FedRSPO+: A Heterogeneity-aware Algorithm for Decision-focused Federated Learning

Konstantinos Ziliaskopoulos
Auburn University
kzz0034@auburn.edu

Alexander Vinel
Auburn University
alexander.vinel@auburn.edu

Jiaqi Wang*
Auburn University
jqwang@auburn.edu

Abstract

Decision-focused learning (DFL) trains predictive models for downstream optimization, but existing methods largely assume centralized data. In cross-silo settings, federated learning is a natural alternative, yet standard federated pipelines optimize prediction quality rather than decision quality and do not account for client heterogeneity in downstream objectives or feasible sets. Heterogeneity is especially problematic for DFL methods: under heterogeneous polyhedral decision problems, small perturbations can cause discontinuous changes in optimal decisions, leading to unstable client updates and aggregation. We propose **FedRSPO+**, a heterogeneity-aware framework for decision-focused federated learning. Our approach is built on **RSPO+**, a regularized predict-then-optimize surrogate that smooths the decision map through projection, enabling stable optimization even when clients face different objectives and constraints. We show that RSPO+ upper bounds downstream decision error and regret of the regularized decision and, under exact regularization and consistent LP solution selection, the original LP decision. We also demonstrate that FedRSPO+ has cross-client heterogeneity bounds that *(i)* scale with both objective and feasible-set heterogeneity, *(ii)* vanish at homogeneity, and *(iii)* do not require strong convexity. We develop an annealed and modular training procedure that is compatible with standard federated personalization and aggregation methods. We evaluate FedRSPO+ on three experimental settings: controlled synthetic knapsack instances that isolate multiple axes of heterogeneity, a shortest path DFL benchmark, and an energy pricing case study based on real data. Across these experiments, we compare against prediction-only federated learning and DFL frameworks under varying heterogeneity and communication budgets. Together, our results suggest that smoothing is a useful ingredient for stable collaborative decision learning and provide a heterogeneity-aware foundation for federated DFL[1].

1 Introduction

Decision-focused learning (DFL) has emerged as a common approach to tackle data-driven contextual optimization [1, 2, 3, 4]. Among DFL methods, surrogate loss functions allow practitioners to take advantage of deep learning innovations in operational contexts, with the Smart Predict-then-Optimize (SPO) framework [5] being widely adopted. Its convex surrogate loss, SPO+, provides a tractable upper bound on the downstream decision loss and has seen broad use [6, 7, 8, 9]. However, cross-silo applications of DFL face a key constraint: collaborative analytics are often prohibited due to regulatory, privacy, and competitive considerations.

*Corresponding author.

[1]The code for the paper is available at https://github.com/kziliask/RSPO_Plus.

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

Federated learning (FL), which enables collaborative model training without centralizing data, offers a natural training procedure for addressing these constraints [10, 11, 12]. In cross-silo FL, a global model is trained across multiple organizations without centralizing data: each client keeps its local dataset and shares only model updates (or related statistics), which are aggregated to update a global model [13]. There exists an extensive FL literature on mitigating data, model, system, and task heterogeneity [14, 15, 16, 17, 18, 19, 20]. In the DFL setting, though, when organizations collaboratively train DFL models while keeping data local, heterogeneity in client objectives and constraints creates a fundamental challenge: standard federated aggregation becomes brittle because the optimal solution map can be discontinuous across clients. Let client i have covariates $x \sim \mathcal{X}_i$, a client-specific true objective map $c_i : \mathcal{X}_i \to \mathbb{R}^d$, and a shared federated predictor $g \in \mathcal{H}$ used in the downstream optimization model $\min_{w \in S_i} g(x)^T w$. Here, the objective map $c_i(\cdot)$ and feasible region S_i can differ across clients, which induces *objective* and *constraint* heterogeneity that direct naive implementation of a decision-focused federated learning (DFFL) does not address (Fig. 1). For example, SPO+, the most prominent DFL approach, relies on an optimal solution map that can be discontinuous when clients have different linear objectives over polyhedral feasible sets. Near boundaries where the optimal basis changes, arbitrarily small perturbations can switch the optimizer to a different extreme point, producing a $O(\mathrm{diam}(S_i))$ change in the optimal decision.

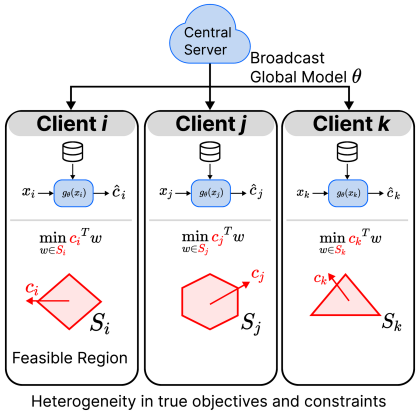


Figure 1: DFFL setting: a central server broadcasts a global model to heterogeneous clients, each with its own objective function and feasible region.

We propose RSPO+, a smoothed surrogate that replaces the discontinuous optimal solution map with a projection-based decision, and FedRSPO+, a federated training framework built on RSPO+ with per-client calibration and co-annealed scheduling. Our contributions are summarized as follows:

- We introduce **RSPO+** and **FedRSPO+**. RSPO+ is a regularized surrogate loss that smooths the decision map while retaining a provable connection to the original downstream regret. FedRSPO+ builds on RSPO+ to provide a modular federated algorithm compatible with standard FL personalization and aggregation methods [14, 16, 21, 22].
- We formally show why naive federation of SPO+ can fail and how FedRSPO+ avoids the same issues, developing point-wise cross-client loss bounds that degrade gracefully with heterogeneity, vanish at homogeneity, and require no strong convexity.
- We evaluate the approach on three tasks [23, 24, 25]: synthetic knapsack, shortest path, and PJM energy pricing. We compare against FedAvg, naive SPO+ federation and perturbation methods across communication budgets and controlled heterogeneity, demonstrating strong improvements in convex resource-allocation settings and competitive performance under structural mismatch.

2 Related Work

Decision-Focused Learning and Contextual Optimization Decision-focused federated learning is largely unexplored. The closest prior work is [26], which analyzed federating SPO+ under client-level objective and feasible-set heterogeneity. Their results showed that naive SPO+-based federation can be brittle in heterogeneous polyhedral problems: optimizer instability induces point-wise heterogeneity bounds with an irreducible diameter term, while stronger geometric assumptions such as strong convexity yield tighter stability guarantees. They also proposed Interp-DFFL, which interpolates between local and federated models using client validation data, which is complementary to our method, and we compare against it in Section 4. We build on this foundation by introducing a smoothed surrogate, RSPO+, and a modular federated training procedure that replaces discontinuous optimizer jumps with stable projection-based decisions, thereby removing the diameter penalty and yielding bounds that vanish at homogeneity.

More broadly, DFL is a sub-topic of the broader area of contextual optimization [2, 3]. SPO+ is among the most widely used surrogates in this literature, with strong theoretical support, including

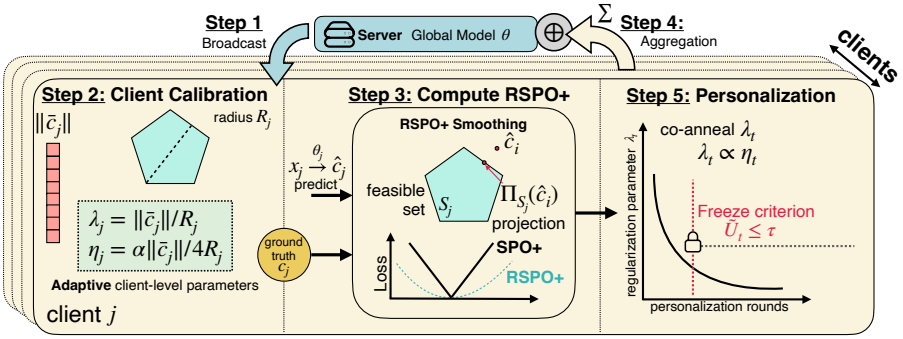


Figure 2: **FedRSPO+: A heterogeneity-aware federated decision-focused learning pipeline.** The server broadcasts a global model, clients perform personalized calibration, local prediction, and projection-based smoothing (RSPO+), and updates are aggregated. During personalization, co-annealed regularization enables stable training under heterogeneous objectives and feasible sets.

Table 1: Overview of the four stages of FedRSPO+, along with the frequency and setting of each stage. We elaborate on each component throughout Section 3.

Component	Execution	Overview	Communication
(1) Calibration	client, once	set λ_j, η_j (Eq. (11))	–
(2) Federated training	server, R_{fed} rounds	SGD on $F_j(\theta; \lambda_j)$, FedAvg ($\propto n_j$)	model weights θ
(3) Personalization	client, R_{per} rounds	SGD on $F_j(\theta; \lambda_{j,t})$, co-anneal λ, η	–
(4) Freezing	client	stop annealing when $\tilde{U}_{j,t} \leq \tau$ (Eq. (6))	–

generalization bounds [27] and calibration guarantees [23, 28]. Other decision-aware surrogates include perturbation-based gradient losses [29] and Fenchel–Young losses [30], while black-box and implicit differentiation methods differentiate through the solver directly rather than introducing an explicit surrogate [24, 31, 32, 33]. Recent work has also begun to study DFL across multiple downstream tasks and changing objectives, suggesting a growing interest in DFL beyond the single-problem setting, though federated DFL still remains underexplored [34, 35, 36].

Federated Learning under Heterogeneity FL, in turn, includes extensive literature on heterogeneity, including proximal penalties [14], control variates [37], localized or partially personalized models [38, 39, 40, 41, 42, 43], and regularized bilevel formulations [21]. However, these methods are designed primarily for heterogeneity in data distributions, model heterogeneity, or predictive objectives, not for heterogeneity in downstream optimization structure. This distinction is especially important in DFL, where small perturbations can induce large changes in the optimal decision. Notably, [21] use Moreau envelopes as a personalization regularizer, whereas we use Moreau smoothing on the decision-focused surrogate itself to stabilize the decision map. More generally, Bregman divergences are widely used in machine learning and optimization as geometry-aware discrepancy measures, supporting methods such as mirror descent, clustering, and structured prediction [44, 45, 46]. We leverage this to construct an adaptive surrogate aligned with each feasible set. Taken together, prior work addresses either decision-aware learning in centralized settings or heterogeneity-aware federated learning for predictive objectives, but not federated DFL with heterogeneous downstream feasible sets.

3 Method

Fig. 2 and Table 1 provide an overview of the four components of the FedRSPO+ pipeline. We begin by defining the notation used throughout this paper, the tools we use to derive our methods, and by stating the assumptions for the client feasible sets and ground truth objectives, which are typical in DFL literature [5, 33, 47, 48]. All proofs are given in the Appendix.

We use $\|\cdot\| = \|\cdot\|_2$ throughout this study. Each feasible set S has a diameter and radius, $D = \sup_{w,w' \in S} \|w - w'\|$, $R = \sup_{w \in S} \|w\|$. The downstream optimization problem is defined as $z_{\text{LP}} =$

$\min_{w\in S} c^T w$, and the quadratically regularized linear program (QLP) as $z_{\text{QLP}} = \min_{w\in S} c^T w + \frac{\lambda}{2}\|w\|^2$. The support function ξ of feasible set S is defined as $\xi_S(u) = \sup_{v\in S}\langle u, v\rangle$. Recall the generalized Bregman divergence is defined as $D_\phi^s(x, y) = \phi(x) - \phi(y) - s^T(x - y)$, where $s \in \partial\phi(y)$. The Moreau envelope of a function with smoothing parameter λ is defined as $M_{\lambda f}(v) = \inf_{u\in\mathbb{R}^d}\left\{f(u) + \frac{1}{2\lambda}\|u - v\|^2\right\}$. The SPO+ loss function [5] is defined as $\ell_{\text{SPO+}}(c, \hat{c}) = \xi_S(c - 2\hat{c}) - 2w^\star(c)(c - \hat{c}) + z^\star(c)$, where S is the feasible set of the downstream optimization problem, $\hat{c} = g(x)$ is the model output, c the ground truth cost vector, $w^\star = \arg\min_{w\in S} c^T w$ and $z^\star$ the optimal objective function value.

Assumption 1 (Bounded, non-empty feasible sets)**.** *The feasible sets of interest are nonempty and compact, so $\xi_S(\cdot)$ is finite and maxima/minima in the sets are attained.*

Assumption 2 (Convex feasible sets)**.** *For each instance, the feasible sets of interest $S \subseteq \mathbb{R}^d$ are convex.*

Assumption 3 (Bounded, non-zero objective)**.** *The objective function vector's norm $0 < C_{\min} \leq \|c\| \leq C_{\max}$ is bounded above and strictly bounded away from zero, while the prediction vector's norm $\|\hat{c}\| \leq B$ is also bounded above.*

3.1 RSPO+

Proposition 1 (SPO+ as a Bregman divergence)**.** *The SPO+ loss admits the representation $\ell_{\text{SPO+}}(c, \hat{c}) = D_\xi^s(c - 2\hat{c}, -c)$, where $s = w^\star(c) \in \partial\xi_S(-c)$ [26].*

Motivated by Proposition 1, we define a smoothed variant by replacing ξ_S with its λ-Moreau envelope, $M_{\lambda\xi_S}(u) = \dfrac{1}{2\lambda}\|u\|^2 - \dfrac{\lambda}{2}\, d_S^2\Big(\dfrac{u}{\lambda}\Big)$, where $d_S^2(u) = \inf_{s\in S}\|u - s\|^2$. We call the resulting loss **RSPO+**:

$$\ell_{\text{RSPO+}}(c, \hat{c}; \lambda) := D_{M_{\lambda\xi_S}}(c - 2\hat{c}, -c).$$

Proposition 2 (Closed form of RSPO+)**.** *Let $u = (c - 2\hat{c})/\lambda$, $v = -c/\lambda$, and Π_S denote projection onto S. Then*

$$\boxed{\ell_{\text{RSPO+}}(c, \hat{c}; \lambda) = \frac{\lambda}{2}\Big(\|u - \Pi_S(v)\|^2 - \|u - \Pi_S(u)\|^2\Big).}$$

Remark 1 (Derivation of RSPO+)**.** *RSPO+ admits an equivalent derivation from decision-focused optimization perspective by applying the SPO+ construction to the QLP, resulting in $\ell_{\text{RSPO+}}$.*

Corollary 1 (Gradient of Smoothed Loss)**.** *From Proposition 2, the gradient of RSPO+ is*

$$\boxed{\nabla_{\hat{c}}\ell_{\text{RSPO+}}(c, \hat{c}; \lambda) = -2\Big(\Pi_S(u) - \Pi_S(v)\Big).}$$

Proposition 3 (Characteristics of RSPO+)**.** *Define $\ell_{\text{RSPO+}}$ from Proposition 2. It has the following characteristics:*

- *$\ell_{\text{RSPO+}} \geq 0$ as a divergence.*
- *$\ell_{\text{RSPO+}}$ is globally convex with respect to $\hat{c}$.*
- *$\ell_{\text{RSPO+}}$ is 2D-Lipschitz continuous under compactness of S.*
- *$\ell_{\text{RSPO+}}$ is $4/\lambda-$smooth, and therefore $\lambda/4$-cocoercive.*

