Reduced Cost Influence Functions for Predict-then-Optimize under Noisy Data

William Zhang
Operations Research Center
Massachusetts Institute of Technology
Cambridge, MA 02139
wzhang64@mit.edu

Saurabh Amin
Operations Research Center
Massachusetts Institute of Technology
Cambridge, MA 02139
amins@mit.edu

Georgia Perakis
Operations Research Center
Massachusetts Institute of Technology
Cambridge, MA 02139
georgiap@mit.edu

Abstract

Machine learning models deployed in decision-making pipelines, from vehicle routing to dynamic pricing to clinical decision support, must remain reliable in the face of low-quality training data. We study a predict-then-optimize setting in which an upstream model predicts the cost vector of a downstream linear program, with a large pool of noisy training data supplemented by a small curated set of clean data. We introduce Reduced Cost Influence Functions (RCIFs), which combine influence functions from robust statistics with the geometry of linear optimization to estimate how each training point shifts downstream decisions. Building on this, we propose a reweighting algorithm that uses RCIFs to downweight harmful training points and reduce downstream regret. Experiments on random linear programs and shortest-path problems show that our approach reduces regret, and accompanying theory characterizes properties of RCIFs and identifies regimes favorable to our method relative to baselines.

1 Introduction

Machine learning models are increasingly used as components of larger decision-making pipelines: pricing engines, routing systems, clinical decision support tools. In these pipelines, predictions are not the end of the line: they are passed to a downstream optimization procedure whose solution carries the real consequences. This shift makes the training of predictive models a more delicate problem, since small errors under standard prediction loss can produce suboptimal downstream decisions. Compounding this, machine learning models are well known to be vulnerable to low-quality training data [Biggio and Roli, 2018], and a decision pipeline that is brittle to training-time corruption can produce costly errors downstream. Thus, how can we improve robustness of the downstream decision against corrupted training data via the upstream training process?

We propose a downstream-aware reweighting of the training data that leaves both the predictor architecture and the downstream optimization method unchanged. Our method assumes access to a small clean validation set, a standard assumption in robust learning literature [Ren et al., 2018, Hendrycks et al., 2018], in addition to a larger, potentially corrupted training set. We introduce the core technical object, the *reduced cost influence function* (RCIF), a decision-aware signal that propagates the effect of upweighting each training point through the upstream predictor onto the reduced costs

40th Conference on Neural Information Processing Systems (NeurIPS 2026).

of the downstream LP, directly connecting each training point to the outcome decision. The RCIF measures how each training point destabilizes or stabilizes the downstream decision boundary; we use it to reweight the training data using the validation set, so that the refit predictor produces decisions that are robust to training corruption.

Our method is best suited to predict-then-optimize problems in which (i) prediction errors frequently cause decision flips, (ii) training data is corrupted, and (iii) a small clean validation set exists but is too small to fit the predictor on its own. We demonstrate gains on random linear programs and on shortest-path routing, a standard benchmark for decision-making pipelines [Elmachtoub and Grigas, 2022], and provide theoretical guarantees identifying validation-size regimes favorable to RCIF reweighting relative to both the standard, unweighted estimator and validation-only estimator.

2 Related Work

Decision-aware learning. A large body of work studies training predictive models for downstream tasks. End-to-end approaches train the upstream predictor with respect to downstream decision quality, most notably via surrogate losses such as SPO+ [Elmachtoub and Grigas, 2022] and the broader decision-aware learning literature [Hu et al., 2022, Shah et al., 2022, Qi et al., 2025, Donti et al., 2017, Bertsimas and Kallus, 2020]. Our approach is complementary: rather than modifying the training loss, we reweight the training data using a downstream-aware signal, leaving both the predictor and the LP unchanged. This is similar to Lawless and Zhou [2022], which directly reweights data via regret but does not consider corrupted data.

Robustness in decision-aware learning. A parallel line of work addresses robustness in decision-aware learning. Work such as Sun et al. [2023], Yeh et al. [2024], Patel et al. [2024] consider robustness via sharper uncertainty sets using conformal prediction, while Schutte et al. [2023], Im et al. [2025] propose surrogate losses for robust decision-focused learning. Instead, we directly reweight the training data, which preserves the original pipeline and focuses on robustness to the training phase rather than at deployment via robust downstreams.

Reweighting under noisy data. Mitigating label corruption given access to a small clean set (and reweighting data) is prevalent in robust learning [Hendrycks et al., 2018, Zhang et al., 2020, Liu et al., 2021]. Other work considers corruption detection such as [Northcutt et al., 2021, Diakonikolas et al., 2019, Sheikholeslami et al., 2021]. Notably, Ren et al. [2018] derive weights from influence functions (popularized in [Koh and Liang, 2017, Broderick et al., 2020]) for robustness. We adopt the influence-function machinery but propagate effects onto reduced costs of the downstream LP rather than onto prediction loss, which differentiates our approach from a direct adaptation of validation-loss influence.

3 Setting

We consider a predict-then-optimize pipeline in which an upstream predictive model produces an estimate of an unknown cost vector, which is then passed to a downstream linear program whose solution is the final decision.

Downstream optimization. The downstream component solves a linear program of the form

$$x^*(c) = \arg\min_{x \in \mathcal{X}} c^\top x,$$

where $c \in \mathbb{R}^d$ is the cost vector and $\mathcal{X} = \{x \in \mathbb{R}^d : Ax = b,\, x \geq 0\}$ is a closed, bounded polyhedron with diameter L defined by a constraint matrix $A \in \mathbb{R}^{a \times d}$ and a right-hand side $b \in \mathbb{R}^a$. We assume throughout that the LP is nondegenerate, so that each optimal vertex corresponds to a unique basis.

Upstream predictor and data. The upstream component is a regression model $f_\theta : \mathcal{Z} \to \mathbb{R}^d$ mapping features $z \in \mathcal{Z} \subseteq \mathbb{R}^q$ to predicted cost vectors, parameterized by $\theta \in \Theta$ with true parameter $\theta^\star \in \Theta$. We have access to two datasets: training data $\mathcal{S} = \{(z_i, \tilde{c}_i)\}_{i=1}^n$ with possibly corrupted labels, and clean validation data $\mathcal{V} = \{(z_v, c_v)\}_{v=1}^m$ with uncorrupted labels, where typically $m \ll n$.

The label model is

$$c_i \;=\; f_{\theta^\star}(z_i) + \varepsilon_i,$$

with ε_i zero-mean σ_ε-subgaussian and mutually independent. A subset $\mathcal{C} \subseteq \{1, \ldots, n\}$ of size $|\mathcal{C}| = \beta n$, $\beta \in (0, 1/2)$, indexes the corrupted training points:

$$\tilde{c}_i \;=\; c_i + \eta_i \quad (i \in \mathcal{C}), \qquad \tilde{c}_i \;=\; c_i \quad (i \notin \mathcal{C}),$$

where η_i is an additional perturbation. The validation set $\mathcal{V}$ contains the clean costs c_j by assumption. Throughout the paper, $\tilde{c}_i$ denotes the *observed* training label at index i, which equals the clean label c_i when $i \notin \mathcal{C}$. Equivalently, $\eta_i = 0$ for $i \notin \mathcal{C}$, so the decomposition $\tilde{c}_i = c_i + \eta_i$ holds for all i.

Fitting algorithm. Given a loss function $\ell : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, the standard empirical risk minimizer (ERM) on the training data is

$$\hat{\theta} = \arg\min_{\theta \in \Theta} \sum_{i=1}^{n} \ell(f_\theta(z_i), \tilde{c}_i).$$

For a given set of weights $p_1, \ldots, p_n$, one can instead obtain the weighted ERM

$$\hat{\theta}_p = \arg\min_{\theta \in \Theta} \sum_{i=1}^{n} p_i \, \ell(f_\theta(z_i), \tilde{c}_i),$$

with larger weights placing greater emphasis on the i-th training point. We assume throughout that the empirical risk is strictly convex in θ and twice differentiable [Koh and Liang, 2017], so that influence-function approximations are well defined.

4 Preliminaries

Bases, reduced costs, and optimality cones. A *basis* $B \subseteq \{1, \ldots, d\}$ is a set of indices such that the submatrix A_B is invertible and $A_B^{-1} b \geq 0$ is feasible, and $N(B) = \{1, \ldots, d\} \setminus B$ denotes the corresponding nonbasic set. For each nonbasic index $k \in N(B)$, the *edge direction* is a d-dimensional vector

$$\delta(B, k) = \begin{bmatrix} -A_B^{-1} A_k \\ e_k \end{bmatrix},$$

where e_k is the elementary unit vector in nonbasic coordinates corresponding to nonbasic index k. The edge direction represents the direction along which the nonbasic variable x_k can increase from zero while maintaining feasibility [Bennouna et al., 2026, Ye et al., 2026]. The *reduced cost* of nonbasic variable k under cost vector c and basis B is

$$s_k(c, B) = c^\top \delta(B, k),$$

and B is optimal for c if and only if $s_k(c, B) \geq 0$ for all $k \in N(B)$. The set of cost vectors for which B is optimal forms the *optimality cone*

$$\Lambda(B) = \{c \in \mathbb{R}^d : s_k(c, B) \geq 0 \text{ for all } k \in N(B)\},$$

and the cones $\{\Lambda(B)\}_B$ cover $\mathbb{R}^d$ with overlaps only on measure-zero boundaries. Two cost vectors in the same cone induce the same optimal decision.

Regret. The quality of a predicted cost vector $\hat{c}$ relative to the true cost c is measured by the downstream *regret*:

$$\mathrm{Regret}(\hat{c}, c) = c^\top x^*(\hat{c}) - c^\top x^*(c),$$

the excess true cost incurred by acting optimally under $\hat{c}$ rather than c. Regret is zero whenever $\hat{c}$ and c lie in the same optimality cone, regardless of error $\|\hat{c} - c\|$.

Influence functions. Influence functions [Koh and Liang, 2017] provide a first-order approximation to the effect of perturbing the training distribution on the learned parameters. Consider the estimator obtained by upweighting training point $w_j = (z_j, \tilde{c}_j)$ by ϵ:

$$\hat{\theta}_{\epsilon,w_j} = \arg\min_{\theta} \frac{1}{n} \sum_{i=1}^{n} \ell(f_\theta(z_i), \tilde{c}_i) + \epsilon\, \ell(f_\theta(z_j), \tilde{c}_j).$$

Under our risk assumptions, the derivative of $\hat{\theta}_{\epsilon,w_j}$ with respect to ϵ at $\epsilon = 0$ admits the closed form

$$\mathcal{I}_{\hat{\theta}}(w_j) := \left. \frac{d\hat{\theta}_{\epsilon,w_j}}{d\epsilon} \right|_{\epsilon=0} = -H_{\hat{\theta}}^{-1} \nabla_\theta \ell(f_{\hat{\theta}}(z_j), \tilde{c}_j),$$

where $H_{\hat{\theta}} = \frac{1}{n} \sum_{i=1}^{n} \nabla_\theta^2 \ell(f_{\hat{\theta}}(z_i), \tilde{c}_i)$ is the empirical Hessian of the training loss at $\hat{\theta}$. We refer to $\mathcal{I}_{\hat{\theta}}(w_j)$ as the *influence* of training point w_j on $\hat{\theta}$.

5 Reduced cost influence function

Recall that downstream regret is zero whenever the predicted cost $\hat{c}$ and the true cost c lie in the same optimality cone. Thus the prediction need not be accurate in norm; it need only lie in the correct cone. Conversely, predictions near the boundary of $\Lambda(B)$ are *unstable*: small perturbations of $\hat{c}$ can push the prediction into an adjacent cone and induce a discontinuous jump in downstream cost.

The reduced cost $s_k(c, B)$ measures the slack in the k-th constraint defining $\Lambda(B)$. Geometrically, the distance from c to the k-th cone boundary equals $s_k(c, B)/\|\delta(B, k)\|$, and the overall margin is given by the minimum such distance over $k \in N(B)$ (this margin coincides with distance to degeneracy [El Balghiti et al., 2019, Liu et al., 2023] under nondegeneracy). Our goal is to quantify the influence of each training point on this margin, identify harmful points (particularly in $\mathcal{C}$), and downweight them. To this end we introduce reduced cost influence functions (RCIFs).

Fix a validation input $z_v \in \mathcal{V}$ and let $\hat{c}_v = f_{\hat{\theta}}(z_v)$ be the predicted cost at z_v. Let B be the optimal basis for $\hat{c}_v$, and let $k \in N(B)$ be a nonbasic index of interest. We wish to quantify how upweighting training point w_j changes the reduced cost $s_k(\hat{c}_v, B)$ at the validation point z_v.

Viewing s_k as a function of θ through the predictor, we have $s_k(f_\theta(z_v), B) = \delta(B, k)^\top f_\theta(z_v)$. Applying the chain rule and evaluating at $\theta = \hat{\theta}$ in the influence direction $\mathcal{I}_{\hat{\theta}}(w_j)$ yields the following definition.

Definition 1 (Reduced Cost Influence Function)**.** *Let $\hat{\theta}$ be the ERM hypothesis, let z_v be a validation input, let $\hat{c}_v = f_{\hat{\theta}}(z_v)$, and let B be the optimal basis for $\hat{c}_v$. For a nonbasic index $k \in N(B)$ and a training point w_j, the* reduced cost influence function *is*

$$\mathrm{RCIF}_{k,\hat{\theta}}(z_v, w_j) := \delta(B, k)^\top \nabla_\theta f_{\hat{\theta}}(z_v)\, \mathcal{I}_{\hat{\theta}}(w_j).$$

The RCIF admits a direct interpretation: it is the first-order approximation of the change in the reduced cost $s_k(\hat{c}_v, B)$ at validation point z_v when training point w_j is upweighted by a unit of ϵ. A positive RCIF means upweighting w_j increases the reduced cost in direction k, pushing $\hat{c}_v$ deeper into $\Lambda(B)$ along that facet. This signed, direction-aware signal is the basis for the reweighting schemes introduced in the next section.

5.1 RCIF score

We use the RCIF to construct aggregate scores for reweighting and refitting the predictor. We present two score variants: a natural two-sided score and a second, more refined directed score.

5.1.1 Two-sided weighting scheme

For each validation point $w_v \in \mathcal{V}$, let $\hat{c}_v = f_{\hat{\theta}}(z_v)$ be the predicted cost, B_v the optimal basis for $\hat{c}_v$, and define the most fragile nonbasic direction as

$$k_v^* = \arg\min_{k \in N(B_v)} \frac{s_k(\hat{c}_v, B_v)}{\|\delta(B_v, k)\|},$$

i.e., the nonbasic index with the smallest reduced cost margin under $\hat{c}_v$. This is the direction along which $\hat{c}_v$ is closest to the boundary of $\Lambda(B_v)$, and thus the direction most vulnerable to a basis flip.

We say validation point w_v is *correct* if B_v is also the optimal basis under the true cost c_v, i.e., if $\hat{c}_v$ and c_v lie in the same optimality cone, and *incorrect* otherwise. Let $\sigma_v = 2 \cdot \mathbf{1}_{\text{correct}}(w_v) - 1 \in \{-1, +1\}$ encode this status.

The two-sided aggregate score for training point w_j is

$$\Psi_j^{\text{TS}} = \frac{1}{m} \sum_{w_v \in \mathcal{V}} \sigma_v \cdot \text{RCIF}_{k_v^*, \hat{\theta}}(z_v, w_j).$$

The sign σ_v aligns two cases: at correct validation points we want to push $\hat{c}_v$ deeper into $\Lambda(B_v)$, and at incorrect points, we wish to destabilize the current basis so that $\hat{c}_v$ may exit the wrong cone.

5.1.2 Directed weighting scheme

The two-sided score destabilizes incorrect validation points along the direction of smallest current margin, but this direction is not necessarily the one that would move $\hat{c}_v$ toward the correct cone. We have access to the true cost c_v and can identify the directions along which the true c_v indicates a basis change is warranted: specifically, those nonbasic indices $k \in N(B_v)$ for which $s_k(c_v, B_v) < 0$. These are the directions in which the true cost violates the optimality condition of the current basis B_v, and hence the directions whose nonbasic variables should enter the basis under c_v.

We refine the choice of k_v^* accordingly. For correct validation points, we keep the most fragile direction as before. For incorrect validation points, we restrict the argmin to the set of true-cost-violating directions:

$$k_v^* = \begin{cases} \displaystyle\arg\min_{k \in N(B_v)} \frac{s_k(\hat{c}_v, B_v)}{\|\delta(B_v, k)\|}, & \text{if } v \text{ is correct,} \\ \displaystyle\arg\min_{k \in N(B_v) \,:\, s_k(c_v, B_v) < 0} \frac{s_k(\hat{c}_v, B_v)}{\|\delta(B_v, k)\|}, & \text{if } v \text{ is incorrect.} \end{cases}$$

The directed aggregate score uses the same outer formula as the two-sided score but with this refined choice of k_v^*:

$$\Psi_j^{\text{D}} = \frac{1}{m} \sum_{w_v \in \mathcal{V}} \sigma_v \cdot \text{RCIF}_{k_v^*, \hat{\theta}}(z_v, w_j).$$

The sign convention is unchanged. The refinement only changes *which* direction is destabilized. These scores can be used as additive weight adjustments directly, but an exponential refinement can improve performance in practice.

5.2 Exponential reweighting

Given an aggregate score Ψ_j (either Ψ_j^{TS} or Ψ_j^{D}) for each training point, we convert scores to training weights via an exponential family with temperature parameter $\lambda > 0$:

$$p_j = \frac{\exp(\lambda \Psi_j)}{\sum_{i=1}^{n} \exp(\lambda \Psi_i)}.$$

We then refit the predictor by solving the weighted ERM problem with weights $\{p_j\}$. The temperature hyperparameter λ controls the aggressiveness of the reweighting: $\lambda = 0$ recovers uniform weighting, while large λ concentrates mass on the highest-scoring training points. For experiments, we set $\lambda = 0.05$.

5.3 Iterative reweighting

A single round of RCIF reweighting evaluates the influence of each training point at the original ERM estimator $\hat{\theta}$. After the refit, however, the predictor has changed, and so have the optimal bases at validation points and the most fragile (or directed) nonbasic directions. We can sharpen the signal by recomputing the scores at the refitted estimator and accumulating them across iterations.

