, we assume that there are no coupling constraints, i.e., $\mathbf{H} = \mathbf{0}_{p\times m}$. Following Henke et al. (2025), we note that this assumption is made without loss of generality (at least for bilevel LPs), as any coupling constraints can be incorporated into the leader's objective function through appropriate penalty terms.

To proceed, we fix $\delta \geq 0$ and $\mathbf{P} \in \mathbb{R}^{k\times q}$, for $k \in [q]$, and introduce the following set:

$$\begin{aligned} \mathcal{S}(\mathbf{P}, \delta) := \text{argmax} \;\; & \mathbf{a}^\top \mathbf{x} + \mathbf{d}^\top \mathbf{y} \\ \text{s.t.} \;\; & \mathbf{G}\mathbf{x} \leq \mathbf{h}\,,\; \mathbf{y} \in \mathcal{Y}(\mathbf{x}) \\ & \mathbf{c}^\top \mathbf{y} \leq \varphi(\mathbf{x}, \mathbf{P}) + \delta \\ & \mathbf{x} \in \{0,1\}^{n_1} \times \mathbb{R}_+^{n_2}\,,\; \mathbf{y} \in \mathbb{R}_+^m, \end{aligned} \tag{7}$$

which might be empty for small values of δ. Accordingly, if the underlying optimization problem is infeasible, then we adopt the convention $\mathcal{S}(\mathbf{P}, \delta) = \varnothing$. This auxiliary problem is closely related to the original leader's formulation (1). Let $\mathbf{x}^*$ denote an optimal leader's decision of (1) and set $\delta := \varphi(\mathbf{x}^*) - \varphi\big(\mathbf{x}^*, \mathbf{P}\big)$. With this choice, the optimal bilevel pair $\big(\mathbf{x}^*, \mathbf{y}^*(\mathbf{x}^*)\big)$ belongs to $\mathcal{S}(\mathbf{P}, \delta)$.

Then, given some arbitrary gap parameter $\delta \geq 0$, and projector $\mathbf{P} \in \mathbb{R}^{k\times q}$, for some $k \in [q]$, we define the *feasibility problem* as follows:

$$s(\mathbf{P}, \delta) := \max_{\mathbf{x}, \mathbf{y}^*(\mathbf{x})} \left\{ \mathbf{a}^\top \mathbf{x} + \mathbf{d}^\top \mathbf{y}^*(\mathbf{x}) \;:\; (\mathbf{x}, \mathbf{y}) \in \mathcal{S}(\mathbf{P}, \delta) \right\}, \tag{8}$$

where $\mathbf{y}^*(\mathbf{x})$ is the follower's rational response in (3), which is a subtile yet important point. We refer to $s(\mathbf{P}, \delta)$ as the *feasibility bound.* If $\mathcal{S}(\mathbf{P}, \delta) = \emptyset$, then we define $s(\mathbf{P}, \delta) := -\infty$ by convention.

The feasibility problem (8) can be interpreted as a reformulation of the leader's problem (1) in terms of the follower's value function. However, we replace the exact follower's value function with its projected counterpart and add a non-negative tolerance δ to guarantee feasibility. Consequently, the resulting feasibility bound satisfies $s(\mathbf{P}, \delta) \leq z^*$ for all $\delta \geq 0$ and $\mathbf{P} \in \mathbb{R}^{k\times q}$, given $k \in [q]$. Importantly, we could derive a similar hierarchy of bounds to that presented in Section 3.1, but we omit it for brevity. Following similar arguments to the ones from Section 3.2, we obtain:

$$\max_{\substack{\mathbf{P}\in\mathbb{R}^{k\times q}\\ \delta\in\mathbb{R}_+}} s(\mathbf{P}, \delta) = \max_{i\in[K]} s\big(\mathbf{e}^{(i)}, 0\big) = z^*,$$

where $\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(K)}$ are the extreme points of the feasible region of the follower's dual problem.

3.4 First Sandwich Theorem

Building on the previous discussion, we establish a *first Sandwich Theorem* that provides bounds for the leader's optimal objective function value z^*. Since the result is a direct implication of our earlier arguments, we omit the formal proof for brevity. Specifically:

Theorem 1 (First Sandwich Theorem)**.** *Given $k \in [q]$, the following inequalities are satisfied:*

$$s(\mathbf{P}, \delta) \;\leq\; z^* \;\leq\; s(\mathbf{P}),$$

for any projector $\mathbf{P} \in \mathbb{R}^{k\times q}$ and any relaxation parameter $\delta \geq 0$.

Example 1. We provide an instance of the leader's problem (1) in which the sandwich inequality is tight, provided δ is chosen appropriately. Consider the bilevel linear program of the form:

$$\max_{\mathbf{x},\mathbf{y}} \left\{ y_1^*(\mathbf{x}) : x_1 + x_2 = 1,\ x_1, x_2 \in [0, 1],\ y_2^*(\mathbf{x}) = \tfrac{1}{2} \right\},$$

where, for a leader's feasible decision $\mathbf{x} = (x_1, x_2)^\top$, the follower solves the following linear program:

$$\left(y_1^*(\mathbf{x}), y_2^*(\mathbf{x})\right)^\top \in \arg\min \left\{ y_1 : y_1 = 1 - x_1,\ y_2 = 1 - x_2,\ y_1, y_2 \in [0, 1] \right\}.$$

The leader's optimal choice is $\mathbf{x}^* = \left(\frac{1}{2}, \frac{1}{2}\right)^\top$ with the leader's optimal objective function value $z^* = 1/2$. Then, fix the projector $\mathbf{P} = \big(1, 1\big)^\top$. We obtain $s(\mathbf{P}) = 1/2$ and $s(\mathbf{P}, \delta) = 1/2$ for every $\delta \geq 1/2$ so that the inequalities from the First Sandwich Theorem are tight for this projector. ■

Key takeaway. The two formulations developed in Section 3 offer a new theoretical lens on surrogate duality for BLPs. Their computational value, however, remains less clear. Although the sandwich result can, in principle, bound the leader's optimal objective function value z^*, the practical tightness of the resulting bounds has yet to be assessed. In particular, closing the gap between the feasibility-based lower bound $s(\mathbf{P}, \delta)$ and the surrogate dual upper bound $s(\mathbf{P})$ hinges on selecting a good projector $\mathbf{P}$ (and an accompanying tolerance δ). Identifying such a projector appears to be at least as difficult as solving the original leader's problem (1).

4 Surrogate bounds via random projection

As emphasized in the previous section, explicitly identifying the "best" projector under the criteria in Section 3.4 is itself challenging. Instead, we adopt a *randomized* approach and leverage the results by Vu et al. (2018) on dimensionality reduction for single-level linear programs. In this approach, the projector is not optimized but sampled from a carefully specified distribution, and performance guarantees follow from the geometric and concentration properties of these projectors.

4.1 Background and notations

Random projection. A *random projector* is a linear map from $\mathbb{R}^q$ to a lower-dimensional space $\mathbb{R}^k$ for some $k \leq q$. The projector used throughout this paper is mainly the Gaussian projector $\mathbf{P} = (P_{ij})_{(i,j)\in[k]\times[q]}$ whose entries are drawn identically and independently as $P_{ij} \sim \mathcal{N}(0, 1/k)$.

Although Gaussian embeddings offer "clean" theoretical guarantees, other constructions may provide substantial computational advantages. To name a few, Achlioptas (2003), Matoušek (2008), Kane and Nelson (2014), and Dasgupta et al. (2010) propose sparse random projectors in which each column contains only a few non-zero entries, greatly reducing the cost of matrix–vector multiplications. Some variants add sign-consistency constraints so that all non-zero entries in a column share the same sign, a property exploited in compressed-sensing applications (Allen-Zhu et al. 2014).

Concentration results. Random-projection methods trace back to the seminal paper by Johnson and Lindenstrauss (1984), which shows that any finite set of vectors can be embedded into a lower-dimensional space while approximately preserving Euclidean distances.

Theorem 2 (Johnson and Lindenstrauss, 1984). *Given $\epsilon \in (0,1)$ and a matrix $\mathbf{F} \in \mathbb{R}^{q\times m}$, there exists a matrix $\mathbf{P} \in \mathbb{R}^{k\times q}$ such that the following inequalities are satisfied:*

$$\forall\, 1 \leq i < j \leq m \qquad \big(1-\varepsilon\big)\|\mathbf{F}_i - \mathbf{F}_j\|_2 \leq \|\mathbf{P}\mathbf{F}_i - \mathbf{P}\mathbf{F}_j\|_2 \leq \big(1+\varepsilon\big)\|\mathbf{F}_i - \mathbf{F}_j\|_2,$$

where $\mathbf{F}_i$ denotes the i-th column of $\mathbf{F}$, and where $k \equiv \mathcal{O}\left(\varepsilon^{-2}\log(m)\right)$.

In the original proof of the Johnson-Lindenstrauss lemma, the existence of projectors that approximately preserve the norm is established by construction using Gaussian random matrices. In our study, we adopt a broader notion of a *random projector* and allow any linear embedding whose columns are sub-Gaussian random vectors with unit variance (Matoušek 2008).

Additional notation. Assumption **A2** guarantees that the feasible regions of both the leader and the follower are bounded; recall our discussion in Section 1. This boundedness enters our analysis through the constant θ from Assumption **A2**, which caps the follower's aggregate decision variables. Given a projector $\mathbf{P} \in \mathbb{R}^{k\times q}$ with $k \leq q$, we define the *projected follower's problem* as:

$$\varphi_\theta(\mathbf{x}, \mathbf{P}) := \min\left\{\mathbf{c}^\top \mathbf{y} : \mathbf{PFy} = \mathbf{P}\big(\mathbf{f} - \mathbf{Lx}\big)\,,\ \mathbf{1}_m^\top \mathbf{y} \leq \theta\,,\ \mathbf{y} \in \mathbb{R}^m_+\right\}, \tag{9}$$

and define the corresponding feasible region by $\mathcal{Y}_\theta(\mathbf{x}, \mathbf{P})$. To parallel the notations from Section 3, we henceforth write $s_\theta(\mathbf{P})$, $\mathcal{S}_\theta(\mathbf{P}, \delta)$ and $s_\theta(\mathbf{P}, \delta)$ instead of $s(\mathbf{P})$, $\mathcal{S}(\mathbf{P}, \delta)$ and $s(\mathbf{P}, \delta)$, when referring to (5), (7), and (8), respectively. This change highlights the dependence of these quantities on the parameter θ. Throughout this section, whenever we refer to (5), (7), and (8), we assume that $\mathcal{Y}(\mathbf{x}, \mathbf{P})$ and $\varphi(\mathbf{x}, \mathbf{P})$ are replaced by $\mathcal{Y}_\theta(\mathbf{x}, \mathbf{P})$ and $\varphi_\theta(\mathbf{x}, \mathbf{P})$, respectively. Similarly, to maintain consistency and emphasize the dependence of $\varphi(\mathbf{x})$ on θ in Assumption **A2**, we write $\varphi_\theta(\mathbf{x})$ instead of $\varphi(\mathbf{x})$.

Uniqueness assumption. Throughout, we assume that for every leader's feasible decision $\mathbf{x} \in \mathcal{X}$, the projected follower's problem in (9) has a *unique* optimal solution. This assumption is a mild nondegeneracy requirement. Indeed, under Gaussian perturbations, LPs are almost surely infeasible, unbounded, or have unique primal-dual optimal solution (Spielman and Teng 2003). Precise probabilistic quantification of this event under absolutely continuous perturbations and in the context of random-projection is beyond our scope and remains a worthwhile direction for future work.

4.2 Random projection of the follower's feasible region

We first recall a random-projection result established for single-level LPs. Building on this foundation, we extend the same ideas to the bilevel setting, obtaining both a feasibility bound and a dual bound for the leader's optimal objective function value. Our exposition relies on Vu et al. (2018); their main theorem, restated below in our notations, provides the cornerstone for this section.

Theorem 3 (Vu et al., 2018)**.** *Let $\mathbf{P} \in \mathbb{R}^{k\times q}$ be a random projector and $\mathbf{x} \in \mathcal{X}$ be a leader's feasible decision. Given $\delta \in \big(0, \varphi_\theta(\mathbf{x})\big)$, we have that the following inequality is satisfied:*

$$\mathbb{P}\Big[\varphi_\theta(\mathbf{x}) - \delta \leq \varphi_\theta(\mathbf{x}, \mathbf{P}) \leq \varphi_\theta(\mathbf{x})\Big] \geq 1 - 4me^{-C\left(\varepsilon^2 - \varepsilon^3\right)k},$$

for $\varepsilon = \mathcal{O}\big(\frac{\delta}{(1+\theta)\|\mathbf{u}^*(\mathbf{x})\|_2\|\mathbf{f}-\mathbf{Lx}\|_2}\big)$, where $\mathbf{u}^*(\mathbf{x})$ denotes the optimal solution of minimal ℓ_2-norm to the dual problem of the follower's linear program in (2)*.*

The original result by Vu et al. (2018) is stated for right-hand sides normalized to have unit norm, an assumption that is without loss of generality in their case. However, it introduces leader-dependent terms when applied to our setting. These terms can be replaced by a leader-independent constant; the proof is identical to that of Vu et al. (2018), so we omit it. Briefly, the proof consists to project the columns of the follower's constraint matrix into $\mathbb{R}^k$ by exploiting the Johnson-Lindenstrauss lemma. Then, it uses feasibility-based arguments to bound the induced error.

Theorem 3 implies that the follower's projected value function can closely approximate the original one in (2) for a broad class of random projectors. The result stands in sharp contrast to the surrogate duality approach in Section 3, where the "optimal" projector is decision-dependent and hard to compute in general, whereas here it can be sampled directly from a suitably chosen distribution.

Figure 2: (Left) Plot of $g(\varepsilon) = \varepsilon^3 - \varepsilon^2$. (Right) Plot of $h(\varepsilon) = 1 - \exp\left(10(\varepsilon^3 - \varepsilon^2)\right)$. The curves depict the dependence of the concentration bound of Theorem 3 on ε, and, consequently, on δ.

Figure 2 illustrates the dependence on ε of the probability that Theorem 3's guarantees hold for the follower's projected value function. Whenever $\varepsilon > 1$, we have that $\exp(-C(\varepsilon^2 - \varepsilon^3)k) > 1$, and the inequality conveys no information. Also, $\delta \to 0$ forces $\varepsilon \to 0$, which renders the bound effectively vacuous. Thus, δ must be carefully chosen: if δ is too large, then the approximation guarantee disappears; if it is too small, then the guarantee, though formally valid, may become too weak to be useful.

Coupling constraints. In Sections 4.2.1 and 4.2.2, we use Theorem 3 to refine the reformulation from Section 3. Theorem 3 ensures that an optimal solution to the projected follower's problem can often approximate the follower's fully rational response. Yet, even when the projection error is small, coupling constraints remain troublesome. Although Section 3.3 argues that they can be absorbed via penalties in the leader's objective function (Henke et al. 2025), selecting those penalties in practice is delicate. We therefore analyze the leader's problem (1) *without* coupling constraints, i.e., set $\mathbf{H} = \mathbf{0}_{p\times m}$, throughout Sections 4.2.1 and 4.2.2. Given their importance, Section 4.3 analyzes the relation between coupling-constraints violation and projection error, and presents practical remedies.

4.2.1 On a probabilistic dual bound with approximation guarantees

We derive a dual bound on the leader's optimal objective function value by combining the random-projection approach introduced earlier with the surrogate-duality theory from Section 3. All proofs from this section are relegated to Appendix B.1. The key idea is to exploit Theorem 3 to replace the follower's value function with its projected counterpart in the surrogate problem (5).

To that end, we propose a new formulation to replace the surrogate dual function:

$$
\begin{aligned}
d_\theta(\mathbf{P},\delta) := \max \quad & \mathbf{a}^\top \mathbf{x} + \mathbf{d}^\top \mathbf{y} \\
\text{s.t.} \quad & \mathbf{G}\mathbf{x} \le \mathbf{h}\,,\ \mathbf{y} \in \mathcal{Y}_\theta(\mathbf{x},\mathbf{P}), \\
& \varphi_\theta(\mathbf{x},\mathbf{P}) \le \mathbf{c}^\top \mathbf{y} \le \varphi_\theta(\mathbf{x},\mathbf{P}) + \delta, \\
& \mathbf{x} \in \{0,1\}^{n_1} \times \mathbb{R}_+^{n_2}\,,\ \mathbf{y} \in \mathbb{R}_+^m.
\end{aligned}
\tag{10}
$$

We refer to (10) as the *adjusted surrogate problem.* It is obtained from the original surrogate problem (5) by replacing the follower's value function with its projected counterpart and introducing a penalty parameter δ. Next, we demonstrate that $d_\theta(\mathbf{P},\delta)$ constitutes a valid upper bound on the leader's optimal objective function value z^* with high probability. Formally:

Lemma 1. *Let $\mathbf{P} \in \mathbb{R}^{k\times q}$ be a random projector. Then, given $\delta \in (0, \min\{\varphi_\theta(\mathbf{x}) \,:\, \mathbf{x} \in \mathcal{X}\})$, we have that $z^* \le d_\theta(\mathbf{P},\delta)$ with probability at least $1 - 4me^{-C\left(\varepsilon^2-\varepsilon^3\right)k}$, where $\varepsilon = \mathcal{O}\big(\frac{\delta}{1+\theta}\big)$.*

Lemma 1 does not rely on any specific assumption on the coupling constraints. The same property holds for the optimal objective function value $d_\theta\,(\mathbf{P},\delta)$ of the adjusted surrogate problem (10) whenever it contains coupling constraints. In particular, with high probability, $d_\theta(\mathbf{P},\delta)$ is an upper bound on z^*. Theorem 4 complements this lemma by showing that $d_\theta(\mathbf{P},\delta)$ remains close to z^*, with an approximation error that can be controlled through δ and with probabilistic guarantees.

Theorem 4. *Let $\mathbf{P} \in \mathbb{R}^{k\times q}$ be a random projector, and define the random variable:*

$$\Delta \equiv \Delta(\mathbf{PF},\mathbf{c}) = \max\Big\{\big|\det(\mathbf{M})\big| \,:\, \mathbf{M} \textit{ is a square submatrix of } [\,(\mathbf{PF})^\top \mid -(\mathbf{PF})^\top \mid -\mathbf{c} \,|\, \mathbf{c} \,|\, \mathbf{1}_m\,]\Big\},$$

induced by $\mathbf{P}$. Let denote by F_Δ the distribution of $\max\{\Delta(\mathbf{PF},\mathbf{c})\,,\, \|\mathbf{PL}\|_\infty\}$. Then, for all $\delta \in (0, \min\{\varphi_\theta(\mathbf{x}) \,:\, \mathbf{x} \in \mathcal{X}\})$ and for all $\rho_\Delta \ge 0$, we have that the following inequality is satisfied:

$$\mathbb{P}\Big[d_\theta(\mathbf{x},\mathbf{P}) \le z^* + 3\rho_\Delta \|\mathbf{d}\|_1 m\delta\Big] \ge F_\Delta(\rho_\Delta) - \kappa e^{-C\left(\varepsilon^2-\varepsilon^3\right)k} \equiv p_\Delta(\delta, C),$$

where $\varepsilon \equiv \mathcal{O}\left(\frac{\delta}{1+\theta}\right)$, $\kappa \equiv \mathcal{O}\left(m\left(1 + \frac{\theta\rho_\Delta^2 m}{\delta}\right)^n\right)$ and $C > 0$ is some positive constant.

When the random projector has bounded entries, such as in the construction by Achlioptas (2003) where components take values in $\{-1,0,1\}$, Hadamard's inequality (Brenner and Cummings 1972) can be used to control ρ_Δ directly. In that case, one can select ρ_Δ large enough so that $F_\Delta(\rho_\Delta) = 1$, and the probability bound in Theorem 4 simplifies to $1 - \kappa e^{-C(\varepsilon^2-\varepsilon^3)k}$. This approach comes with a looser approximation guarantee, since Hadamard's inequality is conservative, but it removes the dependence on the distribution of $\Delta(\mathbf{PF},\mathbf{c})$ and $\|\mathbf{PL}\|_\infty$. This simplification is restricted to bounded projectors and cannot be applied to unbounded ones such as Gaussian projectors.

Theorem 4 also makes explicit a trade-off between approximation accuracy and probabilistic confidence. The approximation error increases linearly in the tolerance through the term $3\rho_\Delta\|\mathbf{d}\|_1 m\delta$, while the confidence level improves as k grows and deteriorates if δ is chosen too small. Decreasing δ tightens the bound on $d_\theta(\mathbf{x}, \mathbf{P})$, but it also decreases ε and increases κ, so that achieving a comparable confidence requires a larger projector's dimension k. To build intuition about this balance, we illustrate the behavior of the bound in Example 2 below on a randomly generated instance.

By the definition of κ, the bound worsens as n increases: the difficulty progressively shifts from the follower's problem to the leader's decision space. This definition for κ comes from a deliberately generic argument based on an η-covering of the leader's feasible region with ℓ_1-balls, which we then exploit to control the randomness induced by $\mathbf{P}$ in (10); see the details in the proof provided in Appendix B.1. We believe that more refined geometric tools, or the exploitation of the polyhedral structure of the problem, could substantially tighten this dependence on n. Keeping all other quantities fixed, selecting k of order $\mathcal{O}\left((\delta^2 - \delta^3)^{-1}\big(\log m + n\log(1+\delta^{-1})\big)\right)$ is enough to make the term $\kappa e^{-C(\varepsilon^2-\varepsilon^3)k}$ negligible independently of q, so that tight approximation guarantees can still be obtained with high probability as q grows. We emphasize that these bounds are *worst-case* and that exploring more sophisticated geometric arguments to improve κ is left for future research.

Re-parametrization. Note that finding an appropriate value or range for δ such that Theorem 3 applies may not be intuitive as it requires solving a potentially challenging mixed-integer linear program of the form $\min\{\varphi_\theta(\mathbf{x}) \,:\, \mathbf{x} \in \mathcal{X}\}$ to get an absolute scale for δ. To circumvent this difficulty, we reformulate the surrogate problem (10) so that it instead takes as input a relative tolerance parameter $\tilde{\delta} > 0$. The *re-parametrized adjusted surrogate problem* takes the following form:

$$\tilde{d}_\theta(\mathbf{P}, \tilde{\delta}) := \max_{\mathbf{x},\mathbf{y}} \left\{ \mathbf{a}^\top\mathbf{x} + \mathbf{d}^\top\mathbf{y} \,:\, \mathbf{G}\mathbf{x} \le \mathbf{h}\,,\, \mathbf{y} \in \mathcal{Y}_\theta(\mathbf{x}, \mathbf{P})\,,\, 0 \le \mathbf{c}^\top\mathbf{y} - \varphi_\theta(\mathbf{x}, \mathbf{P}) \le \tilde{\delta}\,|\varphi_\theta(\mathbf{x}, \mathbf{P})| \right\}.$$

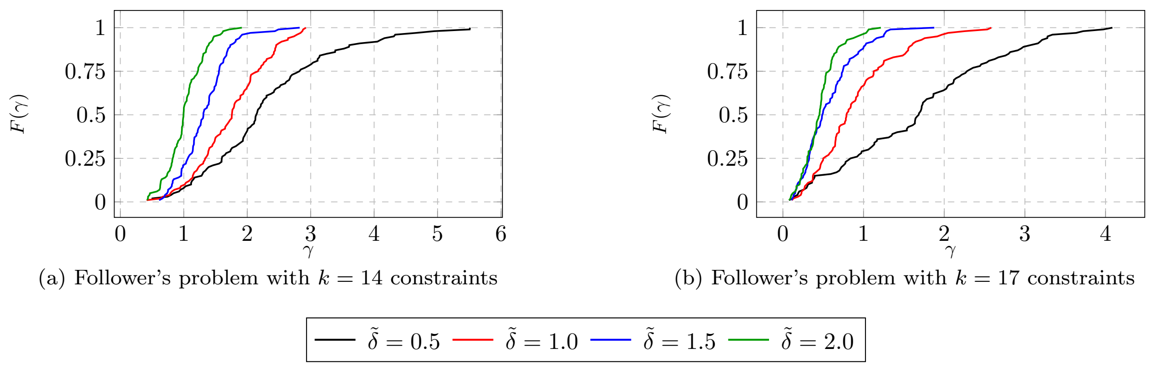
An analogous result to Theorem 4 holds for this re-parametrized problem. Formally:

Corollary 1. *Let $\mathbf{P} \in \mathbb{R}^{k\times q}$ be a random projector. Then, there exists a sufficiently small value $\tilde{\delta} \in \left(0, \left(\|\mathbf{c}\|_\infty\theta\right)^{-1} \min\left\{\varphi_\theta(\mathbf{x}) \,:\, \mathbf{x} \in \mathcal{X}\right\}\right)$ such that the following inequality is satisfied:*

$$\tilde{d}_\theta(\mathbf{x}, \mathbf{P}) \le z^* + 3\rho_\Delta\|\mathbf{d}\|_1 m\|\mathbf{c}\|_\infty\theta\tilde{\delta},$$

with probability at least $F_\Delta(\rho_\Delta) - \kappa e^{-C\left(\varepsilon^2-\varepsilon^3\right)k}$, for $\varepsilon \equiv \mathcal{O}\big(\frac{\tilde{\delta}\|\mathbf{c}\|_\infty\theta}{1+\theta}\big)$ and $\kappa \equiv \mathcal{O}\left(m\left(1 + \frac{\theta\rho_\Delta^2 m}{\tilde{\delta}\|\mathbf{c}\|_\infty\theta}\right)^n\right)$.

The condition on $\tilde{\delta}$ in Corollary 1 restricts $\tilde{\delta}$ to lie in the interval $(0, 1)$, although the bound may in fact hold for a broader set of values. We impose this restriction mainly for clarity, as it makes the corollary an immediate consequence of Theorem 4.



(a) Follower's problem with $k = 14$ constraints

(b) Follower's problem with $k = 17$ constraints

Figure 3: Empirical cumulative distribution $F(\gamma)$ of the gap $\gamma \equiv \gamma(\tilde{d}_\theta\,;\mathbf{P},\tilde{\delta}) \equiv \frac{1}{\tilde{\delta}}\big(\tilde{d}_\theta(\mathbf{P},\tilde{\delta}) - z^*\big)$ for the instance described in Example 2, where $\mathbf{P} \in \mathbb{R}^{k\times q}$ is a random projector with i.i.d. entries $P_{ij} \sim \mathcal{N}(0, 1/k)$; results are shown for different numbers of follower's constraints $k \in \{14, 17\}$ and values of $\tilde{\delta} \in \{0.5, 1.0, 1.5, 2.0\}$.

Example 2. We illustrate numerically the trade-offs from Theorem 4 and Corollary 1. Our goal is not to supplant the more complete computational experiments in Section 6, but rather to provide intuition behind our probabilistic results. To this end, we consider a bilevel program of the form:

$$\max\Big\{\min\Big\{\mathbf{c}^\top\mathbf{y}\ :\ \mathbf{x}+\mathbf{y}=\mathbf{f}\,,\ \mathbf{1}_n^\top\mathbf{y}\leq\theta\,,\ \mathbf{y}\geq\mathbf{0}_n\Big\}\ :\ \mathbf{x}+\mathbf{y}\leq n\mathbf{1}_n\,,\ \mathbf{x}\geq\mathbf{0}_n\Big\}\,,$$

which models a relatively simple zero-sum resource-allocation game between a leader and a follower.

Instance description. We generate the problem data as follows. Set $n = m = q = 20$. Each entry of the vector $\mathbf{f} \in \mathbb{R}^n$ is drawn i.i.d. from a Bernoulli(1/2) distribution. The cost vector $\mathbf{c} \in \mathbb{R}^n$ has entries sampled i.i.d and uniformly from $[0, 1]$. We fix the bound for the follower's feasible region at $\theta = 20$. In our example, we consider a random projection matrix $\mathbf{P} = (P_{ij})_{ij}$ with entries drawn i.i.d from the Gaussian distribution $P_{ij} \sim \mathcal{N}(0, 1/k)$, for $i \in [k]$, $j \in [q]$.

We vary the projector's dimension $k \in \{14, 17\}$ and the tolerance $\tilde{\delta} \in \{0.5, 1.0, 1.5, 2.0\}$. Then, we compute z^*, draw 100 realizations of $\mathbf{P}$ for each k, and then compute $\tilde{d}_\theta(\mathbf{P},\tilde{\delta})$ for each pair $(\mathbf{P},\tilde{\delta})$. We measure the performance of $\tilde{d}_\theta(\mathbf{P},\tilde{\delta})$ by using the empirical distribution of the random variable induced by the difference between z^* and $\tilde{d}_\theta(\mathbf{P},\tilde{\delta})$. Formally, we consider $\gamma \equiv \gamma(\tilde{d}_\theta\,;\mathbf{P},\tilde{\delta}) \equiv \frac{1}{\tilde{\delta}}\big(\tilde{d}_\theta(\mathbf{P},\tilde{\delta}) - z^*\big)$, and study its empirical distribution over the random draws of the random projector $\mathbf{P}$; see the results in Figure 3.

Performance. Loosely speaking, Theorem 4 and Corollary 1 ensure the existence of constants C_1 and $C_2(\tilde{\delta})$ such that the gap γ (which is a random variable induced by $\mathbf{P}$) is guaranteed to satisfies the following inequality $\mathbb{P}\big[\gamma(\tilde{d}_\theta\,;\mathbf{P},\tilde{\delta}) \leq C_1\big] \geq F_\Delta(\rho_\Delta) - C_2(\tilde{\delta})$, where C_1 is independent of $\tilde{\delta}$, and $C_2(\tilde{\delta})$ is non-increasing with $\tilde{\delta}$ (at least for small values of $\tilde{\delta}$, as displayed in Figure 2); that is, for $\tilde{\delta}_2 < \tilde{\delta}_1$, we have $C_2(\tilde{\delta}_1) \leq C_2(\tilde{\delta}_2)$. Consequently, it follows that $\mathbb{P}\big[\gamma(\tilde{d}_\theta\,;\mathbf{P},\tilde{\delta}_1) \leq C_1\big] \geq F_\Delta(\rho_\Delta) - C_2(\tilde{\delta}_2)$.

Theorem 4 does not provide a direct comparison of the probability $\mathbb{P}[\gamma(\tilde{d}_\theta\,;\mathbf{P},\tilde{\delta}) \le C_1]$ across different values of $\tilde{\delta}$, but rather a lower bound that improves with increasing $\tilde{\delta}$. This property suggest that, given $\tilde{\delta}_2 > 0$, there should exist some $\tilde{\delta}_2 < \tilde{\delta}_1$ large enough for which $\mathbb{P}[\gamma(\tilde{d}_\theta\,;\mathbf{P},\tilde{\delta}_2) \le C_1] \le \mathbb{P}[\gamma(\tilde{d}_\theta\,;\mathbf{P},\tilde{\delta}_1) \le C_1]$. We observe this behavior in Figure 3 as $\mathbb{P}[\gamma(\tilde{d}_\theta\,;\mathbf{P},\tilde{\delta}) \le C_1]$ increases as $\tilde{\delta}$ grows. Moreover, as the projection dimension k increases, the maximum gap decreases, confirming that higher-dimensional projectors improve the approximation quality. ■

4.2.2 On a probabilistic feasibility bound with approximation guarantees

We establish probabilistic approximation guarantees for the feasibility problem (8); see Section 3.3. All proofs for this section are deferred to Appendix B.2. The key step is to invoke Theorem 3 to certify the feasibility of (8) and to translate that certification into an approximation guarantee.

Theorem 5. *Let $\mathbf{P} \in \mathbb{R}^{k\times q}$ be a random projector. Define:*

$$\Delta \equiv \Delta(\mathbf{F},\mathbf{c}) = \max\left\{\left|\det(\mathbf{M})\right| \;:\; \mathbf{M} \text{ is a square submatrix of } [\,\mathbf{F}^\top \,|\, -\mathbf{F}^\top \,|\, \mathbf{c} \,|\, \mathbf{1}_m\,]\right\}.$$

Then, given $\delta \in (0, \min\left\{\varphi_\theta(\mathbf{x}) \,:\, \mathbf{x} \in \mathcal{X}\right\})$, we have that the following inequality is satisfied:

$$z^* - \|\mathbf{d}\|_1 m\Delta(\mathbf{F},\mathbf{c})\delta \le s_\theta(\mathbf{P},\delta) \le z^*,$$

with probability at least $1 - 4me^{-C\left(\varepsilon^2-\varepsilon^3\right)k}$, where $\varepsilon = \mathcal{O}\left(\frac{\delta}{1+\theta}\right)$.

Theorem 5 offers a systematic approach for computing approximate solutions to bilevel linear programs with continuous variables at both levels in polynomial time, albeit with probabilistic guarantees. This result might seem like a "free lunch," since these problems are strongly NP-hard (Hansen et al. 1992). Yet, as we discuss in Section 4.4, there is no free lunch: the guarantees are inherently probabilistic, and simple repetition-and-averaging cannot serve as a derandomization strategy.

In fact, the proof of Theorem 5 provides a stronger guarantee than merely bounding the leader's optimal objective function value. Indeed, it provides a bound for the distance between the approximate solution $(\mathbf{x}^\delta, \mathbf{y}^\delta)$ obtained with (7) and the corresponding pair $\left(\mathbf{x}^\delta, \mathbf{y}^*(\mathbf{x}^\delta)\right)$. In other words, the follower's approximate response remains provably close to the follower's rational response.

Corollary 2. *Let $\delta \in (0, \min\{\varphi_\theta(\mathbf{x}) \,:\, \mathbf{x} \in \mathcal{X}\})$, $\mathbf{P} \in \mathbb{R}^{k\times q}$ be a random projector. Then, with probability at least $1 - 4me^{-C\left(\varepsilon^2-\varepsilon^3\right)k}$, where $\varepsilon = \mathcal{O}\left(\frac{\delta}{1+\theta}\right)$, the set $\mathcal{S}_\theta(\mathbf{P},\delta)$ is nonempty, and,*

$$\min\left\{\|(\mathbf{x}^\delta,\mathbf{y}^\delta) - (\mathbf{x}^\delta, y^*(\mathbf{x}^\delta))\|_\infty \,:\, y^*(\mathbf{x}^\delta) \in \mathcal{R}(\mathbf{x}^\delta)\right\} \le m\Delta\big(\mathbf{F},\mathbf{c}\big)\delta,$$

is satisfied for all $\left(\mathbf{x}^\delta,\mathbf{y}^\delta\right) \in \mathcal{S}_\theta(\mathbf{P},\delta)$.

When computing the optimal objective function value $s_\theta(\mathbf{P},\delta)$ from the feasibility problem (8), there are two sources of infeasibility in general. The first is that the feasible region defined in (7) may itself be empty due to the follower's optimality-type of constraint. The second arises from coupling constraints that jointly involve leader's and follower's decision variables in the upper-level constraints. As these coupling constraints are excluded throughout this section, i.e., $\mathbf{H} \equiv \mathbf{0}_{p\times m}$, feasibility of (8) follows directly from that of (7).

Re-parametrization. For practical purposes similar to Section 4.2.1, we introduce a re-parametrization of $s_\theta(\mathbf{P},\delta)$ which takes as input a relative tolerance $\tilde{\delta} > 0$. The problem is given by:

$$\tilde{\mathcal{S}}_\theta(\mathbf{P},\tilde{\delta}) := \max_{\mathbf{x},\mathbf{y}} \left\{ \mathbf{a}^\top \mathbf{x} + \mathbf{d}^\top \mathbf{y} \,:\, \mathbf{x} \in \mathcal{X}\,,\, \mathbf{y} \in \mathcal{Y}_\theta(\mathbf{x})\,,\, \mathbf{c}^\top \mathbf{y} \le \varphi_\theta(\mathbf{x},\mathbf{P})(1+\tilde{\delta}) \right\}.$$

and the re-parametrized feasibility problem thus becomes:

$$\tilde{s}_\theta(\mathbf{P},\tilde{\delta}) = \max \left\{ \mathbf{a}^\top \mathbf{x} + \mathbf{d}^\top \mathbf{y}^*(\mathbf{x}) \,:\, (\mathbf{x},\mathbf{y}) \in \tilde{\mathcal{S}}_\theta(\mathbf{P},\tilde{\delta}) \right\}.$$

Next, we show that a variant of Theorem 5 applies to the re-parametrized feasibility problem.

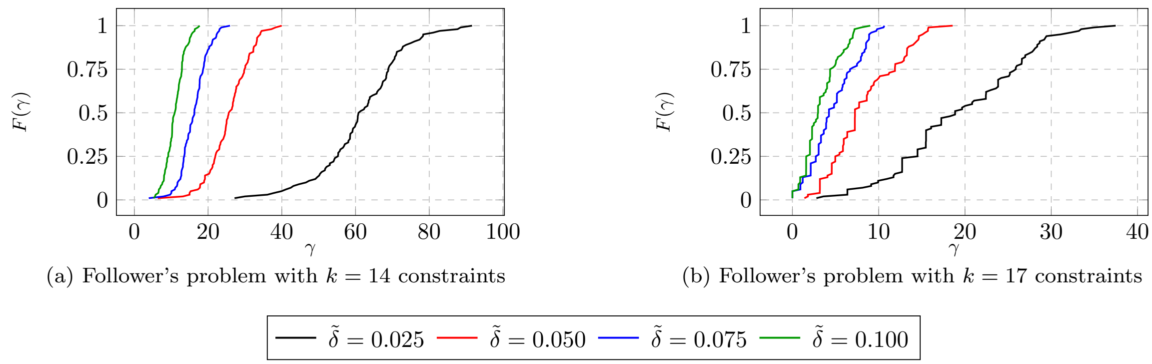
Corollary 3. *Let $\mathbf{P} \in \mathbb{R}^{k\times q}$ be a random projector. Then, there exists a sufficiently small value $\tilde{\delta} \in \left(0, \left(\|\mathbf{c}\|_\infty \theta\right)^{-1} \min\left\{\varphi_\theta(\mathbf{x}) \,:\, \mathbf{x} \in \mathcal{X}\right\}\right)$ such that the following inequality is satisfied:*

$$z^* \le \tilde{s}_\theta(\mathbf{P},\tilde{\delta}) + \|\mathbf{d}\|_1 m \Delta(\mathbf{F},\mathbf{c}) \|\mathbf{c}\|_\infty \theta \tilde{\delta},$$

with probability at least $1 - 4me^{-C\left(\varepsilon^2-\varepsilon^3\right)k}$, where $\varepsilon = \mathcal{O}\left(\frac{\tilde{\delta}}{1+\theta} \min\left\{\varphi_\theta(\mathbf{x}) \,:\, \mathbf{x} \in \mathcal{X}\right\}\right)$.

Then, in Example 3, we present a numerical illustration of the trade-offs identified in Theorem 5 and Corollary 3. Specifically, there exists a trade-off between the size of the gap between the leader's optimal objective function value z^* and the optimal objective function value $s_\theta(\mathbf{P},\delta)$ of the feasibility problem (8), and the probability that the corresponding $\mathcal{O}(\delta)$ bound on this gap holds.

Example 3. We borrow the instance and computational setup from Example 2. We consider a random projection matrix $\mathbf{P} = (P_{ij})_{ij}$ with entries drawn i.i.d from the Gaussian distribution $P_{ij} \sim \mathcal{N}(0, 1/k)$, for $i \in [k]$, $j \in [q]$. We vary the projection dimension $k \in \{14, 17\}$ and the tolerance $\tilde{\delta} \in \{0.025, 0.050, 0.075, 0.100\}$. In particular, we draw 100 realizations of $\mathbf{P}$ and compute both z^* and $\tilde{s}(\mathbf{P},\tilde{\delta})$ for each pair $(k,\tilde{\delta})$. Then, we measure the performance of the re-parametrized projected feasibility problem by using the empirical distribution of the random variable induced by the difference between z^* and $\tilde{s}_\theta(\mathbf{P},\tilde{\delta})$, i.e., $\gamma \equiv \gamma(\tilde{s}_\theta\,;\,\mathbf{P},\tilde{\delta}) \equiv \frac{1}{\tilde{\delta}}\left(z^* - \tilde{s}_\theta(\mathbf{P},\tilde{\delta})\right)$. We study its empirical distribution over the random draws of the random projector $\mathbf{P}$; see the results in Figure 4.



(a) Follower's problem with $k = 14$ constraints

(b) Follower's problem with $k = 17$ constraints

Figure 4: Empirical cumulative distribution $F(\gamma)$ of the gap $\gamma \equiv \gamma(\tilde{s}_\theta\, ; \mathbf{P}, \tilde{\delta}) \equiv \frac{1}{\tilde{\delta}}\big(z^* - \tilde{s}_\theta(\mathbf{P}, \tilde{\delta})\big)$ for the instance describe in Example 3, where $\mathbf{P} \in \mathbb{R}^{k \times q}$ is a random projector with i.i.d. entries $P_{ij} \sim \mathcal{N}(0, 1/k)$; results are shown for different numbers of follower's constraints $k \in \{14, 17\}$ and values of $\tilde{\delta} \in \{0.025, 0.050, 0.075, 0.100\}$.