Theorem 1. *Define $f(c, w) = c^T w + \lambda/2\|w\|^2$, $w^\star = \Pi_S(-c/\lambda)$ the optimal action, and $\hat{w} = \Pi_S(-\hat{c}/\lambda)$ the predicted action. The true regularized regret is $R(c, \hat{w}) = f(c, \hat{w}) - f(c, w^\star)$, therefore,*

$$\ell_{\text{RSPO+}} \geq R(c, \hat{w}) + \lambda\|\hat{w} - w^\star\|^2. \tag{1}$$

Furthermore, because $R(c, \hat{w}) \geq 0$, we also have,

$$\|\hat{w} - w^\star\| \leq \sqrt{\frac{\ell_{\text{RSPO+}}}{\lambda}}. \tag{2}$$

Corollary 2 (Unregularized Regret Bound for RSPO+)**.** *Let $z^\star = \min_{w\in S} c^T w$ be the unregularized objective, and $w^\star_{\text{LP}}$ be the minimizer. For any prediction $\hat{c}$ and regularization parameter λ, the true unregularized regret $R_{\text{unreg}}(c, \hat{w}) = c^T\hat{w} - z^\star$ of the predicted decision $\hat{w} = \Pi_S(-\hat{c}/\lambda)$ satisfies:*

$$R_{\text{unreg}}(c, \hat{w}) \leq \ell_{\text{RSPO+}}(c, \hat{c}; \lambda) + \frac{\lambda}{2}R^2. \tag{3}$$

3.2 FedRSPO+

Section 3.1 established bounds on the *regularized* regret and decision error, with Corollary 2 connecting the RSPO+ loss to the unregularized regret of the regularized decision $\hat{w}_\lambda = \Pi_S(-\hat{c}/\lambda)$. In our experiments, however, evaluation is performed using the original downstream decision rule, $\hat{w}_{\mathrm{LP}}(\hat{c}) \in \arg\min_{w \in S} \hat{c}^\top w$, rather than the regularized decision. Thus, Corollary 2 should be interpreted as a direct guarantee for the regularized decision and as a surrogate-level guarantee for the evaluated LP decision. For polyhedral feasible sets, this distinction disappears once λ is sufficiently small under a problem-dependent exact-regularization threshold: the quadratically regularized solution is then also LP-optimal, with the regularizer selecting a particular LP optimum [49, 50]. Under this condition, the RSPO+ bound also applies to the LP decision used at evaluation. In practice, our freeze rule in Eq. (6) stops annealing before λ reaches this threshold, and we quantify this residual gap in Section C.6 in the Appendix.

More generally, as $\lambda \to 0$, the bias term $\frac{\lambda}{2}R^2$ in Eq. (3) vanishes, motivating an annealing schedule that gradually transitions from a smooth, stable surrogate toward the original downstream LP objective. In this section, we build on this connection by introducing a regularization parameter scheduler, which we refer to as an annealing schedule. Finally, we derive bounds on cross-client loss discrepancy under optimization heterogeneity and compare them with bounds on SPO+ loss to show why FedRSPO+ can mitigate discontinuities from vertex jumping.

Annealing Schedule. Corollary 2 shows that RSPO+ controls the original unregularized regret up to an $O(\lambda R^2)$ bias. Thus, λ induces a bias-stability tradeoff: larger λ smooths the decision map and stabilizes optimization, while smaller λ better aligns the surrogate with the original LP objective. This motivates annealing λ during local training. Specifically, we gradually decrease λ so that the regularized decision tracks the original downstream optimum (i.e., as a homotopy continuation from the regularized to the non-regularized objective). We justify this schedule by analyzing how the RSPO+ loss and gradient change under perturbations of λ in Proposition 4.

Proposition 4 (RSPO+ Loss and Gradient under perturbation of λ)**.** *For $\lambda_1 > \lambda_2 > 0$ regularization parameters of RSPO+ loss, we can bound the difference of loss under parameter shift $\lambda_1 \to \lambda_2$ as,*

$$|\ell_{\mathrm{RSPO+}}(\lambda_1) - \ell_{\mathrm{RSPO+}}(\lambda_2)| \leq R^2|\lambda_1 - \lambda_2| + 2\|c\|\|c - \hat{c}\|\left|\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right|, \tag{4}$$

and difference in RSPO+ gradients as,

$$\|\nabla\ell_{\mathrm{RSPO+}}(\lambda_1) - \nabla\ell_{\mathrm{RSPO+}}(\lambda_2)\| \leq 2\Big(\|c - 2\hat{c}\| + \|c\|\Big)\left|\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right|. \tag{5}$$

Remark 2 (Inverse time scheduling)**.** *Proposition 4 shows that RSPO+ is stable to changes in λ primarily through the reciprocal $\frac{1}{\lambda}$. This motivates parameterizing the schedule linearly in $1/\lambda_t$, i.e. $\lambda_t = \lambda_0/(1 + \kappa t)$ which yields a constant per-step reciprocal increment $|\frac{1}{\lambda_{t+1}} - \frac{1}{\lambda_t}| = \kappa/\lambda_0$ and therefore uniformly controls both its loss and gradient drift.*

Stopping condition. As reducing λ weakens the smoothness properties of RSPO+ in Proposition 3, we devise a stopping condition for the λ-schedule in FedRSPO+ that depends on the regularization bias $c_i^T w_{i,t}^{\star,\mathrm{reg}} - z_i^\star$.

For a subset of the client dataset $\mathcal{V}_c$, at local step t with λ_t, define $\hat{w}_{i,t} = \Pi_{\mathcal{S}_i}(-\hat{c}_{i,t}/\lambda_t)$ and $w_{i,t}^{\star,\mathrm{reg}} = \Pi_{\mathcal{S}_i}(-c_i/\lambda_t)$. Let $z_i^\star = \min_{w \in \mathcal{S}_i} c_i^T w$ (pre-solved once on $\mathcal{V}_c$). We stop annealing when the normalized average regularization bias over $\mathcal{V}_c$ falls below a tolerance τ:

$$\tilde{U}_t = \frac{\frac{1}{|\mathcal{V}_c|}\sum_{i \in \mathcal{V}_c}\left(c_i^T w_{i,t}^{\star,\mathrm{reg}} - z_i^\star\right)}{\frac{1}{|\mathcal{V}_c|}\sum_{i \in \mathcal{V}_c}|z_i^\star| + \varepsilon} \leq \tau. \tag{6}$$

Intuitively, this criterion freezes λ_t once the surrogate regularized target is sufficiently aligned with the true downstream LP objective. Thus, $\tilde{U}_t$ empirically tracks the QP–LP regularization bias discussed above and gives a data-dependent stopping rule. The full FedRSPO+ algorithm is provided in pseudo-code in Algorithm 1 in the Appendix.

Remark 3 (MSE–DFL homotopy)**.** *When $0 \in \mathrm{int}(S)$, there exists $\bar{\lambda}$ such that for all $\lambda \geq \bar{\lambda}$, the projections in Corollary 1 act locally as the identity, and $\nabla_{\hat{c}}\ell_{\mathrm{RSPO+}}$ reduces to a scaled MSE gradient.*

As λ decreases, the projection becomes active and the surrogate becomes increasingly decision-focused. Together with Corollary 2, this gives a continuous bias-stability path from prediction-focused learning toward the original LP objective.

Theorem 2 (FedRSPO+ Convergence Rate with linear predictors)**.** *Under standard regularity assumptions, the weighted average iterate $\hat{\Theta}_T$ of FedRSPO+ satisfies*

$$\mathbb{E}\Big[F_T(\hat{\Theta}_T) - F_T(\Theta^\star)\Big] = O\left(\frac{1}{\sum_{t=0}^{T-1}\eta_t} + \frac{\sum_{t=0}^{T-1}\eta_t^2}{\sum_{t=0}^{T-1}\eta_t} + \frac{\sum_{t=0}^{T-1}\eta_t\Delta_t}{\sum_{t=0}^{T-1}\eta_t}\right),$$

with an E^2-dependent constant from local updates. For a fixed-λ objective and $\eta_t \asymp 1/t$, the bound reduces to $O(1/\log T)$. For the actual annealed objective, the additional tracking term appears, which vanishes after freezing and the bound becomes the standard constant-stepsize form $O(1/(T\eta_f) + \eta_f)$. The formal statement and proof are in the Appendix.

We now characterize how the point-wise RSPO+ loss degrades under both objective shift and feasible set shift. This parallels the SPO+ heterogeneity analysis of [26], but has improved stability due to the smoothness of the RSPO+. We measure objective heterogeneity by $c_d = \|c_\mathrm{i} - c_\mathrm{j}\|$ and feasible-set heterogeneity by the Hausdorff distance $\delta = d_H(S_\mathrm{i}, S_\mathrm{j})$. Unlike SPO+, the origin-centered quadratic penalty in the Moreau envelope breaks translation invariance, so we use Hausdorff distance directly rather than the shape distance δ_N. Let $R_{\max} = \max\{R_\mathrm{i}, R_\mathrm{j}\}$ denote the joint bounding radius. The key difference from SPO+ is in how decisions respond to perturbations. In SPO+, the optimal decision $w^\star(c) \in \arg\min_{w\in S} c^T w$ selects a vertex of S. Under infinitesimal perturbations, this vertex can jump discontinuously across the entire feasible set, introducing a diameter penalty $D_{\min}$ that does not vanish even as heterogeneity tends to zero. In RSPO+, the optimal decision is the Euclidean projection $v^\star_{\lambda,S}(c) = \Pi_S(-c/\lambda)$, which is stable under small changes to the feasible set (measured in Hausdorff distance). For any compact convex S_1, S_2 with $d_H(S_1, S_2) \le \delta$:

$$\|\Pi_{S_1}(p) - \Pi_{S_2}(p)\| \;\le\; \sqrt{2\delta(\|p\|+R_{\max})}. \quad (7)$$

This $\sqrt{\delta}$ stability holds for *all* convex sets, including polyhedrals, without any strong convexity assumption.

Table 2: Heterogeneity bound comparison. $c_d = \|c_{\mathrm{ref}} - c_{\mathrm{cl}}\|$ is the objective shift, δ is the Hausdorff set distance, δ_N is the translation-invariant Hausdorff set distance, and $D_{\min}$ is the minimum set diameter. **Q1**: Rate?; **Q2**: Vanishes at homogeneity?; **Q3**: Requires strong convexity?

	SPO+	**RSPO+ (Ours)**
Q1	$O(c_d + \delta_N + D_{\min})$	$O(c_d/\lambda + \sqrt{\delta/\lambda} + \delta)$
Q2	No ($D_{\min}$ persists)	**Yes**
Q3	Yes, for basic stability	**No**

Example 1 (Why SPO+ struggles with polyhedral clients)**.** *Consider two clients with the same polyhedral feasible set S and cost vectors $c^{(i)}, c^{(j)}$, differing by ϵ. For SPO+, the optimizer can switch between extreme points of S, so the decision gap between clients remains $\Omega(D(S))$ even as $\|c^{(i)} - c^{(j)}\| \to 0$. By contrast, RSPO+ uses projected decisions $v^\star_{\lambda S}(c) = \Pi_S(-c/\lambda)$, and since projection is non-expansive,*

$$\|v^\star_{\lambda S}(c^{(i)}) - v^\star_{\lambda S}(c^{(j)})\| \le \frac{1}{\lambda}\|c^{(i)} - c^{(j)}\|.$$

Thus, RSPO+ removes the discontinuous vertex-switching in polyhedral instances.

Theorem 3 (RSPO+ loss discrepancy under heterogeneity)**.** *Let $(c_\mathrm{i}, S_\mathrm{i})$ and $(c_\mathrm{j}, S_\mathrm{j})$ satisfy our assumptions, with $c_d = \|c_\mathrm{i} - c_\mathrm{j}\|$, $c_{\min} = \min\{\|c_i\|, \|c_j\|\}$, and $\delta = d_H(S_\mathrm{i}, S_\mathrm{j})$. Then for any prediction $\hat{c}$:*

$$|\ell_{S_\mathrm{i}}(\hat{c}, c_\mathrm{i}) - \ell_{S_\mathrm{j}}(\hat{c}, c_\mathrm{j})| \;\le\; T(\hat{c}) + B(\hat{c}), \quad (8)$$

*where the **envelope shift** $T(\hat{c})$ and **decision shift** $B(\hat{c})$ are defined as follows.*

The envelope shift captures how the Moreau envelope values change due to combined objective and set perturbation:

$$T(\hat{c}) = D_i c_d + \delta(\|c_\mathrm{j} - 2\hat{c}\| + \|c_\mathrm{j}\| + 2\lambda R_{\max}), \quad (9)$$

and the decision shift captures how the projection-based decisions drift, amplified by the prediction:

$$B(\hat{c}) = (2\|\hat{c}\| + 2\lambda R_{\max})\left[\frac{c_d}{\lambda} + \sqrt{2\delta\left(\frac{c_{\min}}{\lambda} + R_{\max}\right)}\right]. \quad (10)$$

Combining the regularization bias in Corollary 2 with the explicit λ-dependence in Theorem 3 gives the following informal bias–stability tradeoff:

$$\mathrm{error}(\lambda) \lesssim \lambda R^2 + \frac{c_d}{\lambda} + \sqrt{\frac{\delta}{\lambda}},$$

up to constants depending on the bounded costs, predictions, and feasible-set radii. The first term is the $O(\lambda R^2)$ regularization bias, while the latter two terms come from the objective and feasible-set heterogeneity terms in Theorem 3. Thus, decreasing λ improves alignment with the original LP objective but worsens stability under client heterogeneity. This motivates annealing only until the regularized objective is sufficiently aligned with the original downstream objective, rather than reducing λ to zero and replacing RSPO+ with SPO+. We formalize this relationship in the Appendix.

Calibration of λ and η. From the RSPO+ loss structure, λ scales the cost vectors relative to the feasible set: too large and the loss reduces to MSE, too small and it reverts to SPO+ (see Remark 3). For the loss to alternate between regimes, λ should place the scaled vector near the feasible-set boundary. Combined with the $4/\lambda$-smoothness of RSPO+ (Proposition 3), which suggests a step size of order $\lambda/4$ [51], we initialize per client j:

$$\lambda_j := \frac{\|\bar{c}_j\|}{R_j}, \qquad \eta_j := \frac{\alpha \, \|\bar{c}_j\|}{4 \, R_j}, \tag{11}$$

where $\|\bar{c}_j\|$ is the average cost-vector magnitude in client j's dataset, R_j is the feasible-set radius, and α is a small positive scalar (e.g. $\alpha = 0.01$). When annealing λ, we co-anneal η with $\kappa^{\mathrm{LR}} \propto \kappa^{\lambda}$ to maintain this step-size relationship. This is also validated by the convergence analysis in Section A.1. A sensitivity analysis over κ is in Appendix Section C.

4 Experiment

With each experiment, we test a different aspect of FedRSPO+. Our synthetic fractional knapsack experiment uses controlled heterogeneity across all three axes (feasible set, objective, and data) to test Theorem 3 empirically and the performance of our method against the state-of-the-art. Our PJM experiment uses real price dynamics and heterogeneous feasible sets to test our method on real datasets. Finally, our Warcraft experiment has heterogeneous starts and destinations to vary feasible sets and datasets across clients to evaluate how FedRSPO+ performs when the optimization task relies on the structure of the unregularized objective, and how RSPO+'s regularization affects performance.