For τ iterations, let $\Psi_j^{(t)}$ denote the aggregate score of training point w_j at the t-th iteration, computed using $\hat{\theta}^{(t)}$, the predictor fitted at the start of iteration t (with $\hat{\theta}^{(1)}$ as unweighted ERM estimator). The iterative weight for w_j at iteration τ is

$$p_j^{(\tau)} = \frac{\exp\left(\lambda \sum_{t=1}^{\tau} \Psi_j^{(t)}\right)}{\sum_{i=1}^{n} \exp\left(\lambda \sum_{t=1}^{\tau} \Psi_i^{(t)}\right)}.$$

Each iteration refits the predictor with weights $p_j^{(\tau)}$. We treat τ as a hyperparameter; in our experiments we set $\tau = 5$.

6 Theoretical results

We derive two properties of RCIF reweighting after one iteration of exponential weights: a signed score decrease for corrupted points, and validation-size regimes in which RCIF reweighting admits tighter regret upper bounds than the unweighted ERM and validation-only ERM baselines.

6.1 Setting and assumptions

We consider the same predict-then-optimize setting as defined in Section 3, but with a linear upstream model such that for parameter $\theta \in \mathbb{R}^{d\times q}$, $z \in \mathcal{R}^q$, $f_\theta(z) = \theta z$, and the loss function is square loss $\ell(c, \hat{c}) = \|c - \hat{c}\|^2$. Let $\hat{\theta}$ denote standard ERM on training set $\mathcal{S}$, $\hat{\theta}_p$ denote the one-step RCIF-reweighted ERM, and $\hat{\theta}_{\text{val}}$ denote the validation-only ERM fit on $\mathcal{V}$. Unless stated otherwise, throughout this section, we make the following assumptions.

(D1) Bounded Features. Training and validation points (z, c) are drawn iid from distribution $\mathcal{D}$, with $\|z\| \le R$. Also let empirical design covariance $\hat{\Sigma} := \frac{1}{n}\sum_{i=1}^{n} z_i z_i^\top \in \mathbb{R}^{q\times q}$ be invertible and satisfy $\lambda_{\min}^{\text{eig}}(\hat{\Sigma}) \ge \mu > 0$.

(D2) Labels and noise. For both training and validation, for any i, for true parameter θ^*, $c_i = \theta^\star z_i + \varepsilon_i$ with each entry of ε_i 0-mean σ_ε-sub-Gaussian, independent of each other and features.

(D3) Corruption set. A set $\mathcal{C} \subset \{1, \dots, n\}$ of size $|\mathcal{C}| = \beta n$ is drawn uniformly at random, independently of features and noise, with $\beta \in (0, 1/2)$.

(D4) Bounded directional corruption. For $j \in \mathcal{C}$, the corruption is of the form $\tilde{c}_j = c_j + \eta_j$ where $\eta_j = -\alpha\,\hat{\delta}_j$, and where $\hat{\delta}_j := \delta(B_j^\star, k_j)/\|\delta(B_j^\star, k_j)\|$ is the unit-norm direction at the optimal basis $B_j^\star$ at c_j for non-basic direction $k_j \in N(B_j^\star)$ chosen independently of validation sample or learned estimator.

(D5) Bounded LP geometry. $\|\delta(B, k)\| \le D$ for every basis B and nonbasic k.

6.2 RCIF score decomposition

We present a factorization of the RCIF, which leads to a score decomposition that isolates a "clean" RCIF score and a "corrupted" RCIF score from corruption η_j.

Lemma 1 (RCIF Score Factorization)**.** *We can factor the RCIF of training point w_j on a single validation input z_v on non-basic direction k into two different terms, a downstream residual interaction term, and an input alignment term. Then we have that*

$$\text{RCIF}_{k,\hat{\theta}}(z_v, w_j) = -\underbrace{\delta(B_v, k)^\top r_j}_{\textit{RESIDUAL INTERACTION}} \cdot \underbrace{\kappa(z_v, z_j)}_{\textit{INPUT ALIGNMENT}}$$

where $\kappa(z_v, z_j)$ denotes Mahalanobis inner product $z_v^T \hat{\Sigma}^{-1} z_j$ and r_j denotes the residual of training point j, $r_j \coloneqq \hat{\theta} z_j - \tilde{c}_j$.

Now we may show that the RCIF score can be decomposed into a clean and corrupted term.

Lemma 2 (RCIF Score Decomposition). *For fixed $\hat{\theta}$ and training point $w_j = (z_j, \tilde{c}_j) = (z_j, c_j + \eta_j)$, we can decompose the RCIF score Ψ_j into a clean score and noise term: $\Psi_j = \Psi_j^{\text{clean}} + u_j^\top \eta_j$ where $u_j = \frac{1}{m} \sum_{w_v \in V} \sigma_v \, \kappa(z_j, z_v) \delta(B_v, k_v^*)$, where k^* is the appropriate direction depending on the weighting scheme, and Ψ_j^{clean} is the score if the score were calculated using $w_j = (z_j, c_j)$.*

These lemmas hold without Assumptions (D1)-(D5) (handling any η_j, not just the one in D4). Via this decomposition, we can obtain results that show how scoring on corrupted points differs between clean and corrupted points, with differences in magnitude under zero-mean corruption (see Appendix A.3) as well as signed score decrease for directional corruption as in D4.

6.3 Decreased corruption score

We show that under certain settings, our scoring method will give a lower score to corrupted points. First we define an alignment parameter.

Definition 2 (Validation alignment). *For a fresh sample (z_j, c_j), and fixed $\hat{\theta}$, conditional on $(\hat{\theta}, z_j, c_j)$, define the* alignment parameter

$$\gamma(\hat{\theta}, z_j, c_j) \;:=\; \mathbb{E}_{w_v \sim \mathcal{D}} \Big[\sigma_v \kappa(z_j, z_v) \delta(B_v, k_v^*)^\top \hat{\delta}_j \;\Big|\; \hat{\theta}, z_j, c_j \Big] .$$

Theorem 1 (Signed score decrease under alignment; informal). *Adopt Assumptions (D1)–(D5). Let (z_j, c_j) be a fresh sample drawn from $\mathcal{D}$ independently of $\hat{\theta}$ and $\mathcal{V}$, and let $\tilde{c}_j = c_j + \eta_j$ be its directional corruption (D4). Suppose $\gamma := \gamma(\hat{\theta}, z_j, c_j) > 0$. Then, in expectation over $\mathcal{V}$ conditional on $(\hat{\theta}, z_j, c_j)$, the score on the corrupted point is shifted by $-\alpha\gamma$ relative to its clean counterpart, and with probability at least $1 - \exp(-m\gamma^2/(8M_\gamma^2))$ over the same randomness, for constant M_γ,*

$$\Psi_j(\tilde{c}_j) - \Psi_j(c_j) \;\le\; -\frac{\alpha\gamma}{2}.$$

The full statement, constants, and proof appear as Theorem 4 in Appendix A.4.

Remark 1. *The assumption that $\gamma > 0$ is a natural one. Note that $\sigma_v > 0$ holds if $\hat{\theta}$ is similar to θ^*, i.e. the upstream model makes predictions in the same cone. Furthermore, $\kappa(z_v, z_j) > 0$ is likely positive if the distribution is tight around $\bar{z} = \mathbb{E}[z] \neq 0$, with large n. A tight distribution also makes $\delta(B_v, k_v^*)^\top \hat{\delta}_j > 0$ likely, as it holds when a large proportion of the cost-vector validation distribution falls into a single optimality cone with the same fragile direction.*

Thus, if a freshly sampled point is corrupted via Assumption D4, w.h.p. its RCIF score is decreased compared to its clean counterpart. This requires the point to be out-of-sample, but we show an in-sample counterpart with an additional $O(1/n)$ slack in Appendix A.5.

6.4 Expected regret upper-bound comparison

Under further assumptions, we can obtain regimes on validation set size m in which the RCIF reweighted solution $\hat{\theta}_p$, has tighter upper-bounds on the expected regret than those of the standard ERM solution $\hat{\theta}$, and the validation-only solution $\hat{\theta}_{\text{val}}$. For intuition, we include two relevant assumptions below.

Assumption 1 (High probability alignment). *There exist constants $\gamma_0 > 0$ and $\delta_\gamma \in (0, 1/n)$ such that*

$$\Pr\big[\gamma(\hat{\theta}, z_j, c_j) \ge \gamma_0\big] \;\ge\; 1 - \delta_\gamma,$$

where the probability is over $(\hat{\theta}, z_j, c_j)$ with $\hat{\theta}$ fit on iid samples drawn from $\mathcal{D}$ (D1) and (z_j, c_j) independently drawn from $\mathcal{D}$.

The assumption says the conditional alignment is bounded away from zero with high probability. To obtain a result on regret, we require the following condition. Furthermore, let M be an RCIF score upper-bound $|\Psi_j| \le M$, which holds under (D1)-(D5) with bounded residuals.

Assumption 2 (Margin). *There exists a constant $\omega > 0$ such that*

$$\pi_v(\theta^\star z) \;\ge\; 1/\omega \quad \textit{almost surely under } w = (z, c) \sim \mathcal{D},$$

where $\pi_v(c) := \min_{k \in N(B^\star(c))} s_k(c, B^\star(c)) / \|\delta(B^\star(c), k)\|$ is the reduced-cost margin at c.

Theorem 2 (Validation-size regret regimes; informal)**.** *Adopt Assumptions (D1)–(D5), Assumption 1, and Assumption 2, and additional regularity assumptions listed in Appendix A.6. There exist thresholds $m_\star(\lambda), m^\star(\lambda) \in [1, \infty]$, depending on the problem constants (e.g., $\sigma_\varepsilon, \alpha, \beta, R, \mu, D, \gamma_0, \omega$), such that in the asymptotic regime $n \to \infty$:*

1. *(*Lower threshold $m_\star$*.) For all $m > m_\star(\lambda)$, the RCIF reweighted estimator $\hat{\theta}_p$ admits a tighter expected regret upper bound than the unweighted ERM $\hat{\theta}$. The threshold is finite whenever the contraction factor $\Phi(\lambda, m, n)$ from Lemma 6 eventually drops below $\beta \cdot A$ (for some constant A) in m.*

2. *(*Upper threshold $m^\star$*.) For all $m < m^\star(\lambda)$, $\hat{\theta}_p$ admits a tighter expected regret upper bound than the validation-only ERM $\hat{\theta}_{\mathrm{val}}$. The threshold is finite, since the irreducible corruption bias of $\hat{\theta}_p$ eventually exceeds val-only's $1/m$ variance.*

3. *(*Sweet spot.*) When $m_\star(\lambda) < m^\star(\lambda)$, every m in this interval admits a tighter expected regret upper bound for $\hat{\theta}_p$ than for both baselines simultaneously.*

The thresholds and the sweet-spot condition are characterized explicitly in terms of $\Phi(\lambda, m, n)$ in Theorem 7 (Appendix A.6), where the full statement, constants, and proof appear.

Remark 2. *Theorem 2 compares upper bounds on expected regret, not the regrets themselves. However, the thresholds indicate a transition region in validation size: small m favors training-data methods (unweighted ERM), large m favors validation-only ERM, and RCIF reduces the bias enough to compete with both in the intermediate regime as observed in Section 7.*

7 Experimental results

We evaluate our reweighting scheme across several settings, measuring out-of-sample regret on a clean test set when the training data is corrupted but a small clean validation set is available. To simulate label corruption, we uniformly sample a fraction β (default 0.3) of the training points and add isotropic Gaussian noise to their cost labels; the validation and test sets remain clean.

Methods. We compare ten methods. **Iterative Exp. Directed** (ours): iterative exponential reweighting with directed score Ψ^{D}. **Exp. Two-Sided** (ours): single-round exponential reweighting with Ψ^{TS}. **Directed** (ours): additive reweighting $p_i = 1/n + \phi\, \Psi_i^{\mathrm{D}}$ with $\phi = 0.05$. **Unweighted**: ERM on the corrupted training set. **RC Weight**: single-round exponential reweighting using the raw reduced-cost margin $\Psi_j^{\mathrm{RC}} = \min_{k \in N(B_j)} s_k(\hat{c}_j, B_j)$ as score. **Validation Only**: ERM on the clean validation set alone. **SPO Influence Weight**: reweighting via influence functions on the SPO+ surrogate loss rather than reduced costs; the closest analog to our method, isolating the contribution of reduced costs over a surrogate loss (Appendix B.7.4). **SPO+ Validation Only**: an end-to-end predictor trained by subgradient descent on the SPO+ loss on the validation set alone, initialized at the validation-only least-squares fit. **SPO+ Finetune**: unweighted ERM on the noisy train-plus-validation pool, fine-tuned by SPO+ subgradient descent on the validation set. **Clean ERM (oracle)**: ERM on the full uncorrupted dataset, a reference lower bound on achievable regret. For all methods except the validation-only ones, we fold the clean validation set into the second-stage fit, assigning each validation point the maximum weight given to any training point, so every method using training data sees the same total data. Additional experiments and details appear in Appendix B.

7.1 Random linear program

We sample a random constraint matrix A and right-hand side b defining the downstream LP, and a random linear map θ^* from features to costs. Training, validation, and test inputs are drawn i.i.d., with cost labels $c = \theta^* z + \varepsilon$. The upstream predictor is a linear model fit by squared loss. Table 1 reports results with upstream dimension 20, LP dimension 10, 20 constraints, 200 training points, 50 validation points, 150 test points, and corruption rate $\beta = 0.3$. Iterative Exp. Directed achieves the lowest regret, the gap to unweighted is modest in this setting but significant under a two-sided paired Wilcoxon signed-rank test over seeds (p = 0.007).

Method	Random LP	Shortest Path	NN Upstream
Iterative Exp. Directed	**0.3451 ± 0.0133**	**0.9437 ± 0.0172**	1.4511 ± 0.0187
Exp. Two-Sided	0.3535 ± 0.0129	1.0009 ± 0.0175	1.4211 ± 0.0196
Directed	0.3531 ± 0.0129	0.9825 ± 0.0171	**1.3958 ± 0.0217**
Unweighted	0.3564 ± 0.0130	1.0029 ± 0.0169	1.4317 ± 0.0198
RC Weight	0.3549 ± 0.0131	1.0035 ± 0.0169	1.4255 ± 0.0227
Validation Only	0.4184 ± 0.0191	0.9989 ± 0.0181	1.4208 ± 0.0228
SPO Influence Weight	0.3509 ± 0.0124	0.9560 ± 0.0158	1.4626 ± 0.0183
Clean ERM (oracle)	0.2785 ± 0.0109	0.8513 ± 0.0171	1.3452 ± 0.0190
SPO+ Val Only	0.4854 ± 0.0216	1.0914 ± 0.0182	–
SPO+ Finetune	0.3666 ± 0.0140	1.0042 ± 0.0176	–

Table 1: Mean test regret over 100 seeds under random LP setting, shortest path setting with a linear upstream, and shortest path setting with a neural-network upstream. Best per column in bold.

7.2 Shortest path setting

For a more structured downstream LP, we consider the shortest-path problem on a grid network, following Elmachtoub and Grigas [2022]: traverse from the northwest to the southeast corner of the grid. We use the PyEPO library [Tang and Khalil, 2022] to generate the dataset with an upstream linear model fit via squared loss. Table 1 reports results on a 4×4 grid with upstream input dimension 30, 400 training points, 100 validation points, 200 test points, and $\beta = 0.2$. On this more structured problem, RCIF reweighting yields substantially larger gains: Iterative Exp. Directed again performs best, by a much wider margin than in the random LP setting (paired $p < 0.048$ for all methods). Furthermore, though the raw magnitude difference between methods is small, the recovery relative to Clean ERM is large: our method recovers 39% of the error induced by noise (0.152 between Clean ERM and Unweighted) whereas the next best method SPO Influence Weight only recovers 30%.

7.3 Neural network upstream

To demonstrate that the method retains value when the upstream predictor is nonlinear, we replace the linear model with a two-layer fully-connected network with hidden dimension 8 and Tanh activation as seen in the last column of Table 1. We use the shortest path setting again, with a 5×4 grid, upstream input dimension 30, 200 training points, 100 validation points, 200 test points, and $\beta = 0.3$. We compare methods that leave the training procedure unchanged (reweighting and the clean-label reference); the SPO+ Finetune and SPO+ Validation Only baselines are defined for the linear upstream and require a specialized pipeline, so we omit them. Here the additive Directed variant performs best. We conjecture that nonlinearity makes the iterations of Iterative Exp. Directed harmful, since influence-function approximation error compounds across iterations, consistent with the known fragility of influence functions in nonlinear settings [Basu et al., 2020]. Correspondingly, the non-iterative Exp. Two-Sided performs relatively better here than in the linear settings, and the Directed variant yields strong improvements through more targeted scoring without added error.

7.4 Validation size study

Now we connect back to our theoretical result in Theorem 2, which provides conditions on the validation set size for which our method has an improved regret upper bound by providing a study on validation set size on shortest path problems (4×4 grid, upstream input dimension 20, 500 training points, 200 test points, corruption rate $\beta = 0.3$, 100 seeds with validation set varied over $\{5, 10, 20, 50, 100, 150\}$) and observing that it translates to real regret improvements. As the size of the validation set increases, at first, the unweighted method $\hat{\theta}$ outperforms our method, but then for sufficient validation data, our method $\hat{\theta}_{\mathrm{p}}$ wins, but then finally at large validation, $\hat{\theta}_{\mathrm{val}}$ wins. We can visually see this sweet spot where our method wins in Figure 1 on moderate validation set sizes, demonstrating its value in these regimes. See Appendix B.2 for details.

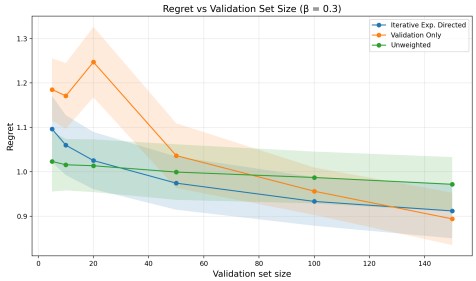


Figure 1: Test regret vs. validation size for shortest path on 4×4 grid (setup in text).