Performance. Although Theorem 5 does not directly compare the probabilities $\mathbb{P}[\gamma(\tilde{s}_\theta; \mathbf{P}, \tilde{\delta}) \leq C_1]$ for different values of $\tilde{\delta}$, it provides a lower bound which itself becomes larger as $\tilde{\delta}$ increases. Empirically, Figure 4 illustrates that this probability grows with $\tilde{\delta}$: larger tolerances shift the distribution of γ to the left, reducing the mass in the upper tail. Conversely, very small tolerances (e.g., $\tilde{\delta} = 0.025$) make the gap's distribution left-skewed, making large deviations more likely. Moreover, increasing the projector's dimension k decreases the maximum observed gap. In turn, this observation confirms that projectors of higher dimension improve the approximation quality. ■

4.3 On the challenge of coupling constraints

In Sections 4.2.1 and 4.2.2, we develop probabilistic guarantees when using random projection techniques to approximately solve the leader's problem (1). To isolate the effect of projection on solution quality, we assume that the leader's problem contains no coupling constraints. Recall that a coupling-constraint violation occurs when a leader's decision and the corresponding follower's rational response violates the upper-level constraints in (1). Appendix B.3 exhibits instances in which the leader's decision obtained from the feasibility problem (8) and the adjusted surrogate problem (10) do not satisfy the coupling constraints; the appendix also contains the proofs for this section.

The remainder of this section tackles the coupling-constraints issue for the feasibility problem studied in Section 4.2.2. Specifically, Section 4.3.1 quantifies the probability of violating the coupling constraints, and Section 4.3.2 shows how to preemptively modify (8) to keep those violations under control. Moreover, analogous concerns arise for the adjusted surrogate dual problem approach from Section 4.2.1. The techniques developed below, and their associated trade-offs, extend to that setting; we therefore omit a formal treatment to avoid repetition and for brevity.

4.3.1 Probability of violating coupling constraints

We adopt a probabilistic approach to quantify the likelihood with which the upper-level coupling constraints are violated. Throughout, the matrix $\mathbf{H} \in \mathbb{R}^{p \times m}$ in (1) is treated as random with independent and identically uniformly distributed entries $H_{j\ell} \overset{\text{i.i.d}}{\sim} \mathcal{U}([-a, a])$ for some $a > 0$ on a probability space $(\Omega, \mathcal{B}, \mathbb{P})$. This setting echoes the analysis in Bertsimas and Sim (2004), where similar assumptions are used to study constraint-violation probabilities in robust optimization.

To proceed with our analysis, we denote by $\left(\mathbf{x}^\delta, \mathbf{y}^\delta\right)$ the optimal solution to (7). We aim to bound the probability that $(\mathbf{x}^\delta, \mathbf{y}^*(\mathbf{x}^\delta))$ does not satisfy the coupling constraints. For that, we fix a random projector $\mathbf{P}$ and make the following assumptions:

- **A4 :** $\mathbf{H}$ and $\mathbf{P}$ are independent.
- **A5 :** $\mathbb{P}[z^* < +\infty] = \mathbb{P}[s_\theta(\mathbf{P}, \delta) < +\infty] = 1$.
- **A6 :** $\mathcal{R}(\mathbf{x}) = \{\mathbf{y}^*(\mathbf{x})\}$, for all $\mathbf{x} \in X$

While Assumption **A4** is mild as one can pick a random projector independent of $\mathbf{H}$, Assumption **A5** is admittedly restrictive. We require this additional assumption to simplify our analysis as they ensure the existence of a solution to both the leader's problem (1) and (7). Moreover, Assumption **A6** ensures that the follower's problem has a unique optimal solution for each leader's decision, thereby simplifying the presentation of our results. This assumption allows us to avoid dealing with multiple follower's responses while still capturing the main trade-offs. We believe that relaxing this assumption is of great interest for future research in a broader context than our approach to better understand the effect of uncertainty on coupling constraints in bilevel programming.

Next, we derive an upper bound on the probability of violating the coupling constraints. The next result lays out the main trade-offs between projection error and coupling constraint violation.

Proposition 1. *Let $\mathbf{P} \in \mathbb{R}^{k \times q}$ be a random projector. Fix $\delta \in (0, \min\{\varphi_\theta(\mathbf{x}) : \mathbf{x} \in \mathcal{X}\})$ and define the minimum coupling slack $\Delta^\delta := \min\left\{h_j - \left(\mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^\delta\right) : j \in [p]\right\}$. Assume that, conditional on Δ^δ, the entries $\{H_{j\ell}\}_{j\in[p],\,\ell\in[m]}$ are independent and identically distributed with $H_{j\ell} \sim \mathcal{U}([-a, a])$. Then, the probability that at least one coupling constraint is violated satisfies:*

$$\mathbb{P}\left[\exists\, j \in [p]\,,\, \mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) > h_j\right] \leq p \cdot \mathbb{E}_{\Delta^\delta}\left[\exp\left(-\frac{\left(\Delta^\delta\right)^2}{2m^3a^2\Delta(\mathbf{F},\mathbf{c})^2 \cdot \delta^2}\right) + 1 - p_\Delta(\delta, \tilde{C})\right],$$

where $p_\Delta(\delta, \tilde{C})$ is as defined in Theorem 4 with another positive constant $\tilde{C} > 0$.

The random variable Δ^δ in Proposition 1 measures the slack of the solution $\left(\mathbf{x}^\delta, \mathbf{y}^\delta\right)$ to (7) with respect to the upper-level constraints. By construction, Δ^δ is non-negative $\mathbb{P}$-almost surely, and it is equal to zero if and only if at least one constraint is tight. Hence, Δ^δ serves as a direct indicator of how close the solution $\left(\mathbf{x}^\delta, \mathbf{y}^\delta\right)$ is to violating an upper-level constraint.

Proposition 1 uses a strong assumption. Because Δ^δ is computed from the coupling slacks, conditioning on Δ^δ restricts $\mathbf{H}$ to values that, together with $\mathbf{y}^\delta$, produce that slack level. This conditioning creates dependence between $\mathbf{H}$ and $\mathbf{y}^\delta$ and can shift the conditional mean, so that $\mathbb{E}[\mathbf{H} \mid \Delta^\delta] \neq 0$. To avoid this issue, we assume that, given Δ^δ, the entries of $\mathbf{H}$ are i.i.d. uniformly distributed. This assumption preserves unbiasedness and entrywise independence after conditioning, and it allows us to apply Hoeffding-type concentration bounds. The assumption is admittedly restrictive and is used to highlight the main trade-offs; below we state a corollary under a more interpretable condition.

If $\mathbb{P}\left[\Delta^\delta = 0\right] = 1$, then at least one upper-level constraint is always tight. In that case, unsurprisingly, the bound in Proposition 1 becomes uninformative. Suppose that one of the constraints (say, indexed by $j \in [p]$) is always active at $\left(\mathbf{x}^\delta, \mathbf{y}^\delta\right)$. Then, when replacing $\mathbf{y}^\delta$ with the follower's rational response $\mathbf{y}^*(\mathbf{x}^\delta)$, it is unclear whether the upper-level constraint $\mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) \leq h_j$ remains satisfied. Indeed, the term $\mathbf{H}_j\left(\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\right)$ might take either sign, regardless of the value of δ.

Proposition 1 highlights three main drivers of the probability of upper-level constraints' violation. First, provided that the dependence of k on δ is carefully chosen (to balance the trade-off for $p_\Delta(\delta, \tilde{C})$ from Theorem 4), then the bound decreases as δ becomes smaller. Indeed, the closer the estimated response is to the true one, the lower the risk of infeasibility, especially if the constraint slack at the upper level is non-negligible. Second, a larger value of Δ^δ directly improves the bound, since more slack makes it more likely that the constraints remain satisfied for the follower's rational response $\mathbf{y}^*(\mathbf{x}^\delta)$, especially if it is closed to $\mathbf{y}^\delta$. Third, the bound worsens with larger values of a, reflecting higher uncertainty in the random matrix $\mathbf{H}$. Viewed differently, higher uncertainty in $\mathbf{H}$ necessitates a smaller deviation between $\mathbf{y}^\delta$ and $\mathbf{y}^*(\mathbf{x}^\delta)$ in order to maintain the same quality of the bound.

Next, we aim to reduce the dependence of the bound in Proposition 1 on the random variable Δ^δ, whose distribution is generally unknown and seemingly difficult to characterize. If the gap Δ^δ is bounded bellow $\mathbb{P}$-almost surely, so that the upper-level constraints are never tight at $\left(\mathbf{x}^\delta, \mathbf{y}^\delta\right)$, then the bound in Proposition 1 can be simplified. Formally:

Corollary 4. *Let δ and Δ^δ be as defined in Proposition 1. Assume that there exists $\eta > 0$ that satisfies $\mathbb{P}\left[\Delta^\delta > \eta\right] = 1$. Then, the probability that at least one coupling constraint is violated satisfies:*

$$\mathbb{P}\left[\exists\, j \in [p]\,,\, \mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) > h_j\right] \leq p \cdot \left(\exp\left(-\frac{\eta^2}{8m^3a^2\Delta(\mathbf{F}, \mathbf{c})^2 \cdot \delta^2}\right) + 1 - p_\Delta(\delta, \tilde{C})\right).$$

Proposition 1 relies on the assumption that, conditional on Δ^δ, the entries $\{H_{j\ell}\}_{j\in[p],\,\ell\in[m]}$ are independent and identically distributed with $H_{j\ell} \sim \mathcal{U}([-a,a])$. To avoid this assumption, Corollary 4 replaces it with the margin assumption $\mathbb{P}[\Delta^\delta > \eta] = 1$, which is more interpretable and gives a comparable constraint-violation bound. It is possible to obtain a similar bound without this assumption. Indeed, one may bypass the assumption $\mathbb{P}[\Delta^\delta > \eta] = 1$ by carefully choosing the parameter η instead; see Proposition 2 bellow. As a first step, we establish an intermediate lemma of independent interest that somehow characterizes the distribution of Δ^δ. Formally:

Lemma 2. *Fix an upper-level constraint's index j in $[p]$ and $\delta \in \big(0, \min\{\varphi_\theta(\mathbf{x}) : \mathbf{x} \in \mathcal{X}\}\big)$. Next, we assume that $h_j - \max\{\mathbf{G}_j\mathbf{x} : \mathbf{x} \in \mathcal{X}\} > 0$. Then, for all $\alpha \in (0,1)$, there exists some finite constant $\tau_j \equiv \tau_j(\alpha,\theta,a)$ such that the following inequality is satisfied:*

$$\mathbb{P}\Big[\mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^\delta > h_j - \tau_j\Big] \leq \alpha.$$

Moreover, there exists some $\alpha_j \equiv \alpha_j(\theta,a)$ such that, for all $\alpha > \alpha_j$, we have $\tau_j(\alpha,\theta,a) > 0$.

Lemma 2 is the key ingredient in proving Proposition 2, which quantifies the probability that at least one coupling constraint is violated. The argument rests on a careful use of the Irwin–Hall distribution (Irwin 1927; Hall 1927). Formally, we show:

Proposition 2. *Let $\mathbf{P} \in \mathbb{R}^{k\times q}$ be a random projector and assume that $h_j - \max\{\mathbf{G}_j\mathbf{x} : \mathbf{x} \in \mathcal{X}\} > 0$. Fix $\delta \in \big(0, \min\{\varphi_\theta(\mathbf{x}) : \mathbf{x} \in \mathcal{X}\}\big)$. Also, let $\alpha \in (0,1)$ be the smallest value that satisfies:*

$$\tau \equiv \tau(\alpha) := \min\big\{h_j - \max\{\mathbf{G}_j\mathbf{x} : \mathbf{x} \in \mathcal{X}\} - a\theta q^{IH}_{1-\alpha} \,:\, j \in [p]\big\} > 0,$$

where $q^{IH}_{1-\alpha}$ is the $(1-\alpha)$-quantile of the Irwan-Hall distribution. Then, the probability that at least one coupling constraint is violated satisfies:

$$\mathbb{P}\Big[\exists\, j \in [p]\,,\, \mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) > h_j\Big] \leq p \cdot \left(\alpha p + \exp\left(-\frac{\tau^2}{2m^3a^2\Delta(\mathbf{F},\mathbf{c})^2\delta^2}\right) + 1 - p_\Delta(\delta,\tilde{C})\right),$$

where $p_\Delta(\delta,\tilde{C})$ is as defined in Theorem 4 with another positive constant $\tilde{C} > 0$.

Proposition 2 restates Proposition 1 without any explicit reference to the distribution of Δ^δ. By using Lemma 2, it replaces the expectation over Δ^δ with two terms that depend solely on the support of $\mathbf{H}$, i.e., a. Here, τ is chosen as the $(1-\alpha)$-quantile of the minimal slack across all upper–level constraints, capturing the worst-case residual gap. Decreasing α raises this quantile (increasing τ) and thus shrinks the first term of the bound. At the same time, a smaller α drives up the quantile $q^{IH}_{1-\alpha}$, which in turn reduces τ and hence increases the second term. Consequently, there remains a trade-off: lowering α simultaneously tightens the α-dependent term and enlarges the τ-dependent term.

4.3.2 Preemptive correction for coupling constraints

Section 4.2.2 gives high-confidence control of violations of the upper-level coupling constraints but does not provide a solution approach whenever a violation occurs. A simple preventive correction is to tighten (by subtracting) every right-hand side $\mathbf{h}$ in the upper-level constraints by a uniform margin $\zeta > 0$, solve the corrected problem (provided it is feasible) to find $\left(\mathbf{x}^{\zeta}, \mathbf{y}^{\zeta}\right)$, and then compute the follower's rational response at the resulting leader's decision to obtain the pair $\left(\mathbf{x}^{\zeta}, \mathbf{y}^{*}(\mathbf{x}^{\zeta})\right)$.

Choosing $\zeta = \|\mathbf{H}\|_1 \, m\Delta\left(\mathbf{F}, \mathbf{c}\right) \delta$ ensures that coupling constraints are satisfied. Indeed, the follower's rational response $\mathbf{y}^{*}(\mathbf{x}^{\zeta})$ is within the $m \, \Delta\left(\mathbf{F}, \mathbf{c}\right) \delta$ neighborhood of its projected counterpart $\mathbf{y}^{\zeta}$ with probability at least $1 - 4m \exp\left(-C(\varepsilon^2 - \varepsilon^3)k\right)$, where $\varepsilon = \mathcal{O}\left(\frac{\delta}{1+\theta}\right)$. On that event,

$$\mathbf{G}\mathbf{x}^{\zeta} + \mathbf{H}\mathbf{y}^{*}(\mathbf{x}^{\zeta}) \;\leq\; \mathbf{h} - \zeta + \|\mathbf{H}\|_1 \left\|\mathbf{y}^{*}(\mathbf{x}^{\zeta}) - \mathbf{y}^{\zeta}\right\|_{\infty} \;\leq\; \mathbf{h},$$

so the resulting pair is bilevel feasible. The price of this correction is an $\mathcal{O}(\delta)$ loss in the leader's objective, which we bound by decomposing the error into (i) the follower-response proximity and (ii) the effect of correction. We defer the full details of the preemptive correction and the corresponding proofs to Appendix B.3.4, where we also relate the tightening step to classical proximity bounds for (mixed-)integer linear programs (Blair and Jeroslow 1977; Mangasarian and Shiau 1987).

4.4 Price of random projection in bilevel linear programming

In this section, we analyze the impact of randomness on the follower's projected value function (4). We show that, under some suitable but admittedly restrictive assumptions, the follower's projected value function follows a Cauchy distribution (Robert 2007). This observation implies that the expected value of the follower's projected value function may be undefined. Such a phenomenon has both theoretical and practical implications for the effectiveness of these methods. Specifically, although random projection techniques may often produce good approximate solutions, the same mechanisms can also generate arbitrarily poor solutions with non-negligible probability.

Our analysis relies on the probabilistic properties of the optimal solutions to random linear programs. These random problems have been studied mainly in the context of sensitivity analysis, yet such studies remain sparse and highly dependent on some strong assumptions and approximations (Babbar 1955; Prékopa 1966). We use these ideas to investigate the probabilistic behavior of randomly projected linear programs and discuss the corresponding implications for BLPs.

Throughout the section, we consider random projectors $\mathbf{P} \in \mathbb{R}^{k \times q}$ with independent and identically normally distributed entries, i.e., $P_{ij} \sim \mathcal{N}(0, 1)$ for all $i \in [k]$ and all $j \in [q]$. Then, we fix some leader's feasible decision $\mathbf{x} \in \mathcal{X}$. Also, to streamline the subsequent discussion, we relax the second part of Assumption **A2** (recall Section 1), namely, the assumption that the follower's decision variables in (2) are bounded by a constant θ.

Next, we consider the dual problem of the projected follower's problem (4). Formally:

$$\eta(\mathbf{x}, \mathbf{P}) := \max \left\{ \big(\mathbf{f} - \mathbf{L}\mathbf{x}\big)^\top \mathbf{P}^\top \mathbf{u} \, : \, \mathbf{F}^\top \mathbf{P}^\top \mathbf{u} \leq \mathbf{c} \, , \, \mathbf{u} \in \mathbb{R}^k \right\}.$$

Then, we introduce $\mathcal{P} := \big\{ \mathbf{v} \in \mathbb{R}^q \, : \, \mathbf{F}^\top \mathbf{v} \leq \mathbf{c} \big\}$ and re-parametrize this dual problem to obtain:

$$\eta(\mathbf{x}, \mathbf{P}) := \max \left\{ \big(\mathbf{f} - \mathbf{L}\mathbf{x}\big)^\top \mathbf{v} \, : \, \mathbf{v} \in \mathcal{P} \, , \, \mathbf{v} = \mathbf{P}^\top \mathbf{u} \, , \, \mathbf{u} \in \mathbb{R}^k \right\}.$$

By using the Minkowski–Weyl Theorem (Bertsimas and Tsitsiklis 1997), there exist extreme points $\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(t)}$ and extreme rays of $\mathcal{P}$, denoted by $\mathbf{w}^{(1)}, \ldots, \mathbf{w}^{(r)}$ such that $\mathcal{P}$ can be represented by:

$$\mathcal{P} = \Big\{ \sum_{i=1}^{t} \lambda_i \mathbf{e}^{(i)} + \sum_{j=1}^{r} \mu_j \mathbf{w}^{(j)} \, : \, \lambda_i \geq 0 \, , \, \mu_j \geq 0 \, , \, \mathbf{1}_t^\top \boldsymbol{\lambda} = 1 \Big\},$$

and we reformulate the follower's dual problem as follows:

$$\eta(\mathbf{x}, \mathbf{P}) := \max_{\substack{\mathbf{u} \in \mathbb{R}^k \\ (\boldsymbol{\lambda}, \boldsymbol{\mu}) \in \mathbb{R}_+^{t+r}}} \left\{ \big(\mathbf{f} - \mathbf{L}\mathbf{x}\big)^\top \mathbf{P}^\top \mathbf{u} \, : \, \mathbf{P}^\top \mathbf{u} - \sum_{i=1}^{t} \lambda_i \mathbf{e}^{(i)} - \sum_{j=1}^{r} \mu_j \mathbf{w}^{(j)} = \mathbf{0}_q \, , \, \mathbf{1}_t^\top \boldsymbol{\lambda} = 1 \right\}. \quad (11)$$

Consequently, by taking the dual of (11), we obtain the following problem:

$$\begin{aligned} \varphi(\mathbf{x}, \mathbf{P}) := \min \quad & z \\ \text{s.t.} \quad & \mathbf{P}\mathbf{s} = \mathbf{P}\big(\mathbf{f} - \mathbf{L}\mathbf{x}\big), \\ & z - \mathbf{s}^\top \mathbf{e}^{(i)} \ \geq \ 0 \qquad i = 1, \ldots, t, \\ & -\mathbf{s}^\top \mathbf{w}^{(j)} \ \geq \ 0 \qquad j = 1, \ldots, r, \\ & \mathbf{s} \in \mathbb{R}^q, \ z \in \mathbb{R}. \end{aligned}$$

Therefore, we somehow manage to "simplify" the constraints for the projected follower's problem (4) in the sense that the matrix $\mathbf{F}$ does not appear explicitly in the problem. To derive explicit probabilistic insights, we adopt the following simplifying assumptions:

- **A7:** $\mathbf{w}^{(j)} = \mathbf{0}_q$ for all $j \in [r]$, and $\mathbf{e}^{(i)} = \mathbf{0}_q$ for all $i \in [t] \setminus \{1\}$.
- **A8:** The optimal basis $B(\mathbf{P}) \subseteq [q]$, $|B(\mathbf{P})| = k$ from $\{\mathbf{s} \in \mathbb{R}^q \, : \, \mathbf{P}\mathbf{s} = \mathbf{P}\left(\mathbf{f} - \mathbf{L}\mathbf{x}\right)\}$ does not depend on the realization of $\mathbf{P}$, i.e., $B(\mathbf{P}) = B_0$. Also, we assume that $B_0 = [k]$.

Consequently, Assumptions **A7** ensures that the projected follower's problem can be reformulated as the following random linear program:

$$\varphi(\mathbf{x},\mathbf{P}) := \min\left\{\left(\mathbf{e}^{(1)}\right)^{\top}\mathbf{s} \,:\, \mathbf{P}\mathbf{s} = \mathbf{P}\left(\mathbf{f}-\mathbf{L}\mathbf{x}\right)\,,\, \mathbf{s}\in\mathbb{R}^{q}\right\}. \tag{13}$$

Before proceeding, we discuss the assumptions above. The first part of Assumption **A7** ensures that the feasible region of the follower's dual problem is bounded. Moreover, we assume that only a single extreme point is non-zero in the decomposition. While admittedly restrictive, this assumption simplifies the analysis by eliminating the piecewise linear structure that otherwise arises in the objective function. Extending our results to the general case, where multiple extreme points are active, appears possible, but would likely require a more intricate application of the polyhedral theory.

Assumption **A8** is unquestionably the most restrictive. This assumption is common in the limited body of work that analyzes the distribution of the optimal solution to random linear systems (Babbar 1955; Prékopa 1966). Without this assumption, each realization of the random projector $\mathbf{P}$ may induce a different optimal basis $B(\mathbf{P})$, and consequently, the distribution of the optimal solution becomes a mixture over potentially many distinct regimes. Assumption **A8** thus serves to simplify the algebraic derivations that follow by ensuring a unimodal distribution for the solution to the projected system. If this assumption is relaxed, then the resulting distribution may become multi-modal, significantly complicating the analysis. Extending the observations bellow to this more general case remains an exciting but challenging direction for future research.

Note that, as the optimal basis is assumed to be given by $B_0 = [k]$ and we denote by NB_0 the indexes $\{k+1,\ldots,q\}$. Therefore, we decompose our projector accordingly with $\mathbf{P} = \left[[\mathbf{P}]_{B_0} \quad [\mathbf{P}]_{NB_0}\right]$. Hence, we have that the optimal solution $\mathbf{s}^*(\mathbf{x},\mathbf{P})$ to (13) can be formulated as follows:

$$\mathbf{s}^*(\mathbf{x},\mathbf{P}) := \begin{pmatrix} \mathbf{I}_{k\times k}[\mathbf{f}-\mathbf{L}\mathbf{x}]_{B_0} + [\mathbf{P}]_{B_0}^{-1}[\mathbf{P}]_{NB_0}[\mathbf{f}-\mathbf{L}\mathbf{x}]_{NB_0} \\ \mathbf{0}_{q-k} \end{pmatrix} \equiv \begin{pmatrix} \mathbf{s}^*(\mathbf{x},\mathbf{P}\,;\,B_0) + \mathbf{s}^*(\mathbf{x},\mathbf{P}\,;\,NB_0) \\ \mathbf{0}_{q-k} \end{pmatrix},$$

where the non-basic variables are all null, and the basic variables can be decomposed into two components. In particular, we are interested in deriving the distribution of the second component:

$$\mathbf{s}^*(\mathbf{x},\mathbf{P}\,;\,NB_0) := [\mathbf{P}]_{B_0}^{-1}[\mathbf{P}]_{NB_0}[\mathbf{f}-\mathbf{L}\mathbf{x}]_{NB_0}.$$

To proceed, $\mathbf{s}^*(\mathbf{x},\mathbf{P}\,;\,NB_0)$ is the solution to the following random linear system:

$$[\mathbf{P}]_{B_0}\mathbf{s} = [\mathbf{P}]_{NB_0}[\mathbf{f}-\mathbf{L}\mathbf{x}]_{NB_0}.$$

Moreover, observe that:

$$[\mathbf{P}]_{B_0} \sim \mathcal{N}\left(\mathbf{0}_{B_0\times B_0}, \mathbf{I}_{(B_0\times B_0)\times(B_0\times B_0)}\right) \text{ and } [\mathbf{P}]_{NB_0}[\mathbf{f}-\mathbf{L}\mathbf{x}]_{NB_0} \sim \mathcal{N}\left(\mathbf{0}_{B_0}, \mathbf{R}_k(\mathbf{x})\right),$$

where $[\mathbf{P}]_{B_0}$ and $[\mathbf{P}]_{NB_0}[\mathbf{f}-\mathbf{Lx}]_{NB_0}$ are independent; furthermore, $\mathbf{R}_k(\mathbf{x})$ is the covariance matrix induced by the linear mapping $[\mathbf{f}-\mathbf{Lx}]_{NB_0} \to [\mathbf{P}]_{NB_0}[\mathbf{f}-\mathbf{Lx}]_{NB_0}$. Hence, using the Theorem 14.1.1 by Girko (2012), we have that $\mathbf{s}^*(\mathbf{x}, \mathbf{P}\,; NB_0)$ follows a distribution F_s with density f_s given by:

$$f_s(\mathbf{s}) := c \cdot (\det(\mathbf{R}_k(\mathbf{x}))^{1/2} \left(1 + \left(\mathbf{s}^\top \mathbf{R}_k(\mathbf{x})\mathbf{s}\right)\right), \quad \mathbf{s} \in \mathbb{R}^{q-k},$$

for some normalizing constant $c \in \mathbb{R}$.

The distribution F_s has two salient features (Girko 2012). It is heavy-tailed, with power-law rather than exponential decay; and its expectation is often not well defined in the usual (Lebesgue) sense. In the special case $k = q - 1$, F_s reduces to the Cauchy distribution, which has no finite expectation or variance. Accordingly, the law of large numbers fails as the average of i.i.d. Cauchy random variables remains Cauchy distributed (Cohen 2012).

This central observation in our analysis has significant implications. First, the follower's projected value function may not possess a finite expectation due to heavy-tailed behavior. Therefore, derandomization approaches based on repeatedly sampling random projection matrices and averaging the resulting solutions are not theoretically supported in a general sense. Put differently, this means that randomization through projections may not provide a computational shortcut or "free lunch" to circumvent the inherent computational complexity of bilevel programs.

Second, with non-negligible explicit probability, the projected follower's value function may be arbitrarily far away from the original one. In other words, although random projection may often perform acceptably, including in our computational experiments in Section 6, it explicitly carries a significant risk of producing highly inaccurate solutions. As such, practitioners should apply this randomization-based methods with caution, fully mindful of its heavy-tailed behavior and the absence of a well-defined expectation established in this analysis.

5 Second Sandwich Theorem

In this section, we build upon the surrogate duality theory from Section 3, and in particular the first Sandwich Theorem (Theorem 1 in Section 3.4), to construct a family of projectors that provide approximation guarantees. Specifically, we establish a two-sided bound that quantifies the deviation from the leader's optimal objective function value z^*. Under mild conditions relating the parameters k and δ, there exist constants $C_1, C_2 > 0$ and a nonempty family of projectors $\mathcal{P}_\delta^{(k)} \subseteq \mathbb{R}^{k\times q}$ such that, for every $\mathbf{P} \in \mathcal{P}_\delta^{(k)}$, the feasibility-based lower bound $s_\theta(\mathbf{P}, \delta)$ and the surrogate-dual upper bound $s_\theta(\mathbf{P})$ approximate z^* with errors bounded by $C_1\delta$ and $C_2\delta$, respectively.

Consequently, as $\delta \to 0$ with k increased appropriately (so that tighter tolerances are paired with a richer set of projectors), both bounds converge to z^*. As in Section 3, we work with no coupling constraints (i.e., $\mathbf{H} \equiv \mathbf{0}_{p\times m}$), which simplifies the analysis without much loss of generality (Henke et al. 2025). The construction of $\mathcal{P}_\delta^{(k)}$, the precise link between k and δ (including how k scales as δ decreases), and the full proof are deferred to Appendix C, with the main intuition provided here-bellow.

Theorem 6 (Second Sandwich Theorem)**.** *If k and δ are chosen so that the projector family $\mathcal{P}_\delta^{(k)}$ is nonempty, then there exist constants $C_1, C_2 > 0$ (independent of k, δ) such that,*

$$z^* - C_1\delta \;\le\; s_\theta(\mathbf{P}, \delta) \;\le\; z^* \;\le\; s_\theta(\mathbf{P}) \;\le\; z^* + C_2\delta,$$

for all projectors $\mathbf{P} \in \mathcal{P}_\delta^{(k)}$.

If the leader's feasible integer decisions are fixed, then the leader's problem (1) is a parametrized bilevel LP with continuous variables at both levels. Hence, the leader's optimal solution corresponds to an extreme point of the polyhedron defined by the parameterized single-level relaxation. This relaxation is obtained by dropping the follower's optimality conditions (Bard and Falk 1982; Bialas and Karwan 1982). Because the leader's feasible region is a bounded polyhedron (recall Assumption **A2** in Section 1), it has finitely many extreme points (Bertsimas and Tsitsiklis 1997). Consequently, there is a finite set $\Lambda \subseteq \mathcal{X}$ whose convex hull contains every leader's optimal decision.

To control the approximation error, we require δ-accuracy of the follower's projected value function on Λ, namely, $\varphi_\theta(\mathbf{x}) - \delta \le \varphi_\theta(\mathbf{x}, \mathbf{P}) \le \varphi_\theta(\mathbf{x})$ for all $\mathbf{x} \in \Lambda$. Random Gaussian projectors satisfy this property simultaneously for every $\mathbf{x} \in \Lambda$ with high probability provided k is chosen large enough; notably, the needed dimension grows only *logarithmically* in $|\Lambda|$. This uniform guarantee suffices as any leader's optimal decision can be recovered from Λ. Propagating the δ-error control to the leader's objective function via proximity theory then delivers the desired optimality gap of order $\mathcal{O}(\delta)$. The argument establishes existence but is not constructive; developing deterministic and computable constructions of $\mathcal{P}_\delta^{(k)}$ remains an open direction for future research.

6 Toward a numerical understanding of random projection ideas

While a full computational study is beyond this paper's scope, we present preliminary numerical results highlighting the potential of these methods and directions for future computational work. Section 6.1 first describes the experimental design. Section 6.2 compares several projection schemes across different types of constraints. Finally, Section 6.3 synthesizes the findings and situates them within the broader literature on random-projection techniques in mathematical programming.

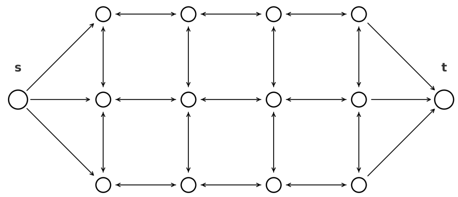


Figure 5: Directed grid network with $n_1 = 3$ rows and $n_2 = 4$ columns, including a source node s (left) and a sink node t (right), generated using variant V3 described in Royset and Wood (2007). For the full formal details of the generation procedure for grids V1, V2 and V3, we refer to their study.

6.1 Experimental setup

We center our experiments on the min-cost flow interdiction problem (Smith and Song 2020), where a leader removes edges from a graph, and a follower solves the min-cost flow problem on the remaining graph. This setting is attractive because (i) the follower's flow constraints are ideal for gauging how random projectors may distort any exploitable structure, and (ii) the interdiction constraints are mutually orthogonal, so any projector compresses sparse constraints into denser ones.

Minimum cost flow interdiction problem. Let $G = (V, E)$ be a directed graph with source s and sink t. For each edge $e \in E$, let u_e be the corresponding capacity. The leader chooses $\mathbf{x} \in \{0,1\}^{|E|}$, with $x_e = 1$ if edge e is interdicted, and interdiction budget $\sum_{e\in E} x_e \leq R$. The follower, in turn, solves the min-cost flow problem over the surviving network. Formally:

$$
\begin{aligned}
z^* := \max_{\mathbf{x}\in\{0,1\}^{|E|}} \quad & \min_{\mathbf{f}} \sum_{e\in E} c_e f_e \\
\text{s.t.} \quad & \sum_{e\in E} x_e \leq R \\
& \sum_{e\in\delta^+(v)} f_e - \sum_{e\in\delta^-(v)} f_e = \mathbf{1}(v = s) - \mathbf{1}(v = t) \quad \forall v \in V \\
& 0 \leq f_e \leq u_e(1 - x_e) \quad \forall e \in E,
\end{aligned}
$$

where $\delta^+(v)$ and $\delta^-(v)$ are the edge sets of the form $(v, i) \in E$, and $(i, v) \in E$, respectively, for $i \in V$.

Test instances. We use the grid networks from Royset and Wood (2007) introduced for the maximum-flow interdiction problem; see Figure 5 for an example. Specifically, a grid with n_1 rows and n_2 columns contains $|V| = n_1 n_2 + 2$ nodes, where the additional two nodes are the source s and sink t. Edges e are independently assigned capacities and costs in the form $u_e \overset{i.i.d.}{\sim} \mathcal{U}(\{1, \ldots, 50\})$ and $c_e \overset{i.i.d.}{\sim} \mathcal{U}(\{1, \ldots, 10\})$, respectively, where $\mathcal{U}$ is the uniform distribution. Also, we fix the interdiction budget at $R = 0.1 \times |E|$. As in Royset and Wood (2007), we consider three variants (V1 to V3) of these grid networks, which differ only in their generation procedure.

Hardware and software. All experiments are implemented in Python 3.10 and executed on a cluster comprising 32 Intel(R) Xeon(R) Gold 6126 CPUs, each running at 2.60GHz under Ubuntu 22.04.3. All mixed-integer programs are solved using Gurobi 11.0.0.

Random projectors. In our computational experiments, we employ two types of random projectors, denoted by $\mathbf{P} \equiv (P_{ij})_{i\in[k],j\in[q]}$. The first, and perhaps the most classical one, assigns entries P_{ij} as independent and identically distributed random variables with $P_{ij} \sim \mathcal{N}(0, 1/k)$ for $i \in [k]$, and $j \in [q]$. Importantly, the theoretical results developed so far assume that the follower's problem is subject only to equality constraints. In particular, applying the above projector to inequality constraints may alter the direction of the inequalities.

To address this issue, we also consider a second class of projectors introduced by Poirion et al. (2023). These projectors have the distinctive property of preserving the direction of inequalities: letting $S_{ij} \sim \mathcal{N}(0, 1)$, we define $P_{ij} := \frac{1}{k}S_{ij}^2$, so that $\mathbf{P} \geq \mathbf{0}$ entrywise. When applied to linear programs with inequality constraints, these projectors admit theoretical guarantees that parallel Theorem 3. Consequently, one can derive probabilistic guarantees analogous to those in Section 4 whenever the follower's problem includes inequality constraints. As these extensions closely parallel our results, we omit formal statements and proofs to streamline the exposition and due to space limitation.

Benchmark. For each projection matrix $\mathbf{P}$, and given gaps $\delta_f > 0$ and $\delta_d > 0$, we compute two bounds: the *feasibility bound* $\tilde{s}_\theta(\mathbf{P}, \delta_f)$ by solving the re-parametrized version of (8), and the *dual bound* $\tilde{d}_\theta(\mathbf{P}, \delta_d)$ by solving the re-parametrized version of (10); see Sections 4.2.2 and 4.2.1, respectively. In both cases, we fix $\theta := \|\mathbf{u}\|_1$, which represents the total capacity in the network.

In addition, we compute the leader's optimal objective function value z^* by solving (1). All three problems are solved using the commercial MILP solver Gurobi (2024). In particular, they are solved via single-level reformulations in which the follower's (or projected follower's) problem is replaced by its optimality conditions, following the strong duality-based approach (Zare et al. 2019). To avoid introducing additional parameters, we do not linearize the resulting constraints (i.e., we avoid big-M formulations). Consequently, the reformulated problems involve quadratic constraints, which are handled directly using Gurobi's quadratic programming capabilities.

As a reference point for the *dual bound*, we compute an upper bound by solving a single-level (also known as high-point) relaxation in which the follower's optimality conditions are omitted. Formally:

$$z_{\text{relax}} := \max_{\mathbf{x},\mathbf{y}} \left\{ \mathbf{a}^\top \mathbf{x} + \mathbf{d}^\top \mathbf{y} \;:\; \mathbf{G}\mathbf{x} + \mathbf{H}\mathbf{y} \leq \mathbf{f},\; \mathbf{y} \in \mathcal{Y}_\theta(\mathbf{x}),\; \mathbf{x} \in \{0,1\}^{n_1} \times \mathbb{R}_+^{n_2} \right\},$$

and we denote the corresponding optimal solution by $(\mathbf{x}^*, \mathbf{y}^*)$. Naturally, we have $z^* \leq z_{\text{relax}}$.

To construct a *feasibility bound*, we evaluate $z_{\text{feas}} := \mathbf{a}^\top \mathbf{x}^* + \mathbf{d}^\top \mathbf{y}^*(\mathbf{x}^*)$, where $\mathbf{y}^*(\mathbf{x}^*)$ is the follower's rational response, given the leader's decision $\mathbf{x}^*$. Note that, in general, if $\mathbf{H}$ has some non-zero entries, then z_{feas} may not constitute a valid lower bound for z^*, since the pair $(\mathbf{x}^*, \mathbf{y}^*(\mathbf{x}^*))$ might violate the coupling constraints. However, in the specific problem considered here, there are no such coupling constraints, so that $z_{\text{feas}} \leq z^*$ is guaranteed to hold.

Performance metrics. We assess the effectiveness of random projection techniques by analyzing the quality of both lower and upper bounds in comparison to the leader's optimal objective function value. Specifically, we compute the feasibility gap $\gamma_f(t) = (z^* - t)/z^*$ for values of t taken from $\{z_{\text{feas}}, \tilde{s}_\theta(\mathbf{P}, \delta_f)\}$. Similarly, we evaluate the dual gap as $\gamma_d(t) = (t - z^*)/z^*$, where t belongs to $\{z_{\text{relax}}, \tilde{d}_\theta(\mathbf{P}, \delta_d)\}$. We also report the runtime (in seconds) required to solve each instance. Loosely speaking, one should compare $\tilde{s}_\theta(\mathbf{P}, \delta_f)$ against z_{feas}, and $\tilde{d}_\theta(\mathbf{P}, \delta_d)$ against z_{relax}.

6.2 Avoiding blind projection of constraints

A natural question arising from the theoretical results in Section 4 is whether random projection techniques reduce runtimes and, if so, at what cost in the optimality gap. To sharpen this question, we further ask whether different projection schemes are, in some sense, equivalent in their impact.

To that end, we examine three variants of random projection schemes: (*i*) a **naive** approach in which all constraints are projected simultaneously using a sign-preserving projector; (*ii*) a selective projection approach in which only the **interdiction** constraints are projected using a sign-preserving projector; (*iii*) a variant in which only the **flow** constraints are projected, using a standard Gaussian projector. The results below provide a snapshot of our broader computational study. Additional findings for the network variants V2 and V3 are reported in Appendix D.

Specifically, we consider grid networks with $n_1 \in \{2, 3, 4, 5\}$ rows and $n_2 = 5$ columns. For each network size, we run 10 independent trials under three different projection schemes: (*i*) the *naive* approach with parameters $k = 10$, relative gap $\delta_f = 2 \cdot 10^6$, and $\delta_d = 3 \cdot 10^7$; (*ii*) the *interdiction-only* projection with $k = 15$ and $\delta_f = \delta_d = 2.0$; and (*iii*) the *flow-only* projection with $k = 5$, $\delta_f = 1.5$, and $\delta_d = 3.5$. For each scheme and for each generated network, we generate 10 random projector.

For each lower and upper bound, we compute the corresponding gaps γ_f and γ_d, respectively, and report both the mean and standard deviation in Table 1. For each problem, we also report the runtime mean and the corresponding standard deviation, in Table 2.