For all experiments, we use a federated warm-start followed by local personalization. Our model is a shallow neural network with 64 hidden units for the fractional knapsack and PJM tasks, and a truncated ResNet-18 for the Warcraft task following prior work [1]. We report relative regret in Table 3 and the optimality gap in Table 6. While FedRSPO+ is trained on a regularized objective, we report on the actual downstream regret. We also compare with popular perturbation DFL methods, perturbed Fenchel-Young Loss (PFYL, [30]) and Differentiable Perturbation Optimizers (DPO, [52]). We also compare with Interp-DFFL [26] and an OptNet-style baseline [48]. Implementation details, hyperparameters, and solver configurations are provided in the Appendix C, and additional centralized and local-only baselines in Appendix Section C.5.

4.1 Fractional Knapsack

To test FedRSPO+ against different axes of heterogeneity in the client dataset, we implement a fractional knapsack experiment, using a synthetic dataset for feature size $p = 8$, cost vector size $d = 50$, and 20 clients. We test 3 different levels of heterogeneity, represented by $(\eta_{\mathrm{o}}, \eta_{\mathrm{c}}, \eta_{\mathrm{d}})$ in Table 3: no heterogeneity, moderate heterogeneity, and extreme heterogeneity. Overall, RSPO+ has the lowest relative regret across the board. The largest gap is during federation, meaning RSPO+ succeeds in developing an effective model for all clients despite heterogeneity. Across all methods, data heterogeneity was most detrimental to model performance, with constraint heterogeneity moderately affecting regret and objective heterogeneity minimally. RSPO+ was found resilient to all forms of heterogeneity compared to the baseline methods.

We also test additional baselines. As shown in Table 4, FedRSPO+ has lower regret than Interp-DFFL in the homogeneous, objective-heterogeneous, and feasible-set-heterogeneous settings. Under pure

Table 3: Mean client relative regret (%) ↓ across heterogeneity configurations. Each axis (η_o:objective, η_c:constraint, η_d:data) is varied independently while others are fixed at 0. "Fed": global federated model; "+P": after local personalization. Results: mean ± std over 5 seeds.

	MSE		PFYL		DPO		SPO+		RSPO+ (Ours)	
(η_o, η_c, η_d)	Fed	+P	Fed	+P	Fed	+P	Fed	+P	Fed	+P
$(0, 0, 0)$	15.9 ± 3.3	10.5 ± 2.8	8.5 ± 1.5	6.1 ± 1.1	16.3 ± 2.7	12.6 ± 2.2	9.7 ± 2.1	5.6 ± 1.4	$\mathbf{2.3 \pm 0.4}$	$\mathbf{1.9 \pm 0.3}$
$(0.5, 0, 0)$	16.0 ± 3.4	10.6 ± 2.8	8.8 ± 1.5	6.3 ± 1.2	16.7 ± 2.5	12.8 ± 2.3	10.0 ± 2.1	5.8 ± 1.5	$\mathbf{2.5 \pm 0.3}$	$\mathbf{1.9 \pm 0.3}$
$(1, 0, 0)$	16.4 ± 3.4	10.8 ± 2.9	9.9 ± 1.6	6.9 ± 1.3	17.4 ± 2.5	13.4 ± 2.2	10.9 ± 2.2	6.3 ± 1.6	$\mathbf{3.0 \pm 0.3}$	$\mathbf{2.0 \pm 0.3}$
$(0, 0.5, 0)$	13.2 ± 3.5	9.8 ± 2.7	13.3 ± 2.4	8.1 ± 1.9	19.6 ± 3.6	13.1 ± 3.1	11.9 ± 2.5	7.3 ± 2.1	$\mathbf{3.9 \pm 0.6}$	$\mathbf{2.3 \pm 0.4}$
$(0, 1, 0)$	13.3 ± 2.3	10.5 ± 2.0	15.0 ± 2.3	9.1 ± 1.5	20.8 ± 3.3	14.1 ± 2.4	12.9 ± 2.0	8.1 ± 1.5	$\mathbf{4.6 \pm 0.5}$	$\mathbf{2.6 \pm 0.4}$
$(0, 0, 0.5)$	17.5 ± 3.0	9.6 ± 2.2	14.0 ± 1.8	8.2 ± 1.2	20.6 ± 3.2	14.4 ± 2.4	14.1 ± 2.0	6.9 ± 1.3	$\mathbf{6.5 \pm 0.8}$	$\mathbf{3.8 \pm 0.2}$
$(0, 0, 1)$	19.1 ± 2.8	8.4 ± 1.7	18.0 ± 1.9	8.8 ± 0.7	22.9 ± 2.9	14.1 ± 1.8	17.7 ± 2.1	6.9 ± 0.8	$\mathbf{9.3 \pm 1.2}$	$\mathbf{3.4 \pm 0.2}$

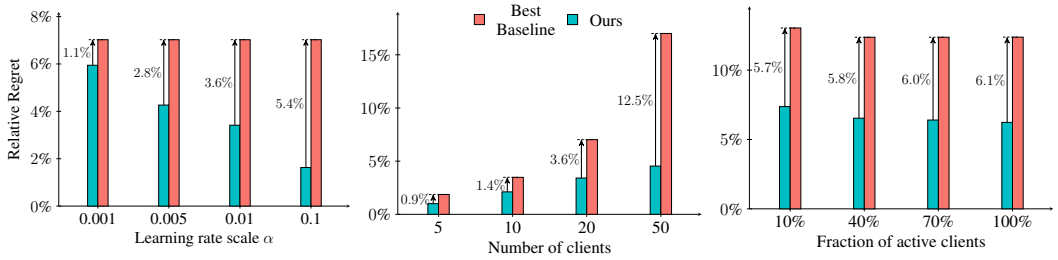


Figure 3: Relative regret (%) of the best baseline SPO+ and our method RSPO+ across different values of the learning rate scale α from Eq. (11), number of clients, and communication budgets. **Lower is better (↓)**. Numbers above the bars show the improvement in relative regret of RSPO+ relative to SPO+.

data heterogeneity $(0, 0, 1)$, however, Interp-DFFL is stronger (2.47% vs. 3.37%), consistent with its explicit local–federated interpolation, which targets statistical heterogeneity. As RSPO+ is a loss and Interp-DFFL an aggregation rule, they are not mutually exclusive. The OptNet-style baseline, which uses the same regularized QP with a different training signal, is substantially weaker (9.4–11.3%), suggesting that the gains come from the RSPO+ surrogate, not just quadratic regularization.

We also run a sweep across relevant parameters, shown in Fig. 3, to test the performance of our method against SPO+ through different fractions of active clients per round, number of clients, and different values of α. We see that the communication budget has a modest impact on the performance of FedRSPO+, with a general improvement over SPO+ as the fraction increases. The parameter α was moderately more impactful on FedRSPO+ performance, with relative improvement ranging from 1.1% for $\alpha = 10^{-3}$ to 5.4% for $\alpha = 0.1$. However, the most surprising result is for increasing number of clients. There FedRSPO+ dramatically surpasses the federated SPO+ setup, with a difference of 12.5% at 50 clients. Extra sweep results for κ^{λ}, κ^{LR}, and τ are available in Appendix Section C. In all cases, FedRSPO+ outperforms the SPO+ implementation.

Table 4: Comparison of FedRSPO+ against Interp-DFFL and OptNet across fractional knapsack heterogeneity settings and PJM cases. Values are mean ± standard deviation across training seeds of mean relative regret (absolute for PJM).

Setting	Interp-DFFL	OptNet-style	FedRSPO+
$(0, 0, 0)$	2.42 ± 0.30	9.4 ± 1.9	$\mathbf{1.86 \pm 0.25}$
$(1, 0, 0)$	2.42 ± 0.33	9.7 ± 2.0	$\mathbf{2.02 \pm 0.26}$
$(0, 1, 0)$	3.51 ± 0.77	11.3 ± 2.0	$\mathbf{2.58 \pm 0.38}$
$(0, 0, 1)$	$\mathbf{2.47 \pm 0.39}$	11.3 ± 1.9	3.37 ± 0.22
PJM ($)	69.6 ± 1.7	69.8 ± 1.3	$\mathbf{60.7 \pm 0.8}$

4.2 PJM Case Study

We evaluate our approach on hourly day-ahead electricity prices from the PJM Interconnection for the year 2025 [25]. Each client represents an aggregator or a large consumer with a fixed daily electricity demand D_j that forecasts pricing in advance to inform their bids [53]. Their goal is to minimize total procurement cost. In effect, the downstream optimization task is a top-k selection problem, where the cheapest $k = D_j$ hours are chosen to fulfill demand. We use weather data from the Open-Meteo API

[54] as input data. We use public price and weather data for reproducibility, which provides naturally occurring client-level price variation and input data heterogeneity. Additional experimental info can be found in the Appendix Section C.4.

Table 5 presents an ablation of the FedRSPO+ components on this task. Starting from MSE (175.5 mean absolute regret), switching to SPO+ reduces regret to 81.7, and RSPO+ further improves to 72.9. Each additional component: personalization, per-client λ/η calibration (Eq. (11)), and co-annealed scheduling yields consistent gains, with the full FedRSPO+ pipeline achieving 60.7 ($\pm$0.8), a 65% reduction over MSE and 26% over naive SPO+ federation. PFYL (67.8 ± 1.2 regret) and DPO (74.7 ± 2.2 regret) performed well in this case study compared to SPO+. This extends to the Interp-DFFL and OptNet-style baselines, as shown in Table 4, where FedRSPO+ outperforms both. We also incorporate popular FL algorithms to showcase FedRSPO+'s modularity. FedProx achieves similar mean regret, with a modest reduction in variance. Ditto performs similarly to RSPO+ with personalization, which coincides with our expectations. Ditto introduces proximal regularization towards the global model, which conflicts with RSPO+'s gradual scheduling and its transition from the regularized to the unregularized objective. This suggests that FL methods adding regularization at the personalization stage require careful coordination with RSPO+'s own scheduling, which is an interesting direction for future work. Finally, we perform additional experiments, including fixed demand across the federation and data-scarcity regimes, and showcase the results in the Appendix.

Table 5: Ablation study for FedRSPO+ on PJM data. Values are mean $\pm$ standard deviation across training seeds of mean absolute regret. "w/ P" denotes the personalization phase, λ_j/η_j denotes per-client parameters, and Full additionally includes time-inverse scheduling.

Method	Mean absolute regret ↓
MSE	175.453 ± 7.260
SPO+	81.676 ± 1.768
RSPO+	72.890 ± 4.231
RSPO+ *w/ P*	68.136 ± 1.402
RSPO+ *w/ P* + λ_j/η_j	63.006 ± 1.059
RSPO+(Full)	$\mathbf{60.654} \pm 0.796$
RSPO+(Full)+Ditto	69.482 ± 2.932
RSPO+(Full)+FedProx	$\mathbf{60.602} \pm 0.623$

4.3 Warcraft Shortest Path

To characterize RSPO+'s behavior under structural mismatch, we evaluate on the 12x12 Warcraft shortest-path benchmark [24] with 10,000 training maps. We define 44 heterogeneous clients by assigning each border cell as an origin and its 180°-rotated opposite border cell as the destination, with additional terrain-cost perturbations. We use Dijkstra's algorithm [55, 56] as the linear oracle.

We explicitly chose shortest path because regularization alters the downstream optimization problem. Importantly, adding the quadratic term destroys the shortest path's total unimodularity and allows fractional flows across the grid. RSPO+ still achieves a 2.571% optimality gap (Table 6) and outperforms perturbation methods and MSE, losing out marginally to SPO+. Annealing narrows the gap between the regularized flow objective and the true shortest-path objective, yielding a $\sim$66% improvement during personalization, the largest personalization gain across all three experiments. Overall, this highlights a limitation of the proposed FedRSPO+ method, as it is most effective when the downstream problem structure is compatible with quadratic regularization.

Table 6: Mean optimality gap (↓) on the Warcraft Shortest Path benchmark. Results are mean $\pm$ std over 5 seeds.

Method	Federated (%)	+P (%)
DPO	11.309 ± 2.244	11.018 ± 1.718
MSE	14.784 ± 2.491	7.241 ± 1.336
PFYL	8.959 ± 1.688	11.449 ± 1.048
SPO+	5.978 ± 1.842	1.850 ± 0.377
Ours	7.504 ± 1.929	2.571 ± 0.901

4.4 Computational overhead

While FedRSPO+ has no additional communication costs compared with FedAvg, it has additional computational overhead, since solving QLPs is more expensive than LP or shortest path problems. We observed 1.35×–1.77× slower solves than LPs with Gurobi and up to $\sim$50× slower solves than Dijkstra. However, oracle-only timings overstate the end-to-end cost. In the Warcraft experiment, MSE and SPO+ both took about 67 seconds per round, whereas RSPO+ took 133 seconds, i.e., a 2.0× increase in training time. The difference is mainly due to solving the QLP and computing the freezing criterion. The remaining model training is shared across methods and dominates most of the runtime. Moreover,

the additional QLP oracle calls are parallelizable across clients and samples, suggesting that further engineering could reduce the observed wall-clock overhead.

5 Conclusions

In this paper, we connected the fields of FL and DFL by addressing the brittleness of current methods under optimization heterogeneity through our proposed method, FedRSPO+. We showed how smoothing via Moreau envelopes replaces discontinuous vertex-switching with $\sqrt{\delta}$-stable projected decisions and achieves heterogeneity bounds that vanish at homogeneity without strong convexity. Our principled per-client calibration and co-annealed scheduling make the method practical without expensive tuning, and we demonstrated compatibility with existing FL methods such as FedProx and Ditto. Across our synthetic knapsack and PJM energy pricing experiments, RSPO+ consistently outperformed both prediction-only FL and naive SPO+ federation, with the largest gains appearing during the federation phase. It also outperformed Interp-DFFL and an OptNet-style baseline under objective and feasible-set heterogeneity. Even under the mismatched downstream objective in the Warcraft experiment, FedRSPO+ achieved close to optimal performance for clients, with significant gains over MSE and popular DFL methods. Together, these results position projection-based smoothing as a key ingredient for stable collaborative decision learning.

Limitations. Our goal is to introduce a heterogeneity-aware surrogate and training methodology for DFFL, and to evaluate it against representative prediction-only FL, SPO+, personalization and aggregation variants, and standard DFL baselines where they are most applicable. A broader benchmark spanning additional FL optimizers, privacy mechanisms, combinatorial DFL methods, and how they fit in the field of DFFL is an important direction for future work. Our regret and decision-error guarantees are for the regularized decision, and apply to the LP decision only under exact regularization, which our stopping criterion does not enforce. Furthermore, our convergence analysis applies to linear estimators, while our main experiments use neural networks. We examine FedRSPO+ performance with linear estimators in the Appendix. RSPO+ replaces each LP solve with a QP of the same dimension. Specialized QP solvers can match simplex-based solvers on this structure [57, 58], and warm-starting can further reduce overhead. Furthermore, RSPO+ introduces hyperparameters $\alpha, \kappa^{\mathrm{LR}}, \kappa^{\lambda}$ beyond SPO+ and MSE. While we provide principled defaults (Eq. (11)), tuning for new domains may add overhead. Our convexity assumptions (Assumptions 1 to 3) are standard in DFL but exclude mixed-integer programs. In Section 4.3 we demonstrate that FedRSPO+ is not appropriate for all downstream objectives, and extending FedRSPO+ to other domains and problem types via relaxations or decompositions is a natural direction.