Method	deg = 1	deg = 2	deg = 4
Iterative Exp. Directed	**0.9437 ± 0.0172**	**0.4856 ± 0.0135**	**0.2899 ± 0.0115**
Exp. Two-Sided	1.0009 ± 0.0175	0.5682 ± 0.0152	0.3393 ± 0.0127
Directed	0.9825 ± 0.0171	0.5347 ± 0.0132	0.3208 ± 0.0122
Unweighted	1.0029 ± 0.0169	0.5715 ± 0.0145	0.3483 ± 0.0135
RC Weight	1.0035 ± 0.0169	0.5713 ± 0.0145	0.3496 ± 0.0138
Validation Only	0.9989 ± 0.0181	0.5284 ± 0.0146	0.3782 ± 0.0147
SPO Influence Weight	0.9560 ± 0.0158	0.4999 ± 0.0119	0.2975 ± 0.0117
Clean ERM (oracle)	0.8513 ± 0.0171	0.3972 ± 0.0118	0.2340 ± 0.0113
SPO+ Val Only	1.0914 ± 0.0182	0.6857 ± 0.0165	0.4410 ± 0.0160
SPO+ Finetune	1.0042 ± 0.0176	0.5406 ± 0.0151	0.3200 ± 0.0123

Table 2: Mean test regret, 50 seeds, shortest path across polynomial degree for feature-cost mapping.

7.5 Misspecification

Similarly to the neural network setting, we test what happens when the upstream model is misspecified. We keep the shortest-path setting (4×4 grid, input dimension 30, $400/100/200$ train/validation/test points, $\beta = 0.2$, 50 seeds) and raise the polynomial degree of PyEPO's feature-to-cost map from 1 to 2 and 4, so that the true costs are increasingly nonlinear in the features while the predictor remains linear and is fit by squared loss. We report results in Table 2. Raw regret values are not comparable across degrees, since PyEPO rescales the costs with the degree, so each column should be read on its own (degree 1 coincides with Table 1). The gains of RCIF reweighting are preserved and, relative to the corruption damage, grow proportionally with degree: Iterative Exp. Directed attains a test regret of 0.4856 ± 0.0135 at degree 2 and 0.2899 ± 0.0115 at degree 4, against 0.5715 ± 0.0145 and 0.3483 ± 0.0135 for unweighted ERM, and it has lower regret than every other competing method. RCIF scores only require the LP geometry at the predicted cost and the first-order response of the predictor, neither of which directly depends on the model being correctly specified, so the mechanism continues to identify harmful training points when the linear fit is only an approximation.

8 Conclusion

We introduced reduced cost influence functions, a downstream-aware signal that propagates training-point influence through the upstream predictor onto the reduced costs of the LP. Building on a score decomposition that isolates clean from corruption-induced contributions, we showed that RCIF scoring decreases the score of directionally corrupted points with high probability, and that there exist validation-size regimes in which RCIF reweighting admits tighter expected-regret upper bounds than unweighted and validation-only ERM. Experiments on random LPs, shortest-path routing, and a neural-network upstream are consistent with these regimes empirically and show consistent gains, particularly for shortest path. The neural-network setting also reveals a limitation: iterating exponential reweighting compounds the approximation error when the upstream is nonlinear, but the non-iterative variant still maintains strong performance. Furthermore, the downstream problem is currently limited to LPs, with assumed nondegeneracy for theoretical results. Promising directions include relaxing the nondegeneracy and connecting with the distance to degeneracy concept, extending the framework to MILPs and convex programs, and combining RCIF scoring with robust regression and active learning techniques for data selection.

References

Samyadeep Basu, Philip Pope, and Soheil Feizi. Influence functions in deep learning are fragile. *arXiv preprint arXiv:2006.14651*, 2020.

Omar Bennouna, Amine Bennouna, Saurabh Amin, and Asuman Ozdaglar. What data enables optimal decisions? an exact characterization for linear optimization. *Advances in Neural Information Processing Systems*, 38:166020–166049, 2026.

Dimitris Bertsimas and Nathan Kallus. From predictive to prescriptive analytics. *Management Science*, 66(3):1025–1044, 2020.

Battista Biggio and Fabio Roli. Wild patterns: Ten years after the rise of adversarial machine learning. In *Proceedings of the 2018 ACM SIGSAC conference on computer and communications security*, pages 2154–2156, 2018.

Tamara Broderick, Ryan Giordano, and Rachael Meager. An automatic finite-sample robustness metric: when can dropping a little data make a big difference? *arXiv preprint arXiv:2011.14999*, 2020.

Ilias Diakonikolas, Gautam Kamath, Daniel Kane, Jerry Li, Jacob Steinhardt, and Alistair Stewart. Sever: A robust meta-algorithm for stochastic optimization. In *International Conference on Machine Learning*, pages 1596–1606. PMLR, 2019.

Priya Donti, Brandon Amos, and J Zico Kolter. Task-based end-to-end model learning in stochastic optimization. *Advances in neural information processing systems*, 30, 2017.

Othman El Balghiti, Adam N Elmachtoub, Paul Grigas, and Ambuj Tewari. Generalization bounds in the predict-then-optimize framework. *Advances in neural information processing systems*, 32, 2019.

Adam N Elmachtoub and Paul Grigas. Smart "predict, then optimize". *Management Science*, 68(1): 9–26, 2022.

Dan Hendrycks, Mantas Mazeika, Duncan Wilson, and Kevin Gimpel. Using trusted data to train deep networks on labels corrupted by severe noise. *Advances in neural information processing systems*, 31, 2018.

Yichun Hu, Nathan Kallus, and Xiaojie Mao. Fast rates for contextual linear optimization. *Management Science*, 68(6):4236–4245, 2022.

Hyungki Im, Wyame Benslimane, and Paul Grigas. Smart surrogate losses for contextual stochastic linear optimization with robust constraints. *arXiv preprint arXiv:2505.22881*, 2025.

Pang Wei Koh and Percy Liang. Understanding black-box predictions via influence functions. In *International conference on machine learning*, pages 1885–1894. PMLR, 2017.

Connor Lawless and Angela Zhou. A note on task-aware loss via reweighing prediction loss by decision-regret. *arXiv preprint arXiv:2211.05116*, 2022.

Feng Liu, Bo Han, Tongliang Liu, Chen Gong, Gang Niu, Mingyuan Zhou, Masashi Sugiyama, et al. Probabilistic margins for instance reweighting in adversarial training. *Advances in neural information processing systems*, 34:23258–23269, 2021.

Mo Liu, Paul Grigas, Heyuan Liu, and Zuo-Jun Max Shen. Active learning in the predict-then-optimize framework: A margin-based approach. *arXiv preprint arXiv:2305.06584*, page 2, 2023.

Curtis Northcutt, Lu Jiang, and Isaac Chuang. Confident learning: Estimating uncertainty in dataset labels. *Journal of Artificial Intelligence Research*, 70:1373–1411, 2021.

Yash P Patel, Sahana Rayan, and Ambuj Tewari. Conformal contextual robust optimization. In *International Conference on Artificial Intelligence and Statistics*, pages 2485–2493. PMLR, 2024.

Meng Qi, Paul Grigas, and Zuo-Jun Shen. Integrated conditional estimation-optimization. *Operations Research*, 2025.

Mengye Ren, Wenyuan Zeng, Bin Yang, and Raquel Urtasun. Learning to reweight examples for robust deep learning. In *International conference on machine learning*, pages 4334–4343. PMLR, 2018.

Noah Schutte, Krzysztof Postek, and Neil Yorke-Smith. Robust losses for decision-focused learning. *arXiv preprint arXiv:2310.04328*, 2023.

Sanket Shah, Kai Wang, Bryan Wilder, Andrew Perrault, and Milind Tambe. Decision-focused learning without decision-making: Learning locally optimized decision losses. In S. Koyejo, S. Mohamed, A. Agarwal, D. Belgrave, K. Cho, and A. Oh, editors, *Advances in Neural Information Processing Systems*, volume 35, pages 1320–1332. Curran Associates, Inc., 2022.

Fatemeh Sheikholeslami, Ali Lotfi, and J Zico Kolter. Provably robust classification of adversarial examples with detection. In *International Conference on Learning Representations*, 2021.

Chunlin Sun, Linyu Liu, and Xiaocheng Li. Predict-then-calibrate: A new perspective of robust contextual lp. *Advances in neural information processing systems*, 36:17713–17741, 2023.

Bo Tang and Elias B Khalil. Pyepo: A pytorch-based end-to-end predict-then-optimize library for linear and integer programming. *arXiv preprint arXiv:2206.14234*, 2022.

Yuhan Ye, Saurabh Amin, and Asuman Ozdaglar. Learning decision-sufficient representations for linear optimization. *arXiv preprint arXiv:2603.18551*, 2026.

Christopher Yeh, Nicolas Christianson, Alan Wu, Adam Wierman, and Yisong Yue. End-to-end conformal calibration for optimization under uncertainty. *arXiv preprint arXiv:2409.20534*, 2024.

Jingfeng Zhang, Jianing Zhu, Gang Niu, Bo Han, Masashi Sugiyama, and Mohan Kankanhalli. Geometry-aware instance-reweighted adversarial training. *arXiv preprint arXiv:2010.01736*, 2020.

A Technical appendices and supplementary material

A.1 Proof of Lemma 1

RCIF score factorization

Proof. By definition of the influence function for linear models, the Hessian of the average training loss, viewed as an operator on $\mathbb{R}^{d\times q}$ (vectorized in row-major order), is

$$H_{\hat{\theta}} = \frac{1}{n}\sum_{i=1}^{n} \nabla_\theta^2 \|\theta z_i - \tilde{c}_i\|^2 = 2\, I_d \otimes \hat{\Sigma},$$

since the $(l_1, m_1; l_2, m_2)$ entry of the Hessian is $2\,\mathbf{1}[l_1 = l_2]\,(z_i)_{m_1}(z_i)_{m_2}$, and summing gives $2\,\mathbf{1}[l_1 = l_2]\,\hat{\Sigma}_{m_1 m_2}$. The inverse is $H_{\hat{\theta}}^{-1} = \frac{1}{2}\, I_d \otimes \hat{\Sigma}^{-1}$. Let vec denote row-major vectorization so $\mathrm{vec}(r_j z_j^\top) = r_j \otimes z_j$. The influence function is

$$\begin{aligned}
\mathrm{vec}\big(I_{\hat{\theta}}(w_j)\big) &= -H_{\hat{\theta}}^{-1}\,\mathrm{vec}\big(\nabla_\theta \ell\big) \\
&= -\tfrac{1}{2}(I_d \otimes \hat{\Sigma}^{-1}) \cdot 2\big[r_j \otimes z_j\big] \\
&= -(I_d\, r_j) \otimes (\hat{\Sigma}^{-1} z_j) \\
&= -r_j \otimes (\hat{\Sigma}^{-1} z_j),
\end{aligned}$$

where we used the Kronecker product rule $(A \otimes B)(u \otimes v) = Au \otimes Bv$. Reshaping back to $\mathbb{R}^{d\times q}$ gives

$$I_{\hat{\theta}}(w_j) = -r_j\,(\hat{\Sigma}^{-1} z_j)^\top.$$

Thus by definition of the RCIF, we have

$$\begin{aligned}
\mathrm{RCIF}_{k,\hat{\theta}}(z_v, w_j) &= \delta(B_v, k)^\top\, \nabla_\theta f_{\hat{\theta}}(z_v)\, \mathcal{I}_{\hat{\theta}}(w_j) \\
&= \delta(B_v, k)^\top\, \mathcal{I}_{\hat{\theta}}(w_j) z_v \\
&= -\delta(B_v, k)^\top\, r_j\,(\hat{\Sigma}^{-1} z_j)^\top z_v \\
&= -\delta(B_v, k)^\top\, r_j \kappa(z_v, z_j)
\end{aligned}$$

where the second equality follows by the directional derivative of a linear function, and the third equality follows by plugging in the derived influence function. □

A.2 Proof of Lemma 2

RCIF score decomposition

Proof. By Lemma 1, we can write the RCIF score on corrupted point w_j:

$$\Psi_j \;:=\; \frac{1}{m}\sum_{w_v \in V} \sigma_v\, \mathrm{RCIF}_{k_v^*,\hat{\theta}}\big(z_v, w_j\big) \;=\; -\frac{1}{m}\sum_{w_v \in V} \sigma_v\, \delta(B_v, k_v^*)^\top r_j\; \kappa(z_j, z_v).$$

Note that $r_j = \hat{\theta} z_j - c_j - \eta_j = r_j^{\text{clean}} - \eta_j$ where r_j^{clean} is the residual if the point was uncorrupted thus

$$\begin{aligned}
-\frac{1}{m}\sum_{w_v \in V} \sigma_v\, \delta(B_v, k_v^*)^\top r_j\; \kappa(z_j, z_v) &= -\frac{1}{m}\sum_{w_v \in V} \sigma_v\, \delta(B_v, k_v^*)^\top (r_j^{\text{clean}} - \eta_j)\; \kappa(z_j, z_v) \\
&= -\frac{1}{m}\sum_{w_v \in V} \sigma_v\, \delta(B_v, k_v^*)^\top r_j^{\text{clean}}\; \kappa(z_j, z_v) + \frac{1}{m}\sum_{w_v \in V} \sigma_v\, \delta(B_v, k_v^*)^\top \eta_j\; \kappa(z_j, z_v) \\
&= \Psi_j^{clean} + \left(\frac{1}{m}\sum_{w_v \in V} \sigma_v\, \kappa(z_j, z_v)\delta(B_v, k_v^*)\right)^\top \eta_j \\
&= \Psi_j^{clean} + u_j^\top \eta_j
\end{aligned}$$

so we see that the score on a corrupted point decomposes into an uncorrupted score and an additional noise term. □

A.3 Score difference in magnitude for corrupted points

Instead of the directed corruption paradigm in Assumption D4, we use the following:

For $j \in \mathcal{C}$, $\tilde{c}_j = c_j + \eta_j$ where η_j is an additional bounded perturbation such that η_j is zero-mean and independent from $\mathcal{S}, \mathcal{V}$ and across indices.

We show that the RCIF score magnitude for a fresh sample (independent of $\mathcal{S}, \mathcal{V}$) that has been corrupted $(z_j, \tilde{c}_j) = (z_j, c_j + \eta_j)$ is larger than the magnitude for uncorrupted version on (z_j, c_j).

Theorem 3 (Magnitude separation in expectation)**.** *Let (z_j, c_j) be a fresh sample with $c_j = \theta^* z_j + \varepsilon_j$, independent of $\mathcal{S}$ used to fit $\hat{\theta}$. Assume the corruption η_j satisfies*

$$\mathbb{E}[\eta_j \mid \mathcal{G}] = 0, \qquad \mathit{Cov}(\eta_j \mid \mathcal{G}) \succeq \sigma_\eta^2 I,$$

where σ-algebra $\mathcal{G} = \sigma(\hat{\theta}, \mathcal{V}, z_j, c_j)$. Let Ψ_j^{clean} be the score evaluated at the clean label (z_j, c_j) and let Ψ_j be the score evaluated at $(z_j, \tilde{c}_j) = (z_j, c_j + \eta_j)$. Then

$$\mathbb{E}\big[\Psi_j^2 - (\Psi_j^{\text{clean}})^2\big] \;\geq\; \sigma_\eta^2\, \mathbb{E}\|u_j\|^2 \;\geq\; 0.$$

where under most distributions the last inequality is strict positivity.

Proof. By the score decomposition (Lemma 2), $\Psi_j = \Psi_j^{\text{clean}} + u_j^\top \eta_j$. Let $\Delta_j = u_j^\top \eta_j$, so

$$\Psi_j^2 - (\Psi_j^{\text{clean}})^2 \;=\; 2\,\Psi_j^{\text{clean}}\,\Delta_j + \Delta_j^2.$$

Since Ψ_j^{clean} and u_j are $\mathcal{G}$-measurable and $\mathbb{E}[\eta_j \mid \mathcal{G}] = 0$,

$$\mathbb{E}[\Psi_j^{\text{clean}}\,\Delta_j \mid \mathcal{G}] \;=\; \Psi_j^{\text{clean}}\,u_j^\top \mathbb{E}[\eta_j \mid \mathcal{G}] \;=\; 0,$$

and by the covariance hypothesis,

$$\mathbb{E}[\Delta_j^2 \mid \mathcal{G}] \;=\; u_j^\top \mathrm{Cov}(\eta_j \mid \mathcal{G})\, u_j \;\geq\; \sigma_\eta^2\, \|u_j\|^2.$$

Taking outer expectation and applying the tower rule gives $\mathbb{E}[\Psi_j^2 - (\Psi_j^{\text{clean}})^2] \geq \sigma_\eta^2\, \mathbb{E}\|u_j\|^2 \geq 0$. However, we note that $\mathbb{E}\|u_j\|^2 = 0$ only if it is a.s. 0, which is a measure-0 event under reasonable data distributions. □

A.4 Statement and proof of Theorem 1

Theorem 4 (Decreased corruption score under alignment)**.** *Adopt Assumptions (D1)–(D5). Let (z_j, c_j) be a fresh sample drawn from $\mathcal{D}$ independently of $\hat{\theta}$ (fit on $\mathcal{S}$) and of $\mathcal{V}$, and let $\tilde{c}_j = c_j + \eta_j$ be its directional corruption (D4). Suppose $\gamma := \gamma(\hat{\theta}, z_j, c_j) > 0$ from Definition 2. Then conditioned on $\mathcal{F} = \sigma(\hat{\theta}, z_j, c_j)$:*

$$\mathbb{E}[\Psi_j - \Psi_j^{\text{clean}} \mid \mathcal{F}] \;=\; -\alpha\,\gamma \;<\; 0.$$

where Ψ_j^{clean} is as defined in Lemma 2. For every $t \in (0, \gamma)$,

$$P\big(\Psi_j - \Psi_j^{\text{clean}} \;\geq\; -\alpha(\gamma - t) \,\big|\, \mathcal{F}\big) \;\leq\; \exp\!\left(-\frac{m\,t^2}{2M_\gamma^2}\right),$$

where $M_\gamma = DR^2/\mu \;\geq\; |\sigma_v \kappa(z_j, z_v)\delta(B_v, k_v^*)^\top \hat{\delta}_j|$. In particular, with probability $\geq 1 - \exp(-m\gamma^2/(8M_\gamma^2))$ (taking $t = \gamma/2$):

$$\Psi_j - \Psi_j^{\text{clean}} \;\leq\; -\frac{\alpha\,\gamma}{2} \;<\; 0.$$

Proof. The w_v are iid, and conditional on $\mathcal{F} = \sigma(\hat{\theta}, z_j, c_j)$, so are the bounded scalars $W_v = \sigma_v \kappa(z_j, z_v)\delta(B_v, k_v^*)^\top \hat{\delta}_j \in [-M_\gamma, M_\gamma]$ with mean $\mathbb{E}[W_v] = \gamma$. By Assumption D4 and Lemma 2,

$$\Psi_j - \Psi_j^{\text{clean}} \;=\; -\frac{\alpha}{m} \sum_{v \in \mathcal{V}} W_v.$$

(1) Expectation. Linearity gives $\mathbb{E}[\Psi_j - \Psi_j^{\text{clean}} \mid \mathcal{F}] = -\alpha\,\mathbb{E}[W_v] = -\alpha\,\gamma$.