The **gap analysis** reveals that the performance guarantees of random projection exhibit distinct behaviors depending on the type of constraints to which they are applied. We observe the following:

	Projection	Network V1 (optimality gap γ)			
		2×3	3×3	4×3	5×3
z_{feas}		0.2 (0.2)	0.4 (0.2)	0.5 (0.2)	0.5 (0.2)
z_{relax}		68.3 (10.4)	72.5 (17.9)	110.5 (22.6)	132.9 (14.7)
$\tilde{s}_\theta(\mathbf{P}, \delta_f)$	flow	0.1 (0.2)	0.4 (0.2)	0.4 (0.2)	0.5 (0.2)
	interdiction	0.2 (0.2)	0.5 (0.2)	0.5 (0.2)	0.6 (0.2)
	naive	0.2 (0.2)	0.5 (0.1)	0.6 (0.2)	0.6 (0.1)
$\tilde{d}_\theta(\mathbf{P}, \delta_d)$	flow	1.8 (1.0)	1.1 (0.8)	0.5 (0.5)	0.2 (0.4)
	interdiction	1.4 (0.4)	0.5 (0.2)	0.4 (0.3)	0.2 (0.3)
	naive	3.0 (22.7)	76.9 (91.9)	129.5 (137.3)	157.5 (164.5)

Table 1: **Network V1 (optimality gap).** Mean ($\pm$ std) optimality gaps for various bounding approaches applied to the min-cost flow interdiction problem on Network V1. We compute the feasibility gap γ_f for both z_{feas} and $\tilde{s}_\theta(\mathbf{P}, \delta_p)$. Also, we compute the dual gap γ_d for both z_{relax} and $\tilde{d}_\theta(\mathbf{P}, \delta_d)$. The table compares: ($i$) the single-level relaxation bounds, and (ii) feasibility and dual bounds obtained via random projections.

• The lower bound derived from the single-level relaxation consistently delivers moderate performance, with mean feasibility gaps $\gamma_f(z_{\text{feas}})$ ranging from 20% to 50%; see Table 1. In contrast, the corresponding dual gap $\gamma_d(z_{\text{relax}})$ deteriorates, reaching up to 13'290% for the (5x3) grid. However, despite the seemingly reasonable performance of the lower bound in this case, it lacks any theoretical guarantee and may, in general, be arbitrarily far from the leader's optimal objective function value.

• When using random projections for feasibility bounds, the gaps $\gamma_f(\tilde{s}_\theta(\mathbf{P}, \delta_f))$ remain roughly comparable to that of the single-level relaxation baseline. In particular, projecting only the flow constraints offers a slight improvement for smaller networks (e.g., a 10% gap for the (2×3) grid) but loses this advantage as the network size increases (up to 50% gap); see Table 1. This observation suggests that random projection does not offer substantial improvements for the current instances.

However, we emphasize that the feasibility bounds from both random projection and single-level relaxation already offer reasonably good results. We attribute this observation to the simplicity of the leader's constraints and the symmetric structure of the problem (min/max). Nevertheless, we expect the single-level relaxation to deteriorate in more complicated settings, whereas the projected feasibility bound benefits from the theoretical guarantees established in Theorem 5.

• As the grid dimension increases, so does the size of the optimization problem. This increase typically leads to worse approximation performance, in line with Theorem 5. Indeed, larger networks leads to more decision variables for both the leader and the follower, which increase quantities such as $\|\mathbf{d}\|_1$ or $\Delta(\mathbf{F}, \mathbf{c})$, both of which impact the theoretical bound in Theorem 5. As expected, we observe in Table 1 that the feasibility gap $\gamma_f(\tilde{s}_\theta(\mathbf{P}, \delta_f))$ increases as the network size grows.

• The dual bound under naive projection performs the worst, with gaps $\gamma_d(\tilde{d}_\theta(\mathbf{P}, \delta_d))$ exceeding 15'000% for larger grids; see Table 1. In contrast, projecting only the flow or only the interdiction constraints leads to significantly tighter dual bounds in comparison to z_{relax}. For example, the interdiction-only projection has a gap of just 20% on the (5×3) grid. This difference can partially be explained by the large difference in the input parameter δ_d (i.e., $3 \cdot 10^7$ against 2.0). Indeed, projecting all the constraints at once deteriorates the quality of the projected follower's problem (4), necessitating in turn a larger tolerance in (10) to produce a valid upper bound on the leader's optimal objective function value.

This outcome is noteworthy, as it highlights that **different classes of constraints behave differently under projection**. Specifically, flow conservation constraints and interdiction constraints induce fundamentally different polyhedral geometries. Projecting both simultaneously amounts to projecting the intersection of these two distinct polyhedra, which can distort their individual structures and lead to information loss. By contrast, projecting one class at a time may preserve more of the underlying structure, which may explain the tighter bounds observed in our experiments.

• We also observe a counterintuitive trend: when projecting either only the flow or only the interdiction constraints, the dual gap $\gamma_d(\tilde{d}_\theta(\mathbf{P}, \delta_d))$ *improves* as the network size increases. This behavior lies outside the scope of our theoretical results, which only provide worst-case guarantees; see Theorem 4. A similar phenomenon is noted by Liberti et al. (2023) for single-level LPs, where empirical performance sometimes diverges from the worst-case prediction. In particular, the actual bound's dependence on the problem parameters may deviate significantly from the theoretical (worst-case) one.

In summary, for the instances considered, projection-based methods appear effective for obtaining dual bounds that are better or comparable to the single-level relaxation. However, this observation is not true for all projection types, especially when selectively applied to certain constraint types (e.g., flow, interdiction, or all together). Although the feasibility bounds might appear less impressive in comparison, they remain theoretically grounded, unlike those derived from the single-level relaxation which could be arbitrarily far away from the leader's optimal objective function value.

The **runtime analysis** reveals extreme behaviors depending on the type of constraints to which the random projection is applied. Our main observations are as follows:

• Solving the original leader's problem exactly becomes computationally expensive as the network grows, requiring over 76 seconds on average for the (5×3) grid instance. In contrast, the runtime for solving the single-level relaxation also increases but remains relatively effective across all grid sizes, solving even the largest instance in under 0.5 seconds, thanks to its simplified structure and the absence of follower's optimality constraints; see Table 2.

• In comparison, random projections show mixed performances for the runtime. The naive

	Projection	Network V1 (runtime)			
		2×3	3×3	4×3	5×3
z^*		0.2 (0.1)	0.5 (0.1)	3.2 (0.2)	76.5 (2.9)
z_{feas}		0.1 (0.1)	0.2 (0.1)	0.4 (0.1)	0.5 (0.1)
z_{relax}		0.1 (0.1)	0.2 (0.1)	0.4 (0.1)	0.5 (0.1)
$\tilde{s}_\theta(\mathbf{P}, \delta_f)$	flow	0.2 (0.1)	0.7 (0.1)	5.0 (0.6)	125.8 (17.4)
	interdiction	0.3 (0.1)	0.6 (0.1)	1.0 (0.2)	1.6 (0.8)
	naive	0.2 (0.1)	0.6 (0.1)	1.0 (0.4)	1.9 (0.9)
$\tilde{d}_\theta(\mathbf{P}, \delta_d)$	flow	0.2 (0.1)	0.7 (0.1)	5.0 (0.7)	131.9 (26.4)
	interdiction	0.3 (0.1)	0.8 (0.3)	1.6 (0.8)	7.1 (7.2)
	naive	0.2 (0.1)	0.4 (0.2)	0.7 (0.7)	4.3 (8.6)

Table 2: **Network V1 (runtime).** Mean (± std) runtime (in sec) for various bounding approaches applied to the min-cost flow interdiction problem on Network V1. The table compares: (*i*) solving the leader's problem exactly, (*ii*) the single-level relaxation approach, and (*iii*) the random projection based approach.

projection reduces runtime (for both feasibility and dual bounds) compared to the original problem, particularly for large instances, since it simplifies "drastically" the size of the problem to $k = 10$ constraints; see Table 2. Similarly, the interdiction-only projection remains computationally light. However, flow-only projection suffers from runtime explosion (e.g., 125.8 seconds for (5x3) for the feasibility bound), exceeding even runtime required to solve the original problem. This observation aligns with an intuitive thought: **projecting flow constraints eliminates network-specific structure that solvers typically exploit**, resulting in denser and harder-to-solve problems.

Overall, while random projection may reduce dimensionality, it can increase runtime, especially when structural properties are lost. Random projection should be used to reduce dimensionality without destroying problem-specific structure that may be exploited by state-of-the-art solvers.

6.3 Summary insights

Our results provide perspective on single-level relaxation (Kleinert et al. 2021) of bilevel programs. This relaxation drops the follower's optimality condition while preserving the polyhedral structure of the follower's feasible region. In contrast, our approach retains a form of optimality information while deliberately modifying (perturbing) the polyhedral structure of the follower's problem. When designed carefully, rather than applied naively, our approach provides tighter bounds with provable guarantees; if handled carelessly and applied naively, however, the follower's polyhedral feasible region can be effectively destroyed, rendering the optimality conditions of little value.

We emphasize the need for caution when applying random projection techniques developed in this paper. **Not all constraints behave equally under projection**; some are more sensitive to distortion than others. As a result, certain classes of constraints, particularly those with an exploitable structure, may require dedicated projection schemes to preserve their computational tractability. For instance, projecting the network flow constraints may compromise the very structure that state-of-the-art MILP solvers are designed to exploit.

Another consideration, previously noted by Liberti et al. (2023) for single-level LPs, is that solvers are typically optimized for sparse instances. In contrast, the use of random projection techniques often replaces such sparse problems with smaller but significantly denser ones, **which may not be solved more effectively in practice**. While dimensionality is reduced, the gain may be offset by increased computational effort required to process dense constraint matrices.

Although our computational study is limited in scope, our findings are consistent with observations in the single-level linear programming literature. Random projections may initially appear to offer a "free lunch" for handling large-scale instances. However, practical outcomes vary considerably, ranging from substantial improvement to, as described by Liberti et al. (2022), "catastrophic" performances. Similar empirical discrepancies are discussed in Liberti et al. (2023), where the observed gaps do not always match the theoretical predictions. This mismatch can be partly attributed to the fact that our theoretical results are worst-case guarantees, which apply uniformly across all problem instances. The specific instances considered in our study may simply not exhibit worst-case behavior. Also, the instance sizes may be insufficient to reveal the asymptotic properties predicted by theory.

Finally, we identify two key factors behind the performance gap and outline corresponding research directions to enhance the computational effectiveness of random projection techniques:

***(i)* Sparsity loss**: Replacing a sparse follower's problem with a denser one can degrade solver performance. Sparse projection techniques, such as those in Allen-Zhu et al. (2014) and Kane and Nelson (2014), may alleviate this issue by preserving sparsity during dimensionality reduction.

***(ii)* Lack of structure preservation**: Sign-preserving projections help extend theoretical guarantees to inequality constraints, but they do not necessarily retain problem-specific structure. Developing **structure-preserving projectors** tailored to flow or other type of commonly used constraints could improve empirical performance. For the problems studied here, extending projection theory to maintain flow conservation structure is, in our view, a particularly promising direction.

7 Conclusion

Surrogate duality. We discuss dimensionality-reduction techniques for a class of bilevel programs whose leader's decision variables are mixed-integer and whose follower's problem is a linear program. Our approach projects the follower's constraints into a compressed representation of a smaller size. We formalize this idea through a surrogate duality theory, which we exploit to derive both a feasibility-based lower bound and a surrogate-dual upper bound on the leader's optimal objective function value. Under some conditions, we prove the existence of a projection scheme for which both bounds coincide, and identify a family of projectors that provide approximation guarantees.

Random projection techniques. Since identifying an "optimal" projection scheme appears to depend on the follower's optimality conditions, we instead turn to random projectors, which exploit probabilistic properties from high-dimensional geometry. By using these random projection schemes, we are able to derive probabilistic approximation guarantees for the bounds obtained from the surrogate duality theory. Yet, these random projectors are not without drawbacks: the expected solution to the projected follower's problem may fail to exist, which preclude using averaging across projections as a derandomization strategy to obtain an approximate solution to the leader's problem.

Using the min-cost flow interdiction problem, we illustrate numerically that random projection can improve the runtime and produce informative bounds, though their effectiveness depends on which constraints they are applied to. Indeed, projection can distort problem's features, such as sparsity or network flow structure, that solvers typically exploit. Still, with judicious constraint selection, projections can preserve sufficient structure to provide meaningful computational gains.

Coupling constraints. We examine dimensionality reduction techniques in the presence of coupling constraints. While projection is relatively effective for settings in which the leader's feasible region does not depend on the follower's rational response, coupling constraints make our theoretical approach fragile. Feasibility of the leader's decision obtained through the projection approach may not persist once decisions are lifted to the original space. This bilevel-specific interplay of feasibility and optimality motivates our analysis of the probability of violating coupling constraints as a function of the "projection's quality." To address this issue, we propose an adjustment to the leader's feasible region that limits such violations while retaining comparable approximation guarantees.

Future research. The main limitation of our study lies in the admittedly narrow scope of its computational study. Bilevel programming would benefit from a more extensive empirical investigation that systematically explores the integration of random projection techniques into exact methods. A key challenge is that random projections may eliminate structural properties, such as sparsity or network constraints, that are crucial for the efficiency of modern solvers. Extending the theoretical results by Vu et al. (2018) to encompass structure-preserving projections represents a promising direction for addressing this issue. An alternative and equally compelling approach is to replace random projections with learned projection schemes, building on ideas by Sakaue and Oki (2024).

References

Achlioptas, D. (2003). “Database-friendly random projections: Johnson-Lindenstrauss with binary coins”. In: *Journal of Computer and System Sciences* 66.4, pp. 671–687.

Akchen, Y.-C. and Mišić, V. V. (2025). “Column-Randomized Linear Programs: Performance Guarantees and Applications”. In: *Operations Research* 73.3, pp. 1366–1383.

Allen-Zhu, Z., Gelashvili, R., Micali, S., and Shavit, N. (2014). “Sparse sign-consistent Johnson-Lindenstrauss matrices: Compression with neuroscience-based constraints”. In: *Proceedings of the National Academy of Sciences* 111.47, pp. 16872–16876.

Aravena, I., Lété, Q., Papavasiliou, A., and Smeers, Y. (2021). “Transmission capacity allocation in zonal electricity markets”. In: *Operations Research* 69.4, pp. 1240–1255.

Audet, C., Hansen, P., Jaumard, B., and Savard, G. (1997). “Links Between Linear Bilevel and Mixed 0–1 Programming Problems”. In: *Journal of Optimization Theory and Applications* 93, pp. 273–300.

Babbar, M. M. (1955). “Distributions of solutions of a set of linear equations (with an application to linear programming)”. In: *Journal of the American Statistical Association* 50.271, pp. 854–869.

Bard, J. F. and Falk, J. E. (1982). “An explicit solution to the multi-level programming problem”. In: *Computers & Operations Research* 9.1, pp. 77–100.

Basu, A., Ryan, C. T., and Sankaranarayanan, S. (2021). “Mixed-integer bilevel representability”. In: *Mathematical Programming* 185, pp. 163–197.

Beck, Y., Ljubić, I., and Schmidt, M. (2023). “A survey on bilevel optimization under uncertainty”. In: *European Journal of Operational Research* 311.2, pp. 401–426.

Bertsimas, D. and Sim, M. (2004). “The price of robustness”. In: *Operations Research* 52.1, pp. 35–53.

Bertsimas, D. and Tsitsiklis, J. N. (1997). *Introduction to linear optimization*. Athena Scientific Belmont, MA.

Bertsimas, D. and Vempala, S. (2004). “Solving convex programs by random walks”. In: *Journal of the ACM (JACM)* 51.4, pp. 540–556.

Bialas, W. and Karwan, M. (1982). “On two-level optimization”. In: *IEEE Transactions on Automatic Control* 27.1, pp. 211–214.

Blair, C. E. and Jeroslow, R. G. (1977). “The value function of a mixed integer program: I”. In: *Discrete Mathematics* 19.2, pp. 121–138.

Brenner, J. and Cummings, L. (1972). “The Hadamard maximum determinant problem”. In: *The American Mathematical Monthly* 79.6, pp. 626–630.

Brown, G., Carlyle, M., Salmerón, J., and Wood, K. (2006). “Defending critical infrastructure”. In: *Interfaces* 36.6, pp. 530–544.

Calafiore, G. and Campi, M. C. (2005). “Uncertain convex programs: randomized solutions and confidence levels”. In: *Mathematical Programming* 102, pp. 25–46.

Campi, M. C. and Garatti, S. (2008). “The exact feasibility of randomized solutions of uncertain convex programs”. In: *SIAM Journal on Optimization* 19.3, pp. 1211–1230.

— (2018). “Wait-and-judge scenario optimization”. In: *Mathematical Programming* 167, pp. 155–189.

Chestnut, S. R. and Zenklusen, R. (2017). “Hardness and approximation for network flow interdiction”. In: *Networks* 69.4, pp. 378–387.

Cohen, M. P. (2012). “Sample means of independent standard Cauchy random variables are standard Cauchy: a new approach”. In: *The American Mathematical Monthly* 119.3, pp. 240–244.

Dan, T. and Marcotte, P. (2019). “Competitive facility location with selfish users and queues”. In: *Operations Research* 67.2, pp. 479–497.

Dasgupta, A., Kumar, R., and Sarlós, T. (2010). “A sparse Johnson-Lindenstrauss transform”. In: *Proceedings of the forty-second ACM symposium on Theory of computing*, pp. 341–350.

De Farias, D. P. and Van Roy, B. (2004). “On constraint sampling in the linear programming approach to approximate dynamic programming”. In: *Mathematics of Operations Research* 29.3, pp. 462–478.

Dempe, S., Kalashnikov, V., Pérez-Valdés, G. A., and Kalashnykova, N. (2015). *Bilevel programming problems*. Vol. 10. 978-3. Springer, pp. 53–56.

Dinkel, J. J. and Kochenberger, G. A. (1978). “An implementation of surrogate constraint duality”. In: *Operations Research* 26.2, pp. 358–364.

Fortuny-Amat, J. and McCarl, B. (1981). “A representation and economic interpretation of a two-level programming problem”. In: *Journal of the Operational Research Society* 32, pp. 783–792.

Gao, W., Ge, D., Sun, C., and Ye, Y. (2023). “Solving linear programs with fast online learning algorithms”. In: *International Conference on Machine Learning*. PMLR, pp. 10649–10675.

Girko, V. L. (2012). *Theory of random determinants*. Vol. 45. Springer Science & Business Media.

Glover, F. (1965). “A multiphase-dual algorithm for the zero-one integer programming problem”. In: *Operations Research* 13.6, pp. 879–919.

Glover, F. (1975). “Surrogate constraint duality in Mathematical Programming”. In: *Operations Research* 23.3, pp. 434–451.

Greenberg, H. J. and Pierskalla, W. P. (1970). “Surrogate Mathematical Programming”. In: *Operations Research* 18.5, pp. 924–939.

Gurobi (2024). *Gurobi Optimization, LLC*. `https://www.gurobi.com`. Accessed: 2024-07-26.

Gutin, E., Kuhn, D., and Wiesemann, W. (2015). “Interdiction games on Markovian PERT networks”. In: *Management Science* 61.5, pp. 999–1017.

Hall, P. (1927). “The distribution of means for samples of size n drawn from a population in which the variate takes values between 0 and 1, all such values being equally probable”. In: *Biometrika* 19, pp. 240–244.

Hansen, P., Jaumard, B., and Savard, G. (1992). “New branch-and-bound rules for linear bilevel programming”. In: *SIAM Journal on Scientific and Statistical Computing* 13.5, pp. 1194–1217.

He, L., Ke, N., Mao, R., Qi, W., and Zhang, H. (2024). “From Curtailed Renewable Energy to Green Hydrogen: Infrastructure Planning for Hydrogen Fuel-Cell Vehicles”. In: *Manufacturing & Service Operations Management* 26.5, pp. 1750–1767.

Henke, D., Lefebvre, H., Schmidt, M., and Thürauf, J. (2025). “On coupling constraints in linear bilevel optimization”. In: *Optimization Letters* 19, pp. 689–697.

IBM (2024). *CPLEX Optimizer*. `https://www.ibm.com/products/ilog-cplex-optimization-studio/cplex-optimizer`. Accessed: 2024-07-26.

Indyk, P. and Naor, A. (2007). “Nearest-neighbor-preserving embeddings”. In: *ACM Transactions on Algorithms (TALG)* 3.3.

Irwin, J. O. (1927). “On the frequency distribution of the means of samples from a population having any law of frequency with finite moments, with special reference to Pearson’s Type II”. In: *Biometrika* 19, pp. 225–239.

Johnson, W. B. and Lindenstrauss, J. (1984). “Extensions of Lipschitz mappings into a Hilbert space”. In: *Contemporary Mathematics* 26, p. 1.

Júdice, J. J. and Faustino, A. M. (1992). “A sequential LCP method for bilevel linear programming”. In: *Annals of Operations Research* 34.1, pp. 89–106.

Kane, D. M. and Nelson, J. (2014). “Sparser Johnson-Lindenstrauss transforms”. In: *Journal of the ACM (JACM)* 61.1, pp. 1–23.

Karmarkar, N. (1984). “A new polynomial-time algorithm for linear programming”. In: *Proceedings of the sixteenth annual ACM symposium on Theory of computing*, pp. 302–311.

Karwan, M. and Rardin, R. (1979). "Some relationships between Lagrangian and surrogate duality in integer programming". In: *Mathematical Programming* 17, pp. 320–334.

Khachiyan, L. G. (1979). "A Polynomial Algorithm in Linear Programming". In: *Doklady Akademii Nauk*. Vol. 244. 5. Russian Academy of Sciences, pp. 1093–1096.

Kleinert, T., Labbé, M., Ljubić, I., and Schmidt, M. (2021). "A survey on mixed-integer programming techniques in bilevel optimization". In: *EURO Journal on Computational Optimization* 9, p. 100007.

Kleinert, T., Labbé, M., Plein, F., and Schmidt, M. (2020). "Technical Note—There's No Free Lunch: On the Hardness of Choosing a Correct Big-M in Bilevel Optimization". In: *Operations Research* 68.6, pp. 1716–1721.

Kuiteing, A. K., Marcotte, P., and Savard, G. (2017). "Network pricing of congestion-free networks: The elastic and linear demand case". In: *Transportation Science* 51.3, pp. 791–806.

Labbé, M., Marcotte, P., and Savard, G. (1998). "A bilevel model of taxation and its application to optimal highway pricing". In: *Management Science* 44, pp. 1608–1622.

Lagos, T. and Prokopyev, O. A. (2023). "On complexity of finding strong-weak solutions in bilevel linear programming". In: *Operations Research Letters* 51.6, pp. 612–617.

Li, Q., Üster, H., and Zhang, Z.-H. (2023). "A bilevel model for robust network design and biomass pricing under farmers' risk attitudes and supply uncertainty". In: *Transportation Science* 57.5, pp. 1296–1320.

Li, X. and Ye, Y. (2022). "Online linear programming: Dual convergence, new algorithms, and regret bounds". In: *Operations Research* 70.5, pp. 2948–2966.

Liberti, L., Manca, B., Poirion, P.-L., et al. (2022). "Practical performance of random projections in linear programming". In: *Leibniz International Proceedings in Informatics* 233.

Liberti, L., Manca, B., and Poirion, P.-L. (2023). "Random projections for Linear Programming: an improved retrieval phase". In: *ACM Journal of Experimental Algorithmics* 28, pp. 1–33.

Liberti, L., Poirion, P.-L., and Vu, K. (2021). "Random projections for conic programs". In: *Linear Algebra and its Applications* 626, pp. 204–220.

Lin, Y. H., Tian, Q., and Zhao, Y. (2024). "Unified framework for choice-based facility location problem". In: *INFORMS Journal on Computing* 36.6, pp. 1436–1458.

Lozano, L. and Smith, J. C. (2017). "A Value-Function-Based Exact Approach for the Bilevel Mixed-Integer Programming Problem". In: *Operations Research* 65.3, pp. 768–786.

Mangasarian, O. L. and Shiau, T.-H. (1987). "Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems". In: *SIAM Journal on Control and Optimization* 25.3, pp. 583–595.

Matoušek, J. (2008). "On variants of the Johnson-Lindenstrauss lemma". In: *Random Structures & Algorithms* 33.2, pp. 142–156.

Müller, B., Muñoz, G., Gasse, M., Gleixner, A., Lodi, A., and Serrano, F. (2022). "On generalized surrogate duality in mixed-integer nonlinear programming". In: *Mathematical Programming* 192.1, pp. 89–118.

Pilanci, M. and Wainwright, M. J. (2015). "Randomized sketches of convex programs with sharp guarantees". In: *IEEE Transactions on Information Theory* 61.9, pp. 5096–5115.

Poirion, P.-L., Lourenço, B. F., and Takeda, A. (2023). "Random projections of linear and semidefinite problems with linear inequalities". In: *Linear Algebra and its Applications* 664, pp. 24–60.

Prékopa, A. (1966). "On the probability distribution of the optimum of a random linear program". In: *SIAM Journal on Control* 4.1, pp. 211–222.

Prokopyev, O. A. and Ralphs, T. K. (2024). *On the Complexity of Finding Locally Optimal Solutions in Bilevel Linear Optimization*. URL: https://optimization-online.org/?p=30665.

Robert, C. P. (2007). *The Bayesian choice: from decision-theoretic foundations to computational implementation*. Springer.

Rodrigues, B., Carvalho, M., Anjos, M. F., and Sugishita, N. (2025). *Unboundedness in bilevel optimization*. URL: https://optimization-online.org/?p=28405.

Royset, J. O. and Wood, R. K. (2007). "Solving the bi-objective maximum-flow network-interdiction problem". In: *INFORMS Journal on Computing* 19.2, pp. 175–184.

Sakaue, S. and Oki, T. (2024). "Generalization Bound and Learning Methods for Data-Driven Projections in Linear Programming". In: *Advances in Neural Information Processing Systems* 37, pp. 12825–12846.

Smith, J. C. and Song, Y. (2020). "A survey of network interdiction models and algorithms". In: *European Journal of Operational Research* 283.3, pp. 797–811.

Spielman, D. A. and Teng, S.-H. (2003). "Smoothed analysis of termination of linear programming algorithms". In: *Mathematical Programming* 97.1, pp. 375–404.

Tavaslıoğlu, O., Prokopyev, O. A., and Schaefer, A. J. (2019). "Solving Stochastic and Bilevel Mixed-Integer Programs via a Generalized Value Function". In: *Operations Research* 67.6, pp. 1659–1677.

Tawfik, C. and Limbourg, S. (2019). "A bilevel model for network design and pricing based on a level-of-service assessment". In: *Transportation Science* 53.6, pp. 1609–1626.

Trapp, A. C. and Prokopyev, O. A. (2015). "A note on constraint aggregation and value functions for two-stage stochastic integer programs". In: *Discrete Optimization* 15, pp. 37–45.

Vempala, S. (2005). *The random projection method*. Vol. 65. American Mathematical Soc.

Vu, K., Poirion, P.-L., and Liberti, L. (2018). "Random projections for linear programming". In: *Mathematics of Operations Research* 43.4, pp. 1051–1071.

Wiesemann, W., Tsoukalas, A., Kleniati, P.-M., and Rustem, B. (2013). "Pessimistic bilevel optimization". In: *SIAM Journal on Optimization* 23.1, pp. 353–380.

Zare, M. H., Borrero, J. S., Zeng, B., and Prokopyev, O. A. (2019). "A note on linearized reformulations for a class of bilevel linear integer problems". In: *Annals of Operations Research* 272, pp. 99–117.

Zenklusen, R. (2010). "Matching interdiction". In: *Discrete Applied Mathematics* 158.15, pp. 1676–1690.

— (2015). "An O(1)-Approximation for Minimum Spanning Tree interdiction". In: *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*. IEEE, pp. 709–728.

A Key notations

Notation	Description	Eq.	Section
z^*	Leader's optimal objective function value	(1)	Section 1
$\mathbf{x}$	Leader's decision	(1)	Section 1
$\mathbf{y}$	Follower's decision	(2)	Section 1
$\mathbf{y}^*(\mathbf{x})$	Follower's rational response	(3)	Section 1
θ	Upper bound on the ℓ_1-norms of $\mathbf{x}$ and $\mathbf{y}$ in Assumption **A2**	-	Section 1
$\mathbf{P}$	Linear projector	-	Section 3
$\varphi(\mathbf{x}) \equiv \varphi_\theta(\mathbf{x})$	Follower's value function given leader's feasible decision $\mathbf{x}$	(2)	-
$\varphi(\mathbf{x}, \mathbf{P})$	Follower's projected value function	(4)	Section 3.1
$\varphi_\theta(\mathbf{x}, \mathbf{P})$	Follower's projected value function when the follower's decision variables are bounded by θ	(9)	Section 4.1
$\mathcal{Y}(\mathbf{x})$	Follower's feasible region	-	Section 1
$\mathcal{Y}_\theta(\mathbf{x})$	Follower's feasible region when the follower's decision variables are bounded by θ	-	Section 4.1
$\mathcal{Y}_\theta(\mathbf{x}, \mathbf{P})$	Projected follower's feasible region with the additional constraints that the decision variables are bounded by θ	-	Section 4.1
$s(\mathbf{P})$	Surrogate dual function	(5)	Section 3.1
$s_\theta(\mathbf{P})$	Surrogate dual function when the follower's decision variables are bounded by θ	-	Section 4.1
$s(\mathbf{P}, \delta)$	Feasibility bound/problem	(8)	Section 3.3
$s_\theta(\mathbf{P}, \delta)$	Feasibility bound/problem when the follower's decision variables are bounded by θ	-	Section 4.1
$\tilde{s}_\theta(\mathbf{P}, \tilde{\delta})$	Re-parametrized feasibility bound/problem when the follower's decision variables are bounded by θ	-	Section 4.2.2
$\mathcal{S}(\mathbf{P}, \delta)$	Feasibility set of $s(\mathbf{P}, \delta)$	(7)	Section 3.3
$\mathcal{S}_\theta(\mathbf{P}, \delta)$	Feasibility set of $s_\theta(\mathbf{P}, \delta)$	-	Section 4.1
$\tilde{\mathcal{S}}_\theta(\mathbf{P}, \tilde{\delta})$	Feasibility set of $\tilde{s}_\theta(\mathbf{P}, \tilde{\delta})$	-	Section 4.2.2
$d_\theta(\mathbf{P}, \delta)$	Adjusted surrogate problem	(10)	Section 4.2.1
$\tilde{d}_\theta(\mathbf{P}, \tilde{\delta})$	Re-parametrized adjusted surrogate problem	-	Section 4.2.1

Table 3: Summary of the most important notations used throughout the study.

B Proofs for Section 4

We formally prove our results from Section 4. Specifically, Section B.1 establishes probabilistic approximation guarantees for the upper bound obtained from the adjusted surrogate problem (10). Section B.2 provides probabilistic approximation guarantees for the lower bound from the feasibility problem (8). Finally, Section B.3 contains the proofs for the results involving coupling constraints.

B.1 Proofs for Section 4.2.1

We now turn to the proofs of the results outlined in Section 4.2.1. First, we present a detailed proof of Lemma 1, which is then exploited to prove Theorem 4.

Proof of Lemma 1. In this proof, we only consider the case without coupling constraints, i.e., $\mathbf{H}_{p\times m} = \mathbf{0}_{p\times m}$, though the same arguments apply to the case which includes these constraints. We show that the optimal solution to the leader's problem (1) is feasible to the adjusted surrogate problem (10) with high probability.

To proceed, we define the following constant term:

$$u^* = \max_{\mathbf{x}\in\mathcal{X}} \|\mathbf{u}^*(\mathbf{x})\|_2 \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2,$$

where, for any given $\mathbf{x} \in \mathcal{X}$, $\mathbf{u}^*(\mathbf{x})$ corresponds to the optimal solution (of minimal ℓ_2-norm) to the follower's dual problem. Next, we fix a leader's feasible decision $\mathbf{x} \in \mathcal{X}$, and some arbitrary $\delta \in \big(0, \min\big\{\varphi_\theta(\mathbf{x}) \,:\, \mathbf{x} \in \mathcal{X}\big\}\big)$, and apply Theorem 3. We obtain the following inequalities:

$$\varphi_\theta(\mathbf{x}, \mathbf{P}) \le \varphi_\theta(\mathbf{x}) \le \varphi_\theta(\mathbf{x}, \mathbf{P}) + \delta$$

with probability at least $1 - 4me^{-C\left(\varepsilon^2-\varepsilon^3\right)k}$, where $\varepsilon = \mathcal{O}\big(\frac{\delta}{(1+\theta)u^*}\big)$. In particular, this inequality is satisfied by the leader's optimal decision obtained when solving the leader's problem (1) so that $z^* \le d_\theta(\mathbf{P}, \delta)$. This feasibility argument, in turn, concludes the proof. ■

Observe that we do not make any use of $\mathbf{H} = \mathbf{0}_{p\times m}$ in the proof of Lemma B.1. Hence, the result naturally extends to the case where $\mathbf{G}\mathbf{x} \le \mathbf{h}$ is replaced by $\mathbf{G}\mathbf{x} + \mathbf{H}\mathbf{y} \le \mathbf{h}$ in the adjusted surrogate problem (10) by using the same feasibility-based argument. Next, we turn to the proof of the main result of Section 4.2.1. Formally:

Proof of Theorem 4. Our goal in this proof is to compare the leader's optimal objective function value z^* to the value $d_\theta(\mathbf{P}, \delta)$. As in the proof above, we first define the following constant term:

$$u^* = \max_{\mathbf{x} \in \mathcal{X}} \|\mathbf{u}^*(\mathbf{x})\|_2 \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2,$$

where, for any given $\mathbf{x} \in \mathcal{X}$, $\mathbf{u}^*(\mathbf{x})$ corresponds to the optimal solution (of minimal ℓ_2-norm) to the follower's dual problem. Then, we fix some arbitrary value for $\delta \in \big(0, \min\big\{\varphi_\theta(\mathbf{x}) \,:\, \mathbf{x} \in \mathcal{X}\big\}\big)$.

Step 1 (Feasibility): By applying Lemma 1, we then obtain the following inequalities:

$$z^* \leq d_\theta(\mathbf{P}, \delta),$$

with probability at least $1 - 4me^{-C\left(\varepsilon^2 - \varepsilon^3\right)k}$, where $\varepsilon = \mathcal{O}\big(\frac{\delta}{(1+\theta)u^*}\big)$. In particular, with probability at least $1 - 4me^{-C\left(\varepsilon^2 - \varepsilon^3\right)k}$, the adjusted surrogate problem (10) is feasible and admits at least one optimal solution. This optimal solution, however, is a random variable determined by the random projector $\mathbf{P}$. This dependence on $\mathbf{P}$ creates an additional layer of difficulty in the subsequent analysis.

Step 2 (Covering): We first state and prove an intermediate lemma. The lemma constructs a cover of the leader's feasible set by balls of radius at most η in the ℓ_1-norm. We use this cover to construct a finite collection of leader's feasible decisions. Then, we apply Theorem 3 to every leader's feasible decision in that set to ensure that the bound holds simultaneously for all of them.

Lemma 3. *Fix $\eta > 0$ arbitrarily. Let $N(\mathcal{X}, \eta\,;\, \|\cdot\|_1)$ denote the η-covering number of $\mathcal{X}$ with respect to the ℓ_1-norm. Then, for every $\eta > 0$, the following inequality is satisfied:*

$$N(\mathcal{X}, \eta\,;\, \|\cdot\|_1) \leq \left(1 + \frac{4\theta}{\eta}\right)^n.$$

Proof. Let $B_1^n := \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\|_1 \leq 1\}$ denote the unit ball in the Euclidian space $(\mathbb{R}^n, \|\cdot\|_1)$. By Assumption **A2**, the leader's feasible region $\mathcal{X}$ is bounded by some θ in the ℓ_1-norm. Therefore, we have $\mathcal{X} \subseteq \theta B_1^n$. Let $N(K, \eta)$ denote the η-covering number of a set K in this norm. Then, by "monotonicity" of the covering number with respect to set inclusion, refer to Proposition 9.6 by Tropp (2023) for more details, the following inequality is satisfied:

$$N(\mathcal{X}, \eta) \leq N(\theta B_1^n, \frac{\eta}{2}) = N\left(B_1^n, \frac{\eta}{2\theta}\right),$$

and the equality comes from the definition of covering number.

Using Corollary 9.14 by Tropp (2023), we have that, on the unit ball B_1^n on $\mathbb{R}^n$ and for arbitrary $\tilde{\eta} > 0$, the following inequality is satisfied:

$$N(B_1^n, \tilde{\eta}) \leq \left(1 + \frac{2}{\tilde{\eta}}\right)^n.$$

Applying this bound with $\tilde{\eta} = \eta/2\theta$ gives the following inequality:

$$N\left(B_1^n, \frac{\eta}{2\theta}\right) \;\leq\; \left(1 + \frac{4\theta}{\eta}\right)^n .$$

This result concludes in turn the proof of Lemma 3. ■

To proceed, fix some $\eta > 0$ and let $N \equiv N\left(\mathcal{X}, \eta\,;\, \|\cdot\|_1\right)$ be as in Lemma 3. Denote by $(C_i^{\eta})_{i\in[N]}$ the corresponding η-cover of $\mathcal{X}$ in the ℓ_1-norm of minimal cardinality. We then construct a finite subset $\mathcal{X}_\eta \subset \mathcal{X}$ of cardinality N by selecting, for each $i \in [N]$, a representative leader's feasible decision $\mathbf{x}_\eta^i \in \mathcal{X} \cap C_i^{\eta}$. This construction guarantees that $|\mathcal{X}_\eta| = N(\mathcal{X}, \eta\,;\, \|\cdot\|_1)$ and that, for every leader's feasible decision $\mathbf{x} \in \mathcal{X}$, there exists some $i \in [N]$ such that $\|\mathbf{x} - \mathbf{x}_\eta^i\|_1 \leq \eta$.

Step 3 (Proximity theory): Next, we define the set $S(\mathbf{x}, u, v; \mathbf{w})$ that is parametrized by some scalars $u, v \in \mathbb{R}$ and a vector $\mathbf{w} \in \mathbb{R}^m$ as follows:

$$S(\mathbf{x}, u, v; \mathbf{w}) := \underset{\mathbf{y} \geq \mathbf{0}_m}{\operatorname{argmax}} \Big\{ \mathbf{w}^\top \mathbf{y} \;:\; \mathbf{PFy} = \mathbf{P}\big(\mathbf{f} - \mathbf{Lx}\big)\,,\; u \leq \mathbf{c}^\top \mathbf{y} \leq v\,,\; \mathbf{1}_m^\top \mathbf{y} \leq \theta \Big\}\,.$$

Our proof relies on proximity theory to bound the difference between the leader's optimal objective value z^* and $d_\theta(\mathbf{P}, \delta)$. Since any optimal solution to the adjusted surrogate problem (10) is a random variable induced by the random projector $\mathbf{P}$, we circumvent this difficulty by working with deterministic leader's feasible decisions selected from an η-cover of $\mathcal{X}$.

To proceed, we fix some leader's feasible decision $\mathbf{x} \in \mathcal{X}$. Then, we apply a classical proximity result for linear programs, and in particular Theorem 5 in Cook et al. (1986). That is, for all $\tilde{\mathbf{y}}(\mathbf{x}, \mathbf{P}, \delta) \in S(\mathbf{x}, \varphi_\theta(\mathbf{x}), \varphi_\theta(\mathbf{x}); \mathbf{0}_m)$, there exists $\mathbf{y}(\mathbf{x}, \mathbf{P}, \delta) \in S(\mathbf{x}, \varphi_\theta(\mathbf{x}, \mathbf{P}), \varphi_\theta(\mathbf{x}, \mathbf{P}); \mathbf{0}_m)$ that satisfies, together, the following inequality:

$$\|\mathbf{y}(\mathbf{x}, \mathbf{P}, \delta) - \tilde{\mathbf{y}}(\mathbf{x}, \mathbf{P}, \delta)\|_\infty \leq m\Delta(\mathbf{PF}, \mathbf{c})\big|\varphi_\theta(\mathbf{x}) - \varphi_\theta(\mathbf{x}, \mathbf{P})\big|, \tag{15}$$

where $\Delta(\mathbf{PF}, \mathbf{c})$ the maximum of the absolute values of the determinants of the square submatrices of $[\,(\mathbf{PF})^\top \mid -(\mathbf{PF})^\top \mid -\mathbf{c} \mid \mathbf{c} \mid \mathbf{1}_m\,]$, as defined in the statement of Theorem 4. As a side remark, for a given realization of $\mathbf{P}$, a tighter Liptschtz constant can be derived as in the study by Mangasarian and Shiau (1987). However, we prefer to use $\Delta(\mathbf{PF}, \mathbf{c})$ as its dependence on $\mathbf{P}$ is somehow explicit.