Broader Impact. FedRSPO+ enables privacy-preserving collaboration for data-driven decision-making in settings where data sharing is restricted by regulation or competition, potentially lowering barriers for smaller organizations that lack the data scale to train effective models independently. However, participants with higher-quality data or more favorable constraint structures may benefit disproportionately from federation, and model updates can still leak information about local objectives, particularly in small-coalition settings. Practitioners should consider pairing FedRSPO+ with differential privacy or secure aggregation mechanisms in sensitive domains.

Acknowledgments We thank the anonymous reviewers and area chair for their feedback, which helped us refine and substantially strengthen this work.

Funding. This work was made possible in part by high-performance computing resources and technical support provided by the Alabama Supercomputer Authority.

Competing interests. The authors declare no competing interests.

References

[1] Jayanta Mandi, James Kotary, Senne Berden, Maxime Mulamba, Victor Bucarey, Tias Guns, and Ferdinando Fioretto. Decision-focused learning: Foundations, state of the art, benchmark and future opportunities. *Journal of Artificial Intelligence Research*, 80:1623–1701, 2024.

[2] Meng Qi and Zuo-Jun Shen. Integrating prediction/estimation and optimization with applications in operations management. In *Tutorials in Operations Research: Emerging and Impactful Topics in Operations*, pages 36–58. INFORMS, 2022.

[3] Utsav Sadana, Abhilash Chenreddy, Erick Delage, Alexandre Forel, Emma Frejinger, and Thibaut Vidal. A survey of contextual optimization methods for decision-making under uncertainty. *European Journal of Operational Research*, 320(2):271–289, 2025.

[4] Velibor V Mišić and Georgia Perakis. Data analytics in operations management: A review. *Manufacturing & Service Operations Management*, 22(1):158–169, 2020.

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

[6] Haoran Zhang and Haoran Liu. Real-time power system optimization under typhoon weather using the smart "predict, then optimize" framework. *Energies*, 18(3):615, 2025.

[7] Go Nam Lui and Soner Demirel. Gradient-based smart predict-then-optimize framework for aircraft arrival scheduling problem. *Journal of Open Aviation Science*, 2(2), 2024.

[8] Xinyu Wang, Yiyang Peng, and Wei Ma. Spo-vcs: An end-to-end smart predict-then-optimize framework with alternating differentiation method for relocation problems in large-scale vehicle crowd sensing. *Transportation Research Part E: Logistics and Transportation Review*, 205:104515, 2026.

[9] Xuecheng Tian, Ran Yan, Yannick Liu, and Shuaian Wang. A smart predict-then-optimize method for targeted and cost-effective maritime transportation. *Transportation Research Part B: Methodological*, 172:32–52, 2023.

[10] Peter Kairouz and H Brendan McMahan. Advances and open problems in federated learning. *Foundations and Trends in Machine Learning*, 14(1-2):1–210, 2021.

[11] Di Chai, Leye Wang, Liu Yang, Junxue Zhang, Kai Chen, and Qiang Yang. A survey for federated learning evaluations: Goals and measures. *IEEE Transactions on Knowledge and Data Engineering*, 36(10):5007–5024, 2024.

[12] Bingyan Liu, Nuoyan Lv, Yuanchun Guo, and Yawen Li. Recent advances on federated learning: A systematic survey. *Neurocomputing*, 597:128019, 2024.

[13] Brendan McMahan, Eider Moore, Daniel Ramage, Seth Hampson, and Blaise Aguera y Arcas. Communication-efficient learning of deep networks from decentralized data. In *Artificial intelligence and statistics*, pages 1273–1282. Pmlr, 2017.

[14] Tian Li, Anit Kumar Sahu, Manzil Zaheer, Maziar Sanjabi, Ameet Talwalkar, and Virginia Smith. Federated optimization in heterogeneous networks. *Proceedings of Machine learning and systems*, 2:429–450, 2020.

[15] Jiaqi Wang, Xingyi Yang, Suhan Cui, Liwei Che, Lingjuan Lyu, Dongkuan DK Xu, and Fenglong Ma. Towards personalized federated learning via heterogeneous model reassembly. *Advances in Neural Information Processing Systems*, 36:29515–29531, 2023.

[16] Tian Li, Shengyuan Hu, Ahmad Beirami, and Virginia Smith. Ditto: Fair and robust federated learning through personalization. In *International Conference on Machine Learning*, pages 6357–6368. PMLR, 2021.

[17] Xi Li, Chen Wu, and Jiaqi Wang. Unveiling backdoor risks brought by foundation models in heterogeneous federated learning. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, pages 168–181. Springer, 2024.

[18] Mang Ye, Xiuwen Fang, Bo Du, Pong C Yuen, and Dacheng Tao. Heterogeneous federated learning: State-of-the-art and research challenges. *ACM Computing Surveys*, 56(3):1–44, 2023.

[19] Yann Fraboni, Richard Vidal, Laetitia Kameni, and Marco Lorenzi. A general theory for federated optimization with asynchronous and heterogeneous clients updates. *Journal of Machine Learning Research*, 24(110):1–43, 2023.

[20] Jiaqi Wang, Chenxu Zhao, Lingjuan Lyu, Quanzeng You, Mengdi Huai, and Fenglong Ma. Bridging model heterogeneity in federated learning via uncertainty-based asymmetrical reciprocity learning. In *Proceedings of the 41st International Conference on Machine Learning*, volume 235 of *Proceedings of Machine Learning Research*, pages 52290–52308, 2024.

[21] Canh T Dinh, Nguyen Tran, and Josh Nguyen. Personalized federated learning with moreau envelopes. *Advances in neural information processing systems*, 33:21394–21405, 2020.

[22] Xu Zhang, Yinchuan Li, Wenpeng Li, Kaiyang Guo, and Yunfeng Shao. Personalized federated learning via variational bayesian inference. In *International Conference on Machine Learning*, pages 26293–26310. PMLR, 2022.

[23] Nam Ho-Nguyen and Fatma Kılınç-Karzan. Risk guarantees for end-to-end prediction and optimization processes. *Management Science*, 68(12):8680–8698, 2022.

[24] Marin Vlastelica Pogančić, Anselm Paulus, Vit Musil, Georg Martius, and Michal Rolinek. Differentiation of blackbox combinatorial solvers. In *International Conference on Learning Representations*, 2020.

[25] PJM Interconnection, L.L.C. Real-time and historical hourly locational marginal pricing (lmp) data. `https://www.energyonline.com/`, 2025. Accessed via EnergyOnline. Data from Jan 1–Dec 31, 2025. Accessed: 2026-02-14.

[26] Konstantinos Ziliaskopoulos and Alexander Vinel. Decision-focused federated learning under heterogeneous objectives and constraints. *arXiv preprint arXiv:2604.20031*, 2026.

[27] 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.

[28] Heyuan Liu and Paul Grigas. Risk bounds and calibration for a smart predict-then-optimize method. *Advances in Neural Information Processing Systems*, 34:22083–22094, 2021.

[29] Vishal Gupta and Michael Huang. Decision-focused learning with directional gradients. *Advances in Neural Information Processing Systems*, 37:79194–79220, 2024.

[30] Mathieu Blondel, André FT Martins, and Vlad Niculae. Learning with fenchel-young losses. *Journal of Machine Learning Research*, 21(35):1–69, 2020.

[31] Subham Sekhar Sahoo, Anselm Paulus, Marin Vlastelica, Vít Musil, Volodymyr Kuleshov, and Georg Martius. Backpropagation through combinatorial algorithms: Identity with projection works. In *The Eleventh International Conference on Learning Representations*, 2023.

[32] Mathieu Blondel, Quentin Berthet, Marco Cuturi, Roy Frostig, Stephan Hoyer, Felipe Llinares-López, Fabian Pedregosa, and Jean-Philippe Vert. Efficient and modular implicit differentiation. *Advances in Neural Information Processing Systems*, 35:5230–5242, 2022.

[33] Akshay Agrawal, Brandon Amos, Shane Barratt, Stephen Boyd, Steven Diamond, and J Zico Kolter. Differentiable convex optimization layers. *Advances in Neural Information Processing Systems*, 32, 2019.

[34] Aymeric Capitaine, Maxime Haddouche, Eric Moulines, Michael I. Jordan, Etienne Boursier, and Alain Oliviero Durmus. Online decision-focused learning. In *The Fourteenth International Conference on Learning Representations*, 2026.

[35] Bo Tang and Elias B Khalil. Multi-task predict-then-optimize. In *International Conference on Learning and Intelligent Optimization*, pages 506–522. Springer, 2023.

[36] Paula Rodriguez-Diaz, Lingkai Kong, Kai Wang, David Alvarez-Melis, and Milind Tambe. What is the right notion of distance between predict-then-optimize tasks? In Silvia Chiappa and Sara Magliacane, editors, *Proceedings of the Forty-first Conference on Uncertainty in Artificial Intelligence*, volume 286 of *Proceedings of Machine Learning Research*, pages 3570–3586. PMLR, 21–25 Jul 2025.

[37] Sai Praneeth Karimireddy, Satyen Kale, Mehryar Mohri, Sashank Reddi, Sebastian Stich, and Ananda Theertha Suresh. Scaffold: Stochastic controlled averaging for federated learning. In *International Conference on Machine Learning*, pages 5132–5143. PMLR, 2020.

[38] Xiaoxiao Li, Meirui JIANG, Xiaofei Zhang, Michael Kamp, and Qi Dou. FedBN: Federated learning on non-IID features via local batch normalization. In *International Conference on Learning Representations*, 2021.

[39] Jiaqi Wang, Shenglai Zeng, Zewei Long, Yaqing Wang, Houping Xiao, and Fenglong Ma. Knowledge-enhanced semi-supervised federated learning for aggregating heterogeneous lightweight clients in IoT. In *Proceedings of the 2023 SIAM International Conference on Data Mining (SDM)*, pages 496–504, 2023.

[40] Manoj Ghuhan Arivazhagan, Vinay Aggarwal, Aaditya Kumar Singh, and Sunav Choudhary. Federated learning with personalization layers. *arXiv preprint arXiv:1912.00818*, 2019.

[41] Jiaqi Wang, Qi Li, Lingjuan Lyu, and Fenglong Ma. pfedclub: Controllable heterogeneous model aggregation for personalized federated learning. *Advances in Neural Information Processing Systems*, 37:54692–54713, 2024.

[42] Liam Collins, Hamed Hassani, Aryan Mokhtari, and Sanjay Shakkottai. Exploiting shared representations for personalized federated learning. In *International conference on machine learning*, pages 2089–2099. PMLR, 2021.

[43] Jae Hoon Oh, Sangmook Kim, and Seyoung Yun. Fedbabu: Toward enhanced representation for federated image classification. In *10th International Conference on Learning Representations, ICLR 2022*. International Conference on Learning Representations (ICLR), 2022.

[44] Arindam Banerjee, Srujana Merugu, Inderjit S Dhillon, and Joydeep Ghosh. Clustering with bregman divergences. *Journal of Machine Learning Research*, 6(Oct):1705–1749, 2005.

[45] Amir Beck and Marc Teboulle. Mirror descent and nonlinear projected subgradient methods for convex optimization. *Operations Research Letters*, 31(3):167–175, 2003.

[46] Garvesh Raskutti and Sayan Mukherjee. The information geometry of mirror descent. *IEEE Transactions on Information Theory*, 61(3):1451–1457, 2015.

[47] Maxime Mulamba, Jayanta Mandi, Michelangelo Diligenti, Michele Lombardi, Victor Bucarey Lopez, and Tias Guns. Contrastive losses and solution caching for predict-and-optimize. In *30th International Joint Conference on Artificial Intelligence (IJCAI-21): IJCAI-21*, pages 2833–2840. International Joint Conferences on Artificial Intelligence, 2021.

[48] Brandon Amos and J Zico Kolter. Optnet: Differentiable optimization as a layer in neural networks. In *International conference on machine learning*, pages 136–145. PMLR, 2017.

[49] Michael P Friedlander and Paul Tseng. Exact regularization of convex programs. *SIAM Journal on Optimization*, 18(4):1326–1350, 2008.

[50] Alberto González-Sanz and Marcel Nutz. Quantitative convergence of quadratically regularized linear programs. *Applied Mathematics & Optimization*, 91(3):68, 2025.

[51] Konstantinos Spiliopoulos, Richard B. Sowers, and Justin Sirignano. *Mathematical Foundations of Deep Learning Models and Algorithms*, volume 252 of *Graduate Studies in Mathematics*. American Mathematical Society, 2025.

[52] Quentin Berthet, Mathieu Blondel, Olivier Teboul, Marco Cuturi, Jean-Philippe Vert, and Francis Bach. Learning with differentiable perturbed optimizers. *Advances in Neural Information Processing Systems*, 33:9508–9519, 2020.

[53] Antonio J Conejo, Miguel A Plazas, Rosa Espinola, and Ana B Molina. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. *IEEE Transactions on Power Systems*, 20(2):1035–1042, 2005.

[54] Patrick Zippenfenig. Open-meteo.com weather api, 2023.

[55] EW Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1):269–271, 1959.

[56] Guillaume Dalle, Léo Baty, Louis Bouvier, and Axel Parmentier. Learning with combinatorial optimization layers: a probabilistic approach. *arXiv preprint arXiv:2207.13513*, 2022.

[57] Daniel Arnström, Alberto Bemporad, and Daniel Axehill. A dual active-set solver for embedded quadratic programming using recursive LDL^T updates. *IEEE Transactions on Automatic Control*, 67(8):4362–4369, 2022.

[58] Bartolomeo Stellato, Goran Banjac, Paul Goulart, Alberto Bemporad, and Stephen Boyd. OSQP: An operator splitting solver for quadratic programs. *Mathematical Programming Computation*, 12(4):637–672, 2020.

[59] Xiang Li, Kaixuan Huang, Wenhao Yang, Shusen Wang, and Zhihua Zhang. On the convergence of FedAvg on non-IID data. In *International Conference on Learning Representations*, 2019.

[60] Hao Yu, Sen Yang, and Shenghuo Zhu. Parallel restarted sgd with faster convergence and less communication: Demystifying why model averaging works for deep learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 5693–5700, 2019.

[61] Heinz H. Bauschke and Patrick L. Combettes. *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. CMS Books in Mathematics. Springer, 2nd edition, 2017.

[62] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

[63] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B Shah. Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1):65–98, 2017.

[64] Mike Innes. Flux: Elegant machine learning with julia. *Journal of Open Source Software*, 3(25):602, 2018.

[65] Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2026.

[66] Miles Lubin, Oscar Dowson, Joaquim Dias Garcia, Joey Huchette, Benoît Legat, and Juan Pablo Vielma. JuMP 1.0: Recent improvements to a modeling language for mathematical optimization. *Mathematical Programming Computation*, 2023.

[67] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[68] Eugene Litvinov, Tongxin Zheng, Gary Rosenwald, and Payman Shamsollahi. Marginal loss modeling in lmp calculation. *IEEE transactions on Power Systems*, 19(2):880–888, 2004.

[69] North American Electric Reliability Corporation. CIP-011-4 – Cyber Security – Information Protection, 2019.