(2) Concentration. By Hoeffding's inequality for iid bounded random variables (range $2M_\gamma$), for any $t > 0$:

$$P\left(\frac{1}{m}\sum_v W_v \le \gamma - t \,\middle|\, \mathcal{F}\right) \;\le\; \exp\left(-\frac{m\,t^2}{2M_\gamma^2}\right).$$

Multiplying by $-\alpha$ (flips the inequality): $\{\frac{1}{m}\sum W_v \le \gamma - t\} = \{\Psi_j - \Psi_j^{\text{clean}} \ge -\alpha(\gamma - t)\}$, which gives the stated bound. Taking $t = \gamma/2$ yields $\Psi_j - \Psi_j^{\text{clean}} \le -\alpha\gamma/2$ with probability $\ge 1 - \exp(-m\gamma^2/(8M_\gamma^2))$.

□

A.5 In-sample signed decreased corruption scoring

Theorem 4 assumes that the corrupted point $(z_j, \tilde{c}_j) = (z_j, c_j + \eta_j)$ (where η_j is a bounded directional corruption as in Assumption D4) is independent of the fitted estimator $\hat{\theta}$. However, we would like to see that if w_j is one of the training points used to fit $\hat{\theta}$, that Theorem 4 still holds. We show that under additional mild regularity assumptions, the conclusion of Theorem 4 continues to hold for in-sample j, with an additive $O(1/n)$ slack.

We replace $\hat{\theta}$ everywhere in the score by the leave-one-out estimator $\hat{\theta}_{-j}$, which is independent of w_j by construction, and control the resulting approximation error.

Notation. Let

$$\hat{\theta}_{-j} \;:=\; \arg\min_\theta \frac{1}{n-1}\sum_{i\neq j} \|\theta z_i - \tilde{c}_i\|^2, \qquad \hat{\Sigma}_{-j} \;:=\; \frac{1}{n-1}\sum_{i\neq j} z_i z_i^\top,$$

let $h_j := \frac{1}{n} z_j^\top \hat{\Sigma}^{-1} z_j$ denote the leverage of w_j, and let

$$r_j \;:=\; \hat{\theta} z_j - \tilde{c}_j, \qquad r_j^- \;:=\; \hat{\theta}_{-j} z_j - \tilde{c}_j.$$

Define the leave-one-out kernel and score

$$\kappa^-(z, z') \;:=\; z^\top \hat{\Sigma}_{-j}^{-1} z', \qquad \Psi_j^{\text{loo}} \;:=\; -\frac{1}{m}\sum_{v\in\mathcal{V}} \sigma_v^-\, \delta(B_v^-, k_v^{*-})^\top r_j^-\, \kappa^-(z_v, z_j),$$

where $B_v^-, k_v^{*-}, \sigma_v^-$ are the basis, fragile direction, and correctness sign computed at $\hat{c}_v^- := \hat{\theta}_{-j} z_v$. Thus Ψ_j^{loo} is the score one would compute by substituting $\hat{\theta}_{-j}$ for $\hat{\theta}$ throughout the score formula.

We make an additional assumption

Assumption 3. *Assume that residuals are bounded by a constant, i.e.* $\|r_j\| \le R_r$ *(and for* r_j^-*).*

Note that by D1, the leverage satisfies $h_j \le R^2/(n\mu)$, so $h_j \to 0$ at rate $1/n$ and $1 - h_j \ge 1/2$ for $n \ge 2R^2/\mu$.

Lemma 3 (Exact LOO identities for linear ERM)**.** *Suppose* $h_j < 1$*. Then*

$$\hat{\theta}_{-j} - \hat{\theta} \;=\; \frac{1}{n(1-h_j)}\, r_j\, (\hat{\Sigma}^{-1} z_j)^\top, \tag{1}$$

$$r_j^- \;=\; \frac{r_j}{1-h_j}, \tag{2}$$

$$\hat{\Sigma}_{-j}^{-1} - \hat{\Sigma}^{-1} \;=\; -\frac{1}{n}\hat{\Sigma}^{-1} + \frac{(n-1)/n^2}{1-h_j}\,\hat{\Sigma}^{-1} z_j z_j^\top \hat{\Sigma}^{-1}. \tag{3}$$

Proof. Set $A := n\hat{\Sigma}$ and $a_j := A^{-1} z_j = \hat{\Sigma}^{-1} z_j / n$, so $z_j^\top a_j = h_j$. By the Sherman–Morrison identity,

$$(A - z_j z_j^\top)^{-1} \;=\; A^{-1} + \frac{a_j a_j^\top}{1-h_j}.$$

Multiplying by $(n-1)$ gives (3) after simplification.

By the closed form of least squares, $\hat{\theta} = (\sum_i \tilde{c}_i z_i^\top)A^{-1}$ and $\hat{\theta}_{-j} = (\sum_{i\neq j} \tilde{c}_i z_i^\top)(A - z_j z_j^\top)^{-1}$, expansion using the identity above yields

$$\hat{\theta}_{-j} \;=\; \hat{\theta} \;+\; \frac{(\hat{\theta} z_j - \tilde{c}_j)\, a_j^\top}{1-h_j} \;=\; \hat{\theta} + \frac{r_j\, a_j^\top}{1-h_j},$$

which is (1). Right-multiplying by z_j and subtracting $\tilde{c}_j$ gives $r_j^- = r_j + \frac{r_j h_j}{1-h_j} = \frac{r_j}{1-h_j}$, which is (2). □

We now make an assumption that the optimal bases of predictions by $\hat{\theta}$ and $\hat{\theta}_{-j}$ are similar.

Assumption 4 (Validation basis stability under LOO)**.** *For every $v \in \mathcal{V}$, the optimal basis B_v, the fragile direction k_v^*, and the correctness sign σ_v are unchanged when $\hat{c}_v$ is replaced by $\hat{c}_v^-$, i.e. $B_v^- = B_v$, $k_v^{*-} = k_v^*$, and $\sigma_v^- = \sigma_v$.*

This is a reasonable assumption which we briefly justify: it holds as n increases.

Remark 3 (One sufficient condition for Assumption 4 with two-sided scoring)**.** *Given (D1)-(D5), and that for every $v \in \mathcal{V}$ the second-smallest normalized reduced cost at $\hat{c}_v$ (by normalized reduced cost we mean $\frac{s_k(\hat{c}_v, B_v)}{\|\delta(B_v,k)\|}$) exceeds the smallest by at least 2ρ, with the smallest value at least $\geq \rho$, where*

$$\rho \;>\; \frac{R^2 R_r}{n\mu(1-h_j)}.$$

Then Assumption 4 holds, assuming two-sided score.

Proof. By Lemma 3, $\|\hat{\theta}_{-j} - \hat{\theta}\|_{\mathrm{op}} \leq \frac{\|r_j\|\,\|\hat{\Sigma}^{-1} z_j\|}{n(1-h_j)} \leq \frac{R R_r}{n\mu(1-h_j)}$, so $\|\hat{c}_v - \hat{c}_v^-\| \leq R^2 R_r/(n\mu(1-h_j))$. For any unit direction $\delta(B,k)/\|\delta(B,k)\|$, by Cauchy-Schwarz

$$\left| s_k(\hat{c}_v^-, B)/\|\delta(B,k)\| - s_k(\hat{c}_v, B)/\|\delta(B,k)\| \right| \;\leq\; \|\hat{c}_v - \hat{c}_v^-\|.$$

Hence each normalized reduced cost shifts by less than ρ. Since the smallest is at least ρ and the gap to the second-smallest exceeds 2ρ, signs of all reduced costs are preserved (so $B_v^- = B_v$) and the argmin is preserved (so $k_v^{*-} = k_v^*$). The sign σ_v is also preserved since $B_v^- = B_v$. □

Assumption 4 is imposed for the chosen scoring rule. Remark 3 gives one sufficient condition for Assumption 4 in the two-sided scoring case.

Lemma 4 (LOO stability of the RCIF score)**.** *Under Assumptions (D1)-(D5) and 4,*

$$|\Psi_j - \Psi_j^{\mathrm{loo}}| \;\leq\; \frac{C}{n}\big(\|r_j\|\big),$$

for a constant $C = C(R, \mu, D)$ depending only on the regularity parameters. In particular, given Assumption 3, we have $|\Psi_j - \Psi_j^{\mathrm{loo}}| \leq C'/n$ for $C' = C'(R, \mu, D, R_r)$.

Proof. Recall the score can be written as $\Psi_j = -u_j^\top r_j$ with

$$u_j \;=\; \frac{1}{m}\sum_{v\in\mathcal{V}} \sigma_v\, \kappa(z_v, z_j)\, \delta(B_v, k_v^*),$$

and analogously $\Psi_j^{\mathrm{loo}} = -(u_j^-)^\top r_j^-$, where u_j^- uses $\sigma_v^-, B_v^-, k_v^{*-}, \kappa^-$. Under Assumption 4, $\sigma_v^- = \sigma_v$, $B_v^- = B_v$, $k_v^{*-} = k_v^*$, so

$$u_j^- - u_j \;=\; \frac{1}{m}\sum_{v\in\mathcal{V}} \sigma_v\, (\kappa^-(z_v, z_j) - \kappa(z_v, z_j))\, \delta(B_v, k_v^*).$$

By Lemma 3 eq (3) and Assumption D1,

$$\|\hat{\Sigma}_{-j}^{-1} - \hat{\Sigma}^{-1}\|_{\mathrm{op}} \;\leq\; \frac{1}{n\mu} + \frac{R^2(n-1)/n^2}{(1-h_j)\mu^2} \;=\; O(1/n),$$

so $|\kappa^-(z_v, z_j) - \kappa(z_v, z_j)| \le R^2 \cdot O(1/n)$, and hence $\|u_j^- - u_j\| \le c_1/n$ for $c_1 = c_1(R, \mu, D)$.

By Lemma 3 eq (2), $r_j^- - r_j = \frac{h_j}{1-h_j} r_j$, so $\|r_j^- - r_j\| \le \frac{R^2}{n\mu(1-h_j)} \|r_j\| = c_2 \|r_j\|/n$.

Combining,

$$\begin{aligned} |\Psi_j - \Psi_j^{\mathrm{loo}}| \;&\le\; |u_j^\top (r_j - r_j^-)| + |(u_j - u_j^-)^\top r_j^-| \\ &\le\; \|u_j\|\, c_2 \|r_j\|/n + (c_1/n)\, \|r_j^-\| \\ &\le\; \frac{C\,\|r_j\|}{n}, \end{aligned}$$

since $\|u_j\| \le R^2 D/\mu$ is bounded under Assumptions D1 and D5 and $\|r_j^-\| \le c_3 \|r_j\|$ by Lemma 3 (2) for some threshold of n, (for example, $c_3 = 2$ for $n \ge 2R^2/\mu$). Then Assumption 3 directly upperbounds $\|r_j\|$ and $\|r_j^-\|$ to provide the stated C'/n form. □

Now we can show the in-sample result.

In-sample setup. For the in-sample variant, the σ-algebra of interest is $\mathcal{F}^- := \sigma(\hat{\theta}_{-j}, z_j, c_j)$, under which corruption $\eta_j = -\alpha\hat{\delta}_j$ is independent of $\hat{\theta}_{-j}$. The score decomposition of Lemma 2 applies directly to Ψ_j^{loo}:

$$\Psi_j^{\mathrm{loo}} \;=\; \Psi_j^{\mathrm{loo,clean}} + (u_j^-)^\top \eta_j, \qquad \Psi_j^{\mathrm{loo,clean}} \;:=\; -(u_j^-)^\top r_j^{-,\mathrm{clean}},$$

where $r_j^{-,\mathrm{clean}} := \hat{\theta}_{-j} z_j - c_j$.

Theorem 5 (Signed score decrease, in-sample variant)**.** *Adopt Assumptions (D1)-(D5) and Assumptions 3, 4. Let $\gamma^- := \gamma(\hat{\theta}_{-j}, z_j, c_j) > 0$ be the validation alignment evaluated at the LOO estimator, and let C, C' be the coefficients from Lemma 4. Then for a point w_j that has been directionally corrupted (D4):*

1. ***(Expectation)***
$$\mathbb{E}[\Delta_j \mid \mathcal{F}^-] \;=\; -\alpha\gamma^- \;+\; O(1/n),$$
where $\Delta_j := \Psi_j - \Psi_j^{clean}$ and the $O(1/n)$ term is bounded by $(2C' + C\alpha)/n$.

2. ***(High probability)** For every $t \in (0, \gamma^-)$, on the event of probability $\ge 1 - \exp(-mt^2/(2M_\gamma^2))$ guaranteed by Theorem 4,*
$$\Delta_j \;\le\; -\alpha(\gamma^- - t) + \frac{2C' + C\alpha}{n}.$$
In particular, taking $t = \gamma^-/2$ and assuming $n \ge (8C' + 4C\alpha)/(\alpha\gamma^-)$,
$$\Delta_j \;\le\; -\frac{\alpha\gamma^-}{4}$$
on an event of probability $\ge 1 - \exp(-m(\gamma^-)^2/(8M_\gamma^2))$.

Proof. Let $\Delta_j^{\mathrm{loo}} := (u_j^-)^\top \eta_j$. Since $\hat{\theta}_{-j}$ is independent of w_j, the proof of Theorem 4 applies verbatim to Δ_j^{loo}, yielding

$$\begin{aligned} \mathbb{E}[\Delta_j^{\mathrm{loo}} \mid \mathcal{F}^-] \;&=\; -\alpha\gamma^-, \\ P\big(\Delta_j^{\mathrm{loo}} \ge -\alpha(\gamma^- - t) \,\big|\, \mathcal{F}^-\big) \;&\le\; \exp(-mt^2/(2M_\gamma^2)). \end{aligned}$$

By Lemma 4 applied at both r_j (full score) and r_j^{clean} (clean score) which incurs an extra corruption η_j term, (via Assumption 3 and $\|\eta_j\| \le \alpha$),

$$|\Delta_j - \Delta_j^{\mathrm{loo}}| \;=\; |\Psi_j - \Psi_j^{\mathrm{clean}} - \Psi_j^{\mathrm{loo}} + \Psi_j^{\mathrm{loo,\,clean}}| \;\le\; \frac{2C' + C\alpha}{n}.$$

Substituting into each statement of Theorem 4 gives parts (1) and (2). □

Remark 4. *For n large enough the in-sample guarantee of Theorem 5 is qualitatively the same as the cross-fit version. The replacement of γ by γ^- is harmless for our purposes since $\gamma^- \to \gamma$ as $n \to \infty$ (because $\hat{\theta}_{-j} \to \hat{\theta}$ at rate $1/n$, and the alignment parameter is continuous in the estimator under Assumption 4).*

A.6 Statement and proof of Theorem 2

In Appendix Subsection A.5 we showed that a single corrupted point receives a lower score with high probability. We now aggregate this across the corruption set to obtain an *unconditional* expected reduction of the worst-case corrupted bias of the weighted ERM, relative to $\hat{\theta}$ and $\hat{\theta}_{\mathrm{val}}$. Under constant-rate, constant-magnitude corruption, no estimator using only the corrupted training data can drive corruption bias to zero, but RCIF reweighting reduces the constant.

Then we may translate the reduction in bias to regret (under the simplifying assumption 2).

A.6.1 Bounding weight ratio between clean and corrupted points

Remark 5. *Under (D1)–(D5) and Assumption 3 let $M = R_r D R^2/\mu$, and observe that $|\Psi_j| < M,\ a.s.\ \forall j$.*

We work with the leave-one-out objects of Appendix A.5: $\hat{\theta}_{-j}$, $\kappa^-(z,z') := z^\top \hat{\Sigma}_{-j}^{-1} z'$, and the LOO score aggregator

$$u_j^{(-1)} \ := \ \frac{1}{m} \sum_{v \in \mathcal{V}} \sigma_v^- \, \kappa^-(z_j, z_v)\, \delta(B_v^-, k_v^{*-}).$$

Quantities of interest

For linear/squared-loss ERM, the parameter error decomposes as

$$\hat{\theta}_p - \theta^\star \ = \ \Big(\sum_{i=1}^n p_i \, \varepsilon_i z_i^\top \ + \ T_w \Big) \hat{\Sigma}_W^{-1}, \qquad \hat{\Sigma}_W := \sum_i p_i z_i z_i^\top,$$

and analogously for the unweighted ERM with $p_i = 1/n$. Define the *corruption-bias matrices*

$$T_0 := \frac{1}{n} \sum_{j \in \mathcal{C}} (-\alpha \, \hat{\delta}_j) z_j^\top, \qquad T_w := \sum_{j \in \mathcal{C}} p_j (-\alpha \, \hat{\delta}_j) z_j^\top,$$

which directly drive the corruption bias of the corresponding estimators.

We note that we can reduce the problem from the corruption set to one corrupt index, due to Assumptions D1 and D3.