As a first observation, we have that $\mathbf{y}^*(\mathbf{x}) \in S(\mathbf{x}, \varphi_\theta(\mathbf{x}), \varphi_\theta(\mathbf{x}); \mathbf{0}_m)$. Accordingly, there exists $\mathbf{y}(\mathbf{x}, \mathbf{P}, \delta) \in S(\mathbf{x}, \varphi_\theta(\mathbf{x}, \mathbf{P}), \varphi_\theta(\mathbf{x}, \mathbf{P}); \mathbf{0}_m)$ that satisfies the following inequality:

$$\|\mathbf{y}^*(\mathbf{x}) - \mathbf{y}(\mathbf{x}, \mathbf{P}, \delta)\|_\infty \leq m\Delta(\mathbf{PF}, \mathbf{c})\big|\varphi_\theta(\mathbf{x}) - \varphi_\theta(\mathbf{x}, \mathbf{P})\big|. \tag{16}$$

Also, recall that we assume that the projected follower's problem has always a unique optimal solution (see our discussion in Section 4.1). By the definition of the set $S(\mathbf{x}, \varphi_\theta(\mathbf{x}, \mathbf{P}), \varphi_\theta(\mathbf{x}, \mathbf{P}); \mathbf{w})$, we have that this set is of cardinality one, i.e., $S(\mathbf{x}, \varphi_\theta(\mathbf{x}, \mathbf{P}), \varphi_\theta(\mathbf{x}, \mathbf{P}); \mathbf{w}) = \{\mathbf{y}(\mathbf{x}, \mathbf{P}, \delta)\}$ for all $\mathbf{w} \in \mathbb{R}^m$. Next, we apply once again the proximity result for linear programs and in particular Theorem 5 in Cook et al. (1986). Accordingly, we have that, for all $\mathbf{y}^*(\mathbf{x}, \mathbf{P}, \delta) \in S(\mathbf{x}, \varphi_\theta(\mathbf{x}, \mathbf{P}), \varphi_\theta(\mathbf{x}, \mathbf{P}) + \delta; \mathbf{d})$, the following inequality is satisfied:

$$\|\mathbf{y}^*(\mathbf{x}, \mathbf{P}, \delta) - \mathbf{y}(\mathbf{x}, \mathbf{P}, \delta)\|_\infty \leq m\Delta(\mathbf{PF}, \mathbf{c})\delta, \tag{17}$$

where we use that $S(\mathbf{x}, \varphi_\theta(\mathbf{x}, \mathbf{P}), \varphi_\theta(\mathbf{x}, \mathbf{P}); \mathbf{d}) = \{\mathbf{y}(\mathbf{x}, \mathbf{P}, \delta)\}$.

We denote by $(\mathbf{x}(\mathbf{P}, \delta), \mathbf{y}^*(\mathbf{x}(\mathbf{P}, \delta), \mathbf{P}, \delta))$ the optimal solution to (10), assuming it exists. Without loss of generality, we assume that this optimal solution is unique as our arguments could be applied to all optimal pairs. By the definition and optimality of $(\mathbf{x}(\mathbf{P}, \delta), \mathbf{y}^*(\mathbf{x}(\mathbf{P}, \delta), \mathbf{P}, \delta))$, we have that $\mathbf{y}^*(\mathbf{x}(\mathbf{P}, \delta), \mathbf{P}, \delta) \in S(\mathbf{x}(\mathbf{P}, \delta), \varphi_\theta(\mathbf{x}(\mathbf{P}, \delta), \mathbf{P}), \varphi_\theta(\mathbf{x}(\mathbf{P}, \delta), \mathbf{P}) + \delta; \mathbf{d})$.

To proceed, we define the follower's rational response to $\mathbf{x}(\mathbf{P}, \delta)$ by $\mathbf{y}^*(\mathbf{x}(\mathbf{P}, \delta))$. Then, we apply the previous proximity results, and in particular the ones from (16) and (17) to $\mathbf{x}(\mathbf{P}, \delta)$. Specifically, we obtain the following inequality:

$$\|\mathbf{y}^*(\mathbf{x}(\mathbf{P}, \delta), \mathbf{P}, \delta) - \mathbf{y}^*(\mathbf{x}(\mathbf{P}, \delta))\|_\infty \leq m\Delta(\mathbf{PF}, \mathbf{c}) \left(|\varphi_\theta(\mathbf{x}(\mathbf{P}, \delta), \mathbf{P}) - \varphi_\theta(\mathbf{x}(\mathbf{P}, \delta))| + \delta\right).$$

Next, we observe that the following sequence of inequalities is satisfied:

$$\begin{aligned} d_\theta(\mathbf{P}, \delta) - z^* &\overset{(a)}{\leq} \mathbf{a}^\top \mathbf{x}(\mathbf{P}, \delta) + \mathbf{d}^\top \mathbf{y}^*(\mathbf{x}(\mathbf{P}, \delta), \mathbf{P}, \delta) - \mathbf{a}^\top \mathbf{x}(\mathbf{P}, \delta) - \mathbf{d}^\top \mathbf{y}^*(\mathbf{x}(\mathbf{P}, \delta)) \\ &= \mathbf{d}^\top \mathbf{y}^*(\mathbf{x}(\mathbf{P}, \delta), \mathbf{P}, \delta) - \mathbf{d}^\top \mathbf{y}^*(\mathbf{x}(\mathbf{P}, \delta)) \\ &\overset{(b)}{\leq} \|\mathbf{d}\|_1 m\Delta(\mathbf{PF}, \mathbf{c}) \left(\delta + D(\varphi, \mathbf{P})\right), \end{aligned}$$

where (a) follows from the feasibility (recall that there is no coupling constraints by assumption) of the leader's decision $\mathbf{x}(\mathbf{P}, \delta)$ and $\mathbf{y}^*(\mathbf{x}(\mathbf{P}, \delta))$ corresponds to the follower's rational response to $\mathbf{x}(\mathbf{P}, \delta)$. Also, (b) follows from the proximity theory results and by the definition of $D(\varphi, \mathbf{P})$, namely, $D(\varphi, \mathbf{P}) \equiv \sup\{|\varphi_\theta(\mathbf{x}, \mathbf{P}) - \varphi_\theta(\mathbf{x})| : \mathbf{x} \in \mathcal{X}\}$.

Moreover, recall that $\Delta(\mathbf{PF}, \mathbf{c})$ is a random variable that is induced by $\mathbf{P}$, which is itself defined over a probability space $(\Omega, \mathcal{B}, \mathbb{P})$. Next, fix some arbitrary $\alpha \geq 0$. By the law of total probability (Jacod and Protter 2012), we obtain the following equality:

$$\begin{aligned} \mathbb{P}\left[\Delta(\mathbf{PF}, \mathbf{c}) \left(\delta + D(\varphi, \mathbf{P})\right) \leq \frac{\alpha}{\|\mathbf{d}\|_1 m}\right] &= \mathbb{P}\left[\Delta(\mathbf{PF}, \mathbf{c}) \left(\delta + D(\varphi, \mathbf{P})\right) \leq \frac{\alpha}{\|\mathbf{d}\|_1 m} \,|\, \mathcal{E}\right] \mathbb{P}\left[\mathcal{E}\right] \\ &\quad + \mathbb{P}\left[\Delta(\mathbf{PF}, \mathbf{c}) \left(\delta + D(\varphi, \mathbf{P})\right) \leq \frac{\alpha}{\|\mathbf{d}\|_1 m} \,|\, \mathcal{E}^c\right] \mathbb{P}\left[\mathcal{E}^c\right], \end{aligned}$$

where we define the event $\mathcal{E} \equiv \{\max\{\Delta(\mathbf{PF}, \mathbf{c}), \|\mathbf{PL}\|_\infty\} \le \rho_\Delta\}$, given the matrix-norm $\|\mathbf{M}\|_\infty := \max_{i\in[k]} \max_{j\in[n]} |M_{ij}|$ for $\mathbf{M} \in \mathbb{R}^{k\times n}$. Recall that the random variable within the event $\mathcal{E}$ defined by $\max\{\Delta(\mathbf{PF}, \mathbf{c}), \|\mathbf{PL}\|_\infty\}$ has a distribution given by F_Δ by assumption.

Now, observe that $\rho_\Delta (\delta + D(\varphi, \mathbf{P})) \le \frac{\alpha}{\|\mathbf{d}\|_1 m}$ implies that $\Delta(\mathbf{PF}, \mathbf{c}) (\delta + D(\varphi, \mathbf{P})) \le \frac{\alpha}{\|\mathbf{d}\|_1 m}$ on the event $\mathcal{E}$. Hence, the following inequality is satisfied:

$$\mathbb{P}\left[\Delta(\mathbf{PF}, \mathbf{c}) (\delta + D(\varphi, \mathbf{P})) \le \frac{\alpha}{\|\mathbf{d}\|_1 m} \,|\, \mathcal{E}\right] \mathbb{P}[\mathcal{E}] \ge \mathbb{P}\left[\rho_\Delta (\delta + D(\varphi, \mathbf{P})) \le \frac{\alpha}{\|\mathbf{d}\|_1 m} \,|\, \mathcal{E}\right] \mathbb{P}[\mathcal{E}].$$

To derive further lower bounds, we introduce another intermediate lemma. Formally:

Lemma 4. *Fix $\delta > 0$ and a random projector $\mathbf{P} \in \mathbb{R}^{k\times q}$. Define the following three constants:*

$$\Delta(\mathbf{PF}) := \max\left\{|\det(\mathbf{M})| \,:\, \mathbf{M} \text{ is a square submatrix of } \left[(\mathbf{PF})^\top \mid -(\mathbf{PF})^\top \mid \mathbf{1}_m\right]\right\}$$
$$\Delta(\mathbf{F}) := \max\left\{|\det(\mathbf{M})| \,:\, \mathbf{M} \text{ is a square submatrix of } \left[\mathbf{F}^\top \mid -\mathbf{F}^\top \mid \mathbf{1}_m\right]\right\}$$
$$\Delta(\mathbf{F}, \mathbf{c}) := \max\left\{ \left|\det(\mathbf{M})\right| \,:\, \mathbf{M} \text{ is a square submatrix of } [\,\mathbf{F}^\top \mid -\mathbf{F}^\top \mid \mathbf{c} \mid \mathbf{1}_m\,]\right\}.$$

Then, for every pair $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$, the following two inequalities are satisfied $\mathbb{P}$-almost surely:

$$|\varphi_\theta(\mathbf{x}, \mathbf{P}) - \varphi_\theta(\mathbf{x}', \mathbf{P})| \le \|\mathbf{c}\|_1 m \Delta(\mathbf{PF}) \|\mathbf{PL}\|_\infty \|\mathbf{x} - \mathbf{x}'\|_1,$$
$$|\varphi_\theta(\mathbf{x}) - \varphi_\theta(\mathbf{x}')| \le \|\mathbf{c}\|_1 m \Delta(\mathbf{F}) \|\mathbf{L}\|_\infty \|\mathbf{x} - \mathbf{x}'\|_1.$$

Thus, the deviation between the follower's projected value function and its exact counterpart $Z(\mathbf{x}, \mathbf{P}) := |\varphi_\theta(\mathbf{x}, \mathbf{P}) - \varphi_\theta(\mathbf{x})|$ is Lipschitz continuous in $\mathbf{x}$. Moreover, one can verify that $\Delta(\mathbf{PF}) \le \Delta(\mathbf{PF}, \mathbf{c})$ and $\Delta(\mathbf{F}) \le \Delta(\mathbf{F}, \mathbf{c})$. Consequently, the following inequality is satisfied $\mathbb{P}$-almost surely:

$$|Z(\mathbf{x}, \mathbf{P}) - Z(\mathbf{x}', \mathbf{P})| \le L_Z(\mathbf{P}) \|\mathbf{x} - \mathbf{x}'\|_1, \tag{20}$$

where $L_Z(\mathbf{P}) := \|\mathbf{c}\|_1 m (\Delta(\mathbf{PF}, \mathbf{c}) \|\mathbf{PL}\|_\infty + \Delta(\mathbf{F}, \mathbf{c}) \|\mathbf{L}\|_\infty)$.

Proof. Fix $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$ and consider the projected follower's problems given these two leader's feasible decision. We apply the proximity result for linear programs and in particular Theorem 5 in Cook et al. (1986). Recall that the projected follower's problem has a unique optimal solution by assumption; see our discussion in Section 4.1. Moreover, the projected follower's problems only differs in their right-hand sides for different values of $\mathbf{x}$ and $\mathbf{x}'$. Therefore, the following inequality is satisfied:

$$|\varphi_\theta(\mathbf{x}, \mathbf{P}) - \varphi_\theta(\mathbf{x}', \mathbf{P})| \le \|\mathbf{c}\|_1 m \Delta(\mathbf{PF}) \|\mathbf{PL}(\mathbf{x} - \mathbf{x}')\|_\infty \le \|\mathbf{c}\|_1 m \Delta(\mathbf{PF}) \|\mathbf{PL}\|_\infty \|\mathbf{x} - \mathbf{x}'\|_1.$$

Then, using a similar argument, we obtain:

$$|\varphi_\theta(\mathbf{x}) - \varphi_\theta(\mathbf{x}')| \leq \|\mathbf{c}\|_1 m\Delta(\mathbf{F})\|\mathbf{L}\left(\mathbf{x}-\mathbf{x}'\right)\|_\infty \leq \|\mathbf{c}\|_1 m\Delta(\mathbf{F})\|\mathbf{L}\|_\infty\|\mathbf{x}-\mathbf{x}'\|_1.$$

Finally, to obtain (20), we simply combine the two Lipschitz bounds. For every $\mathbf{x}, \mathbf{x}' \in \mathcal{X}$,

$$\begin{aligned}|Z(\mathbf{x},\mathbf{P}) - Z(\mathbf{x}',\mathbf{P})| &= \big|\,|\varphi_\theta(\mathbf{x},\mathbf{P}) - \varphi_\theta(\mathbf{x})| - |\varphi_\theta(\mathbf{x}',\mathbf{P}) - \varphi_\theta(\mathbf{x}')|\,\big| \\ &\leq |\varphi_\theta(\mathbf{x},\mathbf{P}) - \varphi_\theta(\mathbf{x}',\mathbf{P})| + |\varphi_\theta(\mathbf{x}) - \varphi_\theta(\mathbf{x}')| \\ &\leq \|\mathbf{c}\|_1 m\big(\Delta(\mathbf{PF})\|\mathbf{PL}\|_\infty + \Delta(\mathbf{P})\|\mathbf{L}\|_\infty\big)\|\mathbf{x}-\mathbf{x}'\|_1,\end{aligned}$$

where the last inequality is obtained through applying the proximity results above. This inequality, in turn, concludes the proof of Lemma 4. ■

Step 4 (Putting all together): Observe that, conditional on the event $\mathcal{E}$, the bound $L_Z(\mathbf{P})$ as initially defined in the statement of Lemma 4 satisfies the following inequality:

$$L_Z(\mathbf{P}) \leq \max\left\{\|\mathbf{c}\|_1 m\left(\rho_\Delta^2 + \Delta(\mathbf{F},\mathbf{c})\|\mathbf{L}\|_\infty\right), \frac{1}{2}\right\} =: L_Z.$$

Moreover, we define $\alpha := \frac{5}{2}\|\mathbf{d}\|_1 m\rho_\Delta\delta$ and $\eta := \delta\,(2L_Z)^{-1}$. Then, we define the finite set $\mathcal{X}_\eta$ obtained from the η-cover of $\mathcal{X}$ in the ℓ_1-norm; recall Step 1.

Hereafter, we work on the event $\mathcal{E}$. We assume that the following inequality is satisfied:

$$\max\left\{Z(\mathbf{x}_\eta^i,\mathbf{P}) \,:\, i \in [N(\mathcal{X},\eta\,;\,\|\cdot\|_1)]\right\} \leq \frac{\alpha}{\|\mathbf{d}\|_1 m\rho_\Delta} - \frac{3\delta}{2}.$$

For all $\mathbf{x} \in \mathcal{X}$, there exists $\mathbf{x}_\eta^i \in \mathcal{X}_\eta$ for some $i \in [N(\mathcal{X},\eta\,;\,\|\cdot\|_1)]$ such that $\|\mathbf{x}-\mathbf{x}_\eta^i\|_1 \leq \eta$. By using the triangle inequality, we obtain the following inequality:

$$|\varphi_\theta(\mathbf{x},\mathbf{P}) - \varphi_\theta(\mathbf{x})| \leq \max\left\{Z(\mathbf{x}_\eta^i,\mathbf{P}) \,:\, i \in [N(\mathcal{X},\eta\,;\,\|\cdot\|_1)]\right\} + |Z(\mathbf{x}_\eta^i,\mathbf{P}) - Z(\mathbf{x},\mathbf{P})|.$$

Next, we use the Lipschitz continuity of $Z(\cdot,\mathbf{P})$ from Lemma 4. In particular, we obtain the following sequence of inequalities:

$$\begin{aligned}|\varphi_\theta(\mathbf{x},\mathbf{P}) - \varphi_\theta(\mathbf{x})| &\leq \frac{\alpha}{\|\mathbf{d}\|_1 m\rho_\Delta} - \delta - \frac{\delta}{2} + L_Z\|\mathbf{x}_\eta^i - \mathbf{x}\|_1 \\ &\leq \frac{\alpha}{\|\mathbf{d}\|_1 m\rho_\Delta} - \delta - \frac{\delta}{2} + L_Z\eta \leq \frac{\alpha}{\|\mathbf{d}\|_1 m\rho_\Delta} - \delta,\end{aligned}$$

where the last inequality comes from the definition of η.

Therefore, we obtain the following sequence of inequalities:

$$\mathbb{P}\left[D(\varphi,\mathbf{P}) \leq \frac{\alpha}{\|\mathbf{d}\|_1 m\rho_\Delta} - \delta \,|\, \mathcal{E}\right] \overset{(a)}{\geq} \mathbb{P}\left[\max\left\{Z(\mathbf{x}_\eta^i,\mathbf{P}) \,:\, i \in [N]\right\} \leq \frac{\alpha}{\|\mathbf{d}\|_1 m\rho_\Delta} - \frac{3\delta}{2} \,|\, \mathcal{E}\right]$$

$$= \mathbb{P}\left[\max\left\{ Z(\mathbf{x}_\eta^i, \mathbf{P}) \,:\, i \in [N] \right\} \leq \delta \,|\, \mathcal{E} \right],$$

where (a) follows from that $D(\varphi, \mathbf{P}) \equiv \max\{|\varphi_\theta(\mathbf{x}, \mathbf{P}) - \varphi_\theta(\mathbf{x})| \,:\, \mathbf{x} \in \mathcal{X}\}$.

Consequently, the following sequence of inequalities is satisfied:

$$\begin{aligned}
\mathbb{P}\left[\Delta(\mathbf{PF}, \mathbf{c})\,(\delta + D(\varphi, \mathbf{P})) \leq \frac{\alpha}{\|\mathbf{d}\|_1 m} \right] &\geq \mathbb{P}\left[\rho_\Delta\,(\delta + D(\varphi, \mathbf{P})) \leq \frac{\alpha}{\|\mathbf{d}\|_1 m} \,|\, \mathcal{E} \right] \mathbb{P}\left[\mathcal{E}\right] \\
&\overset{(a)}{=} \mathbb{P}\left[\max\left\{ Z(\mathbf{x}_\eta^i, \mathbf{P}) \,:\, i \in [N] \right\} \leq \delta \right] \\
&\quad - \mathbb{P}\left[\max\left\{ Z(\mathbf{x}_\eta^i, \mathbf{P}) \,:\, i \in [N] \right\} \leq \delta \,|\, \mathcal{E}^c \right] \mathbb{P}\left[\mathcal{E}^c\right] \\
&\overset{(b)}{\geq} \mathbb{P}\left[\max\left\{ Z(\mathbf{x}_\eta^i, \mathbf{P}) \,:\, i \in [N] \right\} \leq \delta \right] - \mathbb{P}\left[\mathcal{E}^c\right],
\end{aligned}$$

where (a) follows from the law of total probability (Jacod and Protter 2012). Also, (b) follows from $\mathbb{P}[E] \leq 1$ for any event E.

Finally, observe that the following sequence of inequalities is satisfied:

$$\begin{aligned}
\mathbb{P}\left[\max\left\{ Z(\mathbf{x}_\eta^i, \mathbf{P}) \,:\, i \in [N] \right\} \leq \delta \right] &= 1 - \mathbb{P}\left[\left\{ \max\left\{ Z(\mathbf{x}_\eta^i, \mathbf{P}) \,:\, i \in [N\left(\mathcal{X}, \eta\,;\, \|\cdot\|_1\right)] \right\} \leq \delta \right\}^c \right] \\
&\overset{(a)}{=} 1 - \mathbb{P}\left[\bigcup_{i \in [N(\mathcal{X}, \eta\,;\, \|\cdot\|_1)]} \left\{ Z(\mathbf{x}_\eta^i, \mathbf{P}) \leq \delta \right\}^c \right] \\
&\geq 1 - \sum_{i \in [N(\mathcal{X}, \eta\,;\, \|\cdot\|_1)]} \mathbb{P}\left[Z(\mathbf{x}_\eta^i, \mathbf{P}) > \delta \right] \\
&\overset{(b)}{\geq} 1 - 2mN\left(\mathcal{X}, \eta\,;\, \|\cdot\|_1\right)] e^{-C\left(\varepsilon^2 - \varepsilon^3\right)k} \\
&\overset{(c)}{\geq} 1 - 2m\left(1 + \frac{8L_Z\theta}{\delta}\right)^n e^{-C\left(\varepsilon^2 - \varepsilon^3\right)k},
\end{aligned}$$

where (a) follows from a union of bounds. Moreover, (b) follows from Theorem 3, where $\varepsilon \equiv \mathcal{O}\left(\frac{\delta}{1+\theta}\right)$. Finally, (c) follows from Lemma 3 with $\eta = \frac{\delta}{2L_Z}$.

Next, observe that $\mathbb{P}[\mathcal{E}^c] = \mathbb{P}\left[\{\max\{\Delta(\mathbf{PF}, \mathbf{c}), \|\mathbf{PL}\|_\infty\} > \rho_\Delta\right] = 1 - F_\Delta(\rho_\Delta)$ by the definition of the distribution F_Δ in the statement of Theorem 4. Accordingly, since we select $\alpha := \frac{5}{2}\|\mathbf{d}\|_1 m\rho_\Delta\delta \leq 3\delta\|\mathbf{d}\|_1 m\rho_\Delta$, then we obtain the following inequality:

$$\mathbb{P}\left[d_\theta(\mathbf{P}, \delta) - z^* \leq 3\rho_\Delta\|\mathbf{d}\|_1 m\delta \right] \geq F_\Delta(\rho_\Delta) - \kappa e^{-C\left(\varepsilon^2 - \varepsilon^3\right)k},$$

where $\varepsilon \equiv \mathcal{O}\left(\frac{\delta}{1+\theta}\right)$ and where we define $\kappa := 2m\left(1 + \frac{8L_Z\theta}{\delta}\right)^n$. This inequality concludes in turn the proof of Theorem 4. ■

B.2 Proofs for Section 4.2.2

Proof of Theorem 5. To begin, we introduce the following constant term:

$$u^* = \max_{\mathbf{x}\in\mathcal{X}} \|\mathbf{u}^*(\mathbf{x})\|_2 \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2,$$

where for any given $\mathbf{x} \in \mathcal{X}$, $\mathbf{u}^*(\mathbf{x})$ represents the optimal solution (of minimal ℓ_2-norm) to the follower's dual problem. Then, we denote by $\big(\mathbf{x}, \mathbf{y}^*(\mathbf{x})\big)$ the pair of optimal solution to the leader's problem (1), and, introduce auxiliary program:

$$f_\theta(\mathbf{P}, \delta) \ := \ \max\big\{\mathbf{a}^\top\mathbf{x} + \mathbf{d}^\top\mathbf{y} \ : \ (\mathbf{x}, \mathbf{y}) \in \mathcal{S}_\theta(\mathbf{P}, \delta)\big\}, \tag{22}$$

where $\mathcal{S}_\theta(\mathbf{P}, \delta)$ is as defined in (7) in Section 3.3.

Next, we apply Theorem 3 to $\mathbf{x}$ which implies that the following inequality is satisfied:

$$\mathbb{P}\left[\varphi_\theta(\mathbf{x}) \le \varphi_\theta(\mathbf{x}, \mathbf{P}) + \delta \le \varphi_\theta(\mathbf{x}) + \delta\right] \ge 1 - 4me^{-C\left(\varepsilon^2 - \varepsilon^3\right)k},$$

for $\varepsilon = \mathcal{O}\big(\frac{\delta}{(1+\theta)u^*}\big)$.

Hence, we have that $\big(\mathbf{x}, \mathbf{y}^*(\mathbf{x})\big)$ satisfies the constrains of (7) with probability at least $p \equiv 1 - 4me^{-C\left(\varepsilon^2-\varepsilon^3\right)k}$. Consequently, the following inequality is naturally satisfied:

$$z^* \le f_\theta(\mathbf{P}, \delta),$$

with probability at least p, by using a feasibility-based argument.

To proceed, given some arbitrary leader's feasible decision $\tilde{\mathbf{x}} \in \mathcal{X}$ and some $u \in \mathbb{R}_+$, we introduce the set of *approximate follower's responses.* Namely, this set represents all follower's feasible decisions which satisfy a relaxed notion of optimality. Specifically, we introduce:

$$D_\theta(\tilde{\mathbf{x}}, u) := \operatorname*{argmax}_{\mathbf{y}\in\mathbb{R}^m_+} \left\{0 \ : \ \mathbf{F}\mathbf{y} = \mathbf{f} - \mathbf{L}\tilde{\mathbf{x}}\,,\, \mathbf{1}_m^\top\mathbf{y} \le \theta,\, \mathbf{c}^\top\mathbf{y} \le u\right\}.$$

From this definition, one can then verify that $D_\theta(\tilde{\mathbf{x}}, \varphi_\theta(\tilde{\mathbf{x}}))$ simply represents the set of follower's rational responses to the leader's feasible decision $\tilde{\mathbf{x}}$.

Next, we work on the event that $(\mathbf{x}, \mathbf{y}^*(\mathbf{x}))$ is a feasible solution to (7). This event occurs with probability at least $p \equiv 1 - 4me^{-C\left(\varepsilon^2-\varepsilon^3\right)k}$. On this event, consider $\big(\mathbf{x}^\delta, \mathbf{y}^\delta(\mathbf{x}^\delta)\big) \in \mathcal{S}_\theta(\mathbf{P}, \delta)$ a solution in which $f_\theta(\mathbf{P}, \delta)$ is attained. In particular, one can observe that $\mathbf{y}^\delta(\mathbf{x}^\delta) \in D_\theta(\mathbf{x}^\delta, \varphi_\theta(\mathbf{x}^\delta, \mathbf{P}) + \delta)$ by optimality of $\big(\mathbf{x}^\delta, \mathbf{y}^\delta(\mathbf{x}^\delta)\big) \in \mathcal{S}_\theta(\mathbf{P}, \delta)$. Therefore, we have that $\mathbf{y}^\delta(\mathbf{x}^\delta) \in D_\theta(\mathbf{x}^\delta, \varphi_\theta(\mathbf{x}^\delta) + \delta)$ since $\varphi_\theta(\mathbf{x}^\delta, \mathbf{P}) + \delta \le \varphi_\theta(\mathbf{x}^\delta) + \delta$. Moreover, observe that the follower's reaction set $\mathcal{R}\left(\mathbf{x}^\delta\right)$ for the leader's feasible decision $\mathbf{x}^\delta$ satisfies $\mathcal{R}\left(\mathbf{x}^\delta\right) = D_\theta(\mathbf{x}^\delta, \varphi_\theta(\mathbf{x}^\delta))$.

We use proximity results for linear programs, in particular Theorem 5 by Cook et al. (1986). Namely, since $\mathbf{y}^\delta(\mathbf{x}^\delta) \in D_\theta(\mathbf{x}^\delta, \varphi_\theta(\mathbf{x}^\delta) + \delta)$, there exists a follower's rational response $\mathbf{y}^*(\mathbf{x}^\delta) \in D_\theta(\mathbf{x}^\delta, \varphi_\theta(\mathbf{x}^\delta))$ that satisfies the following inequality:

$$\|\mathbf{y}^\delta(\mathbf{x}^\delta) - \mathbf{y}^*(\mathbf{x}^\delta)\|_\infty \leq m\Delta(\mathbf{F}, \mathbf{c})\delta,$$

where $\Delta(\mathbf{F}, \mathbf{c})$ the maximum of the absolute values of the determinants of the square submatrices of $[\,\mathbf{F}^\top \,|\, -\mathbf{F}^\top \,|\, \mathbf{c} \,|\, \mathbf{1}_m\,]$. Note that tighter Liptschtz constant can be derived (in comparison to $\Delta(\mathbf{F}, \mathbf{c})$) as in the study by Mangasarian and Shiau (1987). Accordingly, we have that:

$$\begin{aligned}
|\mathbf{a}^\top \mathbf{x}^\delta + \mathbf{d}^\top \mathbf{y}^*(\mathbf{x}^\delta) - \big(\mathbf{a}^\top \mathbf{x}^\delta + \mathbf{d}^\top \mathbf{y}^\delta(\mathbf{x}^\delta)\big)| &= |\mathbf{d}^\top \big(\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta(\mathbf{x}^\delta)\big)| \\
&\leq \|\mathbf{d}\|_1 \|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta(\mathbf{x}^\delta)\|_\infty \\
&\leq \|\mathbf{d}\|_1 m\Delta(\mathbf{F}, \mathbf{c})\delta.
\end{aligned}$$

Then, with probability at least $1 - 4me^{-C\left(\varepsilon^2 - \varepsilon^3\right)k}$, we obtain the following inequalities:

$$\begin{aligned}
f_\theta(\mathbf{P}, \delta) &= \mathbf{a}^\top \mathbf{x}^\delta + \mathbf{d}^\top \mathbf{y}^\delta(\mathbf{x}^\delta) - \Big(\mathbf{a}^\top \mathbf{x}^\delta + \mathbf{d}^\top \mathbf{y}^*(\mathbf{x}^\delta)\Big) + \Big(\mathbf{a}^\top \mathbf{x}^\delta + \mathbf{d}^\top \mathbf{y}^*(\mathbf{x}^\delta)\Big) \\
&\leq \|\mathbf{d}\|_1 m\Delta(\mathbf{F}, \mathbf{c})\delta + s_\theta(\mathbf{P}, \delta),
\end{aligned}$$

where the last inequality follows by the definition of $s_\theta(\mathbf{P}, \delta)$ as the optimal solution to the feasibility problem (8). In particular, by the definition of $s_\theta(\mathbf{P}, \delta)$, we have that $s_\theta(\mathbf{P}, \delta) \leq z^*$. Finally, the following inequalities are guaranteed to hold:

$$s_\theta(\mathbf{P}, \delta) \leq z^* \leq f_\theta(\mathbf{P}, \delta) \leq s_\theta(\mathbf{P}, \delta) + \|\mathbf{d}\|_1 m\Delta(\mathbf{F}, \mathbf{c})\delta \leq z^* + \|\mathbf{d}\|_1 m\Delta(\mathbf{F}, \mathbf{c})\delta,$$

with probability at least $1 - 4me^{-C\left(\varepsilon^2 - \varepsilon^3\right)k}$, which, in turn, concludes the proof. ■

B.3 Additional results and proofs for Section 4.3

First, Section B.3.1 presents an example showing that the leader's optimal solution to (8) fails to satisfy the coupling constraints. Next, in Section B.3.2, we present another example where the leader's optimal solution to (10) does not satisfy the coupling constraints. Then, Section B.3.3 provides the formal proofs of the coupling-constraints results stated in Section 4.3.

B.3.1 Coupling constraints violation for the feasibility problem

We show that solving (8) might return a leader's decision which, together with its corresponding follower's rational response, violate the upper-level constraints. Specifically, we construct an instance of a bilevel program such that (i) the pair $\left(\mathbf{x}^{\delta}, \mathbf{y}^{\delta}\right)$ satisfies the upper-level constraints, and (ii) the pair $\left(\mathbf{x}^{\delta}, \mathbf{y}^{*}(\mathbf{x}^{\delta})\right)$ does not not satisfy the coupling constraints. To illustrate this phenomena, we consider a small instance and a projector, which exhibit together, the said behavior.

Step 1 (Construction of a bilevel program): We introduce a bilevel program of the form:

$$
\begin{aligned}
\max_{\mathbf{x},\mathbf{y}^*(\mathbf{x})} \quad & x_1 \\
\text{s.t.} \quad & \sum_{i=1}^{n} x_i \leq 1\,, \quad y_3^*\left(\mathbf{x}\right) \leq 0 \\
& \mathbf{y}^*\left(\mathbf{x}\right) \in \operatorname*{argmin}_{\tilde{\mathbf{y}} \in [0,1]^n} \left\{ \tilde{y}_4 \,:\, \tilde{y}_1 + \tilde{y}_2 = 1 + x_1\,,\, \tilde{y}_3 + \tilde{y}_4 = \varepsilon\,,\, \varepsilon \cdot x_i \leq \tilde{y}_i \;\forall\, i \in [n] \setminus \{3\} \right\} \\
& \mathbf{x} \in \{0,1\}^n\,,\, \mathbf{y}^*\left(\mathbf{x}\right) \in \mathbb{R}^n_+\,,
\end{aligned}
\tag{25}
$$

where $\varepsilon \in (0,1)$. The follower's problem is feasible for every leader's feasible decision and its feasible region is bounded so that Assumption **A1** $-$ **A3** are satisfied; recall our discussion in Section 1.

Step 2 (Construction of an instance): Next, rather than deriving the results in its full generality for some family of random projection, we design a projector $\mathbf{P} \in \mathbb{R}^{2n \times (2n+1)}$ that exhibits the desired behavior. Specifically, we introduce:

$$
\mathbf{P} = \begin{pmatrix} \tilde{\mathbf{P}} & \mathbf{0}_{1\times(2n-1)} \\ \mathbf{0}_{(2n-1)\times 2} & \mathbf{I}_{(2n-1)\times(2n-1)} \end{pmatrix},
$$

where the projector $\tilde{\mathbf{P}} \in \mathbb{R}^{1\times 2}$ is defined as follows:

$$
\tilde{\mathbf{P}} = (1\,,\,0)\,.
$$

Note that, the projector $\tilde{\mathbf{P}}$ is sampled with positive probability under the sparse family of random projections introduced by Achlioptas (2003), where entries are drawn from $\{-1, 0, 1\}$ with

$\mathbb{P}[P_{ij} = \pm 1] = 1/6$ and $\mathbb{P}[P_{ij} = 0] = 2/3$. Moreover, the projected follower's problem used inside (8), given some leader's feasible decision $\mathbf{x} \in \mathcal{X}$, takes the form:

$$\varphi(\mathbf{x}, \mathbf{P}) := \min_{\tilde{\mathbf{y}} \in [0,1]^n} \Big\{ \tilde{y}_4 \ : \ \tilde{y}_1 + \tilde{y}_2 = 1 + x_1 \, , \, \varepsilon \cdot x_i \leq \ \tilde{y}_i \ \forall \, i \in [n] \setminus \{3\} \Big\},$$

so $\varphi(\mathbf{x}, \mathbf{P}) = 0$ when $x_1 = 1$, since $\tilde{y}_4 = 0$ in the optimal solution to the projected follower's problem.

Step 3 (Solving the projected problem): One can verify that, for any $\delta \geq \varepsilon$, the pair $\left(\mathbf{x}^\delta, \mathbf{y}^\delta\right)$ defined by the following two vectors:

$$\mathbf{x}^\delta = (1, 0, 0, 0, \ldots, 0) \ , \qquad \mathbf{y}^\delta = (1, 1, 0, \varepsilon \, , \ldots, 0)$$

is feasible for the problem (7) constructed from (25). Since the objective function value of (7) in that case is bounded above by 1, $\left(\mathbf{x}^\delta, \mathbf{y}^\delta\right)$ is an optimal solution to (7). The corresponding follower's rational response for that leader's decision $\mathbf{x}^\delta$ is given by the following vector:

$$\mathbf{y}^* \left(\mathbf{x}^\delta\right) = (1, 1, \varepsilon, 0, 0 \ \ldots, 0) \ ,$$

which is arbitrarily close to $\mathbf{y}^\delta$ as ε goes to 0, but violates the coupling constraint, since $y_3^* \left(\mathbf{x}^\delta\right) = \varepsilon > 0$. Hence $\left(\mathbf{x}^\delta, \mathbf{y}^*(\mathbf{x}^\delta)\right)$ is not bilevel feasible, which concludes the example.

B.3.2 Coupling constraints violation for the adjusted surrogate problem

We show that solving (10) might return a leader's decision, which, together with its corresponding follower's rational response, violate the coupling constraints. Specifically, we construct an instance such that (i) the pair $\left(\mathbf{x}^\delta, \mathbf{y}^\delta\right)$, if it exists, always satisfies the upper-level constraints, but (ii) the pair $\left(\mathbf{x}^\delta, \mathbf{y}^*(\mathbf{x}^\delta)\right)$ does not satisfy the coupling constraints. To illustrate this phenomena, we consider a bilevel program and a projector which exhibit this behavior.

Step 1 (Construction of a bilevel program): Accordingly, we let $n := 2k$ be arbitrary, for some even integer $k > 3$. Then, we introduce the following bilevel LP:

$$\begin{aligned}
\max_{\mathbf{x},\mathbf{y}} \quad & 2\sum_{i=1}^{k} x_i + \sum_{i=k+1}^{n} x_i + \sum_{i=1}^{n} y_i^*(\mathbf{x}) \\
\text{s.t.} \quad & \sum_{i=1}^{n} x_i \leq 2 \, , \quad \sum_{i=1}^{n} y_i^*(\mathbf{x}) \leq \frac{k}{2} \\
& \mathbf{y}^*(\mathbf{x}) \in \operatorname*{argmin}_{\mathbf{y} \in [0,1]^n} \Big\{ \sum_{i=1}^{n} \tilde{y}_i \ : \ \tilde{y}_{2i-1} + \tilde{y}_{2i} = \frac{1}{2} + \frac{x_{2i-1} + x_{2i}}{2} \quad \forall \, i \in [k] \Big\} \\
& \mathbf{x} \in \{0,1\}^n \, , \, \mathbf{y}^*(\mathbf{x}) \in \mathbb{R}^n_+ .
\end{aligned} \tag{26}$$

The bilevel program in (26) above exhibits the following two properties. First, the follower's problem is feasible for every leader's feasible decision. Second, the follower's feasible region is bounded. Accordingly, Assumption **A1** – **A3** are all satisfied; recall our discussion in Section 1.

Step 2 (Construction of an instance): Next, rather than deriving the results in its full generality for some random projection, which would require an extensive probabilistic approach, we design a projector $\mathbf{P} \in \mathbb{R}^{(\frac{k}{2}+n)\times(k+n)}$ that exhibits the desired behavior. Specifically, we introduce:

$$\mathbf{P} = \begin{pmatrix} \tilde{\mathbf{P}} & \mathbf{0}_{\frac{k}{2}\times n} \\ \mathbf{0}_{n\times k} & \mathbf{I}_{n\times n} \end{pmatrix},$$

where $\tilde{\mathbf{P}} \in \mathbb{R}^{\frac{k}{2}\times k}$ defined by:

$$\tilde{\mathbf{P}} = \left(\begin{array}{ccc|c} 1 & 0 & 0 & \\ \vdots & \vdots & \ddots & \mathbf{I}_{\frac{k}{2}\times\frac{k}{2}} \\ 1 & 0 & 0 & \\ 1 & 0 & 0 & \end{array}\right)$$

Note that $\tilde{\mathbf{P}}$ can be sampled from the family of random projections introduced by Achlioptas (2003). Specifically, for these random projectors, each entry is drawn from $\{-1, 0, 1\}$ with $\mathbb{P}[P_{ij} = \pm 1] = 1/6$ and $\mathbb{P}[P_{ij} = 0] = 2/3$. Hence, the probability of sampling $\tilde{\mathbf{P}}$ is positive, though potentially small. Next, we introduce the projected follower's problem. That is, given some leader's feasible decision $\mathbf{x} \in \mathcal{X}$, we introduce:

$$\begin{aligned} \varphi(\mathbf{x}, \mathbf{P}) := \min\ & \mathbf{1}_n^\top \tilde{\mathbf{y}} \\ \text{s.t.}\ & \tilde{y}_1 + \tilde{y}_2 + \tilde{y}_{k+2i-1} + \tilde{y}_{k+2i} = 1 + \frac{x_1 + x_2}{2} + \frac{x_{k+2i-1} + x_{k+2i}}{2}, \quad \forall\, i \in [k/2] \\ & \tilde{\mathbf{y}} \in [0,1]^n. \end{aligned}$$

Step 3 (Solving the projected problem): One can verify that for any $\delta \geq 0$, the pair $(\mathbf{x}, \mathbf{y})$ that is defined by $\mathbf{x} = (1, 1, \mathbf{0}_{n-2})$ and $\mathbf{y} = (1, 1, \mathbf{0}_{n-2})$ is an optimal solution to (10) constructed from (26). Indeed, one can verify that the optimal objective function value is at most 6 and that $(\mathbf{x}, \mathbf{y})$ is feasible for (10) and its corresponding objective function value is 6. The follower's rational response for that leader's decision $\mathbf{x}$ is given by:

$$\mathbf{y}^*(\mathbf{x}) := \big(1, \frac{1}{2}, \underbrace{0, \frac{1}{2}, 0, \frac{1}{2}, \ldots, 0, \frac{1}{2}}_{\text{other } n-2 \text{ elements}}\big).$$

On the other hand, one can verify that $(\mathbf{x}, \mathbf{y}^*(\mathbf{x}))$ is not bilevel feasible as $\sum_{i=1}^n y_i^*(\mathbf{x}) = 1 + k/2 > \frac{k}{2}$. Specifically, the coupling constraints are not satisfied, which concludes the example.