Appendix

A Additional Results

A.1 Convergence of FedRSPO+

Assumption 4 (Model Regularity)**.** *We analyze predictors $\hat{c} = g(x;\Theta)$, including the linear case $g(x;\Theta) = \Theta x$, with bounded features $\|x\| \le X$, bounded initialization $\|\Theta_0 - \Theta^\star\| \le D_\Theta$, and the prediction/cost bounds of Assumption 3: $\|\hat{c}\| \le B$, $\|c\| \le C_{\max}$. For nonlinear predictors, $g(x;\Theta)$ is differentiable in Θ and, for all x and all Θ, Θ' on the training trajectory, the Jacobian satisfies $\|J_\Theta g(x;\Theta)\|_{\mathrm{op}} \le J$, and $\|J_\Theta g(x;\Theta) - J_\Theta g(x;\Theta')\|_{\mathrm{op}} \le \rho\|\Theta - \Theta'\|$, where $\|\cdot\|_{\mathrm{op}}$ is the operator norm [33]. For the linear model, these hold with $J \le X$ and $\rho = 0$. Under this assumption, the convexity and $4/\lambda_t$-smoothness of RSPO+ (Proposition 3) imply that $F_{i,t}(\Theta)$ is convex and L_t-smooth with $L_t = 4X^2/\lambda_t$ in the linear case. Mini-batch stochastic gradients are unbiased estimators of $\nabla F_{i,t}(\Theta)$.*

Assumption 5 (Federated Schedule)**.** *Clients perform E local SGD steps per synchronization round, with co-annealed learning rate $\eta_t = \alpha\lambda_t/4$ for a constant $\alpha > 0$ as in Eq.* (11)*.*

We establish that FedRSPO+ converges under FedAvg aggregation and validate the co-annealed schedule of Eq. (11). Let $F_{i,t}(\Theta) := \mathbb{E}_{(x,c)\sim\mathcal{D}_i}[\ell_{\mathrm{RSPO+}}(c,\, g(x;\Theta);\lambda_t)]$ denote the local objective for client i at round t, and $F_t = \frac{1}{N}\sum_{i=1}^N F_{i,t}$ the global objective. Because the annealed λ_t schedule makes F_t time-varying, we track the deviation from the target objective at the frozen parameter λ_T using Proposition 4:

$$\Delta_t := R^2(\lambda_t - \lambda_T) + 2\, C_{\max}(C_{\max} + B)\left(\frac{1}{\lambda_T} - \frac{1}{\lambda_t}\right). \tag{12}$$

Theorem 4 (FedRSPO+ Convergence, restatement of Theorem 2)**.** *Under Assumptions 4 and 5, running FedRSPO+ with E local SGD steps per round produces global iterates $\bar{\Theta}_t = \frac{1}{N}\sum_i \Theta_t^{(i)}$. The weighted average $\hat{\Theta}_T = \sum_t \eta_t \bar{\Theta}_t / \sum_t \eta_t$ satisfies*

$$\mathbb{E}[F_T(\hat{\Theta}_T) - F_T(\Theta^\star)] \;\le\; \frac{D_\Theta^2 + \sum_{t=0}^{T-1}\eta_t^2\, G^2(1 + 4\gamma E^2) + 4\sum_{t=0}^{T-1}\eta_t\,\Delta_t}{2\sum_{t=0}^{T-1}\eta_t}, \tag{13}$$

where $G = 2DX$ is the uniform gradient bound (Corollary 1), and $\gamma = \eta_t L_t = \alpha X^2$ is the constant *effective condition number enforced by the co-annealed schedule $\eta_t = \alpha\lambda_t/4$.*

We follow a similar proof strategy to [59]. Let $\bar{g}_t = \frac{1}{N}\sum_{i=1}^N g_{i,t}$ denote the aggregated update direction, so $\bar{\Theta}_{t+1} = \bar{\Theta}_t - \eta_t \bar{g}_t$. Expanding the squared distance to $\Theta^\star$:

$$\mathbb{E}\|\bar{\Theta}_{t+1} - \Theta^\star\|^2 = \mathbb{E}\|\bar{\Theta}_t - \Theta^\star\|^2 - 2\eta_t\,\mathbb{E}\langle \bar{g}_t,\, \bar{\Theta}_t - \Theta^\star\rangle + \eta_t^2\,\mathbb{E}\|\bar{g}_t\|^2. \tag{14}$$

Since $\mathbb{E}[\|\bar{g}_t\|^2] \le G^2$ by Assumption 4, it remains to lower-bound the inner product. By the convexity and L_t-smoothness of $F_{i,t}$, we decompose the per-client term:

$$\begin{aligned}
\langle \nabla F_{i,t}(\Theta_t^{(i)}),\, \bar{\Theta}_t - \Theta^\star\rangle &= \langle \nabla F_{i,t}(\Theta_t^{(i)}),\, \Theta_t^{(i)} - \Theta^\star\rangle + \langle \nabla F_{i,t}(\Theta_t^{(i)}),\, \bar{\Theta}_t - \Theta_t^{(i)}\rangle \\
&\ge (F_{i,t}(\Theta_t^{(i)}) - F_{i,t}(\Theta^\star)) + (F_{i,t}(\bar{\Theta}_t) - F_{i,t}(\Theta_t^{(i)}) - \frac{L_t}{2}\|\bar{\Theta}_t - \Theta_t^{(i)}\|^2) \\
&= F_{i,t}(\bar{\Theta}_t) - F_{i,t}(\Theta^\star) - \frac{L_t}{2}\|\bar{\Theta}_t - \Theta_t^{(i)}\|^2.
\end{aligned}$$

Averaging over clients yields:

$$\mathbb{E}\langle \bar{g}_t,\, \bar{\Theta}_t - \Theta^\star\rangle \ge \mathbb{E}[F_t(\bar{\Theta}_t) - F_t(\Theta^\star)] - \frac{L_t}{2N}\sum_{i=1}^N \mathbb{E}\|\bar{\Theta}_t - \Theta_t^{(i)}\|^2. \tag{15}$$

The client drift is bounded by the accumulation of at most E local gradient steps with bounded norm G:

$$\frac{1}{N}\sum_{i=1}^N \mathbb{E}\|\bar{\Theta}_t - \Theta_t^{(i)}\|^2 \le 4E^2\eta_t^2 G^2. \tag{16}$$

Substituting Eqs. (15) and (16) into Eq. (14) and isolating the objective gap:

$$2\eta_t\,\mathbb{E}[F_t(\bar{\Theta}_t) - F_t(\Theta^\star)] \le \mathbb{E}\|\bar{\Theta}_t - \Theta^\star\|^2 - \mathbb{E}\|\bar{\Theta}_{t+1} - \Theta^\star\|^2 + \eta_t^2 G^2 + 4L_t E^2 \eta_t^3 G^2. \tag{17}$$

To relate the transient objective F_t to the target F_T, we apply the tracking bound Δ_t from Eq. (12):

$$F_t(\bar{\Theta}_t) - F_t(\Theta^\star) \ge F_T(\bar{\Theta}_t) - F_T(\Theta^\star) - 2\Delta_t. \tag{18}$$

Substituting into Eq. (17) and applying the co-annealed invariant $\eta_t L_t = \alpha X^2 =: \gamma$:

$$\begin{aligned} 2\eta_t\,\mathbb{E}[F_T(\bar{\Theta}_t) - F_T(\Theta^\star)] \le \mathbb{E}\|\bar{\Theta}_t - \Theta^\star\|^2 - \mathbb{E}\|\bar{\Theta}_{t+1} - \Theta^\star\|^2 \\ + \eta_t^2 G^2(1 + 4\gamma E^2) + 4\eta_t \Delta_t. \end{aligned} \tag{19}$$

Summing from $t = 0$ to $T-1$ telescopes the distance terms. Dropping the non-positive terminal distance $-\mathbb{E}\|\bar{\Theta}_T - \Theta^\star\|^2$, dividing by $2\sum_t \eta_t$, and applying Jensen's inequality for the convex F_T over the weighted average iterate completes the proof.

Corollary 3 (Frozen schedule: $O(1/(T\eta_f) + \eta_f)$ rate)**.** *Under the conditions of Theorem 4, suppose the schedule is frozen at round t_f, so that $\lambda_t = \lambda_f$ and $\eta_t = \eta_f$ for all $t \ge t_f$. Then for $T \gg t_f$:*

$$\mathbb{E}[F_T(\hat{\Theta}_T) - F_T(\Theta^\star)] = O\left(\frac{1}{T\,\eta_f} + \eta_f\right).$$

Proof. After freezing, $\lambda_t = \lambda_T = \lambda_f$ for all $t \ge t_f$, so $\Delta_t = 0$ for $t \ge t_f$. Splitting the sums in Eq. (13) at t_f and noting that the pre-freeze contributions are bounded by constants A_1, A_2, A_3 depending on (t_f, λ_0, η_0):

$$\sum_{t=0}^{T-1} \eta_t = \underbrace{\sum_{t=0}^{t_f-1} \eta_t}_{A_1} + (T - t_f)\,\eta_f,$$

$$\sum_{t=0}^{T-1} \eta_t^2\, G^2(1 + 4\gamma E^2) = \underbrace{\sum_{t=0}^{t_f-1} \eta_t^2\, G^2(1 + 4\gamma E^2)}_{A_2} + (T - t_f)\,\eta_f^2\, G^2(1 + 4\gamma E^2),$$

$$\sum_{t=0}^{T-1} \eta_t\,\Delta_t = \sum_{t=0}^{t_f-1} \eta_t\,\Delta_t =: A_3.$$

Substituting into Eq. (13),

$$\mathbb{E}[F_T(\hat{\Theta}_T) - F_T(\Theta^\star)] \;\le\; \frac{D_\Theta^2 + A_2 + A_3}{(T - t_f)\eta_f} + \frac{\eta_f\, G^2(1 + 4\gamma E^2)}{2}.$$

For T sufficiently large relative to t_f, the transient constants are absorbed into the $O(1/(T\eta_f))$ term, yielding

$$\mathbb{E}[F_T(\hat{\Theta}_T) - F_T(\Theta^\star)] = O\left(\frac{1}{T\,\eta_f} + \eta_f\right).$$

□

Lemma 1 (Neural composition constants)**.** *Suppose $g(x;\theta)$ is differentiable and satisfies Assumption 4. Then*

$$\|\nabla_\theta F_{i,t}(\theta)\| \le 2DJ, \qquad F_{i,t} \text{ is } \left(\frac{4J^2}{\lambda_t} + 2D\rho\right)\text{-smooth}.$$

Proof. By the chain rule, $\nabla_\theta \ell = J_\theta g(x;\theta)^T \nabla_{\hat{c}} \ell$. Proposition 3 gives $\|\nabla_{\hat{c}} \ell\| \le 2D$ and $\nabla_{\hat{c}} \ell$ is $4/\lambda_t$-Lipschitz. Combining these with the Jacobian bound and Jacobian Lipschitzness gives

$$\|\nabla_\theta \ell(\theta) - \nabla_\theta \ell(\theta')\| \le \left(\frac{4J^2}{\lambda_t} + 2D\rho\right)\|\theta - \theta'\|.$$

Taking expectations preserves both bounds and concludes the proof. □

Proposition 5 (Post-freeze nonconvex FedRSPO+)**.** *Suppose Assumptions 1 to 5 hold. After the annealing schedule freezes at λ_f, define*

$$F_f(\theta) := \frac{1}{N}\sum_{i=1}^{N} F_{i,f}(\theta), \qquad L_f^{\mathrm{NN}} := \frac{4J^2}{\lambda_f} + 2D\rho.$$

Assume unbiased stochastic gradients with variance at most σ^2 and post-freeze gradient dissimilarity

$$\frac{1}{N}\sum_{i=1}^{N}\|\nabla F_{i,f}(\theta) - \nabla F_f(\theta)\|^2 \le \zeta^2 \quad \text{for all } \theta.$$

Running FedRSPO+ with E local SGD steps and constant post-freeze step size $\eta_f \le c/(EL_f^{\mathrm{NN}})$ gives, for $s \sim \mathrm{Unif}\{t_f, \dots, T-1\}$,

$$\mathbb{E}\|\nabla F_f(\bar{\theta}_s)\|^2 \le O\left(\frac{F_f(\bar{\theta}_{t_f}) - F_f^{\inf}}{\eta_f E(T-t_f)} + L_f^{\mathrm{NN}}\eta_f\frac{\sigma^2}{N} + (L_f^{\mathrm{NN}})^2\eta_f^2E^2(\sigma^2+\zeta^2)\right).$$

Thus, choosing $\eta_f = \Theta((T-t_f)^{-1/2})$ yields $\mathbb{E}\|\nabla F_f(\bar{\theta}_s)\|^2 = O((T-t_f)^{-1/2})$, up to constants depending on $J, \rho, D, \lambda_f, \sigma, \zeta$, and E.

Additionally, the co-annealed schedule $\eta_t = \alpha\lambda_t/4$ satisfies

$$\eta_t L_t^{\mathrm{NN}} = \alpha J^2 + \frac{\alpha D\rho\lambda_t}{2} \le \alpha\left(J^2 + \frac{D\rho\lambda_{\max}}{2}\right),$$

so choosing $\alpha \le c/(E(J^2 + D\rho\lambda_{\max}/2))$ keeps the effective step size controlled during annealing.

Proof sketch. By Lemma 1, after freezing $\lambda_t = \lambda_f$ the objective F_f is L_f^{NN}-smooth, with $L_f^{\mathrm{NN}} = 4J^2/\lambda_f + 2D\rho$. The stated bound then follows by applying the standard nonconvex FedAvg descent analysis in the literature [59, 60] under unbiased stochastic gradients and bounded client gradient dissimilarity. Summing the one-round descent inequality from t_f to $T-1$ and sampling $s \sim \mathrm{Unif}\{t_f, \dots, T-1\}$ gives the displayed stationary-point bound. The co-annealed stability claim follows by substituting $\eta_t = \alpha\lambda_t/4$ into $\eta_t L_t^{\mathrm{NN}}$. □

Remark 4 (Scope and practical interpretation)**.** *Theorem 4 is stated for the convex linear-model setting, whereas Proposition 5 gives a post-freeze stationary-point analogue for the nonconvex neural predictors used in our experiments. The additional constants J and ρ in Assumption 4 control, respectively, the sensitivity of the predictor to parameter perturbations and the variation of this sensitivity along the training trajectory. For smooth activations with bounded derivatives, these constants are finite on compact parameter regions. For ReLU networks, the Jacobian is only piecewise constant and may jump at activation-boundary crossings, so Assumption 4 only provides a trajectory-level regularity condition.*

The resulting smoothness bound $L_f^{\mathrm{NN}} = 4J^2/\lambda_f + 2D\rho$ can be conservative when J or ρ is large. Therefore, the nonconvex result only adapts a standard FedAvg guarantee to the frozen RSPO+ objective, and the experiments provide empirical support for the same bias–stability mechanism beyond the convex setting. When $\rho = 0$, as in the linear case, the co-annealed condition reduces to controlling the invariant $\eta_t L_t = O(\alpha J^2)$, matching the role of the convex co-annealing condition.

A.2 Unified bias–heterogeneity tradeoff

The results in Theorem 3, Theorem 1, and Corollary 2 can be combined into a single certificate for the unregularized regret of the regularized decision.