Lemma 5 (Exchangeability of corruption)**.** *For any sequence (F_j) such that the joint law of $(F_j, \mathbf{1}_{j \in \mathcal{C}})_{j=1}^n$ is permutation-invariant,*

$$\mathbb{E}\left[\sum_{j \in \mathcal{C}} F_j \right] \ = \ n\beta \, \mathbb{E}[F_1 \mid 1 \in \mathcal{C}].$$

Proof. By (D3), $(\mathbf{1}_{1 \in \mathcal{C}}, \ldots, \mathbf{1}_{n \in \mathcal{C}})$ is exchangeable with marginal mean β. By (D1)–(D2), (z_i, ε_i) are iid, so the joint law of $(w_i, \mathbf{1}_{i \in \mathcal{C}})_{i=1}^n$ is permutation-invariant. Hence $\mathbb{E}[\mathbf{1}_{j \in \mathcal{C}} F_j] = \mathbb{E}[\mathbf{1}_{1 \in \mathcal{C}} F_1]$ for every j, and summing n identical terms with $\Pr(1 \in \mathcal{C}) = \beta$ gives the claim. □

Remark 6 (Deterministic bound on the score sum)**.** *Under (D1)–(D5) and Assumption 3, since $|\Psi_i| \le M$ a.s., the sum of scores*

$$Q := \sum_{i=1}^n e^{\lambda \Psi_i}$$

admits the deterministic lower bound $Q \ge n \, e^{-\lambda M}$ a.s., and hence $1/Q \le e^{\lambda M}/n$ a.s.

Lemma 6 (Weight-ratio contraction for a clean vs corrupted point)**.** *Focusing on the first training index, and under Assumption 4, let*

$$\Phi_{\mathcal{C}}(\lambda) := \mathbb{E}[e^{\lambda \Psi_1} \mid 1 \in \mathcal{C}], \qquad \Phi_{\bar{\mathcal{C}}}(\lambda) := \mathbb{E}[e^{\lambda \Psi_1} \mid 1 \notin \mathcal{C}].$$

Under (D1)–(D5), Assumptions 1 and 3, for every $\lambda > 0$,

$$\frac{\Phi_{\mathcal{C}}(\lambda)}{\Phi_{\bar{\mathcal{C}}}(\lambda)} \ \le \ e^{-\lambda \alpha \gamma_0/4} \ + \ e^{2\lambda M} \big(\delta_\gamma + e^{-m\gamma_0^2/(8M_\gamma^2)} \big) \ + \ O(1/n).$$

We define the contraction factor

$$\Phi(\lambda, m, n) \;:=\; e^{-\lambda\alpha\gamma_0/4} \;+\; e^{2\lambda M}\big(\delta_\gamma + e^{-m\gamma_0^2/(8M_\gamma^2)}\big) \;+\; O(1/n).$$

Proof. Recall the score decomposition (Lemma 2) applied at $\hat{\theta}_{-1}$:

$$\Psi_1 \;=\; \Psi_1^{\mathrm{loo}} + O(1/n) \;=\; \Psi_1^{\mathrm{cf}} - \alpha\, u_1^{(-1)\top}\hat{\delta}_1 \cdot \mathbf{1}_{1\in\mathcal{C}} + O(1/n),$$

where $\Psi_1^{\mathrm{cf}} := -u_1^{(-1)\top}(\hat{\theta}_{-1}z_1 - c_1)$ (which represents the counterfactual score) depends only on $\hat{\theta}_{-1}, \mathcal{V}, z_1, c_1$, using the *clean* cost c_1.

Step 1: pointwise bound on the corruption shift. Define the high-probability event

$$\mathcal{F} \;:=\; \big\{\gamma(\hat{\theta}_{-1}, z_1, c_1) \ge \gamma_0\big\} \cap \big\{u_1^{(-1)\top}\hat{\delta}_1 \ge \gamma_0/2\big\}.$$

By Assumption 1, $\Pr[\gamma(\hat{\theta}_{-1}, z_1, c_1) < \gamma_0] \le \delta_\gamma$. Conditional on $\gamma(\hat{\theta}_{-1}, z_1, c_1) \ge \gamma_0$, Hoeffding's inequality applied to the iid bounded summands W_v defining $u_1^{(-1)\top}\hat{\delta}_1 = \frac{1}{m}\sum_v W_v$ (cf. Theorem 5) gives

$$\Pr[u_1^{(-1)\top}\hat{\delta}_1 < \gamma_0/2 \mid \gamma \ge \gamma_0] \;\le\; e^{-m\gamma_0^2/(8M_\gamma^2)}.$$

By a union bound, $\Pr[\mathcal{F}^c] \le \delta_\gamma + e^{-m\gamma_0^2/(8M_\gamma^2)}$.

Step 2: bound on $\Phi_{\mathcal{C}}$. Conditional on $\{1 \in \mathcal{C}\}$, decompose

$$\mathbb{E}\big[e^{\lambda\Psi_1}\big] \;=\; \mathbb{E}\big[e^{\lambda\Psi_1}\,\mathbf{1}_{\mathcal{F}}\big] \;+\; \mathbb{E}\big[e^{\lambda\Psi_1}\,\mathbf{1}_{\mathcal{F}^c}\big].$$

On $\mathcal{F}$, $\Psi_1 \le \Psi_1^{\mathrm{cf}} - \alpha\gamma_0/2 + O(1/n)$, so

$$\mathbb{E}\big[e^{\lambda\Psi_1}\,\mathbf{1}_{\mathcal{F}} \mid 1 \in \mathcal{C}\big] \;\le\; e^{-\lambda\alpha\gamma_0/2 + O(\lambda/n)}\,\mathbb{E}\big[e^{\lambda\Psi_1^{\mathrm{cf}}} \mid 1 \in \mathcal{C}\big].$$

Off $\mathcal{F}$, $|\Psi_1| \le M$ gives

$$\mathbb{E}\big[e^{\lambda\Psi_1}\,\mathbf{1}_{\mathcal{F}^c} \mid 1 \in \mathcal{C}\big] \;\le\; e^{\lambda M}\big(\delta_\gamma + e^{-m\gamma_0^2/(8M_\gamma^2)}\big).$$

Step 3: counterfactual-MGF identification. We show that

$$\mathbb{E}\Big[e^{\lambda\Psi_1^{\mathrm{cf}}} \mid 1 \in \mathcal{C}\Big] \;=\; \Phi_{\bar{\mathcal{C}}}(\lambda) \;+\; O(1/n). \tag{4}$$

The counterfactual score Ψ_1^{cf} depends on the clean data (z_1, c_1) and the LOO estimator $\hat{\theta}_{-1}$. Its conditional distribution given $\{1 \in \mathcal{C}\}$ differs from its distribution given $\{1 \notin \mathcal{C}\}$ only through $\hat{\theta}_{-1}$: in the first case $\hat{\theta}_{-1}$ is fit on a sample with $|\mathcal{C}| - 1$ corrupted points among $\{2, \dots, n\}$, in the second on a sample with $|\mathcal{C}|$ corrupted points. The marginal of (z_1, c_1) is identical in both cases since corruption assignment is independent of the clean data.

Coupling. We bound the difference between the two conditional expectations via a coupling. Sample uniformly from the clean data $(z_i, c_i)_{i=2}^n$, a subset S of size $|\mathcal{C}| - 1$, and an additional index $j^\star \in \{2, \dots, n\} \setminus S$. Define two parallel LOO estimators on the same probability space:

- $\hat{\theta}^A_{-1}$, fit with corruption set C_S inside $\{2, \dots, n\}$ (corresponding to $1 \in \mathcal{C}$).
- $\hat{\theta}^B_{-1}$, fit with corruption set $C_S \cup \{j^\star\}$ inside $\{2, \dots, n\}$ (corresponding to $1 \notin \mathcal{C}$).

By construction, the marginals of $\hat{\theta}^A_{-1}$ and $\hat{\theta}^B_{-1}$ coincide with the conditional distributions of $\hat{\theta}_{-1}$ given $\{1 \in \mathcal{C}\}$ and $\{1 \notin \mathcal{C}\}$ respectively, and $\hat{\theta}^A_{-1}$, $\hat{\theta}^B_{-1}$ differ only in the corruption status of one training index, $j^\star$. Lemma 3 then gives

$$\big\|\hat{\theta}^A_{-1} - \hat{\theta}^B_{-1}\big\| \;=\; O(1/n) \quad \text{a.s.},$$

and by Lemma 4 yields

$$\big|\Psi_1^{\mathrm{cf},A} - \Psi_1^{\mathrm{cf},B}\big| \;=\; O(1/n) \quad \text{a.s.}$$

Bound on the expected corruption difference. On the coupling space, using $|e^a - e^b| \le e^{\max(a,b)}|a-b|$ and $|\Psi_1^{\mathrm{cf}}| \le M$:

$$\left|e^{\lambda\Psi_1^{\mathrm{cf},A}} - e^{\lambda\Psi_1^{\mathrm{cf},B}}\right| \;\le\; \lambda\, e^{\lambda M}\left|\Psi_1^{\mathrm{cf},A} - \Psi_1^{\mathrm{cf},B}\right| \;=\; O(1/n) \quad \text{a.s.}$$

Taking expectations and applying Jensen's inequality:

$$\left|\,\mathbb{E}[e^{\lambda\Psi_1^{\mathrm{cf},A}}] - \mathbb{E}[e^{\lambda\Psi_1^{\mathrm{cf},B}}]\,\right| \;\le\; \mathbb{E}\left|e^{\lambda\Psi_1^{\mathrm{cf},A}} - e^{\lambda\Psi_1^{\mathrm{cf},B}}\right| \;=\; O(1/n).$$

The marginal-matching property of the coupling identifies the two expectations on the left with the conditional expectations of interest:

$$\left|\,\mathbb{E}[e^{\lambda\Psi_1^{\mathrm{cf}}} \mid 1 \in \mathcal{C}] - \mathbb{E}[e^{\lambda\Psi_1^{\mathrm{cf}}} \mid 1 \notin \mathcal{C}]\,\right| \;=\; O(1/n).$$

See that on the event $\{1 \notin \mathcal{C}\}$ the counterfactual coincides with the actual score, $\Psi_1^{\mathrm{cf}} = \Psi_1$, and therefore $\mathbb{E}[e^{\lambda\Psi_1^{\mathrm{cf}}} \mid 1 \notin \mathcal{C}] = \Phi_{\bar{\mathcal{C}}}(\lambda)$. Substituting yields (4).

Step 4: combine and divide. Substituting (4) into the first conclusion of Step 2 gives

$$\begin{aligned}\Phi_{\mathcal{C}}(\lambda) &\le e^{-\lambda\alpha\gamma_0/2+O(\lambda/n)}\big[\Phi_{\bar{\mathcal{C}}}(\lambda) + O(1/n)\big] \;+\; e^{\lambda M}\big(\delta_\gamma + e^{-m\gamma_0^2/(8M_\gamma^2)}\big)\\ &= e^{-\lambda\alpha\gamma_0/2}\,(1+O(\lambda/n))\,\Phi_{\bar{\mathcal{C}}}(\lambda) \;+\; O(1/n) \;+\; e^{\lambda M}\big(\delta_\gamma + e^{-m\gamma_0^2/(8M_\gamma^2)}\big), \qquad (5)\end{aligned}$$

where the second line uses $e^{-\lambda\alpha\gamma_0/2+O(\lambda/n)} = e^{-\lambda\alpha\gamma_0/2}(1+O(\lambda/n))$ via first order Taylor expansion together with the boundedness $\Phi_{\bar{\mathcal{C}}}(\lambda) \le e^{\lambda M}$ and $e^{-\lambda\alpha\gamma_0/2+O(\lambda/n)} = O(1)$ to absorb the cross-term $e^{-\lambda\alpha\gamma_0/2+O(\lambda/n)}\cdot O(1/n)$ into the $O(1/n)$ term.

To divide both sides by $\Phi_{\bar{\mathcal{C}}}(\lambda)$, we use the deterministic lower bound

$$\Phi_{\bar{\mathcal{C}}}(\lambda) \;=\; \mathbb{E}[e^{\lambda\Psi_1} \mid 1 \notin \mathcal{C}] \;\ge\; e^{-\lambda M}, \qquad (6)$$

which follows from $\Psi_1 \ge -M$ a.s. In particular $1/\Phi_{\bar{\mathcal{C}}}(\lambda) \le e^{\lambda M}$.

Dividing (5) term by term and applying (6):

$$\begin{aligned}\frac{\Phi_{\mathcal{C}}(\lambda)}{\Phi_{\bar{\mathcal{C}}}(\lambda)} &\le e^{-\lambda\alpha\gamma_0/2}\,(1+O(\lambda/n)) \;+\; \frac{O(1/n)}{\Phi_{\bar{\mathcal{C}}}(\lambda)} \;+\; \frac{e^{\lambda M}\big(\delta_\gamma + e^{-m\gamma_0^2/(8M_\gamma^2)}\big)}{\Phi_{\bar{\mathcal{C}}}(\lambda)}\\ &\le e^{-\lambda\alpha\gamma_0/2} \;+\; e^{2\lambda M}\big(\delta_\gamma + e^{-m\gamma_0^2/(8M_\gamma^2)}\big) \;+\; O(1/n),\end{aligned}$$

where we have folded the $O(\lambda/n)$ slack on the leading term into the additive $O(1/n)$ remainder.

Finally, since $e^{-\lambda\alpha\gamma_0/2} \le e^{-\lambda\alpha\gamma_0/4}$ for $\lambda, \alpha, \gamma_0 > 0$, we may relax the leading exponent for notational consistency with Theorem 5, obtaining

$$\frac{\Phi_{\mathcal{C}}(\lambda)}{\Phi_{\bar{\mathcal{C}}}(\lambda)} \;\le\; e^{-\lambda\alpha\gamma_0/4} \;+\; e^{2\lambda M}\big(\delta_\gamma + e^{-m\gamma_0^2/(8M_\gamma^2)}\big) \;+\; O(1/n),$$

which is the conclusion of the lemma. □

A.6.2 Bounding moments of bias corruption term

Theorem 6 (Expected corruption-bias contraction)**.** *Under Assumptions 1, 3, 4, and (D1)–(D5), the expected corruption-bias matrix of the reweighted estimator satisfies*

$$\mathbb{E}\|T_w\|_F \;\le\; \alpha\beta R\, e^{2\lambda M}\,\Phi(\lambda, m, n), \qquad (7)$$

where $\Phi(\lambda, m, n)$ is the weight ratio bound in Lemma 6. The unweighted estimator's corruption bias satisfies the deterministic bound

$$\mathbb{E}\|T_0\|_F \;\le\; \alpha\beta R. \qquad (8)$$

The reweighted bound improves on the unweighted bound by the multiplicative factor $e^{2\lambda M}\,\Phi(\lambda, m, n)$, which is strictly less than 1 whenever $\Phi(\lambda, m, n) < e^{-2\lambda M}$.

Second-moment equivalents.

$$\mathbb{E}\|T_w\|_F^2 \;\le\; \alpha^2\beta R^2\, e^{2\lambda M}\,\Phi(\lambda, m, n), \qquad (9)$$

$$\mathbb{E}\|T_0\|_F^2 \;\le\; \alpha^2\beta^2 R^2. \qquad (10)$$

The reweighted second-moment bound improves on the unweighted one by the multiplicative factor $e^{2\lambda M}\,\Phi(\lambda, m, n)/\beta$, which is strictly less than 1 whenever $\Phi(\lambda, m, n) < \beta\, e^{-2\lambda M}$.

Proof. **First moment.** By the triangle inequality and $\|\hat{\delta}_j z_j^\top\|_F \le R$ a.s.,

$$\|T_w\|_F \;\le\; \sum_{j\in\mathcal{C}} p_j \,\|\alpha\hat{\delta}_j z_j^\top\|_F \;\le\; \alpha R \sum_{j\in\mathcal{C}} p_j \quad \text{a.s.}$$

Taking expectation and applying exchangeability of $(p_j, \mathbf{1}\{j\in\mathcal{C}\})$ across j (Lemma 5),

$$\mathbb{E}\|T_w\|_F \;\le\; \alpha R\,\mathbb{E}\left[\sum_{j=1}^{n} p_j\,\mathbf{1}\{j\in\mathcal{C}\}\right] = \alpha R\, n\beta\,\mathbb{E}[p_1 \mid 1\in\mathcal{C}].$$

By Remark 6, $p_1 = e^{\lambda\Psi_1}/Q \le e^{\lambda\Psi_1}\, e^{\lambda M}/n$ a.s., so

$$\mathbb{E}[p_1 \mid 1\in\mathcal{C}] \;\le\; \frac{e^{\lambda M}}{n}\,\mathbb{E}\big[e^{\lambda\Psi_1}\,\big|\,1\in\mathcal{C}\big] \;=\; \frac{e^{\lambda M}\,\Phi_{\mathcal{C}}(\lambda)}{n}.$$

By Lemma 6 $\Phi_{\mathcal{C}}(\lambda) \le \Phi_{\bar{\mathcal{C}}}(\lambda)\,\Phi(\lambda,m,n)$ together with the deterministic bound $\Phi_{\bar{\mathcal{C}}}(\lambda) \le e^{\lambda M}$ (from $\Psi_1 \le M$ a.s.),

$$\mathbb{E}[p_1 \mid 1\in\mathcal{C}] \;\le\; \frac{e^{2\lambda M}\,\Phi(\lambda,m,n)}{n},$$

and (7) follows.

For T_0: $p_i = 1/n$ deterministically, so $\|T_0\|_F \le \alpha R\sum_{j\in\mathcal{C}} 1/n = \alpha R\beta$ a.s., yielding $\mathbb{E}\|T_0\|_F \le \alpha\beta R$.

Second moment. From $\sum_i p_i = 1$, $\|T_w\|_F \le \alpha R$ a.s., so $\|T_w\|_F^2 \le \alpha R\,\|T_w\|_F$ pointwise. Taking expectation and substituting (7):

$$\mathbb{E}\|T_w\|_F^2 \;\le\; \alpha R\,\mathbb{E}\|T_w\|_F \;\le\; \alpha^2\beta R^2\, e^{2\lambda M}\,\Phi(\lambda,m,n).$$

For T_0, $\|T_0\|_F \le \alpha\beta R$ a.s. as above, so $\mathbb{E}\|T_0\|_F^2 \le \alpha^2\beta^2 R^2$. □

Now we can obtain a comparison in worst-case expected bias from each of $\hat{\theta}, \hat{\theta}_p, \hat{\theta}_{\mathrm{val}}$ under additional assumptions.

Assumption 5. *We assume the equivalent of D1 for the empirical covariance of the validation set, i.e. that the minimum eigenvalue is similarly bounded by $\mu > 0$.*

Assumption 6. *Assume the regularity condition that for each output dimension $a \in [d]$, the row $\varepsilon_a \in \mathbb{R}^n$ of the noise matrix ε satisfies*

$$\mathbb{E}[\varepsilon_a\varepsilon_a^\top \mid P, Z] \preceq \sigma_\varepsilon^2 I_n,$$

i.e., conditioning on the weights does not inflate the noise's second moment. This holds at leading order when the weight–noise interaction is benign (for example under sample splitting, i.e. fitting on one subset of training data, then weighting and refitting on the complement).