B.3.3 Proofs for Section 4.3.1

As a start, we clarify what we mean by "probability" in this section. We first assume that both $\mathbf{H}$ and $\mathbf{P}$ are defined over a probability space, denoted by $\left(\Omega^{(\mathbf{H}}, \mathcal{B}^{(\mathbf{H})}, \mathbb{P}^{(\mathbf{H})}\right)$ and $\left(\Omega^{(\mathbf{P}}, \mathcal{B}^{(\mathbf{P})}, \mathbb{P}^{(\mathbf{P})}\right)$, respectively. To construct the joint probability space on which both $\mathbf{H}$ and $\mathbf{P}$ are defined independently, let $\Omega = \Omega^{(\mathbf{H})} \times \Omega^{(\mathbf{P})}$ and $\mathcal{B} = \mathcal{B}^{(\mathbf{H})} \otimes \mathcal{B}^{(\mathbf{P})}$, where $\mathcal{B}^{(\mathbf{H})} \otimes \mathcal{B}^{(\mathbf{P})}$ denotes the σ-algebra generated by all measurable rectangles $A \times B$ with $A \in \mathcal{B}^{(\mathbf{H})}$ and $B \in \mathcal{B}^{(\mathbf{P})}$. We then define the product measure $\mathbb{P}$ on such rectangles by $\mathbb{P}\left[A \times B\right] = \mathbb{P}^{(\mathbf{H})}[A]\,\mathbb{P}^{(\mathbf{P})}[B]$ and extend it uniquely to all of $\mathcal{B}$ via the extension of Fubini's theorem (Jacod and Protter 2012). The resulting triplet $(\Omega, \mathcal{B}, \mathbb{P})$ is the canonical product probability space whose marginals are $\mathbb{P}^{(\mathbf{H})}$ and $\mathbb{P}^{(\mathbf{P})}$.

In this section, we present the proofs for the results in Section 4.3.1 on the likelihood of violating coupling constraints. We begin with what might be the main result of this section, as it quantifies the probability that the solution obtained by solving (7), together with the corresponding follower's rational response, violates the coupling constraints. Specifically:

Proof of Proposition 1. First, we fix δ as in the proposition's statement and let $\left(\mathbf{x}^\delta, \mathbf{y}^\delta\right)$ denote the optimal solution to (7) whenever it exists. Importantly, $\left(\mathbf{x}^\delta, \mathbf{y}^\delta\right)$ is a random vector that depends on the realization of both $\mathbf{H}$ and the random projector $\mathbf{P}$. We are interested in deriving an upper bound to the following probability:

$$\mathbb{P}\Bigg[\exists\, j \in [p]\,,\ \mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) > h_j\Bigg],$$

where $\mathbf{y}^*(\mathbf{x}^\delta)$ is the follower's rational response, given that the leader's acts according to $\mathbf{x}^\delta$, and $\mathbf{G}_j$ and $\mathbf{H}_j$ corresponds to the j-th row of $\mathbf{G}$ and $\mathbf{H}$, respectively.

Then, we denote by $\mathcal{E}$ the event defined by:

$$\mathcal{E} := \Big\{\omega \in \Omega\ :\ \varphi_\theta(\mathbf{x}^\delta) - \delta \le \varphi_\theta(\mathbf{x}^\delta, \mathbf{P}) \le \varphi_\theta(\mathbf{x}^\delta)\Big\}$$

To begin, observe that the following sequence of inequalities holds:

$$\begin{aligned}
\mathbb{P}\big[\mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) > h_j\big] &= \mathbb{P}\big[\mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{H}_j\mathbf{y}^\delta > h_j - \mathbf{H}_j\mathbf{y}^\delta\big] \\
&\le \mathbb{P}\big[\mathbf{H}_j\big(\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\big) > \Delta^\delta\big] \\
&\le \mathbb{P}\big[\mathbf{H}_j\mathbf{1}_m \cdot \|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty > \Delta^\delta\big] \\
&\overset{(a)}{=} \mathbb{P}\big[\mathbf{H}_j\mathbf{1}_m \cdot \|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty > \Delta^\delta \,|\, \mathcal{E}\big] \cdot \mathbb{P}\left[\mathcal{E}\right] \\
&\quad + \mathbb{P}\big[\mathbf{H}_j\mathbf{1}_m \cdot \|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty > \Delta^\delta \,|\, \mathcal{E}^c\big] \cdot \mathbb{P}\left[\mathcal{E}^c\right] = P_1 + P_2
\end{aligned}$$

where (a) holds by using the law of total probability (Jacod and Protter 2012), and where P_1 and P_2 are defined by two terms respectively:

$$P_1 := \mathbb{P}\big[\mathbf{H}_j\mathbf{1}_m \cdot \|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty > \Delta^\delta \,|\, \mathcal{E}\big] \cdot \mathbb{P}\left[\mathcal{E}\right]$$
$$P_2 := \mathbb{P}\big[\mathbf{H}_j\mathbf{1}_m \cdot \|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty > \Delta^\delta \,|\, \mathcal{E}^c\big] \cdot \mathbb{P}\left[\mathcal{E}^c\right].$$

Step 1 (Bounding P_1): Using Corollary 2 and in particular the proof of Theorem 5, one can verify that, conditional on the event $\mathcal{E}$, the following inequality is satisfied:

$$\|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty \leq m\Delta(\mathbf{F}, \mathbf{c})\delta,$$

where we implicitly use Assumption **A6**, i.e., the optimal solution to the follower's problem is unique.

Consequently, we can bound P_1 as follows. First, we obtain the following sequence of inequalities:

$$\begin{aligned} P_1 = \mathbb{P}\big[\mathbf{H}_j\mathbf{1}_m \cdot \|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty > \Delta^\delta \,|\, \mathcal{E}\big] \cdot \mathbb{P}\left[\mathcal{E}\right] &\leq \mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\Delta^\delta}{m\Delta(\mathbf{F}, \mathbf{c})\delta} \,|\, \mathcal{E}\right] \cdot \mathbb{P}\left[\mathcal{E}\right] \\ &\leq \mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\Delta^\delta}{m\Delta(\mathbf{F}, \mathbf{c})\delta}\right]. \end{aligned}$$

Observe that the entries of $\mathbf{H}$ are independent and identically distributed with $H_{j\ell} \sim \mathcal{U}\big([-a, a]\big)$ for all $j \in [p]$ and $\ell \in [m]$ after conditioning on Δ^δ by assumption. Thus, $H_{j\ell}$ satisfies $\mathbb{E}\big[H_{j\ell} \,|\, \Delta^\delta\big] = 0$ for all $j \in [p]$ and $\ell \in [m]$. By the definition of Δ^δ, we have that $\Delta^\delta \geq 0$ $\mathbb{P}$-almost surely. Using the Hoeffding's inequality (Boucheron et al. 2013), we obtain the following inequalities:

$$\begin{aligned} \mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\Delta^\delta}{2m\Delta(\mathbf{F}, \mathbf{c})\delta}\right] &= \mathbb{E}_{\Delta^\delta}\left[\mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\Delta^\delta}{2m\Delta(\mathbf{F}, \mathbf{c})\delta} \,|\, \Delta^\delta\right]\right] \\ &\leq \mathbb{E}_{\Delta^\delta}\left[\exp\left(-\frac{2\left(\Delta^\delta\right)^2}{4ma^2 \cdot (m\Delta(\mathbf{F}, \mathbf{c})\delta)^2}\right)\right] \\ &\leq \mathbb{E}_{\Delta^\delta}\left[\exp\left(-\frac{\left(\Delta^\delta\right)^2}{2m^3a^2\Delta(\mathbf{F}, \mathbf{c})^2 \cdot \delta^2}\right)\right], \end{aligned}$$

which provides the first part of the statement's inequality.

Step 2 (Bounding P_2): We provide a bound to P_2 by using a rather different approach than in Step 1. Namely, we use the fact that the following inequality is trivially satisfied:

$$P_2 = \mathbb{P}\big[\mathbf{H}_j\mathbf{1}_m \cdot \|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty > \Delta^\delta \,|\, \mathcal{E}^c\big]\mathbb{P}\left[\mathcal{E}^c\right] \leq \mathbb{P}\left[\mathcal{E}^c\right].$$

Next, our goal is to use the "decoupling" trick from the proof of Theorem 4 to obtain the desired bound. We introduce the event $\mathcal{V} \equiv \{\max\{\Delta(\mathbf{PF}, \mathbf{c}), \|\mathbf{PL}\|_\infty\} \leq \rho_\Delta\}$. Observe that, conditional on the event $\mathcal{V}$, the bound $L_Z(\mathbf{P})$ as defined in Lemma 4 satisfies the following inequality:

$$L_Z(\mathbf{P}) \leq \max\left\{\|\mathbf{c}\|_1 m\left(\rho_\Delta^2 + \Delta(\mathbf{F}, \mathbf{c})\|\mathbf{L}\|_\infty\right), \frac{1}{2}\right\} =: L_Z.$$

Moreover, we define $\eta := \delta(2L_Z)^{-1}$. Then, we define the finite set $\mathcal{X}_\eta := \left(\mathbf{x}_\eta^i\right)_{i\in[N]}$, where $N \equiv N(\mathcal{X}, \eta; \|\cdot\|_1)$ is as given in Lemma 3, as the η-cover of $\mathcal{X}$ in the ℓ_1-norm.

Moreover, the following inequality is satisfied:

$$\mathbb{P}[\mathcal{E}^c] = \mathbb{P}\left[\varphi_\theta(\mathbf{x}^\delta) > \varphi_\theta(\mathbf{x}^\delta, \mathbf{P}) + \delta\right]$$

To proceed, assume that the event $\mathcal{E}^c$ occurs. Importantly, $\mathbf{x}^\delta$ is a random variable induced by both $\mathbf{H}$ and $\mathbf{P}$. Accordingly, conditioning on $\mathcal{E}^c$ does not make $\mathbf{x}^\delta$ any less random. However, on $\mathcal{E}^c$, realizations of $\mathbf{x}^\delta$ for which $\varphi_\theta(\mathbf{x}^\delta) > \varphi_\theta(\mathbf{x}^\delta, \mathbf{P}) + \delta$ is satisfied. For any given realization $\mathbf{x}$ of $\mathbf{x}^\delta$, there exists $i \in [N]$ such that $\|\mathbf{x} - \mathbf{x}_\eta^i\|_1 \leq \eta$ by the definition of $\mathcal{X}_\eta$. Consequently, for any given realization $\mathbf{x}$ of $\mathbf{x}^\delta$, there exists $i \in [N]$, which satisfies the following inequality:

$$|Z(\mathbf{x}_\eta^i, \mathbf{P}) - Z(\mathbf{x}, \mathbf{P})| + \varphi_\theta(\mathbf{x}_\eta^i) - \varphi_\theta(\mathbf{x}_\eta^i, \mathbf{P}) \geq \delta$$

where $Z(\cdot, \mathbf{P})$ is as defined in the proof of Theorem 4.

Next, we use the Lipschitz continuity of $Z(\cdot, \mathbf{P})$ from Lemma 4. In particular, we obtain the following inequality:

$$|Z(\mathbf{x}_\eta^i, \mathbf{P}) - Z(\mathbf{x}, \mathbf{P})| \leq L_Z(\mathbf{P})\|\mathbf{x}_\eta^i - \mathbf{x}\|_1 \leq L_Z(\mathbf{P})\eta,$$

where $L_Z(\mathbf{P})$ is as defined in Lemma 4. Therefore, the following inequality holds:

$$\mathbb{P}[\mathcal{E}^c] \leq \mathbb{P}\left[\sup_{i\in[N]}\left\{\varphi_\theta(\mathbf{x}_\eta^i) - \varphi_\theta(\mathbf{x}_\eta^i, \mathbf{P})\right\} \geq \delta - L_Z(\mathbf{P})\eta\right].$$

By applying the law of total probability (Jacod and Protter 2012), we obtain the following sequence of inequalities:

$$\begin{aligned}
\mathbb{P}\left[\varphi_\theta(\mathbf{x}_\eta^i) - \varphi_\theta(\mathbf{x}_\eta^i, \mathbf{P}) \geq \delta - L_Z(\mathbf{P})\eta\right] &= \mathbb{P}\left[\sup_{i\in[N]}\left\{\varphi_\theta(\mathbf{x}_\eta^i) - \varphi_\theta(\mathbf{x}_\eta^i, \mathbf{P})\right\} \geq \delta - L_Z(\mathbf{P})\eta \,|\, \mathcal{V}\right]\mathbb{P}[\mathcal{V}] \\
&\quad + \mathbb{P}\left[\sup_{i\in[N]}\left\{\varphi_\theta(\mathbf{x}_\eta^i) - \varphi_\theta(\mathbf{x}_\eta^i, \mathbf{P})\right\} \geq \delta - L_Z(\mathbf{P})\eta \,|\, \mathcal{V}^c\right]\mathbb{P}[\mathcal{V}^c] \\
&\overset{(a)}{\leq} \mathbb{P}\left[\sup_{i\in[N]}\left\{\varphi_\theta(\mathbf{x}_\eta^i) - \varphi_\theta(\mathbf{x}_\eta^i, \mathbf{P})\right\} \geq \delta - L_Z\eta \,|\, \mathcal{V}\right]\mathbb{P}[\mathcal{V}] + \mathbb{P}[\mathcal{V}^c] \\
&\overset{(b)}{=} \mathbb{P}\left[\sup_{i\in[N]}\left\{\varphi_\theta(\mathbf{x}_\eta^i) - \varphi_\theta(\mathbf{x}_\eta^i, \mathbf{P})\right\} \geq \delta - L_Z\eta\right]
\end{aligned}$$

$$- \mathbb{P}\left[\sup_{i\in[N]}\left\{\varphi_\theta(\mathbf{x}^i_\eta) - \varphi_\theta(\mathbf{x}^i_\eta, \mathbf{P})\right\} \geq \delta - L_Z\eta \,|\, \mathcal{V}^c\right]\mathbb{P}\left[\mathcal{V}^c\right] + \mathbb{P}\left[\mathcal{V}^c\right]$$

$$\overset{(c)}{\leq} \mathbb{P}\left[\sup_{i\in[N]}\left\{\varphi_\theta(\mathbf{x}^i_\eta) - \varphi_\theta(\mathbf{x}^i_\eta, \mathbf{P})\right\} \geq \frac{\delta}{2}\right] + 1 - F_\Delta(\rho_\Delta)$$

$$\overset{(d)}{\leq} 4mNe^{-\tilde{C}(\varepsilon^2-\varepsilon^3)k} + 1 - F_\Delta(\rho_\Delta),$$

where (a) comes from that $L_Z(\mathbf{P}) \leq L_Z$ on $\mathcal{V}$. Also, (b) follows from once again applying the law of total probability. Moreover (c) follows from the definition of η and from the definition of F_Δ. Finally, (d) is a consequence of applying Theorem 3.

Observe that, by construction, we have that $N \leq \left(1 + \frac{4\theta}{\eta}\right)^n = \left(1 + \frac{8\theta L_Z}{\delta}\right)^n$. Additionally, by construction, we have that $L_Z = \max\left\{\|\mathbf{c}\|_1 m\left(\rho_\Delta^2 + \Delta\left(\mathbf{F}, \mathbf{c}\right)\|\mathbf{L}\|_\infty\right), \frac{1}{2}\right\}$. Then, the following sequence of inequalities holds:

$$\mathbb{P}\left[\mathcal{E}^c\right] = \mathbb{P}\left[\varphi_\theta(\mathbf{x}^\delta) > \varphi_\theta(\mathbf{x}^\delta, \mathbf{P}) + \delta\right] \overset{(a)}{\leq} \kappa e^{-\tilde{C}\left(\varepsilon^2-\varepsilon^3\right)k} + 1 - F_\Delta(\rho_\Delta) = 1 - p_\Delta(\delta, \tilde{C}),$$

where (a) follows from that $\varepsilon := \mathcal{O}\left(\frac{\delta}{(1+\theta)u^*}\right)$, where $u^* := \max\left\{\|\mathbf{u}^*(\mathbf{x})\|_2\|\mathbf{f} - \mathbf{L}\mathbf{x} \,:\, \mathbf{x} \in \mathcal{X}\right\}$ and $\mathbf{u}^*(\mathbf{x})$ corresponds to the optimal solution (of minimal ℓ_2-norm) to the follower's dual problem (and provided that δ is chosen small enough, which we omit to explicitly define for brevity). Also, by its construction, κ satisfies $\kappa \equiv \mathcal{O}\left(m\left(1 + \frac{\theta\rho_\Delta^2 m}{\delta}\right)^n\right)$.

Step 3 (Putting all together): We now merge the bounds from Steps 1 and 2 to derive an upper bound on the probability that the coupling constraints are violated. In particular, we obtain the following chain of inequalities by using a union of bounds:

$$\mathbb{P}\Big[\exists\, j \in [p]\,,\, \mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) > h_j\Big] \leq \sum_{j\in[p]} \mathbb{P}\big[\mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) > h_j\big]$$

$$\leq p \cdot \mathbb{E}_{\Delta^\delta}\left[\exp\left(-\frac{\left(\Delta^\delta\right)^2}{2m^3a^2\Delta(\mathbf{F}, \mathbf{c})^2 \cdot \delta^2}\right) + 1 - p_\Delta(\delta, \tilde{C})\right],$$

which, in turn, concludes the proof. ■

Proposition 1 still involves the random variable Δ^δ, whose distribution is not immediately transparent. To overcome this issue, we complement it with three results: Corollary 4, Lemma 2, and its simplified form in Proposition 2. Together, these statements give explicit bounds on the deviation of the upper-level coupling constraints for the optimal solution to problem (7) without requiring detailed knowledge of the law of Δ^δ. We now present the proofs of all three results below.

Proof of Corollary 4. The proof closely follows the one from Proposition 1. First, we fix δ and let $(\mathbf{x}^\delta, \mathbf{y}^\delta)$ denote the optimal solution to (7) whenever it exists. Importantly, $(\mathbf{x}^\delta, \mathbf{y}^\delta)$ is a random vector that depends on the realization of both $\mathbf{H}$ and the random projector $\mathbf{P}$. We are interested in deriving an upper bound to the probability of violating coupling constraints, namely to:

$$\mathbb{P}\left[\exists\, j \in [p]\,,\, \mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) > h_j\right],$$

where $\mathbf{y}^*(\mathbf{x}^\delta)$ is the follower's rational response, given that the leader's acts according to $\mathbf{x}^\delta$, and $\mathbf{G}_j$ and $\mathbf{H}_j$ corresponds to the j-th row of $\mathbf{G}$ and $\mathbf{H}$, respectively.

Then, we denote by $\mathcal{E}$ the event defined by:

$$\mathcal{E} := \left\{\omega \in \Omega \,:\, \varphi_\theta(\mathbf{x}^\delta) - \delta \le \varphi_\theta(\mathbf{x}^\delta, \mathbf{P}) \le \varphi_\theta(\mathbf{x}^\delta)\right\}$$

To begin, similarly to the proof of Proposition 1, observe that the following inequality is satisfied:

$$\mathbb{P}\big[\mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) > h_j\big] \le P_1 + P_2$$

where P_1 and P_2 are defined by:

$$\begin{aligned}
P_1 &:= \mathbb{P}\big[\mathbf{H}_j\mathbf{1}_m \cdot \|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty > \Delta^\delta \,|\, \mathcal{E}\big] \cdot \mathbb{P}\left[\mathcal{E}\right] \\
P_2 &:= \mathbb{P}\big[\mathbf{H}_j\mathbf{1}_m \cdot \|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty > \Delta^\delta \,|\, \mathcal{E}^c\big] \cdot \mathbb{P}\left[\mathcal{E}^c\right].
\end{aligned}$$

Step 1 (Bounding P_1): Using Corollary 2 and in particular the proof of Theorem 5, one can verify that, conditional on the event $\mathcal{E}$, the following inequality is satisfied:

$$\|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty \le m\Delta(\mathbf{F}, \mathbf{c})\delta,$$

where we implicitly use Assumption **A6**, i.e., the optimal solution to the follower's problem is unique.

Consequently, we can bound P_1 as follows. First, we obtain the following sequence of inequalities:

$$\begin{aligned}
P_1 = \mathbb{P}\big[\mathbf{H}_j\mathbf{1}_m \cdot \|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty > \Delta^\delta \,|\, \mathcal{E}\big] \cdot \mathbb{P}\left[\mathcal{E}\right] &\le \mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\Delta^\delta}{m\Delta(\mathbf{F}, \mathbf{c})\delta} \,|\, \mathcal{E}\right] \cdot \mathbb{P}\left[\mathcal{E}\right] \\
&\le \mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\Delta^\delta}{m\Delta(\mathbf{F}, \mathbf{c})\delta}\right] \\
&\le \mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\eta}{m\Delta(\mathbf{F}, \mathbf{c})\delta}\right],
\end{aligned}$$

where we use that $\Delta^\delta > \eta$ almost surely in the last inequality.

Observe that the entries of $\mathbf{H}$ are independent and identically distributed with $H_{j\ell} \sim \mathcal{U}\big([-a, a]\big)$ for all $j \in [p]$ and $\ell \in [m]$. Thus, $H_{j\ell}$ satisfies $\mathbb{E}\big[H_{j\ell}\big] = 0$ for all $j \in [p]$ and $\ell \in [m]$. Using the

Hoeffding's inequality (Boucheron et al. 2013), we obtain the following inequality:

$$\mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\Delta^\delta}{m\Delta(\mathbf{F},\mathbf{c})\delta}\right] \leq \exp\left(-\frac{\eta^2}{2m^3a^2\Delta(\mathbf{F},\mathbf{c})^2\cdot\delta^2}\right).$$

Step 2 (Bounding P_2): We provide an upper bound to P_2 by using the same approach as in the proof of Proposition 1. Since $\mathbf{x}^\delta$ is a random variable, we use the decoupling trick from the proof of Theorem 4 to obtain the bound. Specifically, the following inequality is satisfied:

$$P_2 \leq \kappa e^{-\tilde{C}\left(\varepsilon^2-\varepsilon^3\right)k} + 1 - F_\Delta(\rho_\Delta) = 1 - p_\Delta(\delta,\tilde{C}),$$

where $\varepsilon := \mathcal{O}\left(\frac{\delta}{(1+\theta)u^*}\right)$, with $u^* := \max\{\|\mathbf{u}^*(\mathbf{x})\|_2\|\mathbf{f}-\mathbf{L}\mathbf{x} : \mathbf{x}\in\mathcal{X}\}$ and $\mathbf{u}^*(\mathbf{x})$ corresponds to the optimal solution (of minimal ℓ_2-norm) to the follower's dual problem (and provided that δ is chosen small enough, which we omit to explicitly define). Also, we have that $\kappa \equiv \mathcal{O}\left(m\left(1+\frac{\theta\rho_\Delta^2 m}{\delta}\right)^n\right)$.

Step 3 (Putting all together): We now merge the bounds from Steps 1 and 2 to derive an upper bound on the probability that the coupling constraints are violated. In particular, we obtain the following chain of inequalities by using a union of bounds:

$$\begin{aligned}\mathbb{P}\left[\exists j\in[p],\, \mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) > h_j\right] &\leq \sum_{j\in[p]}\mathbb{P}\left[\mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) > h_j\right]\\ &\leq p\cdot\left(\exp\left(-\frac{\eta^2}{2m^3a^2\Delta(\mathbf{F},\mathbf{c})^2\cdot\delta^2}\right) + 1 - p_\Delta(\delta,\tilde{C})\right).\end{aligned}$$

where $\varepsilon := \mathcal{O}\left(\frac{\delta}{(1+\theta)u^*}\right)$, with $u^* := \max\{\|\mathbf{u}^*(\mathbf{x})\|_2\|\mathbf{f}-\mathbf{L}\mathbf{x} : \mathbf{x}\in\mathcal{X}\}$ and $\mathbf{u}^*(\mathbf{x})$ corresponds to the optimal solution (of minimal ℓ_2-norm) to the follower's dual problem (and provided that δ is chosen small enough, which we omit to explicitly define). Also, we have that $\kappa \equiv \mathcal{O}\left(m\left(1+\frac{\theta\rho_\Delta^2 m}{\delta}\right)^n\right)$. The end proof is immediate by combining all the bounds. ■

Proof of Lemma 2. Fix j, δ and $\alpha\in(0,1)$ as in the statement of the lemma. Let $\left(\mathbf{x}^\delta,\mathbf{y}^\delta\right)$ be the the optimal solution to (7), which is a random vector that is induced by both $\mathbf{H}$ and $\mathbf{P}$. Moreover, we pick $s < \min\{h_j - \mathbf{G}_j\mathbf{x} : \mathbf{x}\in\mathcal{X}\}$ arbitrarily. Assume that the event:

$$\mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^\delta > h_j - s$$

occurs. Then, it follows that the following inequality is satisfied:

$$\mathbf{H}_j\mathbf{y}^\delta > h_j - s - \mathbf{G}_j\mathbf{x}^\delta.$$

Accordingly, we have that the following inequality is satisfied:

$$|\sum_{\ell\in[m]} H_{j\ell}y_\ell^\delta| > h_j - s - \mathbf{G}_j\mathbf{x}^\delta,$$

and, since $s < h_j - \mathbf{G}_j \mathbf{x}^\delta$, both the left-hand side and the right-hand side of the inequality are non-negative. Consequently, the inequality is also satisfied when one takes the absolute value of both side. Hence, we obtain the following inequality:

$$\sum_{\ell \in [m]} |H_{j\ell}| y_\ell^\delta > h_j - s - \mathbf{G}_j \mathbf{x}^\delta,$$

and, if we divide both sides by $a > 0$, then we obtain the following inequality:

$$\sum_{\ell \in [m]} \frac{|H_{j\ell}|}{a} y_\ell^\delta > \frac{h_j - s - \mathbf{G}_j \mathbf{x}^\delta}{a}.$$

Moreover, recall that the follower's problem (in both (2) and (7)) is assumed to be bounded. That is, we have that $\mathbf{y}^\delta$ satisfies $\mathbf{1}_m^\top \mathbf{y}^\delta \leq \theta$. Thus, the following inequality is satisfied:

$$\sum_{\ell \in [m]} \frac{|H_{j\ell}|}{a} > \frac{h_j - s - \mathbf{G}_j \mathbf{x}^\delta}{a\theta}.$$

Next, if we define $Y_{j\ell} := \frac{|H_{j\ell}|}{a}$, for $\ell \in [m]$, then we have that $Y_{j\ell} \sim \mathcal{U}\big([0,1]\big)$ so that $Y_{j1}, \ldots Y_{jm}$ are independent and uniformly distributed. Next, if we define $X_j := \sum_{\ell \in [m]} Y_{j\ell}$, then we have that X_j follows a Irwan-Hall (IH) distribution (Irwin 1927; Hall 1927). Let $q_{1-\alpha}^{IH}$ denotes the $1-\alpha$ quantile of the Irwan-Hall distribution. Next, if we define $s = \tau_j \equiv h_j - \max\big\{\mathbf{G}_j \mathbf{x} \,:\, \mathbf{x} \in \mathcal{X}\big\} - a\theta q_{1-\alpha}^{IH}$, then we have that that the following inequalities are satisfied:

$$\tau_j < h_j - \mathbf{G}_j \mathbf{x}^\delta \quad \text{and} \quad \mathbb{P}\Bigg[\sum_{\ell \in [m]} \frac{|H_{j\ell}|}{a} \geq \frac{h_j - \tau_j - \mathbf{G}_j \mathbf{x}^\delta}{a\theta}\Bigg] = \mathbb{P}\Bigg[\sum_{\ell \in [m]} \frac{|H_{j\ell}|}{a} \geq q_{1-\alpha}^{IH}\Bigg] \leq \alpha,$$

where the first inequality has to be understood as holding $\mathbb{P}$-almost-surely.

Hence, with this choice of $\tau_j \equiv \tau_j(\alpha, \theta, a)$, we have that:

$$\mathbb{P}\Bigg[\mathbf{G}_j \mathbf{x}^\delta + \mathbf{H}_j \mathbf{y}^\delta > h_j - \tau_j\Bigg] \leq \alpha.$$

Additionally, the second part of our result follows from that the Irwan-Hall distribution is a continuous distribution with support in $[0, m]$, which, in turn, concludes the proof. ■

Proof of Proposition 2. The proof closely follows the one from Corollary 4. First, we fix δ and let $\big(\mathbf{x}^\delta, \mathbf{y}^\delta\big)$ denote the optimal solution to (7) whenever it exists. Importantly, $\big(\mathbf{x}^\delta, \mathbf{y}^\delta\big)$ is a random vector that depends on the realization of both $\mathbf{H}$ and the random projector $\mathbf{P}$.

Fix $j \in [p]$. Then, recall that, by Lemma 2, for all α, $\tau_j \equiv \tau_j(\alpha, \theta, a)$ defined by $\tau_j = h_j - \max\big\{\mathbf{G}_j \mathbf{x} \,:\, \mathbf{x} \in \mathcal{X}\big\} - a\theta q_{1-\alpha}^{IH}$, where $q_{1-\alpha}^{IH}$ corresponds to the $1-\alpha$ quantile of the Irwan-Hall distribution, we have that that the following inequality holds:

$$\mathbb{P}\big[\tau_j \geq \mathbf{h}_j - \mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^\delta\big] \leq \alpha.$$

For τ defined as in Proposition 2, we obtain by a union of bounds:

$$\mathbb{P}\big[\tau \geq \Delta^\delta\big] \leq \alpha p,$$

where $\Delta^\delta := \min\big\{h_j - \big(\mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^\delta\big) \,:\, j \in [p]\big\}$.

Next, we are interested in deriving an upper bound to the following probability:

$$\mathbb{P}\Bigg[\exists\, j \in [p]\,,\, \mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) > h_j\Bigg],$$

where $\mathbf{y}^*(\mathbf{x}^\delta)$ is the follower's rational response, given that the leader's acts according to $\mathbf{x}^\delta$, and $\mathbf{G}_j$ and $\mathbf{H}_j$ corresponds to the j-th row of $\mathbf{G}$ and $\mathbf{H}$, respectively.

Then, we denote by $\mathcal{E}$ the event defined by:

$$\mathcal{E} := \Big\{\omega \in \Omega \,:\, \varphi_\theta(\mathbf{x}^\delta) - \delta \leq \varphi_\theta(\mathbf{x}^\delta, \mathbf{P}) \leq \varphi_\theta(\mathbf{x}^\delta)\Big\}$$

To begin, similarly to the proof of Proposition 1, observe that the following inequality is satisfied:

$$\mathbb{P}\big[\mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) > h_j\big] \leq P_1 + P_2$$

where (i) holds by using the law of total probability (Jacod and Protter 2012), and where P_1 and P_2 are defined by:

$$P_1 := \mathbb{P}\big[\mathbf{H}_j\mathbf{1}_m \cdot \|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty > \Delta^\delta \,|\, \mathcal{E}\big] \cdot \mathbb{P}\,[\mathcal{E}]$$
$$P_2 := \mathbb{P}\big[\mathbf{H}_j\mathbf{1}_m \cdot \|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty > \Delta^\delta \,|\, \mathcal{E}^c\big] \cdot \mathbb{P}\,[\mathcal{E}^c]\,.$$

Step 1 (Bounding P_1): Using Corollary 2 and in particular the proof of Theorem 5, one can verify that, conditional on the event $\mathcal{E}$, the following inequality is satisfied:

$$\|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty \leq m\Delta(\mathbf{F}, \mathbf{c})\delta,$$

where we implicitly use Assumption **A6**, i.e., the optimal solution to the follower's problem is unique.

Consequently, we can bound P_1 as follows. First, we obtain the following sequence of inequalities:

$$\begin{aligned} P_1 = \mathbb{P}\big[\mathbf{H}_j\mathbf{1}_m \cdot \|\mathbf{y}^*(\mathbf{x}^\delta) - \mathbf{y}^\delta\|_\infty > \Delta^\delta \,|\, \mathcal{E}\big] \cdot \mathbb{P}\,[\mathcal{E}] &\leq \mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\Delta^\delta}{m\Delta(\mathbf{F}, \mathbf{c})\delta} \,|\, \mathcal{E}\right] \cdot \mathbb{P}\,[\mathcal{E}] \\ &\leq \mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\Delta^\delta}{m\Delta(\mathbf{F}, \mathbf{c})\delta}\right]. \end{aligned}$$

Then, we derive the following sequence of inequalities:

$$\begin{aligned}
\mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\Delta^\delta}{m\Delta(\mathbf{F},\mathbf{c})\delta}\right] &\overset{(a)}{=} \mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\Delta^\delta}{m\Delta(\mathbf{F},\mathbf{c})\delta} \,|\, \tau \geq \Delta^\delta\right]\mathbb{P}\left[\tau \geq \Delta^\delta\right] \\
&\quad + \mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\Delta^\delta}{m\Delta(\mathbf{F},\mathbf{c})\delta} \,|\, \tau < \Delta^\delta\right]\mathbb{P}\left[\tau < \Delta^\delta\right] \\
&\leq \mathbb{P}\left[\tau \geq \Delta^\delta\right] + \mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\Delta^\delta}{m\Delta(\mathbf{F},\mathbf{c})\delta} \,|\, \tau < \Delta^\delta\right]\mathbb{P}\left[\tau < \Delta^\delta\right] \\
&\leq \alpha p + \mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\tau}{m\Delta(\mathbf{F},\mathbf{c})\delta} \,|\, \tau < \Delta^\delta\right]\mathbb{P}\left[\tau < \Delta^\delta\right] \\
&\leq \alpha p + \mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\tau}{m\Delta(\mathbf{F},\mathbf{c})\delta}\right]
\end{aligned}$$

where (a) follows by the law of total expectation (Jacod and Protter 2012).

Observe that the entries of $\mathbf{H}$ are independent and identically distributed with $H_{j\ell} \sim \mathcal{U}\big([-a,a]\big)$ for all $j \in [p]$ and $\ell \in [m]$. Thus, $H_{j\ell}$ satisfies $\mathbb{E}\big[H_{j\ell}\big] = 0$ for all $j \in [p]$ and $\ell \in [m]$. Using the Hoeffding's inequality (Boucheron et al. 2013), we obtain the following inequality:

$$\mathbb{P}\left[\mathbf{H}_j\mathbf{1}_m > \frac{\tau}{m\Delta(\mathbf{F},\mathbf{c})\delta}\right] \leq \exp\left(-\frac{\tau^2}{2m^3a^2\Delta(\mathbf{F},\mathbf{c})^2\cdot\delta^2}\right).$$

Step 2 (Bounding P_2): We provide a bound to P_2 by using the exact same approach as in the proof of Proposition 1. Note that, implicitly, since $\mathbf{x}^\delta$ is a random variable that depends on both $\mathbf{H}$ and $\mathbf{P}$, we use the decoupling trick from the proof of Theorem 4 to obtain the bound. Specifically, the following inequality is satisfied:

$$P_2 \leq \kappa e^{-\tilde{C}\left(\varepsilon^2-\varepsilon^3\right)k} + 1 - F_\Delta(\rho_\Delta) = 1 - p_\Delta(\delta,\tilde{C}),$$

where $\varepsilon := \mathcal{O}\left(\frac{\delta}{(1+\theta)u^*}\right)$, with $u^* := \max\{\|\mathbf{u}^*(\mathbf{x})\|_2\|\mathbf{f}-\mathbf{L}\mathbf{x} \,:\, \mathbf{x}\in\mathcal{X}\}$ and $\mathbf{u}^*(\mathbf{x})$ corresponds to the optimal solution (of minimal ℓ_2-norm) to the follower's dual problem (and provided that δ is chosen small enough, which we omit to explicitly define). Also, we have that $\kappa \equiv \mathcal{O}\left(m\left(1+\frac{\theta\rho_\Delta^2 m}{\delta}\right)^n\right)$.

Step 3 (Putting all together): We now merge the bounds from Steps 1 and 2 to derive an upper bound on the probability that the coupling constraints are violated. In particular, we obtain the following chain of inequalities by using a union of bounds:

$$\begin{aligned}
\mathbb{P}\left[\exists\, j\in[p]\,,\, \mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) > h_j\right] &\leq \sum_{j\in[p]}\mathbb{P}\big[\mathbf{G}_j\mathbf{x}^\delta + \mathbf{H}_j\mathbf{y}^*(\mathbf{x}^\delta) > h_j\big] \\
&\leq p\cdot\left(\alpha p + \exp\left(-\frac{\tau^2}{2m^3a^2\Delta(\mathbf{F},\mathbf{c})^2\cdot\delta^2}\right) + 1 - p_\Delta(\delta,\tilde{C})\right),
\end{aligned}$$

which concludes the proof. ∎

B.3.4 Preemptive correction for coupling constraints

In this section, we introduce a simple preemptive correction of the upper-level constraints: by uniformly reducing each right-hand side h_j in (1) by a carefully chosen margin ζ, we ensure that the resulting leader's decision x^δ together with the follower's response $y^*(x^\delta)$ satisfies all coupling constraints with certainty. We then prove a proximity result, analogous to that of Section 4.2.2, showing that the loss in the leader's objective due to this right-hand-side adjustment is only of order $\mathcal{O}(\delta)$.

To proceed, we first fix some random projector $\mathbf{P}$, $\delta \in \big(0, \min\big\{\varphi_\theta(\mathbf{x}) \,:\, \mathbf{x} \in \mathcal{X}\big\}\big)$, and $\zeta > 0$. Then, we introduce the following adjusted program:

$$\begin{aligned} f_\theta(\mathbf{P}, \delta, \zeta) := \max_{\mathbf{x},\mathbf{y}} \ & \mathbf{a}^\top \mathbf{x} + \mathbf{d}^\top \mathbf{y} \\ \text{s.t.} \ & \mathbf{G}\mathbf{x} + \mathbf{H}\mathbf{y} \leq \mathbf{h} - \zeta \mathbf{1}_p \\ & \mathbf{y} \in \mathcal{Y}_\theta(\mathbf{x}) \\ & \mathbf{c}^\top \mathbf{y} \leq \varphi_\theta(\mathbf{x}, \mathbf{P}) + \delta \\ & \mathbf{x} \in \{0,1\}^{n_1} \times \mathbb{R}_+^{n_2}, \, \mathbf{y} \in \mathbb{R}_+^m, \end{aligned} \tag{45}$$

and denote the corresponding set of optimal solution by $\mathcal{S}_\theta(\mathbf{P}, \delta, \zeta)$. Then, we define:

$$s_\theta(\mathbf{P}, \delta, \zeta) := \max \Big\{ \mathbf{a}^\top \mathbf{x} + \mathbf{d}^\top \mathbf{y}^*(\mathbf{x}) \,:\, \exists \mathbf{y} \in \mathbb{R}_+^m, \, \big(\mathbf{x}, \mathbf{y}\big) \in \mathcal{S}_\theta(\mathbf{P}, \delta, \zeta) \Big\}.$$

Next, we fix $\zeta := \|\mathbf{H}\|_1 m \Delta(\mathbf{F}, \mathbf{c}) \delta$ and we assume that there exists an optimal solution $(\mathbf{x}^*, \mathbf{y}^*(\mathbf{x}^*))$ to (1) which satisfies $\mathbf{G}_j \mathbf{x}^* + \mathbf{H}_j \mathbf{y}^*(\mathbf{x}^*) < h_j$ for all $j \in [p]$. Thus, there exists $\delta > 0$ sufficiently small such that $\mathcal{S}_\theta(\mathbf{I}_{q\times q}, 0, \zeta) \neq \emptyset$. Therefore, we have that $\mathcal{S}_\theta(\mathbf{P}, \delta, \zeta) \neq \emptyset$ is also satisfied with probability at least $p \geq 1 - 4me^{-C\left(\varepsilon^2 - \varepsilon^3\right)k}$, where $\varepsilon = \mathcal{O}\big(\frac{\delta}{(1+\theta)u^*}\big)$ and $u^* := \min\left\{\|\mathbf{u}^*(\mathbf{x})\|_2 \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2 \,:\, \mathbf{x} \in \mathcal{X}\right\}$. Recall that $\mathbf{u}^*(\mathbf{x})$ is the optimal solution (of minimal ℓ_2-norm) to the follower's dual problem by applying Theorem 3 on the leader's optimal decision $\mathbf{x}^*$. Consequently, (45) is feasible with probability at least p. Let $(\mathbf{x}^{\delta,\zeta}, \mathbf{y}^{\delta,\zeta})$ be an optimal solution to (45) whenever it exists.

We proceed under the assumption that the event in Theorem 3 holds for $\mathbf{x}^{\delta,\zeta}$. All subsequent results are conditional on this event, which we subsequently refer to as $\mathcal{E}(\delta)$. Hence, they should be understood as holding with probability at least $p \geq 1 - 4me^{-C\left(\varepsilon^2 - \varepsilon^3\right)k}$, where $\varepsilon = \mathcal{O}\big(\frac{\delta}{(1+\theta)u^*}\big)$.