Proposition 6. *For client i, define*

$$\widehat{w}_{i,\lambda}(\hat{c}) := \Pi_{S_i}(-\hat{c}/\lambda), \qquad \mathrm{Reg}_{i,\lambda}^{\mathrm{LP}}(\hat{c}) := c_i^T\widehat{w}_{i,\lambda}(\hat{c}) - \min_{w\in S_i} c_i^T w.$$

Let $\ell_i^\lambda(\hat{c})$ denote the RSPO+ loss on client i, and let $H_{ij}^\lambda(\hat{c}) := T_{ij}^\lambda(\hat{c}) + B_{ij}^\lambda(\hat{c})$ be the heterogeneity bound from Theorem 3. Then for any aggregation weights $q \in \Delta_N$,

$$\mathrm{Reg}_{i,\lambda}^{\mathrm{LP}}(\hat{c}) \le \sum_{j=1}^{N} q_j\ell_j^\lambda(\hat{c}) + \sum_{j=1}^{N} q_j H_{ij}^\lambda(\hat{c}) + \frac{\lambda}{2}R_i^2.$$

Consequently,

$$\sum_{i=1}^{N} q_i \mathrm{Reg}_{i,\lambda}^{\mathrm{LP}}(\hat{c}) \le \sum_{i=1}^{N} q_i \ell_i^{\lambda}(\hat{c}) + \frac{\lambda}{2} \sum_{i=1}^{N} q_i R_i^2 + \sum_{i,j=1}^{N} q_i q_j H_{ij}^{\lambda}(\hat{c}).$$

Under uniform bounds $\|c_i\| \le C$, $\|\hat{c}\| \le B$, $R_i \le R$, $D_i \le D$, $\|c_i - c_j\| \le c_\Delta$, *and* $d_H(S_i, S_j) \le \delta$, *this gives*

$$\overline{\mathrm{Reg}}_{\lambda}^{\mathrm{LP}} \le \bar{\ell}_{\lambda} + O\left(\lambda R^2 + \frac{c_\Delta}{\lambda} + \sqrt{\frac{\delta}{\lambda}} + \delta\right),$$

for fixed bounded constants and $0 < \lambda \le 1$. *Thus smaller* λ *reduces regularization bias but amplifies objective and feasible-set heterogeneity.*

The proposition can be viewed as the bias–stability tradeoff for the regularized decision $\widehat{w}_{i,\lambda}$ targeted by the annealing and freezing rule. When the exact-regularization condition discussed above holds, the same certificate applies to the LP decision selected at evaluation.

Proof. For each client i, write

$$\ell_i^{\lambda}(\hat{c}) := \ell_{\mathrm{RSPO+}}(c_i, \hat{c}; S_i, \lambda) = D_{M_\lambda \xi_{S_i}}(c_i - 2\hat{c}, -c_i),$$

and define the regularized decision

$$\widehat{w}_{i,\lambda}(\hat{c}) := \Pi_{S_i}(-\hat{c}/\lambda).$$

The unregularized regret of this regularized decision is

$$\mathrm{Reg}_{i,\lambda}^{\mathrm{LP}}(\hat{c}) := c_i^T \widehat{w}_{i,\lambda}(\hat{c}) - z_i^{\star}, \qquad z_i^{\star} := \min_{w \in S_i} c_i^T w.$$

For two clients i, j, define

$$c_{ij} := \|c_i - c_j\|, \qquad \delta_{ij} := d_H(S_i, S_j), \qquad R_{ij} := \max\{R_i, R_j\}.$$

Let

$$H_{ij}^{\lambda}(\hat{c}) := T_{ij}^{\lambda}(\hat{c}) + B_{ij}^{\lambda}(\hat{c}),$$

where the two terms are the symmetric versions of the quantities in Theorem 3:

$$T_{ij}^{\lambda}(\hat{c}) := \min\Big\{D_i c_{ij} + \delta_{ij}(\|c_j - 2\hat{c}\| + \|c_j\| + 2\lambda R_{ij}),\ D_j c_{ij} + \delta_{ij}(\|c_i - 2\hat{c}\| + \|c_i\| + 2\lambda R_{ij})\Big\},$$

and

$$B_{ij}^{\lambda}(\hat{c}) := (2\|\hat{c}\| + 2\lambda R_{ij}) \left[\frac{c_{ij}}{\lambda} + \min\left\{\sqrt{2\delta_{ij}\left(\frac{\|c_j\|}{\lambda} + R_{ij}\right)}, \sqrt{2\delta_{ij}\left(\frac{\|c_i\|}{\lambda} + R_{ij}\right)}\right\}\right].$$

By Theorem 3,

$$|\ell_i^{\lambda}(\hat{c}) - \ell_j^{\lambda}(\hat{c})| \le H_{ij}^{\lambda}(\hat{c}).$$

Equivalently,

$$\ell_i^{\lambda}(\hat{c}) \le \ell_j^{\lambda}(\hat{c}) + H_{ij}^{\lambda}(\hat{c}) \qquad \text{for every } j.$$

Let $q \in \Delta_N$ be any distribution over clients. Multiplying the last inequality by q_j and summing over j yields

$$\ell_i^{\lambda}(\hat{c}) = \sum_{j=1}^{N} q_j \ell_i^{\lambda}(\hat{c}) \le \sum_{j=1}^{N} q_j \ell_j^{\lambda}(\hat{c}) + \sum_{j=1}^{N} q_j H_{ij}^{\lambda}(\hat{c}).$$

By Corollary 2, the unregularized regret of the regularized decision satisfies

$$\mathrm{Reg}_{i,\lambda}^{\mathrm{LP}}(\hat{c}) \le \ell_i^{\lambda}(\hat{c}) + \frac{\lambda}{2} R_i^2.$$

Combining the previous two displays gives, for every target client i,

$$\mathrm{Reg}_{i,\lambda}^{\mathrm{LP}}(\hat{c}) \le \sum_{j=1}^{N} q_j \ell_j^{\lambda}(\hat{c}) + \sum_{j=1}^{N} q_j H_{ij}^{\lambda}(\hat{c}) + \frac{\lambda}{2} R_i^2.$$

This proves the client-wise statement.

For the aggregate statement, multiply the client-wise inequality by q_i and sum over i:

$$\sum_{i=1}^{N} q_i \mathrm{Reg}_{i,\lambda}^{\mathrm{LP}}(\hat{c}) \le \sum_{i=1}^{N} q_i \sum_{j=1}^{N} q_j \ell_j^{\lambda}(\hat{c}) + \sum_{i=1}^{N} q_i \sum_{j=1}^{N} q_j H_{ij}^{\lambda}(\hat{c}) + \frac{\lambda}{2} \sum_{i=1}^{N} q_i R_i^2.$$

Since $\sum_i q_i = 1$, the first term simplifies:

$$\sum_{i=1}^{N} q_i \sum_{j=1}^{N} q_j \ell_j^{\lambda}(\hat{c}) = \sum_{j=1}^{N} q_j \ell_j^{\lambda}(\hat{c}).$$

Therefore,

$$\sum_{i=1}^{N} q_i \mathrm{Reg}_{i,\lambda}^{\mathrm{LP}}(\hat{c}) \le \sum_{j=1}^{N} q_j \ell_j^{\lambda}(\hat{c}) + \frac{\lambda}{2} \sum_{i=1}^{N} q_i R_i^2 + \sum_{i,j=1}^{N} q_i q_j H_{ij}^{\lambda}(\hat{c}).$$

It remains only to derive the simplified asymptotic form. Suppose

$$\|c_i\| \le C, \qquad \|\hat{c}\| \le B, \qquad R_i \le R, \qquad D_i \le D,$$

and for all client pairs,

$$\|c_i - c_j\| \le c_\Delta, \qquad d_H(S_i, S_j) \le \delta.$$

Then $R_{ij} \le R$, and the envelope term satisfies

$$T_{ij}^{\lambda}(\hat{c}) \le D c_\Delta + \delta(2C + 2B + 2\lambda R).$$

Similarly, the decision term satisfies

$$B_{ij}^{\lambda}(\hat{c}) \le (2B + 2\lambda R) \left[\frac{c_\Delta}{\lambda} + \sqrt{2\delta \left(\frac{C}{\lambda} + R \right)} \right].$$

Thus

$$H_{ij}^{\lambda}(\hat{c}) \le D c_\Delta + \delta(2C + 2B + 2\lambda R) + (2B + 2\lambda R) \left[\frac{c_\Delta}{\lambda} + \sqrt{2\delta \left(\frac{C}{\lambda} + R \right)} \right].$$

Substituting this bound into the aggregate inequality and using $\sum_i q_i R_i^2 \le R^2$ gives

$$\overline{\mathrm{Reg}}_{\lambda}^{\mathrm{LP}} \le \overline{\ell}_\lambda + \frac{\lambda}{2} R^2 + D c_\Delta + \delta(2C + 2B + 2\lambda R) + (2B + 2\lambda R) \left[\frac{c_\Delta}{\lambda} + \sqrt{2\delta \left(\frac{C}{\lambda} + R \right)} \right].$$

For fixed bounded B, C, D, R and $0 < \lambda \le 1$, this is summarized as

$$\overline{\mathrm{Reg}}_{\lambda}^{\mathrm{LP}} \le \overline{\ell}_\lambda + O\left(\lambda R^2 + \frac{c_\Delta}{\lambda} + \sqrt{\frac{\delta}{\lambda}} + \delta \right).$$

□

Remark 5 (Connection to LP evaluation.)**.** *Proposition 6 bounds the LP regret of the regularized decision $\widehat{w}_{i,\lambda}(\hat{c})$. If, for a given client i, prediction $\hat{c}$, and regularization level λ, the exact-regularization condition holds so that*

$$\Pi_{S_i}(-\hat{c}/\lambda) \in \arg\min_{w \in S_i} \hat{c}^T w,$$

then the same bound applies to the LP decision selected at evaluation. For polyhedral feasible sets this condition holds for sufficiently small λ [49, 50]. Without such a condition, Proposition 6 should be interpreted as a certificate for the regularized decision and as a surrogate-level justification for LP evaluation.

B Proofs and Derivations

Proof of Proposition 1: SPO+ as a Bregman Divergence

Proof. Recall the SPO+ loss from Section 3:

$$\ell_{\text{SPO+}}(c, \hat{c}) = \xi_S(c - 2\hat{c}) - 2w^\star(c)^T(c - \hat{c}) + z^\star(c).$$

Since $w^\star(c) \in \arg\min_{w \in S} c^T w = \arg\max_{w \in S} (-c)^T w = \partial \xi_S(-c)$ and $z^\star(c) = \min_{w \in S} c^T w = -\xi_S(-c)$, and moreover if the optimizer is unique $\nabla \xi_S(-c) = w^\star(c)$ (by Danskin's theorem, as S is compact), we can write the linearization term as

$$-2w^\star(c)^T(c - \hat{c}) + z^\star(c) = -\xi_S(-c) - \nabla \xi_S(-c)^T[(c - 2\hat{c}) - (-c)].$$

Substituting into the loss yields

$$\ell_{\text{SPO+}}(c, \hat{c}) = \xi_S(c - 2\hat{c}) - \xi_S(-c) - \nabla \xi_S(-c)^T[(c - 2\hat{c}) - (-c)] = D_{\xi_S}(c - 2\hat{c},\, -c),$$

which is the Bregman divergence of the support function evaluated at $x = c - 2\hat{c}$, $y = -c$. □

Proof of Proposition 2: Smoothed SPO+ Loss

Proof. **Step 1: Moreau envelope of the support function.** By convex duality, the conjugate of ξ_S is the indicator function I_S of the set S. The Moreau envelope can therefore be written as the biconjugate of the regularized indicator:

$$F_{\lambda,S}(u) = M_{\lambda \xi_S}(u) = \Big(I_S + \frac{\lambda}{2}\|\cdot\|^2\Big)^*(u) = \sup_{s \in S}\Big\{\langle u, s\rangle - \frac{\lambda}{2}\|s\|^2\Big\}.$$

Completing the square inside the supremum gives

$$\langle u, s\rangle - \frac{\lambda}{2}\|s\|^2 = \frac{1}{2\lambda}\|u\|^2 - \frac{\lambda}{2}\|s - u/\lambda\|^2.$$

The first term is independent of s, so taking the supremum over $s \in S$ is equivalent to minimizing $\|s - u/\lambda\|^2$ over S:

$$F_{\lambda,S}(u) = \frac{1}{2\lambda}\|u\|^2 - \frac{\lambda}{2}\, d_S^2(u/\lambda), \tag{20}$$

where $d_S^2(v) = \inf_{s \in S}\|v - s\|^2$. The supremum is attained at $s^\star = \Pi_S(u/\lambda)$, and Danskin's theorem yields

$$\nabla F_{\lambda,S}(u) = \Pi_S(u/\lambda). \tag{21}$$

Step 2: Bregman divergence. We evaluate $D_{F_{\lambda,S}}(x, y) = F_{\lambda,S}(x) - F_{\lambda,S}(y) - \langle \nabla F_{\lambda,S}(y),\, x - y\rangle$ with $x = c - 2\hat{c}$ and $y = -c$. Write $\bar{v} = \Pi_S(y/\lambda) = \Pi_S(-c/\lambda)$. Expanding the y-dependent terms using (20) and (21):

$$\begin{aligned} F_{\lambda,S}(y) + \langle \nabla F_{\lambda,S}(y),\, x - y\rangle &= \frac{1}{2\lambda}\|y\|^2 - \frac{\lambda}{2}\|y/\lambda - \bar{v}\|^2 + \langle \bar{v},\, x - y\rangle \\ &= \langle y, \bar{v}\rangle - \frac{\lambda}{2}\|\bar{v}\|^2 + \langle \bar{v},\, x - y\rangle \\ &= \langle x, \bar{v}\rangle - \frac{\lambda}{2}\|\bar{v}\|^2, \end{aligned}$$

where the second equality uses $\frac{1}{2\lambda}\|y\|^2 - \frac{\lambda}{2}\|y/\lambda - \bar{v}\|^2 = \langle y, \bar{v}\rangle - \frac{\lambda}{2}\|\bar{v}\|^2$, which follows from expanding the squared norm. Subtracting from $F_{\lambda,S}(x)$:

$$\begin{aligned} D_{F_{\lambda,S}}(x, y) &= \frac{1}{2\lambda}\|x\|^2 - \frac{\lambda}{2}\, d_S^2(x/\lambda) - \langle x, \bar{v}\rangle + \frac{\lambda}{2}\|\bar{v}\|^2 \\ &= \frac{\lambda}{2}\|x/\lambda - \bar{v}\|^2 - \frac{\lambda}{2}\, d_S^2(x/\lambda). \end{aligned}$$

Setting $u = (c - 2\hat{c})/\lambda$ and $v = -c/\lambda$, noting $\bar{v} = \Pi_S(v)$ and $d_S^2(u) = \|u - \Pi_S(u)\|^2$, we obtain

$$\ell_{\text{RSPO+}}(c, \hat{c}; \lambda) = D_{F_{\lambda,S}}(c - 2\hat{c},\, -c) = \frac{\lambda}{2}\|u - \Pi_S(v)\|^2 - \frac{\lambda}{2}\|u - \Pi_S(u)\|^2.$$

□

Remark 6 (QLP equivalence, Remark 1)**.** *An equivalent derivation applies the original SPO+ construction of [5] to the quadratically regularized LP* $z_{\mathrm{QLP}} = \min_{w \in S} c^T w + \frac{\lambda}{2}\|w\|^2$ *from Section 3. The QLP optimizer is* $w^\star_{\mathrm{reg}} = \Pi_S(-c/\lambda)$ *and the optimal value is* $z^\star_{\mathrm{QLP}} = -F_{\lambda,S}(-c)$*. Substituting into the SPO+ template and simplifying reproduces exactly the loss in Proposition 2.*

Proof of Corollary 1: Gradient of RSPO+

Proof. From Proposition 2, $\ell_{\mathrm{RSPO+}}(c, \hat{c}; \lambda) = D_{F_{\lambda,S}}(x, y)$ with $x(\hat{c}) = c - 2\hat{c}$ and $y = -c$ (independent of $\hat{c}$). The gradient of the Bregman divergence with respect to its first argument is $\nabla_x D_F(x, y) = \nabla F(x) - \nabla F(y)$. By the chain rule, $\nabla_{\hat{c}}\, x = -2I$, so

$$\nabla_{\hat{c}}\, \ell_{\mathrm{RSPO+}}(c, \hat{c}; \lambda) = -2(\nabla F_{\lambda,S}(c - 2\hat{c}) - \nabla F_{\lambda,S}(-c)).$$

Substituting $\nabla F_{\lambda,S}(u) = \Pi_S(u/\lambda)$ from (21) yields

$$\nabla_{\hat{c}}\, \ell_{\mathrm{RSPO+}}(c, \hat{c}; \lambda) = -2\Big(\Pi_S(\frac{c - 2\hat{c}}{\lambda}) - \Pi_S(-\frac{c}{\lambda})\Big).$$

□

Proof of Proposition 3: Characteristics of RSPO+

Proof. We prove each property in turn.