A.6.3 Bounding MSE error

Lemma 7 (Parameter MSE rates)**.** *Under (D1)–(D5) and Assumption 1, 3, 4, 5, 6 with $\gamma_0 > 0$, there exist absolute constants $A_1^{\mathrm{unw}}, A_1^{\mathrm{rcif}}, A_1^{\mathrm{val}}, A_2^{\mathrm{unw}}, A_2^{\mathrm{rcif}}$ depending only on R and λM (with $M = R_r D R^2/\mu$ from Remark 5) such that*

$$\mathbb{E}\|\hat{\theta} - \theta^\star\|_F^2 \;\le\; A_1^{\mathrm{unw}}\,\frac{\sigma_\varepsilon^2 d}{n\mu^2} + A_2^{\mathrm{unw}}\,\frac{\alpha^2\beta^2R^2}{\mu^2}, \tag{11}$$

$$\mathbb{E}\|\hat{\theta}_p - \theta^\star\|_F^2 \;\le\; A_1^{\mathrm{rcif}}\,\frac{\sigma_\varepsilon^2 d}{n\mu^2} + A_2^{\mathrm{rcif}}\,\frac{\alpha^2\beta R^2\,\Phi(\lambda,m,n)}{\mu^2}, \tag{12}$$

$$\mathbb{E}\|\hat{\theta}_{\mathrm{val}} - \theta^\star\|_F^2 \;\le\; A_1^{\mathrm{val}}\,\frac{\sigma_\varepsilon^2 d}{m\mu^2}, \qquad (m \ge q), \tag{13}$$

with the standing relationships

$$A_1^{\mathrm{val}} \;=\; A_1^{\mathrm{unw}}, \qquad A_1^{\mathrm{rcif}} \;=\; e^{6\lambda M}\,A_1^{\mathrm{unw}}, \qquad A_2^{\mathrm{rcif}} \;=\; e^{6\lambda M}\,A_2^{\mathrm{unw}}.$$

Concretely $A_1^{\mathrm{unw}} = A_1^{\mathrm{val}} = 2R^2$ and $A_2^{\mathrm{unw}} = 2$.

Proof. Write $Y = \theta^\star Z + E + C$, where $Z = [z_1, \dots, z_n] \in \mathbb{R}^{q\times n}$ is the design, $E = [\varepsilon_1, \dots, \varepsilon_n] \in \mathbb{R}^{d\times n}$ collects the noise, and $C := -\alpha\, U\, \mathrm{diag}(1_{j\in\mathcal{C}}) \in \mathbb{R}^{d\times n}$ collects the corruption shifts with $U := [\hat\delta_1, \dots, \hat\delta_n]$. Throughout we use the elementary inequalities $\|AB\|_F \le \|A\|_F\, \|B\|_{\mathrm{op}}$ and $(a+b)^2 \le 2a^2 + 2b^2$.

Bound (11) **(unweighted).** The OLS estimator $\hat\theta = YZ^\top(ZZ^\top)^{-1}$ satisfies

$$\hat\theta - \theta^\star \;=\; (\bar\varepsilon + T_0)\,\hat\Sigma^{-1}, \qquad \bar\varepsilon := \tfrac{1}{n}EZ^\top,\; T_0 := \tfrac{1}{n}CZ^\top.$$

By (D1), $\|\hat\Sigma^{-1}\|_{\mathrm{op}} \le 1/\mu$ a.s. Then

$$\|\hat\theta - \theta^\star\|_F^2 \;\le\; \|\hat\Sigma^{-1}\|_{\mathrm{op}}^2 \cdot 2\big(\|\bar\varepsilon\|_F^2 + \|T_0\|_F^2\big) \;\le\; \frac{2}{\mu^2}\big(\|\bar\varepsilon\|_F^2 + \|T_0\|_F^2\big) \quad \text{a.s.}$$

Taking expectations:

- *Variance.* Each entry $(EZ^\top)_{ab} = \sum_i \varepsilon_{ai} z_{i,b}$ has variance $\le \sigma_\varepsilon^2 \sum_i z_{i,b}^2$ by mutual independence of noise. Summing over $a \in [d]$ and $b \in [q]$,

  $$\mathbb{E}\|EZ^\top\|_F^2 \;=\; d\,\sigma_\varepsilon^2 \sum_i \|z_i\|^2 \;\le\; d\,\sigma_\varepsilon^2\, nR^2,$$

  so $\mathbb{E}\|\bar\varepsilon\|_F^2 \le d\sigma_\varepsilon^2 R^2/n$, contributing $2R^2\,\sigma_\varepsilon^2 d/(n\mu^2)$ to the bound.
- *Bias.* By Theorem 6 (deterministic part), $\|T_0\|_F^2 \le \alpha^2\beta^2R^2$ a.s., contributing $2\alpha^2\beta^2R^2/\mu^2$.

Combining yields (11) with $A_1^{\mathrm{unw}} = 2R^2$ and $A_2^{\mathrm{unw}} = 2$.

Bound (12) **(RCIF).** The weighted estimator $\hat\theta_p = YPZ^\top(ZPZ^\top)^{-1}$ with $P = \mathrm{diag}(\tilde p_1, \dots, \tilde p_n)$ and $\tilde p_i := np_i$ (so $\sum_i \tilde p_i = n$) satisfies

$$\hat\theta_p - \theta^\star \;=\; (\bar\varepsilon_w + T_w)\,\hat\Sigma_P^{-1}, \qquad \bar\varepsilon_w := \tfrac{1}{n}EPZ^\top,\; T_w := \tfrac{1}{n}CPZ^\top,\; \hat\Sigma_P := \tfrac{1}{n}ZPZ^\top.$$

Lower bound on $\hat\Sigma_P$. Since $|\Psi_i| \le M$ a.s. $e^{\lambda\Psi_i} \in [e^{-\lambda M}, e^{\lambda M}]$ and $Q = \sum_i e^{\lambda\Psi_i} \le n\,e^{\lambda M}$. Hence the deterministic uniform lower bound

$$p_i = \frac{e^{\lambda\Psi_i}}{Q} \;\ge\; \frac{e^{-\lambda M}}{n\,e^{\lambda M}} \;=\; \frac{e^{-2\lambda M}}{n} \quad \text{a.s.},$$

i.e., $\tilde p_i \ge e^{-2\lambda M}$. Therefore

$$\hat\Sigma_P \;=\; \frac{1}{n}\sum_i \tilde p_i\, z_i z_i^\top \;\succeq\; \frac{e^{-2\lambda M}}{n}\sum_i z_i z_i^\top \;=\; e^{-2\lambda M}\hat\Sigma \;\succeq\; e^{-2\lambda M}\mu I,$$

and $\|\hat\Sigma_P^{-1}\|_{\mathrm{op}} \le e^{2\lambda M}/\mu$ a.s. Then

$$\|\hat\theta_p - \theta^\star\|_F^2 \;\le\; \frac{2e^{4\lambda M}}{\mu^2}\big(\|\bar\varepsilon_w\|_F^2 + \|T_w\|_F^2\big) \quad \text{a.s.}$$

Variance. Each entry $(EPZ^\top)_{ab} = \sum_i \varepsilon_{ai}\,\tilde p_i\, z_{i,b}$ has conditional variance $\le \sigma_\varepsilon^2 \sum_i \tilde p_i^2\, z_{i,b}^2$ given (P, Z). Summing and applying Assumption 6,

$$\mathbb{E}\big[\|EPZ^\top\|_F^2 \,\big|\, P, Z\big] \;\le\; d\,\sigma_\varepsilon^2 \sum_i \tilde p_i^2\, \|z_i\|^2 \;\le\; d\,\sigma_\varepsilon^2\, R^2 \sum_i \tilde p_i^2.$$

By $\sum_i \tilde p_i^2 \le (\max_i \tilde p_i)\sum_i \tilde p_i \le e^{2\lambda M}\cdot n$,

$$\mathbb{E}\|\bar\varepsilon_w\|_F^2 \;=\; \tfrac{1}{n^2}\mathbb{E}\|EPZ^\top\|_F^2 \;\le\; \frac{d\,\sigma_\varepsilon^2\, R^2\, e^{2\lambda M}}{n}.$$

Multiplying by $2e^{4\lambda M}/\mu^2$ contributes $2R^2\, e^{6\lambda M}\,\sigma_\varepsilon^2 d/(n\mu^2)$ to the bound.

Bias. By Theorem 6 (second-moment bound (9)), $\mathbb{E}\|T_w\|_F^2 \le \alpha^2\beta R^2\, e^{2\lambda M}\, \Phi(\lambda, m, n)$. Multiplying by $2e^{4\lambda M}/\mu^2$ contributes

$$\frac{2e^{6\lambda M}\,\alpha^2\beta R^2\,\Phi(\lambda, m, n)}{\mu^2}.$$

Combining yields (12) with $A_1^{\mathrm{rcif}} = 2R^2 e^{6\lambda M}$ and $A_2^{\mathrm{rcif}} = 2\,e^{6\lambda M}$, giving the standing relationships $A_1^{\mathrm{rcif}} = e^{6\lambda M}\,A_1^{\mathrm{unw}}$ and $A_2^{\mathrm{rcif}} = e^{6\lambda M}\,A_2^{\mathrm{unw}}$.

Bound (13) **(validation-only).** The validation-only estimator $\hat{\theta}_{\mathrm{val}}$ is OLS on the clean validation sample of size $m \ge q$. The unweighted argument applies verbatim with no corruption term and on the validation design, giving

$$\mathbb{E}\|\hat{\theta}_{\mathrm{val}} - \theta^\star\|_F^2 \;\le\; \frac{2}{\mu^2}\,\mathbb{E}\|\bar{\varepsilon}_{\mathrm{val}}\|_F^2 \;\le\; \frac{2R^2\,\sigma_\varepsilon^2 d}{m\mu^2},$$

which is (13) with $A_1^{\mathrm{val}} = 2R^2 = A_1^{\mathrm{unw}}$. □

Now we must translate the above result on parameter MSE into one about regret.

Lemma 8 (Regret–MSE bridge)**.** *Suppose Assumption 2 holds, the LP diameter $L := \sup_{x,x'\in\mathcal{X}} \|x - x'\|$ is finite, and the irreducible noise ε is σ_ε-sub-Gaussian (so $\mathbb{E}\|\varepsilon\|^2 \le d\,\sigma_\varepsilon^2$). For any estimator θ independent of the test sample $w = (z, c)$,*

$$\mathbb{E}[\mathrm{Regret}(\theta)] \;\le\; C_1\,\mathbb{E}\,\|\theta - \theta^\star\|_F^2 \;+\; C_2.$$

where $C_1 = 2L\omega R^2$ and $C_2 = 8L\omega\, d\,\sigma_\varepsilon^2$.

Proof. Write $\hat{c} := \theta z$, $\hat{x} := x^\star(\hat{c})$, $x^\star := x^\star(c)$, so that $c - \hat{c} = \varepsilon - (\theta - \theta^\star)z$ and $\|c - \hat{c}\| \le \|(\theta - \theta^\star)z\| + \|\varepsilon\|$.

From $\langle \hat{c}, \hat{x}\rangle \le \langle \hat{c}, x^\star\rangle$ (optimality of $\hat{x}$ at cost $\hat{c}$),

$$\mathrm{Regret} = \langle c - \hat{c},\, \hat{x} - x^\star\rangle + \langle \hat{c},\, \hat{x} - x^\star\rangle \le \|c - \hat{c}\| \cdot \|\hat{x} - x^\star\| \le L\,\|c - \hat{c}\|.$$

The bound is vacuously zero when $\hat{x} = x^\star$.

The event $\{\hat{x} \ne x^\star\}$ implies $\hat{c}$ and c lie in different optimality cones, hence $\|c - \hat{c}\| \ge \pi_v(c)$. Since π_v is the distance to a fixed closed union of hyperplanes, it is 1-Lipschitz, so

$$\pi_v(c) = \pi_v(\theta^\star z + \varepsilon) \ge \pi_v(\theta^\star z) - \|\varepsilon\| \ge 1/\omega - \|\varepsilon\|.$$

Combining with $\|c - \hat{c}\| \le \|(\theta - \theta^\star)z\| + \|\varepsilon\|$,

$$\{\hat{x} \ne x^\star\} \subseteq \big\{\|(\theta - \theta^\star)z\| + 2\|\varepsilon\| \ge 1/\omega\big\}.$$

Let $X := \|(\theta - \theta^\star)z\| + 2\|\varepsilon\|$. Then by the above calculations

$$\mathrm{Regret} \le L\,\|c - \hat{c}\|\,\mathbf{1}\{X \ge 1/\omega\} \le L\,X\,\mathbf{1}\{X \ge 1/\omega\} \le L\omega\,X^2,$$

using $x\,\mathbf{1}\{x \ge 1/\omega\} \le \omega x^2$ for $x \ge 0$. By $(a+b)^2 \le 2a^2 + 2b^2$, $X^2 \le 2\,\|(\theta - \theta^\star)z\|^2 + 8\|\varepsilon\|^2$. Taking expectation and using $\|(\theta - \theta^\star)z\| \le R\,\|\theta - \theta^\star\|_F$ together with $\mathbb{E}\|\varepsilon\|^2 \le d\,\sigma_\varepsilon^2$:

$$\mathbb{E}[\mathrm{Regret}] \le L\omega\big(2\,\mathbb{E}\|(\theta - \theta^\star)z\|^2 + 8\,\mathbb{E}\|\varepsilon\|^2\big) \le 2L\omega R^2\,\mathbb{E}\|\theta - \theta^\star\|_F^2 + 8L\omega\, d\,\sigma_\varepsilon^2.$$

which give C_1, C_2. □

A.6.4 Formal regret theorem statement

Theorem 7 (Validation-size regret regimes)**.** *Adopt Assumptions (D1)–(D5) and Assumptions 1-6. Let $C_1 := 2L\omega R^2$ and $C_2 := 8L\omega d\sigma_\varepsilon^2$ be the constants from Lemma 8, and let*

$A_1^{\mathrm{unw}}, A_1^{\mathrm{rcif}}, A_1^{\mathrm{val}}, A_2^{\mathrm{unw}}, A_2^{\mathrm{rcif}}$ be the constants from Lemma 7. Define the unconditional expected regret upper bounds

$$R_{\mathrm{unw}}(n) := C_1\left(A_1^{\mathrm{unw}}\,\frac{\sigma_\varepsilon^2 d}{n\mu^2} + A_2^{\mathrm{unw}}\,\frac{\alpha^2\beta^2R^2}{\mu^2}\right) + C_2,$$
$$R_w(\lambda,m,n) := C_1\left(A_1^{\mathrm{rcif}}\,\frac{\sigma_\varepsilon^2 d}{n\mu^2} + A_2^{\mathrm{rcif}}\,\frac{\alpha^2\beta R^2\,\Phi(\lambda,m,n)}{\mu^2}\right) + C_2,$$
$$R_{\mathrm{val}}(m) := C_1\left(A_1^{\mathrm{val}}\,\frac{\sigma_\varepsilon^2 d}{m\mu^2}\right) + C_2,$$

where $\Phi(\lambda,m,n)$ is the weight ratio bound in Lemma 6. Define the thresholds

$$m_\star(\lambda) := \inf\left\{m \ge 1 \,:\, \Phi(\lambda,m,n) < \beta\,\frac{A_2^{\mathrm{unw}}}{A_2^{\mathrm{rcif}}}\right\},$$
$$m^\star(\lambda) := \sup\left\{m \ge 1 \,:\, m\cdot\Phi(\lambda,m,n) < C\,\frac{A_2^{\mathrm{unw}}}{A_2^{\mathrm{rcif}}}\right\},\quad C := \frac{A_1^{\mathrm{unw}}\,\sigma_\varepsilon^2 d}{A_2^{\mathrm{unw}}\,\alpha^2\beta R^2}.$$

Then in the asymptotic regime $n\to\infty$ with m,β,α,λ held fixed, the following hold:

*(1) (*Lower bound on validation size: tighter expected regret bound for RCIF vs. unweighted ERM.*) For all $m > m_\star(\lambda)$ and n sufficiently large,*

$$R_w(\lambda,m,n) < R_{\mathrm{unw}}(n) + o(1).$$

The threshold $m_\star(\lambda)$ is finite whenever $e^{-\lambda\alpha\gamma_0/4} < \beta\,A_2^{\mathrm{unw}}/A_2^{\mathrm{rcif}}$, where γ_0 is the minimum-margin constant from Assumption 1. With the relationship $A_2^{\mathrm{rcif}} = e^{6\lambda M}A_2^{\mathrm{unw}}$ from Lemma 7, this reads $\lambda(\alpha\gamma_0/4 - 6M) > \ln(1/\beta)$, which holds for sufficiently large λ provided $\alpha\gamma_0 > 24M$.

*(2) (*Upper bound on validation size: tighter expected regret bound for RCIF vs. validation-only ERM.*) For all $m < m^\star(\lambda)$ and n sufficiently large,*

$$R_w(\lambda,m,n) < R_{\mathrm{val}}(m) + o(1).$$

The threshold $m^\star(\lambda)$ is finite, since the bias term in R_w stabilizes to a positive constant as $m\to\infty$ while $R_{\mathrm{val}}(m)\to C_2$.

*(3) (*Sweet spot.*) If*

$$\Phi(\lambda,m,n) < \beta\,\frac{A_2^{\mathrm{unw}}}{A_2^{\mathrm{rcif}}}\quad \textit{and}\quad m\cdot\Phi(\lambda,m,n) < C\,\frac{A_2^{\mathrm{unw}}}{A_2^{\mathrm{rcif}}},$$

then $m_\star(\lambda) < m^\star(\lambda)$, and for every $m\in(m_\star(\lambda),m^\star(\lambda))$ and n sufficiently large,

$$R_w(\lambda,m,n) < \min\big(R_{\mathrm{unw}}(n),\,R_{\mathrm{val}}(m)\big) + o(1).$$

Proof. By Lemma 8, for any of the three estimators $\hat\theta_\bullet \in \{\hat\theta, \hat\theta_p, \hat\theta_{\mathrm{val}}\}$,

$$\mathbb{E}[\mathrm{Regret}(\hat\theta_\bullet)] \le C_1\,\mathbb{E}\|\hat\theta_\bullet - \theta^\star\|_F^2 + C_2.$$

Combining with the method-specific MSE upper bounds of Lemma 7 yields the regret upper bounds $R_{\mathrm{unw}}, R_w, R_{\mathrm{val}}$ in the statement.