Then, by using a proximity argument similar to Corollary 2, the following inequality is satisfied:

$$\|\big(\mathbf{x}^{\delta,\zeta}, \mathbf{y}^{\delta,\zeta}\big) - \big(\mathbf{x}^{\delta,\zeta}, \mathbf{y}^*(\mathbf{x}^{\delta,\zeta})\big)\|_\infty \leq m\Delta(\mathbf{F}, \mathbf{c})\delta,$$

Also, by its feasibility, we have that $\big(\mathbf{x}^{\delta,\zeta}, \mathbf{y}^{\delta,\zeta}\big)$ satisfies the upper-level constraints. To proceed, we first show that $(\mathbf{x}^{\delta,\zeta}, \mathbf{y}^*(\mathbf{x}^{\delta,\zeta}))$ is bilevel feasible. Then, we provide performance guarantees.

Step 1 (Bilevel feasibility to the original problem): To establish bilevel feasibility, we leverage the proximity result introduced above. By applying this result, we obtain the following chain of inequalities:

$$
\begin{aligned}
\mathbf{G}\mathbf{x}^{\delta,\zeta} + \mathbf{H}\mathbf{y}^*(\mathbf{x}^{\delta,\zeta}) &= \mathbf{G}\mathbf{x}^{\delta,\zeta} + \mathbf{H}\mathbf{y}^{\delta,\zeta} + \mathbf{H}\mathbf{y}^*(\mathbf{x}^{\delta,\zeta}) - \mathbf{H}\mathbf{y}^{\delta,\zeta} \\
&\leq \mathbf{h} - \zeta\mathbf{1}_p + \mathbf{H}\big(\mathbf{y}^*(\mathbf{x}^{\delta,\zeta}) - \mathbf{y}^{\delta,\zeta}\big) \\
&\leq \mathbf{h} - \zeta\mathbf{1}_p + \|\mathbf{H}\|_1 \|\mathbf{y}^*(\mathbf{x}^{\delta,\zeta}) - \mathbf{y}^{\delta,\zeta}\|_\infty \leq \mathbf{h} - \zeta\mathbf{1}_p + \|\mathbf{H}\|_1 m\Delta(\mathbf{F},\mathbf{c})\delta\mathbf{1}_p = \mathbf{h},
\end{aligned}
$$

where the last inequality follows by the definition of ζ. Hence, we have that $\big(\mathbf{x}^{\delta,\zeta}, \mathbf{y}^*(\mathbf{x}^{\delta,\zeta})\big)$ is a bilevel feasible solution to (1). Note that, this feasibility argument implies that $s_\theta(\mathbf{P},\delta,\zeta) \leq z^*$.

Step 2 (Value function based reformulation): We formulate an adjusted leader's problem (1) by embedding the follower's rational response into its upper-level constraints (i.e., to include the coupling constraints) and then adjusting the right-hand side of the upper-level constraints. We first fix some leader's feasible decision $\mathbf{x} \in \mathcal{X}$. Then, recall the dual of the follower's problem (9) defines exactly the value function we need. Formally:

$$
\varphi_\theta(\mathbf{x}) := \varphi_\theta(\mathbf{x}, \mathbf{I}_{q\times q}) := \max\Big\{ \big(\mathbf{f} - \mathbf{L}\mathbf{x}\big)^\top \mathbf{u} + \theta v \,:\, \mathbf{F}^\top\mathbf{u} + v\mathbf{1}_m \leq \mathbf{c}^\top ,\, \mathbf{u} \in \mathbb{R}^q ,\, v \leq 0 \Big\}.
$$

Note that, by Proposition 3.1 from Blair and Jeroslow (1977), the value function of a linear program is piecewise-convex linear. Therefore, we obtain:

$$
\varphi_\theta(\mathbf{x}) := \max\big\{ \big[\, (\mathbf{f} - \mathbf{L}\mathbf{x})^\top , \theta\big]^\top \mathbf{e}^{(i)} \,:\, i \in [L] \big\}
$$

where $\mathbf{e}^{(1)}, \ldots, \mathbf{e}^{(L)}$ are the extreme points of the polyhedra defined by:

$$
\big\{ \big(\mathbf{u}, v\big) \in \mathbb{R}^{q+1} \,:\, \mathbf{F}^\top\mathbf{u} + v\mathbf{1}_m \leq \mathbf{c}^\top ,\, v \leq 0 \big\}.
$$

Next, we reformulate the leader's problem (1) by substituting the follower's value function with its piecewise-linear representation, and updating the right-hand sides of the upper-level constraints.

$$
\begin{aligned}
f_\theta(\mathbf{I}_{q\times q}, 0, \zeta) := \max_{\mathbf{x},\mathbf{y}^*(\mathbf{x})} \quad & \mathbf{a}^\top\mathbf{x} + \mathbf{d}^\top\mathbf{y}^*(\mathbf{x}) \\
\text{s.t.} \quad & \mathbf{G}\mathbf{x} + \mathbf{H}\mathbf{y}^*(\mathbf{x}) \leq \mathbf{h} - \zeta\mathbf{1}_p \\
& \mathbf{c}^\top\mathbf{y}^*(\mathbf{x}) \leq \big[\, (\mathbf{f} - \mathbf{L}\mathbf{x})^\top , \theta\big]^\top \mathbf{e}^{(i)} \quad \forall\, i \in [L] \\
& \mathbf{x} \in \{0,1\}^{n_1} \times \mathbb{R}^{n_2}_+ ,\, \mathbf{y}^*(\mathbf{x}) \in \mathcal{Y}_\theta(\mathbf{x}).
\end{aligned} \tag{47}
$$

Note that the reformulation in (47) may become exponentially large, since the number of extreme points of the underlying polyhedron can grow exponentially (Bertsimas and Tsitsiklis 1997).

Step 3 (Bound on the adjusted optimal objective function value): By an application of Theorem 5 and in particular Corollary 2 to (47) instead of the original leader's problem (by simply changing the right-hand side of the upper-level constraints), we then have that $\left(\mathbf{x}^{\delta,\zeta}, \mathbf{y}^{\delta,\zeta}\right)$ satisfies:

$$|\mathbf{a}^\top \mathbf{x}^{\delta,\zeta} + \mathbf{d}^\top \mathbf{y}^{\delta,\zeta} - f_\theta(\mathbf{I}_{q\times q}, 0, \zeta)| \leq \|\mathbf{d}\|_1 m\Delta(\mathbf{F}, \mathbf{c})\delta,$$

on the event $\mathcal{E}(\delta)$.

Step 4 (Proximity bound for the bilevel-feasible solution): We bound the deviation of $(\mathbf{x}^{\delta,\zeta},\ \mathbf{y}^*(x^{\delta,\zeta}))$ from the leader's optimal objective function value z^* by considering:

$$|s_\theta(\mathbf{P}, \delta, \zeta) - z^*|.$$

For that, we proceed in three elementary steps. Specifically, by using the triangle inequality, we obtain the following inequality:

$$|s_\theta(\mathbf{P}, \delta, \zeta) - z^*| \leq \Delta_1 + \Delta_2 + \Delta_3,$$

where Δ_1, Δ_2 and Δ_3 are defined by:

$$\begin{aligned}
\Delta_1 &= \left|\mathbf{a}^\top \mathbf{x}^{\delta,\zeta} + \mathbf{d}^\top \mathbf{y}^*(\mathbf{x}^{\delta,\zeta}) - \left(\mathbf{a}^\top \mathbf{x}^{\delta,\zeta} + \mathbf{d}^\top \mathbf{y}^{\delta,\zeta}\right)\right|, \\
\Delta_2 &= \left|\mathbf{a}^\top \mathbf{x}^{\delta,\zeta} + \mathbf{d}^\top \mathbf{y}^{\delta,\zeta} - f_\theta(\mathbf{I}_{q\times q}, 0, \zeta)\right|, \\
\Delta_3 &= \left|f_\theta(\mathbf{I}_{q\times q}, 0, \zeta) - z^*\right|,
\end{aligned}$$

respectively.

By the discussion preceding Step 1, we have that $\Delta_1 \leq \|\mathbf{d}\|_1 m\Delta(\mathbf{F}, \mathbf{c})\delta$, and by Step 3, $\Delta_2 \leq \|\mathbf{d}\|_1 m\Delta(\mathbf{F}, \mathbf{c})\delta$. The last step is to find an upper bound for Δ_3. By using the reformulation in (47), we can use classical proximity results for mixed-integer linear programs. Specifically, by the strong proximity result (Theorem 2.1 by Blair and Jeroslow (1977)), there are constants $C_1, C_2 > 0$ such that for all feasible $\zeta > 0$, the following inequality is satisfied:

$$|f_\theta(\mathbf{I}_{q\times q}, 0, \zeta) - z^*| \leq C_1\zeta + C_2.$$

Note that C_1 and C_2 might be exponentially large due to the number of constraints in (47). Accordingly, in the worst case, a small value of δ might be required to obtain a meaningful bound. However, such choice for δ might reduce the quality of the probabilistic lower bound. Moreover, following Theorem 2.4 by Mangasarian and Shiau (1987), if the leader's decision variables are all continuous, then the constant term C_2 disappears and we have that:

$$|f_\theta(\mathbf{I}_{q\times q}, 0, \zeta) - z^*| \leq C_1\zeta.$$

Note that, the two constants that appear in Theorem 2.1 by Blair and Jeroslow (1977) and Theorem 2.4 by Mangasarian and Shiau (1987) are not the same. However, to simplify our notations, we pick C_1 as the maximum between those two constants. Then, as $\zeta := \|\mathbf{H}\|_1 m\Delta(\mathbf{F}, \mathbf{c})\delta$, we obtain the following two inequalities:

$$|s_\theta(\mathbf{P}, \delta, \zeta) - z^*| \leq \big(2\|\mathbf{d}\|_1 + C_1\|\mathbf{H}\|_1\big)m\Delta(\mathbf{F}, \mathbf{c})\delta + C_2, \text{ if } n_1 > 0,$$
$$|s_\theta(\mathbf{P}, \delta, \zeta) - z^*| \leq \big(2\|\mathbf{d}\|_1 + C_1\|\mathbf{H}\|_1\big)m\Delta(\mathbf{F}, \mathbf{c})\delta, \text{ if } n_1 = 0.$$

As a final remark, while this section enforces exact feasibility of the coupling constraints, alternative approximation approaches are available. Instead of imposing all constraints, one can obtain comparable guarantees with explicit control over allowable constraint violations. Such controlled-violation schemes are standard in approximation algorithms, including bilevel settings, where constraints are relaxed by a tunable margin. We omit these variants, as similar guarantees follow from arguments analogous to those above. Nevertheless, similar techniques can be adapted to produce approximate solutions with explicit violation budgets, which we leave as future research.

C Proof of Theorem 6 in Section 5

In this section, we complement our results from Section 3 to provide further guarantees on the quality of our bounds. We show that by restricting the class of projectors, we are able to somehow provide information on how "bad" the bounds from Section 3 might be. Specifically, we show that, under some conditions on k and δ, there exists some family $\mathcal{P}_\delta^{(k)}$, $C_1 > 0$ and $C_2 > 0$ that satisfies the following inequalities:

$$z^* - C_1\delta \le s_\theta(\mathbf{P}, \delta) \le z^* \le s_\theta(\mathbf{P}) \le z^* + C_2\delta,$$

for all projector $\mathbf{P} \in \mathcal{P}_\delta^{(k)}$.

Assumption. As discussed in Section 4.3, coupling constraints pose challenges for deriving approximation guarantees whenever using an approximate follower's response. Accordingly, in this section we assume $\mathbf{H} \equiv \mathbf{0}_{p\times m}$, thereby eliminating all coupling constraints. We anticipate that analogous results extend to the general case with nonzero $\mathbf{H}$, but we omit that extension here for clarity.

To begin, we introduce $\mathcal{Z} \subseteq \{0,1\}^{n_1}$, the set of first-stage binary vectors $\mathbf{x}^{(1)}$ for which there exists a second-stage vector $\mathbf{x}^{(2)} \in \mathbb{R}^{n_2}$ such that $\mathbf{x} \equiv (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \in \mathcal{X}$. In other words, $\mathcal{Z}$ is defined by:

$$\mathcal{Z} := \left\{\mathbf{x}^{(1)} \in \{0,1\}^{n_1} \;:\; \exists\, \mathbf{x}^{(2)} \in \mathbb{R}^{n_2}\,,\; (\mathbf{x}^{(1)}, \mathbf{x}^{(2)}) \in \mathcal{X}\right\}.$$

For any $\mathbf{x}^{(1)} \in \mathcal{Z}$, we can define a leader's problem analogue to (1) that is parametrized by $\mathbf{x}^{(1)}$ in which the integer part of the leader's decision is constrained to be equal to $\mathbf{x}^{(1)}$. This parametrized problem is a bilevel linear program with continuous decision variables at both levels. Accordingly, there exists an optimal solution to this parametrized bilevel linear program that is attained at an extreme point of the polyhedral (Bard and Falk 1982; Bialas and Karwan 1982):

$$\mathcal{P}(\mathbf{x}^{(1)}) := \left\{\left(\tilde{\mathbf{x}}, \mathbf{y}\right) \in \mathbb{R}^n_+ \times \mathbb{R}^m_+ \;:\; \mathbf{G}\tilde{\mathbf{x}} + \mathbf{H}\mathbf{y} \le \mathbf{h}\,,\; \mathbf{F}\mathbf{y} + \mathbf{L}\tilde{\mathbf{x}} = \mathbf{f}\,,\; \tilde{x}_j = x_j^{(1)} \quad \forall\, j \in [n_1]\right\}.$$

We denote by $\mathcal{E}(\mathbf{x}^{(1)}) \subseteq \mathcal{P}(\mathbf{x}^{(1)})$ the set of leader's feasible decision that are extreme points of the parametrized polyhedral $\mathcal{P}(\mathbf{x}^{(1)})$. Then, we define $\Lambda \subseteq \mathcal{X}$ as the subset of leader's feasible decisions that might potentially be optimal. Specifically, we introduce:

$$\Lambda := \bigcup_{\mathbf{x}^{(1)} \in \mathcal{Z}} \mathcal{E}(\mathbf{x}^{(1)}).$$

To proceed, given some $k \in [q]$, we define $\mathcal{P}_\delta^{(k)}$ the subset of linear mappings from $\mathbb{R}^q$ to $\mathbb{R}^k$ that somehow provide good approximation guarantees for the projected follower's problem. Formally:

$$\mathcal{P}_\delta^{(k)} := \big\{ \mathbf{P} \in \mathbb{R}^{k\times q} \,:\, \varphi_\theta(\mathbf{x}) - \delta \le \varphi_\theta(\mathbf{x}, \mathbf{P}) \le \varphi_\theta(\mathbf{x}) \quad \forall\, \mathbf{x} \in \Lambda \big\}.$$

Step 1 (nonempty set): Next, we provide a condition for both $k \in [q]$ and δ that guarantees that $\mathcal{P}_\delta^{(k)}$ is nonempty by leveraging the results from Section 4, and in particular Theorem 3. For any leader's feasible decision $\mathbf{x} \in \mathcal{X}$, we denote by $\mathbf{u}^*(\mathbf{x})$ the optimal solution to the dual of the follower's problem (2) with the minimal ℓ_2-norm. Recall that, given a random projector $\mathbf{P}$ and $0 < \delta < \min\big\{\varphi_\theta(\mathbf{x}) \,:\, \mathbf{x} \in \mathcal{X}\big\}$, we have that the following inequalities hold:

$$\varphi_\theta(\mathbf{x}) - \delta \le \varphi_\theta(\mathbf{x}, \mathbf{P}) \le \varphi_\theta(\mathbf{x}),$$

with probability at least $p \ge 1 - 4me^{-C\left(\varepsilon^2 - \varepsilon^3\right)k}$, where $\varepsilon(\mathbf{x}) := \mathcal{O}\big(\frac{\delta}{(1+\theta)\|\mathbf{u}^*(\mathbf{x})\|_2 \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2}\big)$. Then, we select $\mathbf{x}^* \in \Lambda$ that minimizes $\varepsilon(\mathbf{x})^2 - \varepsilon(\mathbf{x})^3$ and introduce $u^* := \max\{\|\mathbf{u}^*(\mathbf{x})\|_2 \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2 \,:\, \mathbf{x} \in \mathcal{X}\}$.

In addition, we impose that δ is sufficiently small so that $\varepsilon(\mathbf{x}^*) \equiv \tilde{D} \cdot \frac{\delta}{\theta} < \frac{1}{2}\min\{1, u^*\}$, for some positive constant $\tilde{D} > 0$ that is independent of k, δ and q, so that $\varepsilon(\mathbf{x}^*)^2 - \varepsilon(\mathbf{x}^*)^3$ is guaranteed to be positive. Then, given some random projector $\mathbf{P}$ (e.g., Gaussian), we have that the following sequence of inequalities is satisfied:

$$\begin{aligned}
\mathbb{P}\Big[\varphi(\mathbf{x}) - \delta \le \varphi_\theta(\mathbf{x}, \mathbf{P}) \le \varphi_\theta(\mathbf{x}) \quad \forall\, \mathbf{x} \in \Lambda\Big] &= 1 - \mathbb{P}\Big[\exists\, \mathbf{x} \in \Lambda \,:\, \big\{\varphi_\theta(\mathbf{x}) - \delta \le \varphi_\theta(\mathbf{x}, \mathbf{P}) \le \varphi_\theta(\mathbf{x})\big\}^c\Big] \\
&= 1 - \mathbb{P}\Big[\bigcup_{\mathbf{x}\in\Lambda} \big\{\varphi_\theta(\mathbf{x}) - \delta \le \varphi_\theta(\mathbf{x}, \mathbf{P}) \le \varphi_\theta(\mathbf{x})\big\}^c\Big] \\
&\ge 1 - \sum_{\mathbf{x}\in\Lambda} \Big(1 - \mathbb{P}\Big[\varphi_\theta(\mathbf{x}) - \delta \le \varphi_\theta(\mathbf{x}, \mathbf{P}) \le \varphi_\theta(\mathbf{x})\Big]\Big) \\
&\ge 1 - \sum_{\mathbf{x}\in\Lambda} 4me^{-C\left(\varepsilon(\mathbf{x})^2 - \varepsilon(\mathbf{x})^3\right)k} \\
&\ge 1 - 4m|\Lambda|e^{-C\left(\varepsilon(\mathbf{x}^*)^2 - \varepsilon(\mathbf{x}^*)^3\right)k}.
\end{aligned}$$

In particular, one can verify that, if δ and k are chosen so that they satisfy, together, the conditions $\delta < \frac{\theta}{2\tilde{D}}\min\{1, u^*\}$ and $k\delta^2 > \frac{2\theta^2}{\tilde{D}^2 C}\log(4m|\Lambda|)$, then the following inequality is satisfied:

$$1 - 4m|\Lambda|e^{-C\left(\varepsilon(x^*)^2 - \varepsilon(x^*)^3\right)k} > 0.$$

In that case, there exists at least one realization (and possibly infinitely many) of $\mathbf{P}$ with the desired property (which belongs to $\mathcal{P}_\delta^{(k)}$), as long as k and δ satisfy, together, the following inequalities:

$$\mathcal{O}\left(\theta\sqrt{\frac{\log(m|\Lambda|)}{k}}\right) < \delta < \mathcal{O}\Big(\theta \min\{1, u^*\}\Big).$$

In the worst case, the number of nondegenerate extreme points grows as $O\big((p+q)^{n+m}\big)$, which immediately provides the estimate $\log(|\Lambda|) \leq \mathcal{O}\left((n+m)\log(p+q)\right)$. By choosing $\delta := \frac{\theta}{2\tilde{D}} \min\{1, u^*\}$, one can verify that the condition above forces $k = \Omega\left((n+m)\log(p+q)\right)$. Hence, to secure the desired probability bound, k must grow linearly with the dimension of the decision variables but only logarithmically with the number of constraints.

Step 2 (On the lower bound): By Theorem 5, and notably by following its proof, one can verify that, for all $\mathbf{P} \in \mathcal{P}_\delta^{(k)}$ (provided that the set is nonempty), the following inequality is satisfied:

$$z^* - s_\theta(\mathbf{P}, \delta) \leq \|\mathbf{d}\|_1 m \Delta(\mathbf{F}, \mathbf{c})\delta.$$

Hence, together with the result from Section 3, we have that:

$$z^* - \|\mathbf{d}\|_1 m \Delta(\mathbf{F}, \mathbf{c})\delta \leq s_\theta(\mathbf{P}, \delta) \leq z^*.$$

Step 3 (On the upper bound): We fix a projector $\mathbf{P} \in \mathcal{P}_\delta^{(k)}$, where k and δ are chosen such that the set $\mathcal{P}_\delta^{(k)}$ is nonempty. Then, let $\big(\mathbf{x}^*, \mathbf{y}^*(\mathbf{x}^*)\big)$ be the optimal solution to the original bilevel program (1). Then, by the definition of $\mathbf{P}$, the following inequalities are satisfied:

$$\varphi_\theta(\mathbf{x}^*, \mathbf{P}) \leq \varphi_\theta(\mathbf{x}^*) \leq \varphi_\theta(\mathbf{x}^*, \mathbf{P}) + \delta.$$

Next, for an arbitrary $\tilde{\delta} \in (0, \delta]$, we introduce the set of follower's decisions that are feasible for the projected follower's problem and additionally satisfy some constraints on their objective function value. Formally:

$$\Gamma_{\mathbf{P}}(\tilde{\delta}) := \Big\{\mathbf{y} \in \mathbb{R}^m_+ \,:\, \mathbf{P}\big(\mathbf{L}\mathbf{x}^* + \mathbf{F}\mathbf{y} - \mathbf{f}\big) = \mathbf{0}_k \,,\, \varphi_\theta(\mathbf{x}^*) \geq \mathbf{c}^\top \mathbf{y} \geq \varphi_\theta(\mathbf{x}^*, \mathbf{P}) + \tilde{\delta} \,,\, \mathbf{1}_m^\top \mathbf{y} \leq \theta\Big\}$$

In particular, if we pick $\tilde{\delta} := \varphi_\theta(\mathbf{x}^*) - \varphi_\theta(\mathbf{x}^*, \mathbf{P})$, then we have that $\mathbf{y}^*(\mathbf{x}^*) \in \Gamma_{\mathbf{P}}(\tilde{\delta})$. Hence, by using proximity results for linear programs by Cook et al. (1986) (see Theorem 5 in their paper), there exists some constant $C > 0$ and a solution $\mathbf{y}^{(0)} \in \Gamma_{\mathbf{P}}(0)$ that satisfies the following inequalities:

$$\|\mathbf{y}^{(0)} - \mathbf{y}^*(\mathbf{x}^*)\|_\infty \leq C\tilde{\delta} \leq C\delta.$$

Consequently, we have that the following sequence of inequalities is satisfied:

$$\begin{aligned} z^* = \mathbf{a}^\top \mathbf{x}^* + \mathbf{d}^\top \mathbf{y}^*(\mathbf{x}^*) &= \mathbf{a}^\top \mathbf{x}^* + \mathbf{d}^\top \mathbf{y}^{(0)} + \big(\mathbf{d}^\top \mathbf{y}^*(\mathbf{x}^*) - \mathbf{d}^\top \mathbf{y}^{(0)}\big) \\ &= \mathbf{a}^\top \mathbf{x}^* + \mathbf{d}^\top \mathbf{y}^{(0)} + |\mathbf{d}^\top \mathbf{y}^*(\mathbf{x}^*) - \mathbf{d}^\top \mathbf{y}^{(0)}| \\ &\leq \mathbf{a}^\top \mathbf{x}^* + \mathbf{d}^\top \mathbf{y}^{(0)} + \|\mathbf{d}\|_1 \|\mathbf{y}^*(\mathbf{x}^*) - \mathbf{y}^{(0)}\|_\infty \\ &\leq \mathbf{a}^\top \mathbf{x}^* + \mathbf{d}^\top \mathbf{y}^{(0)} + \|\mathbf{d}\|_1 C\delta \end{aligned}$$

$$\leq s_\theta(\mathbf{P}) + \|\mathbf{d}\|_1 C\delta,$$

where the last inequality follows from the feasibility of $(\mathbf{x}^*, \mathbf{y}^{(0)})$ to (5).

To summarize, if k and δ are chosen such that $\mathcal{P}_\delta^{(k)} \neq \emptyset$, then, there exists some $C_1, C_2 > 0$ such that for all $\mathbf{P} \in \mathcal{P}_\delta^{(k)}$, the following inequalities are satisfied:

$$z^* - C_1\delta \leq s_\theta(\mathbf{P}, \delta) \leq z^* \leq s_\theta(\mathbf{P}) \leq z^* + C_2\delta.$$

Note that we could replace $s_\theta(\mathbf{P})$ by $d_\theta(\mathbf{P}, \delta)$ in the above inequalities to obtain a similar result for the adjusted surrogate problem by applying similar feasibility-based arguments.

The sandwich theorem applies to every projector $\mathbf{P} \in \mathcal{P}_\delta^{(k)}$ as long as $k\delta^2$ exceeds the order of $\mathcal{O}\big(\log(m|\Lambda|)\big)$. Consequently, our main result is driven by the underlying problem structure, most critically, the number of extreme points in the single-level relaxation of (1). Moreover, this perspective offers an alternative description of "good" projectors for the surrogate-based bounds in Section 3 that does not directly rely on the follower's optimality conditions. Instead, one can interpret $S_\delta^{(k)}$ as the set of $\mathcal{O}(\delta)$-isometries from $\mathbb{R}^q$ to $\mathbb{R}^k$, that is, embeddings of the type guaranteed by the Johnson–Lindenstrauss lemma (Johnson and Lindenstrauss 1984).

D Additional computational study

We present additional computational results that complement the findings reported in Section 6.2. Specifically, we consider instances in which the network variant is chosen to be either V2 or V3, as described by Royset and Wood (2007). For both variants, we report the optimality gap and the runtime, using the same computational setting as in Section 6.2.

		Network V2 (optimality gap γ)			
	Projection	2×3	3×3	4×3	5×3
z_{feas}		0.2 (0.2)	0.4 (0.2)	0.4 (0.1)	0.5 (0.1)
z_{relax}		62.2 (24.3)	67.4 (13.7)	94.8 (11.0)	127.6 (23.9)
$\tilde{s}_\theta(\mathbf{P}, \delta_f)$	flow	0.2 (0.2)	0.4 (0.2)	0.5 (0.2)	0.5 (0.2)
	interdiction	0.2 (0.2)	0.4 (0.2)	0.5 (0.1)	0.5 (0.2)
	naive	0.2 (0.2)	0.5 (0.2)	0.5 (0.1)	0.6 (0.1)
$\tilde{d}_\theta(\mathbf{P}, \delta_d)$	flow	1.9 (0.9)	1.0 (0.8)	0.4 (0.5)	0.2 (0.4)
	interdiction	1.4 (0.4)	0.5 (0.4)	0.5 (0.2)	0.3 (0.3)
	naive	13.2 (43.4)	88.4 (91.8)	120.4 (119.8)	163.6 (162.1)

Table 4: **Network V2 (optimality gap).** Mean ($\pm$ std) optimality gaps for various bounding approaches applied to the min-cost flow interdiction problem on Network V2. We compute the feasibility gap γ_f for both z_{feas} and $\tilde{s}_\theta(\mathbf{P}, \delta_p)$. Also, we compute the dual gap γ_d for both z_{relax} and $\tilde{d}_\theta(\mathbf{P}, \delta_d)$. The table compares: ($i$) the single-level relaxation bounds, and (ii) feasibility and dual bounds obtained via random projections.

		Network V2 (runtime)			
	Projection	2×3	3×3	4×3	5×3
z^*		0.2 (0.1)	0.5 (0.1)	3.2 (0.1)	76.7 (2.0)
z_{feas}		0.1 (0.1)	0.2 (0.1)	0.3 (0.1)	0.5 (0.1)
z_{relax}		0.1 (0.1)	0.2 (0.1)	0.3 (0.1)	0.5 (0.1)
$\tilde{s}_\theta(\mathbf{P}, \delta_f)$	flow	0.2 (0.1)	0.8 (0.1)	4.9 (0.5)	129.8 (21.7)
	interdiction	0.3 (0.1)	0.6 (0.1)	0.9 (0.2)	1.4 (0.6)
	naive	0.2 (0.1)	0.6 (0.1)	1.0 (0.3)	2.0 (1.0)
$\tilde{d}_\theta(\mathbf{P}, \delta_d)$	flow	0.2 (0.1)	0.7 (0.1)	5.1 (0.7)	127.4 (22.5)
	interdiction	0.3 (0.1)	0.8 (0.3)	1.5 (0.8)	6.3 (7.6)
	naive	0.2 (0.1)	0.4 (0.2)	0.7 (0.7)	3.5 (5.9)

Table 5: **Network V2 (runtime).** Mean ($\pm$ std) runtime (in sec) for various bounding approaches applied to the min-cost flow interdiction problem on Network V2. The table compares: (i) solving the leader's problem exactly, (ii) the single-level relaxation approach, and (iii) the random projection based approach.

	Projection	Network V3 (optimality gap γ)			
		2×3	3×3	4×3	5×3
z_{feas}		0.2 (0.2)	0.4 (0.2)	0.4 (0.2)	0.5 (0.2)
z_{relax}		62.2 (11.2)	63.3 (16.0)	104.6 (17.6)	127.0 (24.3)
$\tilde{s}_\theta(\mathbf{P}, \delta_f)$	flow	0.2 (0.2)	0.4 (0.2)	0.5 (0.2)	0.5 (0.2)
	interdiction	0.2 (0.2)	0.5 (0.2)	0.5 (0.2)	0.5 (0.1)
	naive	0.3 (0.2)	0.5 (0.2)	0.6 (0.1)	0.6 (0.2)
$\tilde{d}_\theta(\mathbf{P}, \delta_d)$	flow	1.7 (0.9)	0.9 (0.7)	0.5 (0.5)	0.2 (0.4)
	interdiction	1.2 (0.4)	0.5 (0.3)	0.3 (0.3)	0.3 (0.3)
	naive	12.8 (41.2)	77.7 (87.2)	139.1 (133.0)	139.4 (146.3)

Table 6: **Network V3 (optimality gap).** Mean ($\pm$ std) optimality gaps for various bounding approaches applied to the min-cost flow interdiction problem on Network V3. We compute the feasibility gap γ_f for both z_{feas} and $\tilde{s}_\theta(\mathbf{P}, \delta_p)$. Also, we compute the dual gap γ_d for both z_{relax} and $\tilde{d}_\theta(\mathbf{P}, \delta_d)$. The table compares: ($i$) the single-level relaxation bounds, and (ii) feasibility and dual bounds obtained via random projections.

	Projection	Network V3 (runtime)			
		2×3	3×3	4×3	5×3
z^*		0.2 (0.1)	0.5 (0.1)	3.2 (0.2)	78.5 (4.5)
z_{feas}		0.1 (0.1)	0.2 (0.1)	0.3 (0.1)	0.5 (0.1)
z_{relax}		0.1 (0.1)	0.2 (0.1)	0.3 (0.1)	0.5 (0.1)
$\tilde{s}_\theta(\mathbf{P}, \delta_f)$	flow	0.2 (0.1)	0.7 (0.1)	5.0 (0.6)	123.5 (19.6)
	interdiction	0.3 (0.1)	0.6 (0.2)	0.9 (0.2)	1.5 (0.7)
	naive	0.2 (0.1)	0.6 (0.2)	1.0 (0.3)	2.1 (1.2)
$\tilde{d}_\theta(\mathbf{P}, \delta_d)$	flow	0.2 (0.1)	0.7 (0.1)	5.1 (0.8)	130.6 (26.8)
	interdiction	0.3 (0.1)	0.7 (0.3)	1.7 (1.0)	4.1 (4.3)
	naive	0.2 (0.1)	0.4 (0.2)	0.7 (0.6)	3.3 (5.6)

Table 7: **Network V3 (runtime).** Mean ($\pm$ std) runtime (in sec) for various bounding approaches applied to the min-cost flow interdiction problem on Network V3. The table compares: (i) solving the leader's problem exactly, (ii) the single-level relaxation approach, and (iii) the random projection based approach.

References

Achlioptas, D. (2003). "Database-friendly random projections: Johnson-Lindenstrauss with binary coins". In: *Journal of Computer and System Sciences* 66.4, pp. 671–687.

Bard, J. F. and Falk, J. E. (1982). "An explicit solution to the multi-level programming problem". In: *Computers & Operations Research* 9.1, pp. 77–100.

Bertsimas, D. and Tsitsiklis, J. N. (1997). *Introduction to linear optimization.* Athena Scientific Belmont, MA.

Bialas, W. and Karwan, M. (1982). "On two-level optimization". In: *IEEE Transactions on Automatic Control* 27.1, pp. 211–214.

Blair, C. E. and Jeroslow, R. G. (1977). "The value function of a mixed integer program: I". In: *Discrete Mathematics* 19.2, pp. 121–138.

Boucheron, S., Lugosi, G., and Massart, P. (2013). *Concentration Inequalities: A Nonasymptotic Theory of Independence.* Oxford University Press.

Cook, W., Gerards, A. M., Schrijver, A., and Tardos, É. (1986). "Sensitivity theorems in integer linear programming". In: *Mathematical Programming* 34, pp. 251–264.

Hall, P. (1927). "The distribution of means for samples of size n drawn from a population in which the variate takes values between 0 and 1, all such values being equally probable". In: *Biometrika* 19, pp. 240–244.

Irwin, J. O. (1927). "On the frequency distribution of the means of samples from a population having any law of frequency with finite moments, with special reference to Pearson's Type II". In: *Biometrika* 19, pp. 225–239.

Jacod, J. and Protter, P. (2012). *Probability Essentials.* Universitext. Springer Berlin Heidelberg.

Johnson, W. B. and Lindenstrauss, J. (1984). "Extensions of Lipschitz mappings into a Hilbert space". In: *Contemporary Mathematics* 26, p. 1.

Mangasarian, O. L. and Shiau, T.-H. (1987). "Lipschitz continuity of solutions of linear inequalities, programs and complementarity problems". In: *SIAM Journal on Control and Optimization* 25.3, pp. 583–595.

Royset, J. O. and Wood, R. K. (2007). "Solving the bi-objective maximum-flow network-interdiction problem". In: *INFORMS Journal on Computing* 19.2, pp. 175–184.

Tropp, J. A. (2023). *Probability in High Dimensions.*

Electronic Companion

For "Dimensionality Reduction in Bilevel Linear Programming"

In this electronic companion, we provide the formal proof of Theorem 3 from Section 4 for completeness. The proof and the result are originally established in the excellent study by Vu et al. (2018) to whom full credit is due. Their approach relies on the assumption that the right-hand side of the linear system is normalized to have unit norm, which is an innocuous assumption in their setting. However, in our context, this normalization conceals a dependence on the leader's decision, which would appear in the resulting bound and thus render the bound (or the probabilistic guarantee) decision-dependent rather than truly constant. In what follows, we adapt their argument to the *projected* constraints of the follower's problem, whose feasible set depends on the leader's decision.

Consequently, we present the proof in full, so that the reader can identify exactly where the leader's decision enters the analysis. We begin by recalling the auxiliary concentration and geometry results collected by Vu et al. (2018); these results are summarized in Appendix E.C.1. Next, Appendix E.C.2 contains the proofs of several intermediate lemmas for linear programming. Finally, we combine these ingredients to establish the main result in Appendix E.C.3.

E.C.1 Preliminary results

We open the discussion with several preliminaries from random-projection theory that serve to establish approximation bounds for the projected follower's problem. The arguments rely on the Johnson–Lindenstrauss lemma (Johnson and Lindenstrauss 1984).

Theorem 1 (Johnson and Lindenstrauss, 1984)**.** *Given $\epsilon \in (0,1)$ and a matrix $\mathbf{F} \in \mathbb{R}^{q\times m}$, there exists a matrix $\mathbf{P} \in \mathbb{R}^{k\times q}$ such that the following inequalities are satisfied:*

$$\forall\, 1 \leq i < j \leq m \qquad \big(1-\varepsilon\big)\|\mathbf{F}_i - \mathbf{F}_j\|_2 \leq \|\mathbf{P}\mathbf{F}_i - \mathbf{P}\mathbf{F}_j\|_2 \leq \big(1+\varepsilon\big)\|\mathbf{F}_i - \mathbf{F}_j\|_2,$$

where $\mathbf{F}_i$ denotes the i-th column of $\mathbf{F}$, and where $k \equiv \mathcal{O}\left(\varepsilon^{-2}\log(m)\right)$.

The Johnson–Lindenstrauss lemma guarantees that any finite set of vectors admits a linear projection into a lower-dimensional space that almost perfectly preserves pairwise Euclidean distances. We now cite several auxiliary results from Vu et al. (2018), referring the reader to their paper for complete proofs.

Lemma 5 (Vu et al., 2018). *Fix $\epsilon \in (0,1)$ and $\mathbf{u} \in \mathbb{R}^q$ arbitrarily, and let $\mathbf{P}$ be a $k \times q$ random projector. Then, we have that the following inequality is satisfied:*

$$\mathbb{P}\big[(1-\epsilon)\|\mathbf{u}\|_2 \leq \|\mathbf{P}\mathbf{u}\|_2 \leq (1+\epsilon)\|\mathbf{u}\|_2\big] \geq 1 - 2e^{-C\epsilon^2 k},$$

for some constant $C > 0$ (independent of q, k, ϵ) .

Lemma 6 (Vu et al., 2018). *Fix $\epsilon \in (0,1)$ and $\mathbf{u} \in \mathbb{R}^q$ arbitrarily, and let $\mathbf{P}$ be a $k \times q$ random projector. Then, we have that the following inequality is satisfied:*

$$\mathbb{P}\big[(1-\epsilon)\|\mathbf{u}\|_2^2 \leq \|\mathbf{P}\mathbf{u}\|_2^2 \leq (1+\epsilon)\|\mathbf{u}\|_2^2\big] \geq 1 - 2e^{-C(\epsilon^2-\epsilon^3)k},$$

for some constant $C > 0$ (independent of q, k, ϵ) .

Proposition 3 (Vu et al., 2018). *Fix $\epsilon \in (0,1)$ and $\mathbf{u}, \mathbf{v} \in \mathbb{R}^q$ arbitrarily, and let $\mathbf{P}$ be a $k \times q$ random projector. Then, we have that the following inequalities are satisfied:*

$$-\epsilon\|\mathbf{u}\|_2\|\mathbf{v}\|_2 \leq \ \langle \mathbf{P}\mathbf{u}, \mathbf{P}\mathbf{v}\rangle - \langle \mathbf{u}, \mathbf{v}\rangle \ \leq \epsilon\|\mathbf{u}\|_2\|\mathbf{v}\|_2$$

with probability at least $1 - 2e^{-C(\epsilon^2-\varepsilon^3)k}$, where $C > 0$ is a constant independent of q, k, and ϵ.

Corollary 5 (Vu et al., 2018). *Fix $\epsilon \in (0,1)$ and $\mathbf{u} \in \mathbb{R}^q \setminus \{\mathbf{0}_q\}$ arbitrarily, and let $\mathbf{P}$ be a $k \times q$ random projector. Then, we have that the following inequality is satisfied:*

$$\mathbb{P}\big[\mathbf{P}\mathbf{u} \neq \mathbf{0}_k\big] \geq 1 - 2e^{-Ck},$$

for some constant $C > 0$ (independent of q and k).

E.C.2 Intermediate results

In this section, we establish two lemmas that provide the first application of random projection techniques in our context. Each concerns the feasibility of the projected follower's problem and builds on the probabilistic results of the previous section. Specifically:

Lemma 7. *Let $\mathbf{P}$ is a $k \times q$ be random projector. Then, for any $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}_+^n \times \mathbb{R}_+^m$, we have that:*

(i) if $\mathbf{x} \in \mathcal{X}$ and $\mathbf{F}\mathbf{y} = \mathbf{f} - \mathbf{L}\mathbf{x}$, then $\mathbf{P}\mathbf{F}\mathbf{y} = \mathbf{P}(\mathbf{f} - \mathbf{L}\mathbf{x})$,

(ii) if $\mathbf{x} \in \mathcal{X}$ and $\mathbf{F}\mathbf{y} + \mathbf{L}\mathbf{x} - \mathbf{f} \neq \mathbf{0}_q$, then $\mathbb{P}\left[\mathbf{P}(\mathbf{F}\mathbf{y} + \mathbf{L}\mathbf{x} - \mathbf{f}) \neq \mathbf{0}_k\right] \geq 1 - 2e^{-Ck}$,

(iii) if $\mathbf{x} \in \mathcal{X}$ and $\mathbf{F}\mathbf{y} + \mathbf{L}\mathbf{x} - \mathbf{f} \neq \mathbf{0}_q$, $\forall \mathbf{y} \in \mathcal{Y} \subseteq \mathbb{R}_+^m$, for some finite set $|\mathcal{Y}| < +\infty$, then $\mathbb{P}\left[\mathbf{P}(\mathbf{F}\mathbf{y} + \mathbf{L}\mathbf{x} - \mathbf{f}) \neq \mathbf{0}_k \quad \forall \mathbf{y} \in \mathcal{Y}\right] \geq 1 - |\mathcal{Y}| 2e^{-Ck}$,

where the constant $C > 0$ is independent of $\mathbf{x}, m$ and k.