(i) Non-negativity. Non-negativity is immediate since $\ell_{\mathrm{RSPO+}}$ is a Bregman divergence of the convex function $F_{\lambda,S}$.

(ii) Convexity in $\hat{c}$. The Moreau envelope $F_{\lambda,S}$ is convex (as the Moreau envelope of a convex function). A Bregman divergence $D_F(x, y)$ is convex in x for any convex F, since $D_F(x, y) = F(x) - F(y) - \langle \nabla F(y),\, x - y\rangle$ and the last two terms are affine in x. Because $x = c - 2\hat{c}$ is an affine mapping of $\hat{c}$, the composition $\ell_{\mathrm{RSPO+}}(c, \hat{c}; \lambda) = D_{F_{\lambda,S}}(c - 2\hat{c},\, -c)$ is globally convex with respect to $\hat{c}$.

(iii) $2D$-Lipschitz continuity. By Corollary 1, for any $(c, \hat{c})$,

$$\|\nabla_{\hat{c}}\, \ell_{\mathrm{RSPO+}}\| = 2\|\Pi_S(u) - \Pi_S(v)\| \le 2 \sup_{a,b \in S} \|a - b\| = 2D,$$

since both projections land in S. A function whose gradient norm is uniformly bounded by L is L-Lipschitz.

(iv) $4/\lambda$-smoothness and $\lambda/4$-cocoercivity. The Euclidean projection Π_S onto a closed convex set is *firmly non-expansive*:

$$\langle \Pi_S(a) - \Pi_S(b),\, a - b\rangle \ge \|\Pi_S(a) - \Pi_S(b)\|^2 \quad \forall\, a, b.$$

Let $\hat{c}_1, \hat{c}_2$ be two prediction vectors and define $x_k = (c - 2\hat{c}_k)/\lambda$ for $k = 1, 2$. By Corollary 1, $\nabla_{\hat{c}}\, \ell(\hat{c}_k) = -2(\Pi_S(x_k) - \Pi_S(v))$ where $v = -c/\lambda$ is common. Thus

$$\nabla \ell(\hat{c}_1) - \nabla \ell(\hat{c}_2) = -2(\Pi_S(x_1) - \Pi_S(x_2)),$$

and $x_1 - x_2 = -2(\hat{c}_1 - \hat{c}_2)/\lambda$. Evaluating the inner product:

$$\begin{aligned}
\langle \nabla \ell(\hat{c}_1) - \nabla \ell(\hat{c}_2),\, \hat{c}_1 - \hat{c}_2\rangle &= -2\langle \Pi_S(x_1) - \Pi_S(x_2),\, \hat{c}_1 - \hat{c}_2\rangle \\
&= \lambda \langle \Pi_S(x_1) - \Pi_S(x_2),\, x_1 - x_2\rangle \\
&\ge \lambda\, \|\Pi_S(x_1) - \Pi_S(x_2)\|^2 \\
&= \frac{\lambda}{4}\, \|\nabla \ell(\hat{c}_1) - \nabla \ell(\hat{c}_2)\|^2,
\end{aligned}$$

where the inequality uses firm non-expansiveness and the last equality uses $\|\nabla \ell(\hat{c}_1) - \nabla \ell(\hat{c}_2)\| = 2\|\Pi_S(x_1) - \Pi_S(x_2)\|$. This establishes $\lambda/4$-cocoercivity. By the Baillon–Haddad theorem [61], a convex function with a β-cocoercive gradient is $1/\beta$-smooth, so $\ell_{\mathrm{RSPO+}}$ is $4/\lambda$-smooth. □

Proof of Theorem 1: Regret and Decision-Error Bounds

Proof. Define the regularized objective $f(c, w) = c^T w + \frac{\lambda}{2}\|w\|^2$ and let $w^\star = \Pi_S(-c/\lambda)$, $\hat{w} = \Pi_S(-\hat{c}/\lambda)$ denote the optimal and predicted regularized decisions, respectively. The regularized regret is $R(c, \hat{w}) = f(c, \hat{w}) - f(c, w^\star)$.

With $u = (c - 2\hat{c})/\lambda$, Proposition 2 gives

$$\ell_{\mathrm{RSPO+}} = \frac{\lambda}{2}\|u - w^\star\|^2 - \frac{\lambda}{2}\, d_S^2(u).$$

Since $\hat{w} \in S$, the point $\hat{w}$ is feasible for the distance minimization, so $d_S^2(u) \le \|u - \hat{w}\|^2$. Therefore

$$\ell_{\mathrm{RSPO+}} \ge \frac{\lambda}{2}\|u - w^\star\|^2 - \frac{\lambda}{2}\|u - \hat{w}\|^2. \tag{22}$$

Expanding the difference of squares:

$$\begin{aligned}\frac{\lambda}{2}\|u - w^\star\|^2 - \frac{\lambda}{2}\|u - \hat{w}\|^2 &= \lambda\langle u,\, \hat{w} - w^\star\rangle + \frac{\lambda}{2}\|w^\star\|^2 - \frac{\lambda}{2}\|\hat{w}\|^2 \\ &= \langle c - 2\hat{c},\, \hat{w} - w^\star\rangle + \frac{\lambda}{2}\|w^\star\|^2 - \frac{\lambda}{2}\|\hat{w}\|^2.\end{aligned}$$

Rearranging the linear terms:

$$\begin{aligned}&= \underbrace{(c^T\hat{w} + \frac{\lambda}{2}\|\hat{w}\|^2) - (c^T w^\star + \frac{\lambda}{2}\|w^\star\|^2)}_{R(c,\,\hat{w})} + \lambda\|w^\star\|^2 - \lambda\|\hat{w}\|^2 - 2\hat{c}^T(\hat{w} - w^\star) \\ &= R(c, \hat{w}) + 2\Big[(\hat{c}^T w^\star + \frac{\lambda}{2}\|w^\star\|^2) - (\hat{c}^T\hat{w} + \frac{\lambda}{2}\|\hat{w}\|^2)\Big] \\ &= R(c, \hat{w}) + 2[f(\hat{c}, w^\star) - f(\hat{c}, \hat{w})].\end{aligned}$$

Since $f(\hat{c}, \cdot)$ is λ-strongly convex in w and is minimized over S at $\hat{w} = \Pi_S(-\hat{c}/\lambda)$, the suboptimality gap satisfies

$$f(\hat{c}, w^\star) - f(\hat{c}, \hat{w}) \ge \frac{\lambda}{2}\|w^\star - \hat{w}\|^2.$$

Combining with (22) yields Eq. (1):

$$\ell_{\mathrm{RSPO+}}(c, \hat{c}; \lambda) \ge R(c, \hat{w}) + \lambda\|\hat{w} - w^\star\|^2.$$

Since $R(c, \hat{w}) \ge 0$ (by optimality of $w^\star$), the decision-error bound Eq. (2) follows immediately:

$$\|\hat{w} - w^\star\| \le \sqrt{\frac{\ell_{\mathrm{RSPO+}}(c, \hat{c}; \lambda)}{\lambda}}.$$

□

Proof of Corollary 2: Unregularized Regret Bound for RSPO+

Proof. By definition, the regularized regret from Theorem 1 expands as $R(c, \hat{w}) = (c^T\hat{w} + \frac{\lambda}{2}\|\hat{w}\|^2) - (c^T w^\star_{\mathrm{reg}} + \frac{\lambda}{2}\|w^\star_{\mathrm{reg}}\|^2)$. Rearranging to isolate the unregularized cost of the prediction gives:

$$c^T\hat{w} = R(c, \hat{w}) + c^T w^\star_{\mathrm{reg}} + \frac{\lambda}{2}\|w^\star_{\mathrm{reg}}\|^2 - \frac{\lambda}{2}\|\hat{w}\|^2.$$

Because $w^\star_{\mathrm{reg}}$ minimizes the regularized objective over S, its regularized objective is lower than that of the true unregularized optimum $w^\star_{\mathrm{LP}} \in S$:

$$c^T w^\star_{\mathrm{reg}} + \frac{\lambda}{2}\|w^\star_{\mathrm{reg}}\|^2 \le c^T w^\star_{\mathrm{LP}} + \frac{\lambda}{2}\|w^\star_{\mathrm{LP}}\|^2 = z^\star + \frac{\lambda}{2}\|w^\star_{\mathrm{LP}}\|^2.$$

Substituting this upper bound into our expression for $c^T\hat{w}$ and subtracting $z^\star$ from both sides:

$$R_{\mathrm{unreg}}(c, \hat{w}) \le R(c, \hat{w}) + \frac{\lambda}{2}\|w^\star_{\mathrm{LP}}\|^2 - \frac{\lambda}{2}\|\hat{w}\|^2.$$

Substituting $R(c, \hat{w}) \leq \ell_{\text{RSPO+}}(c, \hat{c}; \lambda) - \lambda\|\hat{w} - w^\star_{\text{reg}}\|^2$ from Theorem 1 yields:

$$R_{\text{unreg}}(c, \hat{w}) \leq \ell_{\text{RSPO+}} + \frac{\lambda}{2}\|w^\star_{\text{LP}}\|^2 - \lambda\|\hat{w} - w^\star_{\text{reg}}\|^2 - \frac{\lambda}{2}\|\hat{w}\|^2.$$

To maximize the negative quadratic components, let $x = \hat{w} - w^\star_{\text{reg}}$, so $\|\hat{w}\|^2 = \|x + w^\star_{\text{reg}}\|^2$. The x-dependent terms become $-\lambda\|x\|^2 - \frac{\lambda}{2}\|x + w^\star_{\text{reg}}\|^2 = -\frac{3\lambda}{2}\|x\|^2 - \lambda x^T w^\star_{\text{reg}} - \frac{\lambda}{2}\|w^\star_{\text{reg}}\|^2$. This unconstrained concave quadratic is strictly maximized at $x = -w^\star_{\text{reg}}/3$, yielding a maximum value of $-\frac{\lambda}{3}\|w^\star_{\text{reg}}\|^2$. Substituting this maximum gives us,

$$R_{\text{unreg}}(c, \hat{w}) \leq \ell_{\text{RSPO+}}(c, \hat{c}; \lambda) + \frac{\lambda}{2}\|w^\star_{\text{LP}}\|^2 - \frac{\lambda}{3}\|w^\star_{\text{reg}}\|^2. \tag{23}$$

Since $\|w^\star_{\text{reg}}\|^2 \geq 0$ and $\|w^\star_{\text{LP}}\| \leq R$, this yields the proof. □

Proof of Proposition 4: RSPO+ Stability under Perturbation of λ

Proof. We prove the two bounds separately.

Gradient bound Eq. (5) By Corollary 1, $\nabla_{\hat{c}}\, \ell_{\text{RSPO+}}(\lambda) = -2(\Pi_S(u_\lambda) - \Pi_S(v_\lambda))$, where $u_\lambda = (c - 2\hat{c})/\lambda$ and $v_\lambda = -c/\lambda$. For two parameters $\lambda_1, \lambda_2 > 0$:

$$\|\nabla\ell(\lambda_1) - \nabla\ell(\lambda_2)\| \leq 2(\|\Pi_S(u_{\lambda_1}) - \Pi_S(u_{\lambda_2})\| + \|\Pi_S(v_{\lambda_1}) - \Pi_S(v_{\lambda_2})\|),$$

by the triangle inequality. Since Π_S is 1-Lipschitz (non-expansive),

$$\|\Pi_S(u_{\lambda_1}) - \Pi_S(u_{\lambda_2})\| \leq \|u_{\lambda_1} - u_{\lambda_2}\| = \|c - 2\hat{c}\|\, |\frac{1}{\lambda_1} - \frac{1}{\lambda_2}|,$$

and similarly $\|\Pi_S(v_{\lambda_1}) - \Pi_S(v_{\lambda_2})\| \leq \|c\|\, |\frac{1}{\lambda_1} - \frac{1}{\lambda_2}|$. Combining gives

$$\|\nabla\ell(\lambda_1) - \nabla\ell(\lambda_2)\| \leq 2(\|c - 2\hat{c}\| + \|c\|)\Big|\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\Big|.$$

Loss bound Eq. (4) Write the RSPO+ loss via its Bregman divergence representation:

$$\ell(\lambda) = F_{\lambda,S}(c - 2\hat{c}) - F_{\lambda,S}(-c) - \langle \Pi_S(-c/\lambda),\, 2c - 2\hat{c}\rangle.$$

We bound the variation of each component.