We use the following standing relationships among the constants, established in Lemma 7:

(C1) $A_1^{\mathrm{val}} = A_1^{\mathrm{unw}}$,

(C2) $A_1^{\mathrm{rcif}} = e^{6\lambda M}\,A_1^{\mathrm{unw}}$,

(C3) $A_2^{\mathrm{rcif}} = e^{6\lambda M}A_2^{\mathrm{unw}}$.

Part (1): comparison with $\hat\theta$.

$$R_w(\lambda,m,n) - R_{\mathrm{unw}}(n) = \underbrace{C_1(A_1^{\mathrm{rcif}} - A_1^{\mathrm{unw}})\,\frac{\sigma_\varepsilon^2 d}{n\mu^2}}_{=:(V)} + \underbrace{C_1\,\frac{\alpha^2R^2}{\mu^2}\Big(A_2^{\mathrm{rcif}}\,\beta\,\Phi(\lambda,m,n) - A_2^{\mathrm{unw}}\,\beta^2\Big)}_{=:(B)}.$$



Figure 2: Corruption rate and regret

The variance term $(V) = C_1(e^{6\lambda M} - 1)A_1^{\text{unw}}\sigma_\varepsilon^2 d/(n\mu^2) = O(1/n)$, hence absorbs into $o(1)$. The bias term (B) is strictly negative iff

$$A_2^{\text{rcif}}\,\beta\,\Phi(\lambda, m, n) < A_2^{\text{unw}}\,\beta^2 \iff \Phi(\lambda, m, n) < \beta\,\frac{A_2^{\text{unw}}}{A_2^{\text{rcif}}}.$$

By Lemma 6, $\Phi(\lambda, m, n) \to e^{-\lambda\alpha\gamma_0/4} + o(1)$ as $m \to \infty$, so the set $\{m : \Phi(\lambda, m, n) < \beta A_2^{\text{unw}}/A_2^{\text{rcif}}\}$ is nonempty whenever $e^{-\lambda\alpha\gamma_0/4} < \beta A_2^{\text{unw}}/A_2^{\text{rcif}}$, and $m_\star(\lambda)$ is well-defined. For all $m > m_\star(\lambda)$ and n sufficiently large, the comparison holds strictly, giving conclusion (1) up to $o(1)$.

Part (2): comparison with $\hat{\theta}_{\text{val}}$. Using $A_1^{\text{val}} = A_1^{\text{unw}}$ from (C1),

$$R_w(\lambda, m, n) - R_{\text{val}}(m) = C_1 A_1^{\text{rcif}}\,\frac{\sigma_\varepsilon^2 d}{n\mu^2} + C_1 A_2^{\text{rcif}}\,\frac{\alpha^2\beta R^2\,\Phi(\lambda, m, n)}{\mu^2} - C_1 A_1^{\text{unw}}\,\frac{\sigma_\varepsilon^2 d}{m\mu^2}.$$

For n large enough that $A_1^{\text{rcif}}\sigma_\varepsilon^2 d/n$ is dominated by the other two terms (absorbed into $o(1)$), the comparison reduces to

$$A_2^{\text{rcif}}\,\frac{\alpha^2\beta R^2\,\Phi(\lambda, m, n)}{\mu^2} < A_1^{\text{unw}}\,\frac{\sigma_\varepsilon^2 d}{m\mu^2}.$$

Multiplying both sides by $m\mu^2/A_1^{\text{unw}}$, this is equivalent to

$$m \cdot \Phi(\lambda, m, n) < C\,\frac{A_2^{\text{unw}}}{A_2^{\text{rcif}}}, \quad C := \frac{A_1^{\text{unw}}\,\sigma_\varepsilon^2 d}{A_2^{\text{unw}}\,\alpha^2\beta R^2}.$$

By definition this holds for all $m < m^\star(\lambda)$. To verify $m^\star(\lambda)$ is finite: by Lemma 6, Φ stabilizes to the positive limit $e^{-\lambda\alpha\gamma_0/4}$ as $m \to \infty$, so $f(m) := m \cdot \Phi(\lambda, m, n)$ grows at least linearly in m for large m and eventually exceeds the constant $C\,A_2^{\text{unw}}/A_2^{\text{rcif}}$.

Part (3): sweet spot. Direct from Parts (1) and (2). □

B Additional experiments

B.1 Corruption rate study shortest path

We vary the corruption rate β across [0.1,0.5], observing that there is a similar tradeoff in corruption rate for $\hat{\theta}_p$ vs $\hat{\theta}_{\text{val}}$ as seen in Figure 2. As corruption rate increases, both unweighted and iterative exp. directed suffer in performance, while validation only remains constant. In particular, $\beta = 0.3$ appears to be the crossover point. This is to be expected, as once the dataset becomes sufficiently noisy, it is difficult to fit on, regardless of reweighting or not, but we do observe that our method consistently retains lower regret compared to the unweighted ERM. These experiments were run on shortest path problems with fixed parameters over 50 seeds: 4 $\times$4 grid, upstream input dimension 30, 400 training points, 100 validation points, 200 test points. The full table is given below in Table 3

B.2 Validation study shortest path full details

We provide the performance of all methods across varying validation set sizes, not just the unweighted and validation only baselines. We see that RC Weight very slightly outperforms the unweighted

Method	$\beta = 0.1$	$\beta = 0.2$	$\beta = 0.3$	$\beta = 0.5$
Iterative Exp. Directed	**0.9089 ± 0.0172**	**0.9437 ± 0.0172**	0.9992 ± 0.0211	1.0275 ± 0.0160
Exp. Two-Sided	0.9528 ± 0.0172	1.0009 ± 0.0175	1.0580 ± 0.0191	1.0817 ± 0.0208
Directed	0.9357 ± 0.0170	0.9825 ± 0.0171	1.0433 ± 0.0189	1.0568 ± 0.0187
Unweighted	0.9573 ± 0.0170	1.0029 ± 0.0169	1.0610 ± 0.0194	1.0784 ± 0.0208
RC Weight	0.9575 ± 0.0171	1.0035 ± 0.0169	1.0604 ± 0.0192	1.0780 ± 0.0210
Validation Only	0.9989 ± 0.0181	0.9989 ± 0.0181	**0.9989 ± 0.0181**	**0.9989 ± 0.0181**
SPO Influence Weight	0.9212 ± 0.0172	0.9560 ± 0.0158	0.9998 ± 0.0191	1.0307 ± 0.0209
Clean ERM (oracle)	0.8513 ± 0.0171	0.8513 ± 0.0171	0.8513 ± 0.0171	0.8513 ± 0.0171
SPO+ Val Only	1.0914 ± 0.0182	1.0914 ± 0.0182	1.0914 ± 0.0182	1.0914 ± 0.0182
SPO+ Finetune	1.0015 ± 0.0184	1.0042 ± 0.0176	1.0160 ± 0.0178	1.0233 ± 0.0190

Table 3: Mean test regret as a function of corruption rate β. Best per column in bold.

baseline in performance particularly at smaller validation sizes, whereas our method excels starting in the medium validation set size range, with a larger margin. Results were obtained on the shortest path setting with 100 seeds on a 4×4 grid with upstream input dimension 20, 500 training points, 200 test points, and $\beta = 0.2$, using the base setting from Table 1. We also provide the same setting

Method	|V| = 5	|V| = 10	|V| = 20	|V| = 50	|V| = 100	|V| = 150	|V| = 200
Iterative Exp. Directed	1.0595 ± 0.0140	1.0051 ± 0.0127	0.9707 ± 0.0123	0.9354 ± 0.0120	**0.9088 ± 0.0107**	**0.8751 ± 0.0107**	**0.8688 ± 0.0111**
Exp. Two-Sided	1.0050 ± 0.0109	0.9935 ± 0.0116	0.9789 ± 0.0125	0.9586 ± 0.0135	0.9458 ± 0.0118	0.9228 ± 0.0117	0.9127 ± 0.0124
Directed	1.0015 ± 0.0129	0.9756 ± 0.0111	0.9659 ± 0.0115	0.9520 ± 0.0124	0.9389 ± 0.0117	0.9071 ± 0.0116	0.8964 ± 0.0121
Unweighted	0.9857 ± 0.0119	0.9736 ± 0.0113	0.9714 ± 0.0119	0.9627 ± 0.0125	0.9516 ± 0.0122	0.9282 ± 0.0121	0.9184 ± 0.0124
RC Weight	**0.9855 ± 0.0120**	**0.9732 ± 0.0112**	0.9713 ± 0.0117	0.9634 ± 0.0126	0.9515 ± 0.0122	0.9280 ± 0.0120	0.9179 ± 0.0124
Validation Only	1.1854 ± 0.0137	1.1710 ± 0.0145	1.2471 ± 0.0155	1.0366 ± 0.0142	0.9563 ± 0.0105	0.8940 ± 0.0115	0.8871 ± 0.0107
SPO Influence Weight	1.0009 ± 0.0111	0.9771 ± 0.0124	**0.9506 ± 0.0113**	**0.9257 ± 0.0123**	0.9185 ± 0.0108	0.8864 ± 0.0113	0.8813 ± 0.0115
Clean ERM (oracle)	0.8256 ± 0.0094	0.8265 ± 0.0088	0.8279 ± 0.0108	0.8255 ± 0.0104	0.8129 ± 0.0094	0.8068 ± 0.0107	0.8178 ± 0.0108
SPO+ Val Only	1.1854 ± 0.0137	1.1710 ± 0.0145	1.2471 ± 0.0155	1.1180 ± 0.0133	1.0386 ± 0.0133	0.9778 ± 0.0133	0.9567 ± 0.0117
SPO+ Finetune	0.9982 ± 0.0121	1.0064 ± 0.0124	1.0234 ± 0.0130	0.9997 ± 0.0119	0.9657 ± 0.0111	0.9356 ± 0.0116	0.9248 ± 0.0106

Table 4: Mean test regret as a function of validation set size $|\mathcal{V}|$. Best per column in bold.

with $\beta = 0.3$ for Figure 1, where we demonstrate the crossing over point of Validation Only which is readily apparent at the higher corruption rate.

Method	|V| = 5	|V| = 10	|V| = 20	|V| = 50	|V| = 100	|V| = 150
Iterative Exp. Directed	1.0964 ± 0.0146	1.0600 ± 0.0132	1.0254 ± 0.0126	**0.9745 ± 0.0117**	**0.9334 ± 0.0107**	0.9120 ± 0.0121
Unweighted	**1.0234 ± 0.0133**	**1.0159 ± 0.0113**	**1.0137 ± 0.0116**	0.9995 ± 0.0122	0.9873 ± 0.0114	0.9719 ± 0.0120
Validation Only	1.1854 ± 0.0137	1.1710 ± 0.0145	1.2471 ± 0.0155	1.0366 ± 0.0142	0.9563 ± 0.0105	**0.8940 ± 0.0115**

Table 5: Mean test regret values across $|\mathcal{V}|$ for Fig. 1

B.3 Corruption rate study for random LP

We provide a similar study on corruption rate, now with the random LP downstream. The gaps between the reweighting methods are small: Iterative Exp. Directed and SPO Influence Weight (paired $p \geq 0.18$), and both improve on unweighted ERM by a modest margin that is significant at $\beta = 0.3$ ($p = 0.016$). Validation Only is constant across β and attains the lowest mean regret at $\beta = 0.5$ as much of the data has been corrupted, but the difference to our method is within noise ($p = 0.42$); the clean-ERM ceiling shows that the total damage from corruption is small in this setting (0.025 to 0.08 regret), which bounds what any reweighting can recover. The SPO+ baselines do not help here: fine-tuning on the validation set is worse than unweighted ERM at every β, and training on the validation set alone is far worse. Results were obtained with upstream input dimension 10, LP dimension 10, 20 constraints, train set size 200, validation set size 50, test set size 100, and 100 seeds.

B.4 Validation size study for random LP

We provide a similar study on validation set size, but for random LP downstream. Like we have previously observed, the impact of RCIF reweighting is less apparent in this setting (although still present). In particular, it appears that the range of validation set sizes for which the RCIF scoring

Method	β = 0.1	β = 0.2	β = 0.3	β = 0.5
Iterative Exp. Directed	0.3746 ± 0.0176	0.3865 ± 0.0180	0.3978 ± 0.0195	0.4230 ± 0.0181
Exp. Two-Sided	0.3774 ± 0.0181	0.3893 ± 0.0182	0.4037 ± 0.0185	0.4331 ± 0.0180
Directed	0.3771 ± 0.0182	0.3896 ± 0.0180	0.4010 ± 0.0187	0.4341 ± 0.0184
Unweighted	0.3764 ± 0.0181	0.3896 ± 0.0186	0.4053 ± 0.0183	0.4333 ± 0.0179
RC Weight	0.3772 ± 0.0186	0.3906 ± 0.0190	0.4046 ± 0.0184	0.4333 ± 0.0178
Validation Only	0.4173 ± 0.0197	0.4173 ± 0.0197	0.4173 ± 0.0197	$\mathbf{0.4173 \pm 0.0197}$
SPO Influence Weight	$\mathbf{0.3709 \pm 0.0176}$	$\mathbf{0.3863 \pm 0.0181}$	$\mathbf{0.3957 \pm 0.0193}$	0.4251 ± 0.0192
Clean ERM (oracle)	0.3513 ± 0.0157	0.3513 ± 0.0157	0.3513 ± 0.0157	0.3513 ± 0.0157
SPO+ Val Only	0.6970 ± 0.0337	0.6970 ± 0.0337	0.6970 ± 0.0337	0.6970 ± 0.0337
SPO+ Finetune	0.4357 ± 0.0184	0.4455 ± 0.0201	0.4407 ± 0.0188	0.4694 ± 0.0210

Table 6: Mean test regret as a function of corruption rate β on the random LP setting. Best per column in bold.

performs better than unweighted is smaller, which is likely due to the lack of structure in the downstream problem. Results were obtained with upstream input dimension 10, LP dimension 10, 20 constraints, train set size 200, test set size 100, and corruption rate $\beta = 0.5$.

Method	|V| = 5	|V| = 10	|V| = 20	|V| = 50	|V| = 100	|V| = 150
Iterative Exp. Directed	0.6033 ± 0.0331	0.5311 ± 0.0258	0.4640 ± 0.0221	0.4230 ± 0.0181	0.3693 ± 0.0185	0.3632 ± 0.0160
Exp. Two-Sided	0.4774 ± 0.0255	0.4704 ± 0.0225	$\mathbf{0.4491 \pm 0.0209}$	0.4331 ± 0.0180	0.3848 ± 0.0183	0.3758 ± 0.0164
Directed	0.4766 ± 0.0251	0.4667 ± 0.0227	0.4498 ± 0.0207	0.4341 ± 0.0184	0.3835 ± 0.0183	0.3737 ± 0.0165
Unweighted	$\mathbf{0.4576 \pm 0.0232}$	0.4618 ± 0.0219	0.4503 ± 0.0203	0.4333 ± 0.0179	0.3830 ± 0.0183	0.3752 ± 0.0165
RC Weight	0.4596 ± 0.0236	$\mathbf{0.4609 \pm 0.0218}$	0.4501 ± 0.0208	0.4333 ± 0.0178	0.3838 ± 0.0184	0.3750 ± 0.0168
Validation Only	9.8179 ± 0.4086	7.9994 ± 1.3074	0.6616 ± 0.0284	$\mathbf{0.4173 \pm 0.0197}$	$\mathbf{0.3570 \pm 0.0190}$	$\mathbf{0.3602 \pm 0.0159}$
SPO Influence Weight	0.4712 ± 0.0243	0.4722 ± 0.0223	0.4593 ± 0.0204	0.4251 ± 0.0192	0.3757 ± 0.0185	0.3690 ± 0.0161
Clean ERM (oracle)	0.3447 ± 0.0165	0.3511 ± 0.0152	0.3451 ± 0.0163	0.3513 ± 0.0157	0.3357 ± 0.0175	0.3437 ± 0.0151
SPO+ Val Only	9.8179 ± 0.4086	7.9994 ± 1.3074	0.7749 ± 0.0360	0.6970 ± 0.0337	0.6588 ± 0.0296	0.5773 ± 0.0266
SPO+ Finetune	0.4747 ± 0.0233	0.4775 ± 0.0219	0.4878 ± 0.0219	0.4694 ± 0.0210	0.4141 ± 0.0188	0.4053 ± 0.0185

Table 7: Mean test regret as a function of validation set size $|\mathcal{V}|$ on the random LP setting. Best per column in bold.

B.5 Adversarial noise

Beyond the isotropic Gaussian corruption considered in the main experiments, we also evaluate our method against a more targeted adversary that uses minimum-norm perturbations to flip the optimality cone of corrupted training points. This corruption model is particularly challenging for influence-based methods because it is designed to induce wrong downstream decisions while remaining as close as possible to the clean label.

Construction. For a corrupted index $j \in \mathcal{C}$ with clean cost c_j, let B_j be the optimal basis for c_j and let $N(B_j)$ denote the corresponding nonbasic set. The minimum ℓ_2-norm perturbation that pushes c_j across the boundary of $\Lambda(B_j)$ along nonbasic direction k has magnitude $s_k(c_j, B_j)/\|\delta(B_j, k)\|$. The cheapest direction to flip is

$$k_j^\dagger \;=\; \arg\min_{k \in N(B_j)} \frac{s_k(c_j, B_j)}{\|\delta(B_j, k)\|}.$$

Given a budget $\rho > 0$, the adversarial corruption is

$$\tilde{c}_j \;=\; c_j - \alpha_j \frac{\delta(B_j, k_j^\dagger)}{\|\delta(B_j, k_j^\dagger)\|}, \qquad \alpha_j = \min\left(\rho,\; \frac{(1+\varepsilon)\, s_{k_j^\dagger}(c_j, B_j)}{\|\delta(B_j, k_j^\dagger)\|}\right),$$

with a small slack $\varepsilon > 0$ to ensure the perturbation strictly crosses the boundary when budget allows, which is similar to the directional attack in Assumption D4, but with a variable magnitude. If ρ is insufficient to flip any direction, we use the full budget along the cheapest flip direction without crossing.