Proof of Lemma 7. *(i)* The result follows from the linearity of the equality operator.

(ii) Define $\mathbf{u} = \mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{F}\mathbf{y}$. Then, by our assumption above, we have that $\mathbf{u} \neq \mathbf{0}_q$. Therefore, we obtain the following inequality:

$$\mathbb{P}\left[\mathbf{P}\mathbf{u} \neq \mathbf{0}_k\right] \geq 1 - 2e^{-Ck},$$

which is obtained by using Corollary 5. The result then follows by the definition of $\mathbf{u}$.

(iii) We derive the following sequence of inequalities:

$$\begin{aligned}
\mathbb{P}\big[\mathbf{P}\big(\mathbf{F}\mathbf{y} + \mathbf{L}\mathbf{x} - \mathbf{f}\big) \neq \mathbf{0}_k \quad \forall\, \mathbf{y} \in \mathcal{Y}\big] &= \mathbb{P}\big[\bigcap_{\mathbf{y}\in\mathcal{Y}} \{\mathbf{P}\big(\mathbf{F}\mathbf{y} + \mathbf{L}\mathbf{x} - \mathbf{f}\big) \neq \mathbf{0}_k\}\big] \\
&= 1 - \mathbb{P}\big[\big\{\bigcap_{\mathbf{y}\in\mathcal{Y}} \{\mathbf{P}\big(\mathbf{F}\mathbf{y} + \mathbf{L}\mathbf{x} - \mathbf{f}\big) \neq \mathbf{0}_k\}\big\}^c\,\big] \\
&= 1 - \mathbb{P}\big[\bigcup_{\mathbf{y}\in\mathcal{Y}} \{\mathbf{P}\big(\mathbf{F}\mathbf{y} + \mathbf{L}\mathbf{x} - \mathbf{f}\big) \neq \mathbf{0}_k\}^c\big] \\
&\geq 1 - \sum_{\mathbf{y}\in\mathcal{Y}} \mathbb{P}\big[\{\mathbf{P}\big(\mathbf{F}\mathbf{y} + \mathbf{L}\mathbf{x} - \mathbf{f}\big) \neq \mathbf{0}_k\}^c\big] \\
&\geq 1 - \sum_{\mathbf{y}\in\mathcal{Y}} \mathbb{P}\big[\{\mathbf{P}\big(\mathbf{F}\mathbf{y} + \mathbf{L}\mathbf{x} - \mathbf{f}\big) = \mathbf{0}_k\}\big] \\
&\overset{(a)}{\geq} 1 - \sum_{\mathbf{y}\in\mathcal{Y}} 2e^{-Ck} = 1 - |\mathcal{Y}| 2e^{-Ck},
\end{aligned}$$

where (a) follows from the result in *(ii)*. ■

Lemma 8. *Let us fix a leader's feasible decision $\mathbf{x} \in \mathcal{X}$. Given $\mathbf{F}_1, \ldots, \mathbf{F}_m \in \mathbb{R}^q$, we assume that $\mathbf{f} - \mathbf{L}\mathbf{x} \notin \mathcal{C} \equiv \mathrm{cone}\{\mathbf{F}_1, \ldots, \mathbf{F}_m\}$, i.e., the cone induced by the vectors $\mathbf{F}_1, \ldots, \mathbf{F}_m$. Next, we introduce the following two constants:*

$$c_1 \equiv c_1(\mathbf{x}) := \min_{\mathbf{u}\in C} \|\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{u}\|_2 \quad \textit{and} \quad c_2 \equiv c_2(\mathbf{x}) := \max_{j\in[m]} \max\{\|2\mathbf{f} - 2\mathbf{L}\mathbf{x} - \mathbf{F}_j\|_2, \|\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{F}_j\|_2\}.$$

Next, let $\mathbf{P} : \mathbb{R}^q \to \mathbb{R}^k$ be a random projector and define $\mathbf{P}\mathcal{C} := conv\big(\{\mathbf{P}\mathbf{F}_1, \ldots, \mathbf{P}\mathbf{F}_m\}\big)$, i.e., the convex hull of $\mathbf{P}\mathbf{F}_1, \ldots, \mathbf{P}\mathbf{F}_m$. The following inequality is satisfied:

$$\mathbb{P}\big[\mathbf{P}\big(\mathbf{f} - \mathbf{L}\mathbf{x}\big) \notin \mathbf{P}\mathcal{C}\big] \geq 1 - 2(m+1)^2 e^{-C\left(\varepsilon^2 - \varepsilon^3\right)k},$$

for some constant C independent of $m, n, k, c_1, c_2, \mathbf{x}$ and for ε that satisfies:

$$\varepsilon < \min\left\{c_1^2(\mathbf{x})/c_2^2(\mathbf{x}) \,:\, \mathbf{x} \in \mathcal{X},\, \mathbf{f} - \mathbf{L}\mathbf{x} \notin \mathcal{C}\right\}.$$

Proof of Lemma 8. To begin, we assume that the random projector $\mathbf{P}$ is defined over a probability space $(\Omega, \mathcal{B}, \mathbb{P})$. Then, we introduce the following two events (which are both subsets of Ω):

$$S_\varepsilon^- := \big\{\forall\, \mathbf{u}, \mathbf{v} \in \{\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{F}_j\,,\, j \in [m]\} \cup \{\mathbf{0}_q\} \,:\, (1-\varepsilon)\|\mathbf{u}-\mathbf{v}\|_2^2 \le \|\mathbf{P}(\mathbf{u}-\mathbf{v})\|_2^2 \le (1+\varepsilon)\|\mathbf{u}-\mathbf{v}\|_2^2\big\},$$

$$S_\varepsilon^+ := \big\{\forall\, \mathbf{u}, \mathbf{v} \in \{\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{F}_j\,,\, j \in [m]\} \cup \{\mathbf{0}_q\} \,:\, (1-\varepsilon)\|\mathbf{u}+\mathbf{v}\|_2^2 \le \|\mathbf{P}(\mathbf{u}+\mathbf{v})\|_2^2 \le (1+\varepsilon)\|\mathbf{u}+\mathbf{v}\|_2^2\big\},$$

and we define $S_\varepsilon := S_\varepsilon^- \cap S_\varepsilon^+$.

We assume that the event S_ε occurs. Moreover, we fix $\boldsymbol{\lambda} := (\lambda_j)_{j=1}^m \in \mathbb{R}_+^m$ arbitrarily with the constraint that $\sum_{j=1}^m \lambda_j = 1$. Then, the following sequence of equalities is guaranteed to hold:

$$\begin{aligned}
\|\mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}\big) - \sum_{j=1}^m \lambda_j \mathbf{P}\mathbf{F}_j\|_2^2 &= \|\mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x} - \sum_{j=1}^m \lambda_j \mathbf{F}_j\big)\|_2^2 \\
&\overset{(a)}{=} \|\mathbf{P}\sum_{j=1}^m \big(\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\big)\lambda_j\|_2^2 \\
&= \langle \mathbf{P}\sum_{j=1}^m \big(\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\big)\lambda_j, \mathbf{P}\sum_{j=1}^m \big(\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\big)\lambda_j\rangle \\
&= \sum_{j=1}^m\sum_{i=1}^m \lambda_j\lambda_i \langle \mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\big), \mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_i\big)\rangle \\
&= \sum_{j=1}^m \lambda_j^2 \langle \mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\big), \mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\big)\rangle \\
&\quad + \sum_{j=1}^m \sum_{\substack{i=1\\ i\neq j}}^m \lambda_j\lambda_i \langle \mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\big), \mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_i\big)\rangle \\
&= \sum_{j=1}^m \lambda_j^2 \|\mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\big)\|_2^2 \\
&\quad + 2\sum_{1\le i<j\le m} \lambda_j\lambda_i \langle \mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\big), \mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_i\big)\rangle \\
&\overset{(b)}{=} \sum_{j=1}^m \lambda_j^2 \|\mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\big)\|_2^2 \\
&\quad + \frac{1}{2}\sum_{1\le i<j\le m} \lambda_j\lambda_i \big(\|\mathbf{P}\big(2\mathbf{f}-2\mathbf{L}\mathbf{x}-\mathbf{F}_j-\mathbf{F}_i\big)\|_2^2 - \|\mathbf{P}\big(\mathbf{F}_i-\mathbf{F}_j\big)\|_2^2\big),
\end{aligned}$$

where (a) follows from the definition of $\boldsymbol{\lambda}$, i.e., $\sum_{j=1}^m \lambda_j = 1$. In addition, (b) follows from the following equality: $\langle \mathbf{u}, \mathbf{v}\rangle = \frac{1}{4}\big(\|\mathbf{u}+\mathbf{v}\|_2^2 - \|\mathbf{u}-\mathbf{v}\|_2^2\big)$.

Then, recall that we initially assume that the event S_ε occurs. Consequently, the following two inequalities are guaranteed to hold on event S_ε:

$$\|\mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\big)\|_2^2 \ge \big(1-\varepsilon\big)\|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2, \quad \text{and}$$

$$\|\mathbf{P}\big(2\mathbf{f}-2\mathbf{L}\mathbf{x}-\mathbf{F}_j-\mathbf{F}_i\big)\|_2^2-\|\mathbf{P}\big(\mathbf{F}_i-\mathbf{F}_j\big)\|_2^2\geq\big(1-\varepsilon\big)\|2\mathbf{f}-2\mathbf{L}\mathbf{x}-\mathbf{F}_j-\mathbf{F}_i\|_2^2-\big(1+\varepsilon\big)\|\mathbf{F}_i-\mathbf{F}_j\|_2^2$$

for all $1\leq i<j\leq m$.

Accordingly, we obtain the following inequality:

$$\begin{aligned}\|\mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}\big)-\sum_{j=1}^m\lambda_j\mathbf{P}\mathbf{F}_j\|_2^2\geq\big(1-\varepsilon\big)&\sum_{j=1}^m\lambda_j^2\|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2\\&+\frac{1}{2}\sum_{1\leq i<j\leq m}\lambda_i\lambda_j\big(\big(1-\varepsilon\big)\|2\mathbf{f}-2\mathbf{L}\mathbf{x}-\mathbf{F}_j-\mathbf{F}_i\|_2^2-\big(1+\varepsilon\big)\|\mathbf{F}_i-\mathbf{F}_j\|_2^2\big)\\&=\mathcal{A}^--\varepsilon\mathcal{A}^+,\end{aligned}$$

where $\mathcal{A}^-$ and $\mathcal{A}^+$ are defined as follows:

$$\begin{aligned}\mathcal{A}^-&:=\sum_{j=1}^m\lambda_j^2\|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2+\frac{1}{2}\sum_{1\leq i<j\leq m}\big(\lambda_i\lambda_j\|2\mathbf{f}-2\mathbf{L}\mathbf{x}-\mathbf{F}_j-\mathbf{F}_i\|_2^2-\lambda_i\lambda_j\|\mathbf{F}_i-\mathbf{F}_j\|_2^2\big),\\\mathcal{A}^+&:=\sum_{j=1}^m\lambda_j^2\|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2+\frac{1}{2}\sum_{1\leq i<j\leq m}\big(\lambda_i\lambda_j\|2\mathbf{f}-2\mathbf{L}\mathbf{x}-\mathbf{F}_j-\mathbf{F}_i\|_2^2+\lambda_i\lambda_j\|\mathbf{F}_i-\mathbf{F}_j\|_2^2\big).\end{aligned}$$

Next, we simplify the expressions for both $\mathcal{A}^-$ and $\mathcal{A}^+$ into two consecutive steps. Specifically, we obtain the following sequence of equalities on $\mathcal{A}^-$:

$$\begin{aligned}\mathcal{A}^-&=\sum_{j=1}^m\lambda_j^2\|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2+\frac{1}{2}\sum_{1\leq i<j\leq m}\lambda_i\lambda_j\left[\|2\mathbf{f}-2\mathbf{L}\mathbf{x}-\mathbf{F}_j-\mathbf{F}_i\|_2^2-\|\mathbf{F}_i-\mathbf{F}_j\|_2^2\right]\\&\overset{(a)}{=}\sum_{j=1}^m\lambda_j^2\langle\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j,\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\rangle+\frac{1}{2}\sum_{1\leq i<j\leq m}\lambda_i\lambda_j\left[4\langle\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j,\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_i\rangle\right]\\&=\sum_{j=1}^m\sum_{i=1}^m\lambda_i\lambda_j\langle\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_i,\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\rangle\\&\overset{(b)}{=}\langle\mathbf{f}-\mathbf{L}\mathbf{x}-\sum_{i=1}^m\lambda_i\mathbf{F}_i,\mathbf{f}-\mathbf{L}\mathbf{x}-\sum_{j=1}^m\lambda_j\mathbf{F}_j\rangle\\&=\|\mathbf{f}-\mathbf{L}\mathbf{x}-\sum_{i=1}^m\lambda_i\mathbf{F}_i\|_2^2\overset{(c)}{\geq}c_1(\mathbf{x})^2,\end{aligned}$$

where (a) follows from that $4\langle\mathbf{u},\mathbf{v}\rangle=\|\mathbf{u}+\mathbf{v}\|_2^2-\|\mathbf{u}-\mathbf{v}\|_2^2$, for all $\mathbf{u},\mathbf{v}\in\mathbb{R}^m$. Moreover, (b) follows from the definition of $\boldsymbol{\lambda}$, i.e., $\sum_{j=1}^m\lambda_j=1$. Also, (c) follows from $\sum_{i=1}^m\lambda_i\mathbf{F}_i\in\mathcal{C}$.

Next, we derive the following sequence of equalities on $\mathcal{A}^+$:

$$\begin{aligned}\mathcal{A}^+&=\sum_{j=1}^m\lambda_j^2\|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2+\frac{1}{2}\sum_{1\leq i<j\leq m}\big(\lambda_i\lambda_j\|2\mathbf{f}-2\mathbf{L}\mathbf{x}-\mathbf{F}_j-\mathbf{F}_i\|_2^2+\lambda_i\lambda_j\|\mathbf{F}_i-\mathbf{F}_j\|_2^2\big)\\&=\sum_{j=1}^m\lambda_j^2\|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2+\frac{1}{2}\sum_{1\leq i<j\leq m}\lambda_i\lambda_j\big(\langle2\mathbf{f}-2\mathbf{L}\mathbf{x}-\mathbf{F}_i-\mathbf{F}_j,2\mathbf{f}-2\mathbf{L}\mathbf{x}-\mathbf{F}_i-\mathbf{F}_j\rangle\big)\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{2}\sum_{1\le i<j\le m}\lambda_i\lambda_j\big(\langle\mathbf{F}_i-\mathbf{F}_j,\mathbf{F}_i-\mathbf{F}_j\rangle\big)\\
&=\sum_{j=1}^{m}\lambda_j^2\|\mathbf{f}-\mathbf{Lx}-\mathbf{F}_j\|_2^2+\frac{1}{2}\sum_{1\le i<j\le m}\lambda_i\lambda_j\big(\langle 2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_i,2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_i\rangle\big)\\
&+\frac{1}{2}\sum_{1\le i<j\le m}\lambda_i\lambda_j\big(\langle 2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_i,2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_j\rangle+\langle 2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_j,2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_i\rangle\big)\\
&+\frac{1}{2}\sum_{1\le i<j\le m}\lambda_i\lambda_j\big(\langle 2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_j,2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_j\rangle+\langle\mathbf{F}_i-\mathbf{F}_j,\mathbf{F}_i-\mathbf{F}_j\rangle\big)\\
&=\sum_{j=1}^{m}\lambda_j^2\|\mathbf{f}-\mathbf{Lx}-\mathbf{F}_j\|_2^2+\frac{1}{2}\sum_{1\le i<j\le m}\lambda_i\lambda_j\big(\|2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_i\|_2^2+\|2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_j\|_2^2\big)\\
&+\frac{1}{2}\sum_{1\le i<j\le m}\lambda_i\lambda_j\big(2\langle 2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_i,2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_j\rangle+\langle\mathbf{F}_i-\mathbf{F}_j,\mathbf{F}_i-\mathbf{F}_j\rangle\big)\\
&=\sum_{j=1}^{m}\lambda_j^2\|\mathbf{f}-\mathbf{Lx}-\mathbf{F}_j\|_2^2+\frac{1}{2}\sum_{1\le i<j\le m}\lambda_i\lambda_j\big(\|2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_i\|_2^2+\|2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_j\|_2^2\big)\\
&+\frac{1}{2}\sum_{1\le i<j\le m}\lambda_i\lambda_j\big(2\big(\frac{1}{4}\|4\mathbf{f}-4\mathbf{Lx}-\mathbf{F}_i-\mathbf{F}_j\|_2^2-\frac{1}{4}\|\mathbf{F}_i-\mathbf{F}_j\|_2^2\big)+\|\mathbf{F}_i-\mathbf{F}_j\|_2^2\big)\\
&=\sum_{j=1}^{m}\lambda_j^2\|\mathbf{f}-\mathbf{Lx}-\mathbf{F}_j\|_2^2+\frac{1}{2}\sum_{1\le i<j\le m}\lambda_i\lambda_j\big(\|2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_i\|_2^2+\|2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_j\|_2^2\big)\\
&+\frac{1}{2}\sum_{1\le i<j\le m}\lambda_i\lambda_j\big(\frac{1}{2}\|4\mathbf{f}-4\mathbf{Lx}-\mathbf{F}_i-\mathbf{F}_j\|_2^2+\frac{1}{2}\|\mathbf{F}_i-\mathbf{F}_j\|_2^2\big).
\end{aligned}
$$

Moreover, observe that the following equality holds:

$$
\begin{aligned}
\|4\mathbf{f}-4\mathbf{Lx}-\mathbf{F}_i-\mathbf{F}_j\|_2^2+\|\mathbf{F}_i-\mathbf{F}_j\|_2^2=\|2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_i\|_2^2+\|2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_j\|_2^2\\
+\frac{1}{2}\|4\mathbf{f}-4\mathbf{Lx}-\mathbf{F}_i-\mathbf{F}_j\|_2^2+\frac{1}{2}\|\mathbf{F}_i-\mathbf{F}_j\|_2^2,
\end{aligned}
$$

and therefore, the following equality is also guaranteed to hold:

$$
\|4\mathbf{f}-4\mathbf{Lx}-\mathbf{F}_i-\mathbf{F}_j\|_2^2+\|\mathbf{F}_i-\mathbf{F}_j\|_2^2=2\|2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_i\|_2^2+2\|2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_j\|_2^2,
$$

for all $1\le i<j\le m$.

Accordingly, we derive the following inequality on $\mathcal{A}^+$ as follows:

$$
\begin{aligned}
\mathcal{A}^+&=\sum_{j=1}^{m}\lambda_j^2\|\mathbf{f}-\mathbf{Lx}-\mathbf{F}_j\|_2^2+\sum_{1\le i<j\le m}\lambda_i\lambda_j\big(\|2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_i\|_2^2+\|2\mathbf{f}-2\mathbf{Lx}-\mathbf{F}_j\|_2^2\big)\\
&\le\sum_{j=1}^{m}\lambda_j^2c_2(\mathbf{x})^2+2\sum_{1\le i<j\le m}\lambda_i\lambda_jc_2(\mathbf{x})^2
\end{aligned}
$$

$$= \sum_{j=1}^{m} \lambda_j^2 c_2(\mathbf{x})^2 + \sum_{\substack{i=1 \\ i \neq j}}^{m} \lambda_i \lambda_j c_2(\mathbf{x})^2 = c_2(\mathbf{x})^2.$$

Hence, we obtain the following inequality:

$$\|\mathbf{P}\big(\mathbf{f} - \mathbf{L}\mathbf{x}\big) - \sum_{j=1}^{m} \lambda_j \mathbf{P}\mathbf{F}_j\|_2^2 \geq c_1(\mathbf{x})^2 - \varepsilon c_2(\mathbf{x})^2 > 0, \tag{E.C.11}$$

where the strict inequality comes from the restriction made on ε. Moreover, (E.C.11) holds for all choices of $\boldsymbol{\lambda} \in \mathbb{R}_+^m$ that satisfies $\sum_{j=1}^m \lambda_j = 1$. Therefore, if S_ε is true, then we have that:

$$\mathbf{f} - \mathbf{L}\mathbf{x} \notin conv\big(\{\mathbf{P}\mathbf{F}_1, \ldots, \mathbf{P}\mathbf{F}_m\}\big).$$

To conclude, we derive the following sequence of inequalities:

$$\begin{aligned}
&\mathbb{P}\left[\mathbf{P}\big(\mathbf{f} - \mathbf{L}\mathbf{x}\big) \notin \mathbf{P}\mathcal{C}\right] \geq \mathbb{P}\left[S_\varepsilon\right] \\
&\geq \mathbb{P}\left[\big(1-\varepsilon\big)\|\boldsymbol{\omega}\|_2^2 \leq \|\mathbf{P}\boldsymbol{\omega}\|_2^2 \leq \big(1+\varepsilon\big)\|\boldsymbol{\omega}\|_2^2 \,:\, \boldsymbol{\omega} \in \{\mathbf{u} \pm \mathbf{v}\}\ \forall\, \mathbf{u}, \mathbf{v} \in \{\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{F}_i\,,\, \forall\, i \in [m]\} \bigcup \{0\}\right] \\
&= 1 - \mathbb{P}\left[\bigcup_{\substack{\boldsymbol{\omega} \in \{\mathbf{u} \pm \mathbf{v}\} \\ \mathbf{u}, \mathbf{v} \in \{\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{F}_i\,,\, \forall i \in [m]\} \bigcup \{0\}}} \left\{\big(1-\varepsilon\big)\|\boldsymbol{\omega}\|_2^2 \leq \|\mathbf{P}\boldsymbol{\omega}\|_2^2 \leq \big(1+\varepsilon\big)\|\boldsymbol{\omega}\|_2^2\right\}^c\right] \\
&\geq 1 - \sum_{\substack{\boldsymbol{\omega} \in \{\mathbf{u} \pm \mathbf{v}\} \\ \mathbf{u}, \mathbf{v} \in \{\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{F}_i\,,\, \forall i \in [m]\} \bigcup \{0\}}} \mathbb{P}\left[\left\{\big(1-\varepsilon\big)\|\boldsymbol{\omega}\|_2^2 \leq \|\mathbf{P}\boldsymbol{\omega}\|_2^2 \leq \big(1+\varepsilon\big)\|\boldsymbol{\omega}\|_2^2\right\}^c\right] \\
&\overset{(a)}{\geq} 1 - \sum_{\substack{\boldsymbol{\omega} \in \{\mathbf{u} \pm \mathbf{v}\} \\ \mathbf{u}, \mathbf{v} \in \{\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{F}_i\,,\, \forall i \in [m]\} \bigcup \{0\}}} 2e^{-C\left(\varepsilon^2 - \varepsilon^3\right)k} \geq 1 - 2(m+1)^2 e^{-C\left(\varepsilon^2 - \varepsilon^3\right)k},
\end{aligned}$$

where (a) follows from Lemma 6 and concludes the proof of Lemma 8. ■

Feasibility cone. After projecting the follower's feasible region to a region with fewer constraints, the right-hand-side vector may fall outside the cone generated by the projected columns. We formalize this phenomenon and bound the probability of such infeasibility under random projections.

We fix a leader's feasible decision $\mathbf{x} \in \mathcal{X}$. Then, for $\mathbf{y} \in \text{cone}(\{\mathbf{F}_1, \ldots, \mathbf{F}_m\})$, we define $\|\mathbf{y}\|_F := \min\left\{\sum_{j=1}^m \lambda_j \,:\, \boldsymbol{\lambda} \geq \mathbf{0}_m,\, \mathbf{y} = \sum_{j=1}^m \lambda_j \mathbf{F}_j\right\}$. Also, we say that $\boldsymbol{\lambda} \in \mathbb{R}_+^m$ is a *minimal F-representation* of $\mathbf{y}$ if and only if $\sum_{j=1}^m \lambda_j = \|\mathbf{y}\|_F$. This term is originally used by Vu et al. (2018).

Then, we introduce the following two parameters:

$$\bar{F} := \min\big\{|\langle \mathbf{F}_i, \mathbf{F}_j \rangle| \,:\, i, j \in [m]\big\}, \quad \text{and}$$

$$\mu_F := \max\Big\{\|\mathbf{y}\|_F \,:\, \mathbf{y} \in \text{cone}(\{\mathbf{F}_1, \ldots, \mathbf{F}_m\}),\, \mathbf{y} = \sum_{j=1}^{m} \lambda_j \mathbf{F}_j,\, \|\mathbf{y}\|_2 \leq 1\Big\}.$$

Lemma 9. *Let us fix* $\mathbf{y} \in \operatorname{cone}(\{\mathbf{F}_1, \dots, \mathbf{F}_m\})$ *arbitrarily. Next, let* $\boldsymbol{\lambda} \in \mathbb{R}^m$ *be the corresponding minimal representation of* $\mathbf{y}$*. We have that that the following inequalities are satisfied:*

$$\bar{F}^{-1/2}\|\mathbf{y}\|_2 \leq \|\mathbf{y}\|_F \leq \mu_F\|\mathbf{y}\|_2.$$

Proof of Lemma 9. ***First inequality.*** The following sequence of (in)-equalities is satisfied:

$$\begin{aligned}\left(\bar{F}^{-1/2}\|\mathbf{y}\|_2\right)^2 &= \bar{F}^{-1}\|\mathbf{y}\|_2^2 \\ &= \bar{F}^{-1}\sum_{i=1}^{m}\sum_{j=1}^{m}\lambda_i\lambda_j\langle\mathbf{F}_i, \mathbf{F}_j\rangle \\ &\leq \bar{F}^{-1}\bar{F}\sum_{i=1}^{m}\sum_{j=1}^{m}\lambda_i\lambda_j = \|\mathbf{y}\|_F^2,\end{aligned}$$

which follows by the definition of minimal representation.

Second inequality. If $\|\mathbf{y}\|_2 = 0$, then the second inequality is trivially satisfied. Accordingly, we next assume that $\mathbf{y} \neq \mathbf{0}_m$; hence, $\|\mathbf{y}\|_F \neq 0$. Consequently, the following inequality holds:

$$\|\frac{\mathbf{y}}{\|\mathbf{y}\|_2}\|_F \leq \mu_F$$

since $\mathbf{y}/\|\mathbf{y}\|_2 \in \operatorname{cone}(\{\mathbf{F}_1, \dots, \mathbf{F}_m\})$ and $\|\mathbf{y}/\|\mathbf{y}\|_2\|_2 \leq 1$. Thus, the following inequality is satisfied:

$$\|\mathbf{y}\|_F = \frac{\|\mathbf{y}\|_F}{\|\mathbf{y}\|_2}\|\mathbf{y}\|_2 = \|\frac{\mathbf{y}}{\|\mathbf{y}\|_2}\|_F\|\mathbf{y}\|_2 \leq \mu_F\|\mathbf{y}\|_2,$$

which, in turn, concludes the proof. ■

Theorem 2. *Let* $\mathbf{x} \in \mathcal{X}$*, and assume that* $\mathbf{f} - \mathbf{L}\mathbf{x} \in \operatorname{cone}(\{\mathbf{F}_1, \dots, \mathbf{F}_m\})$*. We introduce some* $k \times q$ *random projector* $\mathbf{P}$*, for* $k \in [q]$*. Moreover, we assume that* $\inf\{\|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2 : \mathbf{x} \in \mathcal{X}\} > 0$*. Next, we define* $g \equiv g(\mathbf{x})$ *and* Δ *as follows:*

$$\begin{aligned} g(\mathbf{x}) &\in \operatorname{argmin}\left\{\|\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{u}\|_2^2 \,:\, \mathbf{u} \in \operatorname{cone}(\{\mathbf{F}_1, \dots, \mathbf{F}_m\})\right\}, \\ \Delta &= \min\left\{1, \inf\left\{\|\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{u}\|_2 \,:\, \mathbf{x} \in \mathcal{X},\, \mathbf{u} \in \operatorname{cone}(\{\mathbf{F}_1, \dots, \mathbf{F}_m\})\right\}\right\}, \\ c(\mathbf{x}) &= \max\{\max\left\{\|\mathbf{f} - \mathbf{L}\mathbf{x} + \mathbf{u}\|_2^2 \,:\, \mathbf{u} \in (\cup_{j=1}\{\mathbf{F}_j\}) \cup (\cup_{j=1}\{-\mathbf{F}_j\}) \cup \{\mathbf{0}_q\}\right\}, \bar{F}\}, \end{aligned}$$

and introduce some $\varepsilon \in \mathbb{R}$ *that can be chosen arbitrarily as long as it satisfies:*

$$0 < \varepsilon < \min\left\{\min_{\mathbf{x}\in\mathcal{X}}\frac{1}{c(\mathbf{x})\mu_F^2}, \min_{\mathbf{x}\in\mathcal{X}}\frac{1}{c(\mathbf{x})} \cdot \frac{\Delta^2}{2\|g(\mathbf{x})\|_2\mu_F + 1 + \mu_F^2\|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2}\right\}.$$

Then, there exists a finite constant C *(independent of* m, n, k, Δ*) that satisfies:*

$$\mathbb{P}\left[\mathbf{P}\left(\mathbf{f} - \mathbf{L}\mathbf{x}\right) \notin \operatorname{cone}(\{\mathbf{P}\mathbf{F}_1, \dots, \mathbf{P}\mathbf{F}_m\})\right] \geq 1 - 2(m+1)(m+2)e^{-C(\varepsilon^2 - \varepsilon^3)k}.$$

Proof of Theorem 2. To begin, we introduce the following two events:

$$S_\varepsilon^- := \big\{\forall\, \mathbf{u}, \mathbf{v} \in \{\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{F}_j\,,\, j \in [m]\} \cup \{0\} \,:\, (1-\varepsilon)\|\mathbf{u}-\mathbf{v}\|_2^2 \le \|\mathbf{P}(\mathbf{u}-\mathbf{v})\|_2^2 \le (1+\varepsilon)\|\mathbf{u}-\mathbf{v}\|_2^2\big\}$$

$$S_\varepsilon^+ := \big\{\forall\, \mathbf{u}, \mathbf{v} \in \{\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{F}_j\,,\, j \in [m]\} \cup \{0\} \,:\, (1-\varepsilon)\|\mathbf{u}+\mathbf{v}\|_2^2 \le \|\mathbf{P}(\mathbf{u}+\mathbf{v})\|_2^2 \le (1+\varepsilon)\|\mathbf{u}+\mathbf{v}\|_2^2\big\}$$

and we define $S_\varepsilon := S_\varepsilon^- \cap S_\varepsilon^+$. By using Lemma 6, we have that, there exists some finite constant C (independent of m, n, k, ε) such that:

$$\mathbb{P}\big[S_\varepsilon\big] \ge 1 - 2\binom{m+1}{2} e^{-C(\varepsilon^2-\varepsilon^3)k} \ge 1 - 2(m+1)(m+2)e^{-C(\varepsilon^2-\varepsilon^3)k}.$$

Next, assume that the event S_ε occurs. Let fix $\mathbf{y} \in \operatorname{cone}(\{\mathbf{F}_1, \ldots, \mathbf{F}_m\})$ and let $\sum_{j=1}^m \lambda_j \mathbf{F}_j$ be a minimal F-representation of $\mathbf{y}$. Then, we introduce some intermediate notations, namely $\mathbf{b} = \mathbf{f} - \mathbf{L}\mathbf{x}$, and we derive the following sequence of inequalities:

$$\begin{aligned}
\|\mathbf{P}\big(\mathbf{f} - \mathbf{L}\mathbf{x}\big) - \mathbf{y}\|_2^2 &= \|\mathbf{P}\mathbf{b} - \sum_{j=1}^m \lambda_j \mathbf{P}\mathbf{F}_j\|_2^2 \\
&= \langle \mathbf{P}\mathbf{b} - \sum_{j=1}^m \lambda_j \mathbf{P}\mathbf{F}_j, \mathbf{P}\mathbf{b} - \sum_{j=1}^m \lambda_j \mathbf{P}\mathbf{F}_j \rangle \\
&= \langle \mathbf{P}\mathbf{b}, \mathbf{P}\mathbf{b}\rangle + \langle \sum_{j=1}^m \lambda_j \mathbf{P}\mathbf{F}_j, \sum_{j=1}^m \lambda_j \mathbf{P}\mathbf{F}_j \rangle - 2\langle \mathbf{P}\mathbf{b}, \sum_{j=1}^m \lambda_j \mathbf{P}\mathbf{F}_j \rangle \\
&= \|\mathbf{P}\mathbf{b}\|_2^2 + \sum_{j=1}^m \sum_{i=1}^m \lambda_i \lambda_j \langle \mathbf{P}\mathbf{F}_i, \mathbf{P}\mathbf{F}_j \rangle - 2 \sum_{j=1}^m \lambda_j \langle \mathbf{P}\mathbf{b}, \mathbf{P}\mathbf{F}_j \rangle \\
&= \|\mathbf{P}\mathbf{b}\|_2^2 + \sum_{j=1}^m \sum_{i \ne j} \lambda_i \lambda_j \langle \mathbf{P}\mathbf{F}_i, \mathbf{P}\mathbf{F}_j \rangle \\
&\quad + \sum_{j=1}^m \lambda_j^2 \|\mathbf{P}\mathbf{F}_j\|_2^2 - 2 \sum_{j=1}^m \lambda_j \langle \mathbf{P}\mathbf{b}, \mathbf{P}\mathbf{F}_j \rangle \\
&= \|\mathbf{P}\mathbf{b}\|_2^2 + \sum_{j=1}^m \lambda_j^2 \|\mathbf{P}\mathbf{F}_j\|_2^2 + \sum_{1 \le i < j \le m} \frac{\lambda_i \lambda_j}{2} \big(\|\mathbf{P}\big(\mathbf{F}_i + \mathbf{F}_j\big)\|_2^2 - \|\mathbf{P}\big(\mathbf{F}_i - \mathbf{F}_j\big)\|_2^2\big) \\
&\quad - \frac{1}{2} \sum_{j=1}^m \lambda_j \big(\|\mathbf{P}\mathbf{b} + \mathbf{P}\mathbf{F}_j\|_2^2 - \|\mathbf{P}\mathbf{b} - \mathbf{P}\mathbf{F}_j\|_2^2\big),
\end{aligned}$$

where the last equality is obtained using $4\langle \mathbf{u}, \mathbf{v}\rangle = \|\mathbf{u} + \mathbf{v}\|_2^2 - \|\mathbf{u} - \mathbf{v}\|_2^2$.

Since S_ε occurs, we have that, for $j \in [m]$:

$$\|\mathbf{P}\big(\mathbf{f} - \mathbf{L}\mathbf{x}\big)\|_2^2 \ge (1-\varepsilon)\|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2\,,\ \text{and,}$$

$$\|\mathbf{P}\mathbf{F}_j\|_2^2 \ge (1-\varepsilon)\|\mathbf{F}_j\|_2^2.$$

Moreover, we obtain the following two inequalities:

$$\|\mathbf{P}\big(\mathbf{F}_i+\mathbf{F}_j\big)\|_2^2-\|\mathbf{P}\big(\mathbf{F}_i-\mathbf{F}_j\big)\|_2^2\geq(1-\varepsilon)\|\mathbf{F}_i+\mathbf{F}_j\|_2^2-(1+\varepsilon)\|\mathbf{F}_i-\mathbf{F}_j\|_2^2,$$
$$\|\mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}+\mathbf{F}_j\big)\|_2^2-\|\mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\big)\|_2^2\leq(1+\varepsilon)\|\mathbf{f}-\mathbf{L}\mathbf{x}+\mathbf{F}_j\|_2^2-(1-\varepsilon)\|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2.$$

Accordingly, the following sequence of inequalities is guaranteed to hold:

$$\begin{aligned}\|\mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}\big)-\mathbf{y}\|_2^2\geq{}&(1-\varepsilon)\|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2+(1-\varepsilon)\sum_{j=1}^m\lambda_j^2\|\mathbf{F}_j\|_2^2\\&+\sum_{1\leq i<j\leq m}\frac{\lambda_i\lambda_j}{2}\big((1-\varepsilon)\|\mathbf{F}_i+\mathbf{F}_j\|_2^2-(1+\varepsilon)\|\mathbf{F}_i-\mathbf{F}_j\|_2^2\big)\\&-\sum_{j=1}^m\frac{\lambda_j}{2}\big((1+\varepsilon)\|\mathbf{f}-\mathbf{L}\mathbf{x}+\mathbf{F}_j\|_2^2-(1-\varepsilon)\|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2\big)\\&=\|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2+\sum_{j=1}^m\lambda_j^2\|\mathbf{F}_j\|_2^2+\sum_{1\leq i<j\leq m}\frac{\lambda_i\lambda_j}{2}\big(\|\mathbf{F}_i+\mathbf{F}_j\|_2^2-\|\mathbf{F}_i-\mathbf{F}_j\|_2^2\big)\\&\quad-\sum_{j=1}^m\frac{\lambda_j}{2}\big(\|\mathbf{f}-\mathbf{L}\mathbf{x}+\mathbf{F}_j\|_2^2-\|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2\big)-\varepsilon\big(\|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2+\sum_{j=1}^m\lambda_j^2\|\mathbf{F}_j\|_2^2\big)\\&\quad-\varepsilon\sum_{1\leq i<j\leq m}\frac{\lambda_i\lambda_j}{2}\big(\|\mathbf{F}_i+\mathbf{F}_j\|_2^2+\|\mathbf{F}_i-\mathbf{F}_j\|_2^2\big)\\&\quad-\varepsilon\sum_{j=1}^m\frac{\lambda_j}{2}\big(\|\mathbf{f}-\mathbf{L}\mathbf{x}+\mathbf{F}_j\|_2^2+\|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2\big)\\&=\mathcal{A}^--\varepsilon\mathcal{A}^+,\end{aligned}$$

where:

$$\begin{aligned}\mathcal{A}^-={}&\|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2+\sum_{j=1}^m\lambda_j^2\|\mathbf{F}_j\|_2^2+\sum_{1\leq i<j\leq m}\frac{\lambda_i\lambda_j}{2}\big(\|\mathbf{F}_i+\mathbf{F}_j\|_2^2-\|\mathbf{F}_i-\mathbf{F}_j\|_2^2\big)\\&-\sum_{j=1}^m\frac{\lambda_j}{2}\big(\|\mathbf{f}-\mathbf{L}\mathbf{x}+\mathbf{F}_j\|_2^2-\|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2\big),\\\mathcal{A}^+={}&\|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2+\sum_{j=1}^m\lambda_j^2\|\mathbf{F}_j\|_2^2+\sum_{1\leq i<j\leq m}\frac{\lambda_i\lambda_j}{2}\big(\|\mathbf{F}_i+\mathbf{F}_j\|_2^2+\|\mathbf{F}_i-\mathbf{F}_j\|_2^2\big)\\&+\sum_{j=1}^m\frac{\lambda_j}{2}\big(\|\mathbf{f}-\mathbf{L}\mathbf{x}+\mathbf{F}_j\|_2^2+\|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2\big).\end{aligned}$$

Next, we introduce the following two constants:

$$\underline{F}:=\min_{i\neq j\in[m]}\langle\mathbf{F}_i,\mathbf{F}_j\rangle\text{ and }\bar{F}:=\max\big\{\max_{i\neq j\in[m]}\langle\mathbf{F}_i,\mathbf{F}_j\rangle,\,1\big\}.$$

Moreover, observe that the following two equalities hold:

$$\|\mathbf{F}_i + \mathbf{F}_j\|_2^2 - \|\mathbf{F}_i - \mathbf{F}_j\|_2^2 = 4\langle \mathbf{F}_i, \mathbf{F}_j \rangle$$
$$\|\mathbf{f} - \mathbf{L}\mathbf{x} + \mathbf{F}_j\|_2^2 - \|\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{F}_j\|_2^2 = 4\langle \mathbf{f} - \mathbf{L}\mathbf{x}, \mathbf{F}_j \rangle.$$

To proceed, we derive a lower bound and an upper bound for $\mathcal{A}^-$ and $\mathcal{A}^+$, respectively. Specifically, we obtain the following sequence of equalities:

$$\begin{aligned}
\mathcal{A}^- &= \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + \sum_{j=1}^m \lambda_j^2 \|\mathbf{F}_j\|_2^2 + \sum_{1 \le i < j \le m} 4\frac{\lambda_i \lambda_j}{2} \langle \mathbf{F}_i, \mathbf{F}_j \rangle - 4 \sum_{j=1}^m \frac{\lambda_j}{2} \langle \mathbf{f} - \mathbf{L}\mathbf{x}, \mathbf{F}_j \rangle \\
&= \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + \sum_{j=1}^m \lambda_j^2 \|\mathbf{F}_j\|_2^2 + \sum_{i=1}^m \sum_{j \ne i} \lambda_i \lambda_j \langle \mathbf{F}_i, \mathbf{F}_j \rangle - 2 \sum_{j=1}^m \lambda_j \langle \mathbf{f} - \mathbf{L}\mathbf{x}, \mathbf{F}_j \rangle \\
&= \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + \sum_{i=1}^m \sum_{j=1}^m \lambda_i \lambda_j \langle \mathbf{F}_i, \mathbf{F}_j \rangle - 2\langle \mathbf{f} - \mathbf{L}\mathbf{x}, \sum_{j=1}^m \lambda_j \mathbf{F}_j \rangle \\
&= \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 - 2\langle \mathbf{f} - \mathbf{L}\mathbf{x}, \mathbf{y} \rangle + \|\mathbf{y}\|_2^2 = \|\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{y}\|_2^2.
\end{aligned}$$

We fix $\alpha = \|\mathbf{y}\|_2^2$. Recall that by definition of $g \equiv g(\mathbf{x})$, we have that g is the orthogonal projection of $\mathbf{f} - \mathbf{L}\mathbf{x}$ into $\text{cone}(\{\mathbf{F}_1, \ldots, \mathbf{F}_m\})$. We next use this definition to demonstrate the following intermediate result:

Lemma 10. $\|\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{y}\|_2^2 \ge \alpha^2 - 2\alpha\|g\|_2^2 + \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2$.

Proof of Lemma 10. If $\mathbf{y} = \mathbf{0}_m$, then the claim is trivially true by the definition of α. Accordingly, we assume that $\mathbf{y} \ne \mathbf{0}_m$. To proceed, we treat two cases, whether $g \ne 0$.

Case 1 ($g(\mathbf{x}) \ne 0$): We define $d(\mathbf{x}) := \min\{\|\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{u}\|_2 : \mathbf{u} \in \mathcal{C}\}$ where we recall $\mathcal{C} := \text{cone}(\{\mathbf{F}_1, \ldots, \mathbf{F}_m\})$, which is attained for $g(\mathbf{x})$. Then, we proceed by showing the following equality $d(\mathbf{x})^2 + \|g(\mathbf{x})\|_2^2 = \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2$. Accordingly, the following equalities are guaranteed to hold:

$$\begin{aligned}
d^2(\mathbf{x}) + \|g(\mathbf{x})\|_2^2 &= \|\mathbf{f} - \mathbf{L}\mathbf{x} - g(\mathbf{x})\|_2^2 + \|g(\mathbf{x})\|_2^2 \\
&= \langle \mathbf{f} - \mathbf{L}\mathbf{x} - g(\mathbf{x}), \mathbf{f} - \mathbf{L}\mathbf{x} - g(\mathbf{x}) \rangle + \langle g(\mathbf{x}), g(\mathbf{x}) \rangle \\
&= \langle g(\mathbf{x}), g(\mathbf{x}) \rangle + \langle \mathbf{f} - \mathbf{L}\mathbf{x}, \mathbf{f} - \mathbf{L}\mathbf{x} \rangle + \langle g(\mathbf{x}), g(\mathbf{x}) \rangle - 2\langle \mathbf{f} - \mathbf{L}\mathbf{x}, g(\mathbf{x}) \rangle \\
&= \langle g(\mathbf{x}), g(\mathbf{x}) \rangle - 2\langle g(\mathbf{x}), g(\mathbf{x}) \rangle + \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + \|g(\mathbf{x})\|_2^2 \\
&\quad - 2\langle \mathbf{f} - \mathbf{L}\mathbf{x} - g(\mathbf{x}), g(\mathbf{x}) \rangle \\
&\overset{(a)}{=} \langle g(\mathbf{x}), g(\mathbf{x}) \rangle - 2\langle g(\mathbf{x}), g(\mathbf{x}) \rangle + \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + \|g(\mathbf{x})\|_2^2 \\
&= \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2,
\end{aligned}$$

where (a) follows by that g is the orthogonal projection of $\mathbf{f} - \mathbf{L}\mathbf{x}$ into the cone $\mathcal{C}$.

Then, we fix $\mathbf{z} = \frac{\|g\|_2}{\alpha}\mathbf{y}$. Accordingly, as $\mathbf{y} \in \mathcal{C}$, we have that $\mathbf{z} \in \text{cone}(\{\mathbf{F}_1, \ldots, \mathbf{F}_m\})$ and $\|\mathbf{z}\|_2^2 = \|g\|_2^2$. Next, we introduce $\delta = \frac{\alpha}{\|g\|_2}$, and we obtain the following sequence of inequalities:

$$\begin{aligned}
\|\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{y}\|_2^2 &= \|\mathbf{f} - \mathbf{L}\mathbf{x} - \delta\mathbf{z}\|_2^2 \\
&= \langle \mathbf{f} - \mathbf{L}\mathbf{x} - \delta\mathbf{z}, \mathbf{f} - \mathbf{L}\mathbf{x} - \delta\mathbf{z}\rangle \\
&= \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + \delta^2\|\mathbf{z}\|_2^2 - 2\delta\langle \mathbf{f} - \mathbf{L}\mathbf{x}, z\rangle \\
&= (1-\delta)\|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 \\
&\quad + \delta\langle \mathbf{f} - \mathbf{L}\mathbf{x}, \mathbf{f} - \mathbf{L}\mathbf{x}\rangle - 2\delta\langle \mathbf{f} - \mathbf{L}\mathbf{x}, \mathbf{z}\rangle + (\delta^2 - \delta)\|\mathbf{z}\|_2^2 + \delta\|\mathbf{z}\|_2^2 \\
&= (1-\delta)\|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + (\delta^2 - \delta)\|\mathbf{z}\|_2^2 \\
&\quad + \delta\langle \mathbf{f} - \mathbf{L}\mathbf{x}, \mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{z}\rangle - \delta\langle \mathbf{f} - \mathbf{L}\mathbf{x}, \mathbf{z}\rangle + \delta\langle \mathbf{z}, \mathbf{z}\rangle \\
&= (1-\delta)\|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + (\delta^2 - \delta)\|\mathbf{z}\|_2^2 \\
&+ \delta\big(\langle \mathbf{f} - \mathbf{L}\mathbf{x}, \mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{z}\rangle - \langle -\mathbf{z}, \mathbf{z}\rangle - \langle \mathbf{f} - \mathbf{L}\mathbf{x}, \mathbf{z}\rangle\big) \\
&= (1-\delta)\|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + (\delta^2 - \delta)\|\mathbf{z}\|_2^2 + \delta\|\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{z}\|_2^2 \\
&\geq (1-\delta)\|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + (\delta^2 - \delta)\|\mathbf{z}\|_2^2 + \delta d(\mathbf{x})^2 \\
&= (1-\delta)\|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + (\delta^2 - \delta)\|\mathbf{z}\|_2^2 + \delta\big(\|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 - \|g(\mathbf{x})\|_2^2\big) \\
&= \delta^2\|g(\mathbf{x})\|_2^2 - 2\delta\|g(\mathbf{x})\|_2^2 + \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 \\
&= \alpha^2 - 2\alpha\|g\|_2 + \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2,
\end{aligned}$$

which conclude the first case of the proof.

Case 2 ($g(\mathbf{x}) = 0$): Assume that $g(\mathbf{x}) = 0$. Then, we show that $\big(\mathbf{f} - \mathbf{L}\mathbf{x}\big)^\top\mathbf{z} \leq 0$, for all $\mathbf{z} \in \text{cone}(\{\mathbf{F}_1, \ldots, \mathbf{F}_m\})$. Fix $\mathbf{y} \in \text{cone}(\{\mathbf{F}_1, \ldots, \mathbf{F}_m\})$ and $\delta > 0$ be chosen arbitrarily. Then, we derive the following inequality and limit:

$$\begin{aligned}
0 &\leq \frac{1}{\delta}\big(\|\mathbf{f} - \mathbf{L}\mathbf{x} - \delta\mathbf{y}\|_2^2 - \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2\big) \\
&= \frac{1}{\delta}\big(\|\delta\mathbf{y}\|_2^2 - 2\delta\langle \mathbf{f} - \mathbf{L}\mathbf{x}, \mathbf{y}\rangle\big) \overset{\delta\downarrow 0}{\to} -2\big(\mathbf{f} - \mathbf{L}\mathbf{x}\big)^\top\mathbf{y}.
\end{aligned}$$

Therefore, the following (in)-equalities are satisfied:

$$\begin{aligned}
\|\mathbf{f} - \mathbf{L}\mathbf{x} - \mathbf{y}\|_2^2 &= \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2 - 2\big(\mathbf{f} - \mathbf{L}\mathbf{x}\big)^\top\mathbf{y} \\
&\geq \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + \|\mathbf{y}\|_2^2 = \alpha^2 + \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 - 2\alpha\|g(\mathbf{x})\|_2,
\end{aligned}$$

as $g(\mathbf{x}) = 0$. This observation, in turn, concludes the proof of Lemma 10. ■

Next, we derive an upper bound on $\mathcal{A}^+$. Thus, we derive the following sequence of inequalities:

$$\begin{aligned}
\mathcal{A}^+ &= \|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2 + \sum_{j=1}^m \lambda_j^2 \|\mathbf{F}_j\|_2^2 + \sum_{1\le i<j\le m} \frac{\lambda_i\lambda_j}{2}\big(\|\mathbf{F}_i+\mathbf{F}_j\|_2^2 + \|\mathbf{F}_i-\mathbf{F}_j\|_2^2\big) \\
&\quad + \sum_{j=1}^m \frac{\lambda_j}{2}\big(\|\mathbf{f}-\mathbf{L}\mathbf{x}+\mathbf{F}_j\|_2^2 + \|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2\big) \\
&\le \|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2 + \sum_{j=1}^m \lambda_j^2 \bar{F} + \sum_{1\le i<j\le m} \frac{\lambda_i\lambda_j}{2}\cdot 4\bar{F} + \sum_{j=1}^m \frac{\lambda_j}{2}\big(\|\mathbf{f}-\mathbf{L}\mathbf{x}+\mathbf{F}_j\|_2^2 + \|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2\big) \\
&\le \|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2 + \bar{F}\sum_{i=1}^m\sum_{j=1}^m \lambda_i\lambda_j + \sum_{j=1}^m \frac{\lambda_j}{2}\big(\|\mathbf{f}-\mathbf{L}\mathbf{x}+\mathbf{F}_j\|_2^2 + \|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2\big) \\
&= \|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2 + \bar{F}\|\mathbf{y}\|_F^2 + \sum_{j=1}^m \frac{\lambda_j}{2}\big(\|\mathbf{f}-\mathbf{L}\mathbf{x}+\mathbf{F}_j\|_2^2 + \|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2\big) \\
&\overset{(a)}{\le} \|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2 + \bar{F}\mu_F\|\mathbf{y}\|_2^2 + \sum_{j=1}^m \frac{\lambda_j}{2}\big(\|\mathbf{f}-\mathbf{L}\mathbf{x}+\mathbf{F}_j\|_2^2 + \|\mathbf{f}-\mathbf{L}\mathbf{x}-\mathbf{F}_j\|_2^2\big) \\
&\overset{(b)}{\le} c(\mathbf{x}) + \bar{F}\mu_F^2\|\mathbf{y}\|_2^2 + \sum_{j=1}^m \frac{\lambda_j}{2}\cdot 2c(\mathbf{x}) \\
&\overset{(c)}{\le} c(\mathbf{x})\left(1+\mu_F^2\|\mathbf{y}\|_2^2 + \mu_F\|\mathbf{y}\|_2\right) = c(\mathbf{x})\left(1+\mu_F^2\alpha^2 + 2\mu_F\alpha\right),
\end{aligned}$$

where (a) follows from Lemma 9. Also, (b) follows from the definition of $c(\mathbf{x})$, namely:

$$c(\mathbf{x}) := \max\{\max\left\{\|\mathbf{f}-\mathbf{L}\mathbf{x}+\mathbf{u}\|_2^2 \,:\, \mathbf{u}\in(\cup_{j=1}\{\mathbf{F}_j\})\cup(\cup_{j=1}\{-\mathbf{F}_j\})\cup\{\mathbf{0}_q\}\right\}, \bar{F}\}.$$

Moreover, (c) follows from both the definition of $c(\mathbf{x})$ and applying Lemma 9.

Therefore, the following inequalities are satisfied:

$$\begin{aligned}
\|\mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}\big)-\mathbf{y}\|_2^2 \ge \alpha^2 - 2\alpha\|g(\mathbf{x})\|_2 + \|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2& \\
-\varepsilon\big(c(\mathbf{x})\big(1+\mu_F^2\alpha^2+2\mu_F\alpha\big)\big)& \\
= \big(1-\varepsilon c(\mathbf{x})\mu_F^2\big)\alpha^2 - 2\big(\|g(\mathbf{x})\|_2 + \varepsilon c(\mathbf{x})\mu_F\big)\alpha + \|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2 - \varepsilon c(\mathbf{x}).&
\end{aligned}$$

Recall that a polynomial of degree two of the form $p(\alpha) = a\alpha^2 + b\alpha + c$ is strictly positive if and only if $a > 0$ and $b^2 - 4ac < 0$. Hence, if ε is chosen such that:

$$\begin{aligned}
1-\varepsilon c(\mathbf{x})\mu_F^2 > 0,& \\
4\big(\|g(\mathbf{x})\|_2 + \varepsilon c(\mathbf{x})\mu_F\big)^2 - 4\big(1-\varepsilon c(\mathbf{x})\mu_F^2\big)\big(\|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2 - \varepsilon c(\mathbf{x})\big) < 0,&
\end{aligned}$$

then we have that $\|\mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}\big)-\mathbf{y}\|_2^2 > 0$. Accordingly, to guarantee that the first inequality holds, we should select ε that satisfies the following inequality:

$$\varepsilon < \frac{1}{c(\mathbf{x})\mu_F^2}.$$

Additionally, ε should be picked so that is satisfies the following condition:

$$\begin{aligned}
& 4\big(\|g(\mathbf{x})\|_2 + \varepsilon c(\mathbf{x})\mu_F\big)^2 - 4\big(1-\varepsilon c(\mathbf{x})\mu_F^2\big)\big(\|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2 - \varepsilon c(\mathbf{x})\big) < 0 \\
\Leftrightarrow\quad & \big(\|g(\mathbf{x})\|_2 + \varepsilon c(\mathbf{x})\mu_F\big)^2 - \big(1-\varepsilon c(\mathbf{x})\mu_F^2\big)\big(\|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2 - \varepsilon c(\mathbf{x})\big) < 0 \\
\Leftrightarrow\quad & \|g(\mathbf{x})\|_2^2 + \varepsilon^2 c(\mathbf{x})^2\mu_F^2 + 2\|g(\mathbf{x})\|_2\varepsilon c(\mathbf{x})\mu_F \\
& \quad - \big(\|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2 - \varepsilon c(\mathbf{x}) - \varepsilon c(\mathbf{x})\mu_F^2\|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2 + \varepsilon^2 c(\mathbf{x})^2\mu_F^2\big) < 0 \\
\Leftrightarrow\quad & \|g(\mathbf{x})\|_2^2 + 2\|g(\mathbf{x})\|_2\varepsilon c(\mathbf{x})\mu_F - \big(\|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2 - \varepsilon c(\mathbf{x})\big(1+\mu_F^2\|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2\big)\big) < 0 \\
\Leftrightarrow\quad & \varepsilon c(\mathbf{x})\big(2\|g(\mathbf{x})\|_2\mu_F + 1 + \mu_F^2\|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2\big) < \|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2 - \|g(\mathbf{x})\|_2^2 \equiv d(\mathbf{x})^2 \\
\Leftrightarrow\quad & \varepsilon < \frac{1}{c(\mathbf{x})}\cdot\frac{d(\mathbf{x})^2}{2\|g(\mathbf{x})\|_2\mu_F + 1 + \mu_F^2\|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2}
\end{aligned}$$

Accordingly, if we pick ε so that it satisfies the following inequality:

$$0 < \varepsilon < \min\{\frac{1}{c(\mathbf{x})\mu_F^2}, \frac{1}{c(\mathbf{x})}\cdot\frac{\Delta^2}{2\|g(\mathbf{x})\|_2\mu_F + 1 + \mu_F^2\|\mathbf{f}-\mathbf{L}\mathbf{x}\|_2^2}\},$$

then we have that $\|\mathbf{P}\big(\mathbf{f}-\mathbf{L}\mathbf{x}\big) - \mathbf{y}\|_2^2 > 0$, for all $\mathbf{y} \in \mathrm{cone}(\{\mathbf{F}_1,\ldots,\mathbf{F}_m\})$. Consequently, we have that $\mathbf{f}-\mathbf{L}\mathbf{x} \notin \mathrm{cone}(\{\mathbf{PF}_1,\ldots,\mathbf{PF}_m\})$.

Hence, we have that:

$$\mathbb{P}\big[\mathbf{P}(\mathbf{f}-\mathbf{L}\mathbf{x}) \notin \mathbf{P}\mathcal{C}\big] \geq \mathbb{P}\big[S_\varepsilon\big] \geq 1 - 2(m+1)(m+2)e^{-C(\varepsilon^2-\varepsilon^3)k},$$

which concludes, in turn, the proof of Theorem 2 . ■

Cone transformation. We introduce the set $\mathcal{K} := \left\{\sum_{j=1}^m v_j\mathbf{C}_j \,:\, \mathbf{v}\in\mathbb{R}_+^m\right\}$, where $\mathbf{C}_1,\ldots,\mathbf{C}_m$ are the column vectors of a $q\times m$ matrix $\mathbf{C}$, i.e., $\mathcal{K} := \mathrm{cone}(\{\mathbf{C}_1,\ldots,\mathbf{C}_m\})$. Then, for $\mathbf{u}\in\mathbb{R}^q$ and θ such that $\mathbf{1}_m^\top\mathbf{y} \leq \theta$ for all follower's feasible decision $\mathbf{y} \in \cup_{\mathbf{x}\in\mathcal{X}}\mathcal{Y}(\mathbf{x})$, we introduce:

$$\phi_{\mathbf{u},\theta}(\mathcal{K}) := \{\sum_{j=1}^m v_j\big(\mathbf{C}_j - \frac{1}{\theta}\mathbf{u}\big) \,:\, \mathbf{v}\in\mathbb{R}_+^m\}$$

$$\mathcal{K}_\theta := \{\sum_{j=1}^m v_j\mathbf{C}_j \,:\, \mathbf{v}\in\mathbb{R}_+^m,\, \sum_{j=1}^m v_j \leq \theta\}.$$

Lemma 11 (Vu et al., 2018)**.** *For any $\mathbf{u}\in\mathbb{R}^q$, we have that $\mathbf{u}\in\mathcal{K}_\theta$ if and only if $\mathbf{u}\in\phi_{\mathbf{u},\theta}(\mathcal{K})$.*

Corollary 6 (Vu et al., 2018)**.** *For any $\mathbf{u}\in\mathbb{R}^q$ and $J\subseteq\{1,\ldots,m\}$, we have that $\mathbf{u}\in\mathcal{K}_\theta^J$ if and* *only if* $\mathbf{u} \in \phi^J_{\mathbf{u},\theta}(\mathcal{K})$*, where:*

$$\mathbf{C}_j^{J\mathbf{u}} = \begin{cases} \mathbf{C}_j - \frac{1}{\theta}\mathbf{u} & \textit{if } j \in J, \\ \mathbf{C}_j & \textit{otherwise;} \end{cases}$$

$$\phi^J_{\mathbf{u},\theta}(\mathcal{K}) := \{\sum_{j=1}^m v_j \mathbf{C}_j^{J\mathbf{u}} \ : \ \mathbf{v} \in \mathbb{R}^m_+\} \equiv \mathrm{cone}(\{\mathbf{C}_j^J\mathbf{u} \ : \ j \in [m]\});$$

$$\mathcal{K}^J_\theta := \{\sum_{j=1}^m v_j \mathbf{C}_j \ : \ \mathbf{v} \in \mathbb{R}^m_+, \ \sum_{j\in J} v_j \leq \theta\}.$$

Then, let $\varphi_\theta(\mathbf{x}, \mathbf{P}) := \min\{\mathbf{c}^\top\mathbf{y} \ : \ \mathbf{PFy} = \mathbf{P}(\mathbf{f} - \mathbf{Lx}), \ \mathbf{1}_m^\top\mathbf{y} \leq \theta, \ \mathbf{y} \in \mathbb{R}^m_+\}$ be the follower's projected value function, with the additional constrain which guarantees that the follower's feasible region is bounded. Moreover, we assume that θ is large enough so that:

$$\max\left\{1, \sup_{\mathbf{x}\in\mathcal{X}} \|\mathbf{f} - \mathbf{Lx}\|_2^2\right\} < \theta.$$

Observation. Let $\mathbf{P}$ be a $k \times q$ random projection. Then, Lemma 7 also applies to $(k+h) \times q$ random projection $\tilde{\mathbf{P}}$ which takes the following form:

$$\tilde{\mathbf{P}} := \left(\begin{array}{c|c} \mathbf{I}_{h\times h} & \mathbf{0}_{h\times(q-h)} \\ \hline \multicolumn{2}{c}{\mathbf{P}} \end{array}\right)$$

Equipped with the preceding lemmas and propositions, we proceed by demonstrating the primary result: namely that the follower's projected value function deviates from the original value function by only a small amount with "high" probability.

E.C.3 Main result

The remainder of this section is devoted to proving Theorem 3. As pointed out at the beginning of the companion, the exposition bellow parallels the approach by Vu et al. (2018). Formally:

Proof of Theorem 3. To proceed, we first fix some leader's feasible decision $\mathbf{x} \in \mathcal{X}$, together with some follower's feasible decision $\mathbf{y} \in \mathcal{Y}_\theta(\mathbf{x})$. Then, we introduce the following matrix and vectors:

$$\tilde{\mathbf{F}} = \left(\begin{array}{c|c} \mathbf{c}^\top & 1 \\ \hline \mathbf{F} & \mathbf{0} \end{array}\right), \quad \tilde{\mathbf{y}} = \begin{pmatrix} \mathbf{y} \\ s \end{pmatrix}, \quad \tilde{\mathbf{b}} = \begin{pmatrix} \varphi_\theta(\mathbf{x}) - \delta \\ \mathbf{f} - \mathbf{Lx} \end{pmatrix},$$

where $s \in \mathbb{R}_{\geq 0}$. Next, given a $k \times q$ random projector $\mathbf{P}$ defined for some $k \in [q]$, we introduce the

$(k+1)\times(q+1)$ random projector $\tilde{\mathbf{P}}$ of the form:

$$\tilde{\mathbf{P}} = \left(\begin{array}{c|c} 1 & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{P} \end{array}\right).$$

Let J be the index set of the first m columns of $\tilde{\mathbf{F}}$. Consider $\phi^J_{\tilde{\mathbf{b}},\theta'}$ for some $\theta' \in (\theta, \theta+1)$ chosen arbitrarily. Then, we introduce the matrix $\mathbf{F}'$ defined as follows:

$$\mathbf{F}' = \left[\tilde{\mathbf{F}}_1 - \frac{1}{\theta'}\tilde{\mathbf{b}}, \ldots, \tilde{\mathbf{F}}_m - \frac{1}{\theta'}\tilde{\mathbf{b}}, \tilde{\mathbf{F}}_{m+1}\right].$$

Since $\mathbf{y}$ is a follower's feasible decision, it cannot have an objective function value lower than $\varphi_\theta(\mathbf{x})$. Hence, the following system of linear equations:

$$\begin{cases} \mathbf{F}\mathbf{y} = \mathbf{f} - \mathbf{L}\mathbf{x}, \\ \varphi_\theta(\mathbf{x}) - \delta \geq \mathbf{c}^\top \mathbf{y}, \\ \mathbf{y} \geq \mathbf{0}, \end{cases}$$

is infeasible. Therefore, we conclude that the following system of linear equations:

$$\begin{cases} \tilde{\mathbf{F}}\,\tilde{\mathbf{y}} = \tilde{\mathbf{b}}, \\ \mathbf{1}_m^\top \mathbf{y} \leq \theta', \\ \tilde{\mathbf{y}} \geq \mathbf{0}, \end{cases}$$

is also infeasible.

Our goal is to show that the following system:

$$\begin{cases} \tilde{\mathbf{P}}\mathbf{F}'\,\tilde{\mathbf{y}} = \tilde{\mathbf{P}}\tilde{\mathbf{b}}, \\ \tilde{\mathbf{y}} \geq \mathbf{0}, \end{cases}$$

is infeasible with small probability.

To proceed, we obtain the following result on $\tilde{\mathbf{b}}$:

$$\tilde{\mathbf{b}} \notin \left\{ \sum_{j=1}^{m} \tilde{v}_j \tilde{\mathbf{F}}_j \;:\; \tilde{\mathbf{v}} \in \mathbb{R}^m_+,\; \mathbf{1}_J \tilde{\mathbf{v}} \leq \theta' \right\}.$$

Accordingly, $\tilde{\mathbf{b}} \notin \mathcal{K}^J_\theta$, and, using Corollary 6, we have that $\tilde{\mathbf{b}} \notin \phi^J_{\tilde{\mathbf{b}},\theta'}(\mathcal{K}) \equiv \operatorname{cone}\left(\mathbf{F}'_1, \ldots, \mathbf{F}'_{m+1}\right)$. Then, we introduce $\mathbf{u}^*(\mathbf{x}) \in \mathbb{R}^q$, as in the statement of the theorem. We let $\mathbf{u}^*(\mathbf{x})$ be the optimal solution to the follower's dual problem with the minimal ℓ_2-norm. By using a strong duality argument (Bertsimas and Tsitsiklis 1997), $\mathbf{u}^*(\mathbf{x})$ is guaranteed to satisfy the following two (in)-equalities:

$$\mathbf{u}^*(\mathbf{x})^\top \mathbf{F} \leq \mathbf{c}^\top \text{ and } \mathbf{u}^*(\mathbf{x})^\top \big(\mathbf{f} - \mathbf{L}\mathbf{x}\big) = \varphi_\theta(\mathbf{x}).$$

To continue, we define $\tilde{\mathbf{u}}^*(\mathbf{x})^\top := \big(1, -\mathbf{u}^*(\mathbf{x})^\top\big)$. In our next step, we show that the following two inequalities are satisfied by $\tilde{\mathbf{u}}^*(\mathbf{x})$:

$$\tilde{\mathbf{u}}^*(\mathbf{x})^\top \mathbf{F}' > \mathbf{0} \text{ and } \tilde{\mathbf{u}}^*(\mathbf{x})^\top \tilde{\mathbf{b}} < 0.$$

By using the dual feasibility of $\mathbf{u}^*(\mathbf{x})$, the following inequalities are guaranteed to hold:

$$\tilde{\mathbf{u}}^*(\mathbf{x})^\top \tilde{\mathbf{F}} = \Big(1, -\mathbf{u}^*(\mathbf{x})^\top\Big) \left(\begin{array}{c|c} \mathbf{c}^\top & 1 \\ \hline \mathbf{F} & \mathbf{0} \end{array}\right) = \big(\mathbf{c}^\top - \mathbf{u}^*(\mathbf{x})^\top \mathbf{F}, 1\big) \geq \mathbf{0}_{m+1}^\top$$

where the last inequality comes from dual feasibility, namely, $\mathbf{u}^*(\mathbf{x})^\top \mathbf{F} \leq \mathbf{c}^\top$.

Moreover, we have that:

$$\tilde{\mathbf{u}}^*(\mathbf{x})^\top \tilde{\mathbf{b}} = \big(1, -\mathbf{u}^*(\mathbf{x})^\top\big)\big(\varphi_\theta(\mathbf{x}) - \delta, (\mathbf{f} - \mathbf{L}\mathbf{x})^\top\big)^\top = \varphi_\theta(\mathbf{x}) - \delta - \mathbf{u}^*(\mathbf{x})^\top\big(\mathbf{f} - \mathbf{L}\mathbf{x}\big) = -\delta < 0.$$

Therefore, the following sequence of (in)equalities holds:

$$\begin{aligned}
\tilde{\mathbf{u}}^*(\mathbf{x})^\top \mathbf{F}' &= \Big(1, -\mathbf{u}^*(\mathbf{x})^\top\Big) \left(\tilde{\mathbf{F}}_1 - \frac{1}{\theta'}\tilde{\mathbf{b}}, \ldots, \tilde{\mathbf{F}}_m - \frac{1}{\theta'}\tilde{\mathbf{b}}, \tilde{\mathbf{F}}_{m+1}\right) \\
&= \Big(1, -\mathbf{u}^*(\mathbf{x})^\top\Big) \left(\begin{array}{c|c} \mathbf{c}^\top - \frac{1}{\theta'}\big(\varphi_\theta(\mathbf{x}) - \delta\big)\mathbf{1}_m^\top & 1 \\ \hline \mathbf{F} - \frac{1}{\theta'}\big(\mathbf{f} - \mathbf{L}\mathbf{x}\big)\mathbf{1}_m^\top & \mathbf{0}_m \end{array}\right) \\
&= \left(\mathbf{c}^\top - \frac{1}{\theta'}\big(\varphi_\theta(\mathbf{x}) - \delta\big)\mathbf{1}_m^\top - \mathbf{u}^*(\mathbf{x})^\top\big(\mathbf{F} - \frac{1}{\theta'}\big(\mathbf{f} - \mathbf{L}\mathbf{x}\big)\mathbf{1}_m^\top\big), 1\right)^\top \\
&= \left(\mathbf{c}^\top - \mathbf{u}^*(\mathbf{x})^\top \mathbf{F} + \frac{\delta}{\theta'}\mathbf{1}_m^\top, 1\right)^\top \geq \left(\frac{\delta}{\theta'}\mathbf{1}_m^\top, 1\right)^\top = \left(\frac{\delta}{1+\theta}\mathbf{1}_m^\top, 1\right)^\top > \mathbf{0}_{m+1}.
\end{aligned}$$

Then, by using Proposition 3, we have that, for all $j \in [m+1]$, the following two inequalities are guaranteed to hold:

$$\begin{gathered}
|\big(\langle \tilde{\mathbf{P}}\tilde{\mathbf{u}}^*(\mathbf{x}), \tilde{\mathbf{P}}\mathbf{F}'\rangle - \tilde{\mathbf{u}}^*(\mathbf{x})^\top \mathbf{F}'\big)_j| \leq \varepsilon \|\tilde{\mathbf{u}}^*(\mathbf{x})\|_2 \|\mathbf{F}_j'\|_2, \text{ and} \\
|\langle \tilde{\mathbf{P}}\tilde{\mathbf{u}}^*(\mathbf{x}), \tilde{\mathbf{P}}\tilde{\mathbf{b}}\rangle - \tilde{\mathbf{u}}^*(\mathbf{x})^\top \tilde{\mathbf{b}}| \leq \varepsilon \|\tilde{\mathbf{u}}^*(\mathbf{x})\|_2 \|\tilde{\mathbf{b}}\|_2,
\end{gathered}$$

each with probability at least $p = 1 - 2e^{-C(\varepsilon^2 - \varepsilon^3)k}$.

Next, we define the parameter $\eta \equiv \eta(\mathbf{x})$ as follows:

$$\eta \equiv \eta(\mathbf{x}) := \|\tilde{\mathbf{u}}^*(\mathbf{x})\|_2 \cdot \max\big\{ \max_{j \in [m+1]} \{\|\mathbf{F}_j'\|_2\}, \|\tilde{\mathbf{b}}\|_2\big\}.$$

By using the union of bound, we have that the following two inequalities:

$$|\langle \tilde{\mathbf{P}}\tilde{\mathbf{u}}^*(\mathbf{x}), \tilde{\mathbf{P}}\mathbf{F}'\rangle - \tilde{\mathbf{u}}^*(\mathbf{x})^\top \mathbf{F}'| \leq \varepsilon\eta \quad \text{and} \quad |\langle \tilde{\mathbf{P}}\tilde{\mathbf{u}}^*(\mathbf{x}), \tilde{\mathbf{P}}\tilde{\mathbf{b}}\rangle - \tilde{\mathbf{u}}^*(\mathbf{x})^\top \tilde{\mathbf{b}}| \leq \varepsilon\eta,$$

hold (the intersection of their respective events) with probability at least $p = 1 - 4me^{-C(\varepsilon^2 - \varepsilon^3)k}$.

Lemma 12. $\eta(\mathbf{x}) := \mathcal{O}\big(\theta\|\mathbf{u}^*(\mathbf{x})\|_2\big)$*, where the constant only depends on* $\mathbf{c}, \mathbf{F}, \mathbf{f}, \mathbf{L}$*.*

Proof of Lemma 12. To begin, the following sequence of inequalities is guaranteed to hold:

$$\begin{aligned}
\|\tilde{\mathbf{b}}\|_2^2 &= \| \Big(\varphi_\theta(\mathbf{x}) - \delta, (\mathbf{f} - \mathbf{L}\mathbf{x})^\top\Big) \|_2^2 \\
&= \|\varphi_\theta(\mathbf{x}) - \delta\|_2^2 + \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 \\
&\overset{(a)}{\leq} \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + 2\varphi_\theta(\mathbf{x})^2 + 2\delta^2 \\
&\overset{(b)}{\leq} \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + 4\varphi_\theta(\mathbf{x})^2 \\
&= \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + 4|\mathbf{c}^\top \mathbf{y}^*(\mathbf{x})|^2 \\
&\overset{(c)}{\leq} \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + 4\|\mathbf{c}\|_\infty^2 \|\mathbf{y}^*(\mathbf{x})\|_1^2 \\
&\overset{(d)}{\leq} \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2^2 + 4\|\mathbf{c}\|_\infty^2 \theta^2 \leq \big(1 + 4\|\mathbf{c}\|_\infty^2\big)\theta^2,
\end{aligned}$$

where (a) follows from that $(u - v)^2 \leq 2u^2 + 2v^2$ for all u, v. Moreover, (b) follows from that $0 < \delta < \varphi_\theta(\mathbf{x})$. Additionally, (c) follows from the Hölder's inequality. Then, (d) follows from that the follower's feasible region is bounded and the definition of θ; recall the end of Section E.C.2.

Therefore, by taking the square root, we obtain $\|\tilde{\mathbf{u}}^*(\mathbf{x})\|_2\|\tilde{\mathbf{b}}\|_2 \leq \|\tilde{\mathbf{u}}^*(\mathbf{x})\|_2\big(1 + \|\mathbf{c}\|_\infty^2\big)^{1/2}\theta$. On the other hand, given some $j \in [m]$, we obtain the following sequence of inequalities:

$$\begin{aligned}
\|\mathbf{F}_j^{'}\|_2^2 &= \|\tilde{\mathbf{F}}_j - \frac{1}{\theta'}\tilde{\mathbf{b}}\|_2^2 \\
&\leq \|\tilde{\mathbf{F}}_j\|_2^2 + \frac{1}{\theta'}\|\tilde{\mathbf{b}}\|_2^2 \\
&\leq \|\big(c_j, \mathbf{F}_j^\top\big)^\top\|_2^2 + \frac{1}{\theta}\big(1 + 4\|\mathbf{c}\|_\infty^2\big)\theta^2 \\
&\leq \|c_j\|_2^2 + \|\mathbf{F}_j\|_2^2 + \big(1 + 4\|\mathbf{c}\|_\infty^2\big)\theta.
\end{aligned}$$

Thus, we have that $\|\mathbf{F}_j^{'}\|_2 \leq \big(\|c_j\|_2^2 + \|\mathbf{F}_j\|_2^2\big(1 + 4\|\mathbf{c}\|_\infty^2\big)\theta\big)^{\frac{1}{2}}$. Moreover, $\|\mathbf{F}_{m+1}^{'}\|_2 = 1$.

Finally, observe that:

$$\eta \leq \|\tilde{\mathbf{u}}^*(\mathbf{x})\|_2 \max\big\{1, \big(1 + 4\|\mathbf{c}\|_\infty^2\big)^{\frac{1}{2}}\theta, \max_{j\in[m]}\{\big(\|c_j\|_2^2 + \|\mathbf{F}_j\|_2^2\big(1 + 4\|\mathbf{c}\|_\infty^2\big)\theta\big)^{\frac{1}{2}}\}\big\},$$

which essentially concludes the proof of Lemma 12. ■

Next, we introduce $\varepsilon \equiv \varepsilon(\mathbf{x}) := \frac{\delta}{2(1+\theta)\eta(\mathbf{x})}$. Accordingly, with probability at least p, we have that the following two inequalities are guaranteed to hold (assuming θ is large enough so that $\delta \leq 1 + \theta$):

$$\Big(\tilde{\mathbf{P}}\tilde{\mathbf{u}}^*(\mathbf{x})\Big)^\top \Big(\tilde{\mathbf{P}}\mathbf{F}^{'}\Big) \geq \tilde{\mathbf{u}}^*(\mathbf{x})^\top \mathbf{F}^{'} - \mathbf{1}_{m+1}\varepsilon\eta \geq \mathbf{1}_{m+1}\frac{\delta}{1+\theta} - \mathbf{1}_{m+1}\varepsilon\eta = \big(\frac{\delta}{1+\theta} - \varepsilon\eta\big)\mathbf{1}_{m+1} \geq 0,$$

$$\left(\tilde{\mathbf{P}}\tilde{\mathbf{u}}^*(\mathbf{x})\right)^\top \left(\tilde{\mathbf{P}}\tilde{\mathbf{b}}\right) \leq \tilde{\mathbf{u}}^*(\mathbf{x})^\top \tilde{\mathbf{b}} + \varepsilon\eta \leq -\delta + \varepsilon\eta < \mathbf{0}_{m+1}.$$

Moreover, by the definition of η, we obtain the following inequalities:

$$\eta(\mathbf{x}) \geq \|\tilde{\mathbf{u}}^*(\mathbf{x})\|_2 \|\tilde{\mathbf{b}}\|_2 \geq \|\mathbf{u}^*(\mathbf{x})\|_2 \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2,$$

which essentially implies that $\varepsilon(\mathbf{x})$ satisfies the following inequality:

$$\varepsilon(\mathbf{x}) \leq \frac{\delta}{2(1+\theta)\|\mathbf{u}^*(\mathbf{x})\|_2 \|\mathbf{f} - \mathbf{L}\mathbf{x}\|_2}.$$

Therefore, we conclude that the following system of linear equation:

$$\begin{cases} \tilde{\mathbf{P}}\mathbf{F}^{'}\tilde{\mathbf{y}} = \tilde{\mathbf{P}}\tilde{\mathbf{b}}, \\ \tilde{\mathbf{y}} \geq \mathbf{0}, \end{cases}$$

is infeasible with high probability by using Farka's Lemma (Bertsimas and Tsitsiklis 1997).

Note that, by the definition of $\mathbf{F}^{'}$ and $\tilde{\mathbf{y}}$, we obtain that:

$$\tilde{\mathbf{P}}\mathbf{F}^{'}\tilde{\mathbf{y}} = \tilde{\mathbf{P}}\big(\tilde{\mathbf{F}}\tilde{\mathbf{y}} - \frac{1}{\theta'}\sum_{j=1}^{m} y_j \tilde{\mathbf{b}}\big),$$

so that $\tilde{\mathbf{b}} \notin \phi^J_{\tilde{\mathbf{b}},\theta'}(\mathcal{K})$, where $\mathcal{K} := \text{cone}(\tilde{\mathbf{P}}\tilde{\mathbf{F}}_1, \ldots, \tilde{\mathbf{P}}\tilde{\mathbf{F}}_{m+1})$. Accordingly, by using Corollary 6, we have that the system of linear equations given by:

$$\begin{cases} \tilde{\mathbf{P}}\tilde{\mathbf{F}}\tilde{\mathbf{y}} = \tilde{\mathbf{P}}\tilde{\mathbf{b}}, \\ \mathbf{1}_m^\top \tilde{\mathbf{y}} \leq \theta^{'}, \\ \tilde{\mathbf{y}} \geq \mathbf{0}, \end{cases}$$

is also infeasible with high probability since $\tilde{\mathbf{b}} \notin \mathcal{K}_\theta$.

Therefore, with probability at least p, the projected follower's program:

$$\varphi_\theta(\mathbf{x}, \mathbf{P}) := \min\left\{\mathbf{c}^\top \mathbf{y} \,:\, \mathbf{P}\mathbf{F}\mathbf{y} = \mathbf{P}\big(\mathbf{f} - \mathbf{L}\mathbf{x}\big)\,,\, \mathbf{1}_m^\top \mathbf{y} \leq \theta^{'}\,,\, \mathbf{y} \in \mathbb{R}^m_+\right\}$$

has its optimal objective function value greater than $\varphi_\theta(\mathbf{x}) - \delta$. Since $\theta^{'} > \theta$, with probability at least p, we have that $\varphi_\theta(\mathbf{x}, \mathbf{P}) \geq \varphi_\theta(\mathbf{x}) - \delta$. This observation, in turn, concludes the proof. ■

References

Bertsimas, D. and Tsitsiklis, J. N. (1997). *Introduction to linear optimization*. Athena Scientific Belmont, MA.

Johnson, W. B. and Lindenstrauss, J. (1984). "Extensions of Lipschitz mappings into a Hilbert space". In: *Contemporary Mathematics* 26, p. 1.

Vu, K., Poirion, P.-L., and Liberti, L. (2018). "Random projections for linear programming". In: *Mathematics of Operations Research* 43.4, pp. 1051–1071.