Note that the perturbation magnitude is *minimal* rather than fixed: the adversary uses just enough perturbation to cross a cone boundary, producing small residuals at the clean ERM estimator and weak first-order influence signals.

Empirical observations. Under cone-flipping corruption, all methods including ours perform comparably to unweighted ERM, with no method showing a significant improvement over unweighted (Table 8), perhaps due to the magnitude of the corruption being small so they perform similarly to unweighted. Diagnostic analysis confirms that the RCIF score does carry information distinguishing corrupted from clean points: the mean score over corrupted indices is consistently more negative than over clean indices but the separation is insufficient for reliable detection. We attribute this to a fundamental property of first-order influence methods: the influence of a training point scales with its perturbation magnitude, and minimum-norm adversarial corruption is designed precisely to minimize this signal. Methods based on residual magnitude or influence function approximations are therefore inherently limited against such adversaries, and we expect that defenses against cone-flipping corruption require either second-order influence information or outlier-detection such as Diakonikolas et al. [2019]. Results were obtained with both random LP and shortest path problems. For random LP, settings were upstream input dimension 10, LP dimension 10, 20 constraints, train set size 200, test set size 100, validation set size 50, $\beta = 0.3$, $\rho = 0.3$. For shortest path settings: grid 5×5, upstream input dimension 30, train set size 300, test set size 200, validation set size 100, $\beta = 0.3$, $\rho = 0.05$. ρ is much smaller for shortest path because the optimality cones appear to be narrower, and thus require less perturbation to flip.

Method	Random LP	Shortest Path
Iterative Exp. Directed	0.3527 ± 0.0158	1.2412 ± 0.0204
Exp. Two-Sided	$\mathbf{0.3514 \pm 0.0155}$	1.2430 ± 0.0196
Directed	0.3514 ± 0.0155	1.2424 ± 0.0195
Unweighted	0.3522 ± 0.0156	1.2412 ± 0.0194
RC Weight	0.3519 ± 0.0155	$\mathbf{1.2410 \pm 0.0194}$
Validation Only	0.4173 ± 0.0197	1.4459 ± 0.0177
SPO Influence Weight	0.3524 ± 0.0157	1.2410 ± 0.0199
Clean ERM (oracle)	0.3513 ± 0.0157	1.2422 ± 0.0196
SPO+ Val Only	0.6970 ± 0.0337	1.5713 ± 0.0207
SPO+ Finetune	0.4264 ± 0.0179	1.3983 ± 0.0203

Table 8: Mean test regret under adversarial corruption (minimal magnitude). Best per column in bold.

B.5.1 Adversarial noise with fixed magnitude and robust regression

Residual-based robust regression is the natural alternative to reweighting when training labels are corrupted, so we compare against two standard instances: Huber regression (threshold 1.35, fit per output coordinate) and iteratively trimmed least squares (three rounds, each refitting after discarding the training points with the largest residual norms). Both are fit on the same train-plus-validation pool as unweighted ERM. For trimmed ERM we set the trim fraction equal to the true corruption rate β, which is an oracle choice that favors the baseline. We evaluate under two corruption models of the same magnitude scale: the isotropic Gaussian corruption of the main experiments and the directional corruption of Assumption (D4), in which each corrupted label is shifted by a fixed amount α along the most fragile direction of its own optimality cone. This is the same adversarial noise direction as Appendix B.5 but with a *fixed* magnitude, not minimal. Both are run on the shortest-path problem (4×4 grid, input dimension 20, $400/100/200$ train/validation/test points, 50 seeds), with $\beta = 0.2$ for the Gaussian setting and $\beta = 0.4$, $\alpha = 1.5$ for the directional setting. Table 9 reports the results under various types of noise.

Under isotropic Gaussian corruption the robust methods are strong, and trimmed ERM is the best method in every column: it attains 0.8544 ± 0.0141 against 0.9085 ± 0.0164 for Iterative Exp. Directed and 0.9600 ± 0.0158 for unweighted ERM, and it also improves on our method similarly for random LP (paired Wilcoxon $p < 0.001$ in each case) while Huber loss is consistent but shows the greatest gains under random LPs. We note that this is the regime these methods are designed for, and furthermore these methods have been given the corruption rate, an advantage our method nor any other method receives. Gaussian corruption of scale 2 produces residuals that are large relative to the

Method	LP, Gauss. $\beta = 0.3$	Shortest, Gauss. $\beta = 0.2$	Shortest, Directional $\beta = 0.4$
Iterative Exp. Directed	0.3451 ± 0.0133	0.9085 ± 0.0164	$\mathbf{0.8509} \pm 0.0164$
Directed	0.3531 ± 0.0129	0.9459 ± 0.0168	0.8547 ± 0.0161
Unweighted	0.3564 ± 0.0130	0.9600 ± 0.0158	0.8591 ± 0.0164
Clean ERM (oracle)	0.2785 ± 0.0109	0.8257 ± 0.0152	0.8257 ± 0.0152
Huber	0.2952 ± 0.0108	0.9266 ± 0.0163	0.8730 ± 0.0156
Trimmed ERM	$\mathbf{0.2865} \pm 0.0118$	$\mathbf{0.8544} \pm 0.0141$	0.8977 ± 0.0158

Table 9: Comparison against robust regression methods; under standard Gaussian isotropic noise they outperform every method, but under directional shifts greatly underperform the unweighted approach, requiring specific noise structure.

label noise, so a corrupted point is visible in residual norm at the unweighted fit, and trimming at the true rate removes most of them; Huber regression down-weights the same points more gently and lands between unweighted ERM and our method on the shortest-path settings (0.9266 ± 0.0163).

Under directional corruption the ordering reverses. The shift $\alpha = 1.5$ is small relative to the label noise and does not stand out in residual norm, but it is aimed at the fragile direction of the cone and flips the optimal decision of every corrupted point. Trimmed ERM now discards clean points as often as corrupted ones and reaches 0.8977 ± 0.0158, worse than unweighted ERM (0.8591 ± 0.0164) on 84% of seeds ($p < 0.001$); Huber reaches 0.8730 ± 0.0156, also worse than unweighted ERM ($p < 0.001$). Iterative Exp. Directed is better at 0.8509 ± 0.0164, improving on unweighted ERM ($p = 0.036$) and on both robust baselines ($p < 0.001$ for each). The total damage is small in this setting (clean ERM attains 0.8257 ± 0.0152), which compresses all gaps, but the sign of the effect is what matters: the robust losses lose ground precisely when the corruption is decision-targeted rather than magnitude-visible.

The two experiments isolate different notions of robustness. Residual-based methods are robust to corruption magnitude and require that corrupted labels be outliers in the regression sense, together with a well-specified model and, for trimming, knowledge of the corruption rate and noise type. RCIF reweighting is robust to decision-targeted corruption: it scores a training point by its first-order effect on the reduced-cost margins at the clean validation points, which responds to where a label pushes the downstream decision rather than to how far it moves the label, and it requires neither the corruption rate nor magnitude-visible residuals. We regard the two approaches as complementary, and combining RCIF scores with a robust training loss is a natural direction for future work.

B.6 Degeneracy

Note that shortest-path and combinatorial problems generally violate primal nondegeneracy. For LPs with generic continuous b and full-row-rank A (such as our random-LP setting, or pricing/portfolio-style problems with continuous coefficients), degeneracy is a measure-zero event, so the assumption holds almost surely. More importantly, the assumption serves the theory rather than the algorithm: the method still functions and applies to degenerate problems, at the cost of some conservatism.

The algorithm itself never uses nondegeneracy: it solves one LP per validation point under the *predicted* cost, takes whichever optimal basis the solver returns, and determines correctness via the cone test $s_k(c_v, B_v) \geq 0 \;\; \forall k \in N(B_v)$, which is well defined under degeneracy. The cost of degeneracy is conservatism: the containment $\Lambda(B_v) \subseteq \mathcal{N}(x^*(\hat{c}_v))$ becomes strict, so the algorithm may mark a correct decision as incorrect when c_v certifies the same vertex through a different basis, and the computed margin is a lower bound on the true decision margin. We note that (i) the correctness certificate never certifies a truly suboptimal point; and (ii) for the directed variant, such a misclassification is harmless at the point in question, because the violated directions there are all degenerate pivots. Its cost is signal efficiency, since part of the validation signal is spent on basis bookkeeping rather than on actual decisions. Empirically, our gains are largest on shortest path, which we take as evidence that the conservative signal remains strong where the assumption fails

B.7 Experimental details

B.7.1 Hardware and software

All experiments were conducted on a personal laptop equipped with an Intel i7 CPU, 16GB RAM, and no dedicated GPU. Implementation required standard Python packages such as PyTorch, scikit-learn, as well as Gurobipy and Gurobi software for the optimization aspects. Each seed (which involves all baselines and optimization calls) runs in approximately 4 minutes. Code for reproducing our experiments is available at: https://github.com/william-zhang64/Reduced-Cost-Influence-Functions.

B.7.2 Parameter and baseline settings

Throughout the experiments, we kept λ fixed at 0.05. The values $\lambda = 0.05$ and $\tau = 5$ were selected during preliminary experimentation across a grid of values $\{0.01, 0.05, 0.1, 0.5\}$ and $\{3, 5, 10\}$ respectively, with a plateau in performance for $\lambda \in \{0.05, 0.1\}$. Qualitatively, larger λ produced overly extreme weights and degraded test regret, consistent with overfitting the small clean validation set, while smaller λ left the weights too close to uniform to counteract the corruption. The number of iterations behaves similarly: beyond roughly $\tau = 5$ we observed a plateau in performance. Given that the pair $(\lambda, \tau) = (0.05, 5)$ was used across the settings in the paper for strong performance, this suggests these values serve as reasonable defaults. Note that the weights were normalized such that unweighted had each weight = 1.0, so the total weights sum to n rather than 1, which does not substantively affect any properties of the algorithm. Similarly, the ϕ parameter for the Directed method was also fixed to 0.05, for similar reasons. We used $\tau = 5$ iterations across the experiments for the Iterative Exp. Directed method. For SPO influence, we set the number of iterations to 5 as well, with learning rate of 0.01. For SPO+ Val Only we run full-batch subgradient descent on the SPO+ loss over the validation set for 100 epochs with normalized subgradients and diminishing step size $0.5/\sqrt{t+1}$, initialized at the validation-only least-squares fit, and keep the iterate with the lowest SPO+ objective. SPO+ Finetune uses the same procedure initialized at the unweighted least-squares fit on the train-plus-validation pool, with 30 epochs and step size $0.1/\sqrt{t+1}$. Huber regression uses threshold 1.35 and is fit per output coordinate on the train-plus-validation pool, and trimmed ERM runs three rounds, each refitting after discarding the fraction β of training points with the largest residual norms. For the random LP data-generating process, we used standard normal distribution for features. For the isotropic noise, we used a covariance matrix of $4\mathcal{I}_d$ and for the directional noise with fixed magnitude we used $\alpha = 1.5$. For the neural network, we used hidden layer of size 8, with learning rate 1e-3.

B.7.3 Computational cost

Each iteration of RCIF reweighting solves the downstream LP once at each of the m validation points, computes the influence functions of the n training points, and refits the predictor by weighted least squares. The LP is never differentiated through and is not solved at training points. For the linear predictor with squared loss the influence has the closed form of Lemma 1: the Hessian is $2I_d \otimes \hat{\Sigma}$, so one $q \times q$ factorization is shared by all training and validation points. Forming $\hat{\Sigma}$, its inverse, and the residuals $r_j = \hat{\theta} z_j - \tilde{c}_j$ costs $O(nq(q+d) + q^3)$ per iteration. At a validation point v with direction δ_k, the scores of all n training points are the elementwise product of $\delta_k^\top r_j$, an $n \times d$ matrix-vector product, and $\kappa(z_v, z_j) = z_j^\top \hat{\Sigma}^{-1} z_v$, an $n \times q$ one, so scoring over the validation set costs $O(mn(q+d))$; the weighted least-squares refit on the $n+m$ points is absorbed into the same terms. The LP solves and basis information cost $O(m\, T_{\mathrm{LP}})$ per iteration, where T_{LP} is the cost of one simplex solve, and the iterative variant repeats all of this τ times. In practice the τm LP solves dominate: on a laptop CPU with Gurobi, in the shortest-path setting of Table 1 ($n = 400$, $m = 100$), one scoring round takes about 0.5 s and the complete Iterative Exp. Directed fit with $\tau = 5$ about 2 s.

B.7.4 SPO Influence

We adapt the influence-based reweighting framework of Ren et al. [2018] to the predict-then-optimize setting by substituting the SPO+ surrogate loss [Elmachtoub and Grigas, 2022] for validation MSE. This isolates the contribution of LP cone geometry in our scoring from the general benefit of influence-based reweighting on a downstream-aware loss.

For predicted cost $\hat{c}_v = f_{\hat{\theta}}(z_v)$ and true cost c_v, SPO+ is a convex upper bound on regret with subgradient

$$g(\hat{c}_v, c_v) \;=\; 2\big(x^*(c_v) - x^*(2\hat{c}_v - c_v)\big),$$

where $x^*(c) = \arg\min_{x \in \mathcal{X}} c^\top x$. Composing with the influence function $\mathcal{I}_{\hat{\theta}}(w_j)$ gives the per-training-point sensitivity of validation loss:

$$\Sigma_j^{\text{SPO+}} \;=\; \frac{1}{m} \sum_{w_v \in \mathcal{V}} g(\hat{c}_v, c_v)^\top \, \nabla_\theta f_{\hat{\theta}}(z_v) \, \mathcal{I}_{\hat{\theta}}(w_j).$$

Starting from uniform weights $p_j^{(0)} = 1$, we iteratively update

$$p_j^{(t+1)} \;\leftarrow\; \max\big(0,\; p_j^{(t)} - \xi \, \Sigma_j^{\text{SPO+}}\big|_{\hat{\theta}^{(t)}}\big),$$

normalize so that $\sum_j p_j^{(t+1)} = n$, and refit $\hat{\theta}^{(t+1)}$ via weighted ERM. We use $\xi = 0.01$ and run for $\tau = 5$ iterations.

This baseline matches RCIF in data usage and influence machinery; the methods differ only in the projection direction for the prediction shift: $g(\hat{c}_v, c_v)$ for SPO+ versus the simplex direction $\delta(B_v, k)$ for RCIF.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: The abstract and introductions contain claims made in the paper and reflect the theoretical (Section 6) and experimental results (Section 7), and we note many assumptions in the main text and appendix necessary for theoretical results to hold and experimental settings.

   Guidelines:

   - The answer [N/A] means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A [No] or [N/A] answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: Our theoretical results clearly discuss assumptions (Appendix A), and we note how our empirical performance does or does not extend well to certain settings, such as deep neural networks (Discussed briefly in Section 7), and we directly state the nondegeneracy assumption and leave relaxing it to future work.

   Guidelines:

   - The answer [N/A] means that the paper has no limitation while the answer [No] means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

   Answer: [Yes]

   Justification: We provide assumptions throughout the paper and although provide informal results in the main body, the full statement, proof, constants and assumptions are in Appendix A.

   Guidelines:

   - The answer [N/A] means that the paper does not include theoretical results.
   - All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
   - All assumptions should be clearly stated or referenced in the statement of any theorems.
   - The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
   - Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
   - Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

   Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

   Answer: [Yes]

   Justification: For each experimental result in Section 7, we document the detailed settings needed to reproduce them, and our algorithm details are easily implementable based on the main text.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - If the paper includes experiments, a [No] answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
   - If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
   - Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
   - While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
     (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
     (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
     (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).

(d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

   Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

   Answer: [Yes]

   Justification: We release the code in a public repository.

   Guidelines:

   - The answer [N/A] means that paper does not include experiments requiring code.
   - Please see the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - While we encourage the release of code and data, we understand that this might not be possible, so [No] is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
   - The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://neurips.cc/public/guides/CodeSubmissionPolicy`) for more details.
   - The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
   - The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
   - At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
   - Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

   Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer) necessary to understand the results?

   Answer: [Yes]

   Justification: In Appendix B, we provide experimental details, hyperparameter settings, and software details.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
   - The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

   Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

   Answer: [Yes]

   Justification: All of our experiments in the paper include standard deviation from averaging over at least 50 seeds.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.

- The authors should answer [Yes] if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.
- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g., negative error rates).
- If error bars are reported in tables or plots, the authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

   Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

   Answer: [Yes]

   Justification: We describe our hardware specs and runtime in Appendix B and note that the experiments do not require any special equipment or resources.

   Guidelines:

   - The answer [N/A] means that the paper does not include experiments.
   - The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
   - The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
   - The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

   Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

   Answer: [Yes]

   Justification: We align with the NeurIPS Code of Ethics.

   Guidelines:

   - The answer [N/A] means that the authors have not reviewed the NeurIPS Code of Ethics.
   - If the authors answer [No], they should explain the special circumstances that require a deviation from the Code of Ethics.
   - The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

    Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

    Answer: [N/A]

Justification: This work studies a general algorithmic component for predict-then-optimize; societal impacts would depend on downstream deployment context.

Guidelines:

- The answer [N/A] means that there is no societal impact of the work performed.
- If the authors answer [N/A] or [No], they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate Deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pre-trained language models, image generators, or scraped datasets)?

Answer: [N/A]

Justification: The paper poses no such risks.

Guidelines:

- The answer [N/A] means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification: We use the PyEPO package which we cite.

Guidelines:

- The answer [N/A] means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.

- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

    Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

    Answer: [N/A]

    Justification: We do not release new assets.

    Guidelines:

    - The answer [N/A] means that the paper does not release new assets.
    - Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
    - The paper should discuss whether and how consent was obtained from people whose asset is used.
    - At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

    Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing or human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
    - Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
    - According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

    Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

    Answer: [N/A]

    Justification: The paper does not involve crowdsourcing nor human subjects.

    Guidelines:

    - The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.

- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

    Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does *not* impact the core methodology, scientific rigor, or originality of the research, declaration is not required.

    Answer: [N/A]

    Justification: The paper does not involve LLMs as any important, original, or non-standard components.

    Guidelines:

    - The answer [N/A] means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
    - Please refer to our LLM policy in the NeurIPS handbook for what should or should not be described.