Envelope terms. Using the supremum form $F_{\lambda,S}(u) = \sup_{s\in S}\{\langle u, s\rangle - \frac{\lambda}{2}\|s\|^2\}$, for any fixed u and $s \in S$:

$$|\langle u, s\rangle - \frac{\lambda_1}{2}\|s\|^2| - |\langle u, s\rangle - \frac{\lambda_2}{2}\|s\|^2| \leq \frac{|\lambda_1 - \lambda_2|}{2}\|s\|^2 \leq \frac{R^2}{2}|\lambda_1 - \lambda_2|.$$

Taking suprema over $s \in S$ on both sides yields $|F_{\lambda_1,S}(u) - F_{\lambda_2,S}(u)| \leq \frac{R^2}{2}|\lambda_1 - \lambda_2|$. Applying this to $u = c - 2\hat{c}$ and $u = -c$ and summing:

$$|F_{\lambda_1,S}(c - 2\hat{c}) - F_{\lambda_2,S}(c - 2\hat{c})| + |F_{\lambda_1,S}(-c) - F_{\lambda_2,S}(-c)| \leq R^2|\lambda_1 - \lambda_2|.$$

Linearization term. The variation in the inner-product term is

$$\begin{aligned}|\langle \Pi_S(-c/\lambda_1) - \Pi_S(-c/\lambda_2),\, 2c - 2\hat{c}\rangle| &\leq 2\|c - \hat{c}\|\, \|\Pi_S(-c/\lambda_1) - \Pi_S(-c/\lambda_2)\| \\ &\leq 2\|c - \hat{c}\| \cdot \|c\|\, |\frac{1}{\lambda_1} - \frac{1}{\lambda_2}|,\end{aligned}$$

by non-expansiveness of Π_S. Combining the envelope and linearization bounds by the triangle inequality:

$$|\ell_{\text{RSPO+}}(\lambda_1) - \ell_{\text{RSPO+}}(\lambda_2)| \leq R^2|\lambda_1 - \lambda_2| + 2\|c\|\, \|c - \hat{c}\|\, \Big|\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\Big|.$$

□

B.1 Proof of Remark 3:MSE-SPO+ homotopy

Proof. Since $0 \in \mathrm{int}(S)$, there exists $r > 0$ such that $B(0,r) := \{x \in \mathbb{R}^d : \|x\| < r\} \subseteq S$. For any p with $\|p\| < r$ we have $p \in S$, and since S is convex and closed the projection satisfies $\Pi_S(p) = p$. Set $u = \dfrac{c - 2\hat{c}}{\lambda}$ and $v = -\dfrac{c}{\lambda}$. Under the boundedness assumptions,

$$\|u\| = \frac{\|c - 2\hat{c}\|}{\lambda} \leq \frac{C_{\max} + 2B}{\lambda} \leq \frac{C_{\max} + 2B}{\bar{\lambda}} = r,$$

and

$$\|v\| = \frac{\|c\|}{\lambda} \leq \frac{C_{\max}}{\lambda} \leq \frac{C_{\max}}{\bar{\lambda}} < r.$$

Hence $u, v \in S$, so $\Pi_S(u) = u$ and $\Pi_S(v) = v$. From Proposition 2, the RSPO+ loss is

$$\ell_{\mathrm{RSPO+}}(c, \hat{c}; \lambda) = \frac{\lambda}{2}(\|u - \Pi_S(v)\|^2 - \|u - \Pi_S(u)\|^2).$$

Substituting $\Pi_S(u) = u$ and $\Pi_S(v) = v$ yields

$$\ell_{\mathrm{RSPO+}} = \frac{\lambda}{2}(\|u - v\|^2 - 0) = \frac{\lambda}{2}\left\|\frac{c - 2\hat{c}}{\lambda} + \frac{c}{\lambda}\right\|^2 = \frac{\lambda}{2} \cdot \frac{4\|c - \hat{c}\|^2}{\lambda^2} = \frac{2}{\lambda}\|c - \hat{c}\|^2.$$

From Corollary 1, $\nabla_{\hat{c}}\, \ell_{\mathrm{RSPO+}} = -2(\Pi_S(u) - \Pi_S(v))$. Substituting the identity projections:

$$\nabla_{\hat{c}}\, \ell_{\mathrm{RSPO+}} = -2(u - v) = -2\left(\frac{c - 2\hat{c}}{\lambda} + \frac{c}{\lambda}\right) = -\frac{4}{\lambda}(c - \hat{c}).$$

This is exactly $\nabla_{\hat{c}}[\dfrac{2}{\lambda}\|c - \hat{c}\|^2]$. □

B.2 Proof of Theorem 3: RSPO+ Heterogeneity Bound

We prove Theorem 3 through three supporting lemmas and an algebraic decomposition of the RSPO+ loss. Throughout, we use $\|\cdot\| = \|\cdot\|$ and write $v^\star_{\lambda,S}(c) := \Pi_S(-c/\lambda)$ for the optimal regularized decision.

Loss decomposition. Recall from Proposition 2 that the RSPO+ loss is the generalized Bregman divergence of the Moreau envelope of the support function:

$$\ell_S(\hat{c}, c) \;=\; F_{\lambda,S}(c - 2\hat{c}) \;-\; F_{\lambda,S}(-c) \;-\; \langle \nabla F_{\lambda,S}(-c),\; (c - 2\hat{c}) - (-c) \rangle. \tag{24}$$

Since the Moreau envelope of the support function satisfies $\nabla F_{\lambda,S}(-c) = \Pi_S(-c/\lambda) = v^\star_{\lambda,S}(c)$ and evaluates to $F_{\lambda,S}(-c) = \langle -c, v^\star_{\lambda,S}(c) \rangle - \dfrac{\lambda}{2}\|v^\star_{\lambda,S}(c)\|^2$, we expand the linearization term as

$$\begin{aligned} -\langle v^\star_{\lambda,S}(c),\; 2c - 2\hat{c} \rangle &= -2c^T v^\star_{\lambda,S}(c) + 2\hat{c}^T v^\star_{\lambda,S}(c) \\ &= 2\Big(F_{\lambda,S}(-c) + \frac{\lambda}{2}\|v^\star_{\lambda,S}(c)\|^2\Big) + 2\hat{c}^T v^\star_{\lambda,S}(c). \end{aligned}$$

Substituting back into (24) yields

$$\ell_S(\hat{c}, c) \;=\; \underbrace{F_{\lambda,S}(c - 2\hat{c}) + F_{\lambda,S}(-c)}_{M_{\lambda,S}(\hat{c},\, c)} \;+\; \underbrace{2\hat{c}^T v^\star_{\lambda,S}(c) + \lambda\|v^\star_{\lambda,S}(c)\|^2}_{B_{\lambda,S}(\hat{c},\, c)}. \tag{25}$$

This decomposition mirrors the SPO+ structure and allows us to bound the envelope term M and the decision term B independently under perturbations to c and S.

Hausdorff distance. Unlike SPO+, the origin-centered quadratic penalty $\frac{\lambda}{2}\|w\|^2$ in the Moreau envelope breaks translation invariance. We therefore measure set perturbations via the standard Hausdorff distance $\delta := d_H(S_{\mathrm{ref}}, S_{\mathrm{cl}})$, and define the joint bounding radius $R_{\max} := \max\{R(S_{\mathrm{ref}}),\, R(S_{\mathrm{cl}})\}$.

Lemma 2 (Sensitivity of the Moreau envelope to set perturbations)**.** *Let $S_1, S_2 \subseteq \mathbb{R}^d$ be nonempty compact convex sets with $\max\{R(S_1), R(S_2)\} \le R_{\max}$ and $\delta = d_H(S_1, S_2)$. Then for any $u \in \mathbb{R}^d$:*

$$|F_{\lambda,S_1}(u) - F_{\lambda,S_2}(u)| \;\le\; \delta(\|u\|+\lambda R_{\max}).$$

Proof. Let $w_1 = \arg\max_{w\in S_1}(u^T w - \dfrac{\lambda}{2}\|w\|^2)$. By definition of the Hausdorff distance, there exists $\tilde{w}_2 \in S_2$ with $\|w_1 - \tilde{w}_2\| \le \delta$. Since w_1 is optimal for S_1 and $\tilde{w}_2$ is feasible for S_2,

$$\begin{aligned} F_{\lambda,S_1}(u) - F_{\lambda,S_2}(u) &\le (u^T w_1 - \frac{\lambda}{2}\|w_1\|^2) - (u^T \tilde{w}_2 - \frac{\lambda}{2}\|\tilde{w}_2\|^2) \\ &= \langle u - \frac{\lambda}{2}(w_1 + \tilde{w}_2),\; w_1 - \tilde{w}_2\rangle \\ &\le (\|u\|+\frac{\lambda}{2}(\|w_1\|+\|\tilde{w}_2\|))\|w_1 - \tilde{w}_2\| \\ &\le (\|u\|+\lambda R_{\max})\,\delta, \end{aligned}$$

where the second line uses the identity $a^T a - b^T b = \langle a+b, a-b\rangle$ applied to the quadratic terms, and the last line uses $\|w_1\|, \|\tilde{w}_2\| \le R_{\max}$. Swapping the roles of S_1 and S_2 and taking the maximum of both directions yields the absolute-value bound. □

Lemma 3 (Diameter-Lipschitzness of the joint envelope term)**.** *Under the compactness assumption of Assumption 1, for any $c_1, c_2, \hat{c} \in \mathbb{R}^d$ and fixed compact convex set S:*

$$|M_{\lambda,S}(\hat{c}, c_1) - M_{\lambda,S}(\hat{c}, c_2)| \;\le\; D(S)\,\|c_1 - c_2\|.$$

Proof. By the maximization definition, $F_{\lambda,S}(x) - F_{\lambda,S}(y) \le \max_{w\in S}(x-y)^T w$ for any x, y (the quadratic terms cancel in the upper bound by using the same maximizer for both terms). Applying this to both components of $M_{\lambda,S}$:

$$F_{\lambda,S}(c_1 - 2\hat{c}) - F_{\lambda,S}(c_2 - 2\hat{c}) \le \max_{w\in S}(c_1 - c_2)^T w,$$

$$F_{\lambda,S}(-c_1) - F_{\lambda,S}(-c_2) \le \max_{v\in S}(c_2 - c_1)^T v \;=\; \max_{v\in S}(c_1 - c_2)^T(-v).$$

Summing these inequalities:

$$\begin{aligned} M_{\lambda,S}(\hat{c}, c_1) - M_{\lambda,S}(\hat{c}, c_2) &\le \max_{w,v\in S}(c_1 - c_2)^T(w - v) \\ &\le \|c_1 - c_2\| \cdot \sup_{w,v\in S}\|w - v\| = \|c_1 - c_2\|\, D(S). \end{aligned}$$

Swapping c_1 and c_2 and taking the maximum gives the absolute-value bound. □

Lemma 4 (Set stability of Euclidean projections)**.** *Let $S_1, S_2 \subseteq \mathbb{R}^d$ be nonempty compact convex sets with $\max\{R(S_1), R(S_2)\} \le R_{\max}$ and $\delta = d_H(S_1, S_2)$. For any $p \in \mathbb{R}^d$:*

$$\|\Pi_{S_1}(p) - \Pi_{S_2}(p)\| \;\le\; \sqrt{2\delta(\|p\|+R_{\max})}.$$

Proof. Let $z_i = \Pi_{S_i}(p)$ for $i = 1, 2$. By the Hausdorff distance bound, there exist proxy points $\tilde{z}_1 \in S_1$ and $\tilde{z}_2 \in S_2$ satisfying $\|z_2 - \tilde{z}_1\| \le \delta$ and $\|z_1 - \tilde{z}_2\| \le \delta$.

The characterization of Euclidean projections onto closed convex sets gives the variational inequalities

$$\langle p - z_1,\; \tilde{z}_1 - z_1\rangle \le 0, \qquad \langle p - z_2,\; \tilde{z}_2 - z_2\rangle \le 0.$$

We expand $\|z_1 - z_2\|^2$ by inserting the proxy points:

$$\begin{aligned} \|z_1 - z_2\|^2 &= \langle z_1 - z_2,\; z_1 - z_2\rangle \\ &= \langle p - z_2,\; z_1 - z_2\rangle + \langle p - z_1,\; z_2 - z_1\rangle \\ &= [\langle p - z_2,\; z_1 - \tilde{z}_2\rangle + \langle p - z_2,\; \tilde{z}_2 - z_2\rangle] \\ &\quad + [\langle p - z_1,\; z_2 - \tilde{z}_1\rangle + \langle p - z_1,\; \tilde{z}_1 - z_1\rangle] \\ &\le \langle p - z_2,\; z_1 - \tilde{z}_2\rangle + 0 + \langle p - z_1,\; z_2 - \tilde{z}_1\rangle + 0 \\ &\le \|p - z_2\|\,\|z_1 - \tilde{z}_2\| + \|p - z_1\|\,\|z_2 - \tilde{z}_1\| \\ &\le \delta(\|p - z_1\| + \|p - z_2\|). \end{aligned}$$

Since $z_i \in S_i$ with $\|z_i\| \leq R_{\max}$, the triangle inequality gives $\|p - z_i\| \leq \|p\| + R_{\max}$ for each i. Therefore,

$$\|z_1 - z_2\|^2 \leq 2\delta(\|p\| + R_{\max}).$$

Taking the square root completes the proof. □

Proof of Theorem 3. Using decomposition (25), the total discrepancy satisfies the triangle inequality:

$$|\ell_{S_{\mathrm{ref}}}(\hat{c}, c_{\mathrm{ref}}) - \ell_{S_{\mathrm{cl}}}(\hat{c}, c_{\mathrm{cl}})| \leq |M_{\lambda,S_{\mathrm{ref}}}(\hat{c}, c_{\mathrm{ref}}) - M_{\lambda,S_{\mathrm{cl}}}(\hat{c}, c_{\mathrm{cl}})| + |B_{\lambda,S_{\mathrm{ref}}}(\hat{c}, c_{\mathrm{ref}}) - B_{\lambda,S_{\mathrm{cl}}}(\hat{c}, c_{\mathrm{cl}})|.$$

Step 1: Bounding $T(\hat{c})$. We add and subtract the intermediate term $M_{\lambda,S_{\mathrm{ref}}}(\hat{c}, c_{\mathrm{cl}})$:

$$\begin{aligned} &|M_{\lambda,S_{\mathrm{ref}}}(\hat{c}, c_{\mathrm{ref}}) - M_{\lambda,S_{\mathrm{cl}}}(\hat{c}, c_{\mathrm{cl}})| \\ &\quad \leq \underbrace{|M_{\lambda,S_{\mathrm{ref}}}(\hat{c}, c_{\mathrm{ref}}) - M_{\lambda,S_{\mathrm{ref}}}(\hat{c}, c_{\mathrm{cl}})|}_{\text{objective shift at fixed } S_{\mathrm{ref}}} + \underbrace{|M_{\lambda,S_{\mathrm{ref}}}(\hat{c}, c_{\mathrm{cl}}) - M_{\lambda,S_{\mathrm{cl}}}(\hat{c}, c_{\mathrm{cl}})|}_{\text{set shift at fixed } c_{\mathrm{cl}}}. \end{aligned}$$

By Lemma 3, the objective-shift term is bounded by $D_{\mathrm{ref}}\, c_d$.

For the set-shift term, recall that $M_{\lambda,S}(\hat{c}, c) = F_{\lambda,S}(c - 2\hat{c}) + F_{\lambda,S}(-c)$. Applying Lemma 2 to each envelope evaluation independently at arguments $u_1 = c_{\mathrm{cl}} - 2\hat{c}$ and $u_2 = -c_{\mathrm{cl}}$:

$$\begin{aligned} |M_{\lambda,S_{\mathrm{ref}}}(\hat{c}, c_{\mathrm{cl}}) - M_{\lambda,S_{\mathrm{cl}}}(\hat{c}, c_{\mathrm{cl}})| &\leq \delta(\|c_{\mathrm{cl}} - 2\hat{c}\| + \lambda R_{\max}) + \delta(\|c_{\mathrm{cl}}\| + \lambda R_{\max}) \\ &= \delta(\|c_{\mathrm{cl}} - 2\hat{c}\| + \|c_{\mathrm{cl}}\| + 2\lambda R_{\max}). \end{aligned}$$

Combining these gives:

$$|M_{\mathrm{ref}} - M_{\mathrm{cl}}| \leq D_{\mathrm{ref}}\, c_d + \delta(\|c_{\mathrm{cl}} - 2\hat{c}\| + \|c_{\mathrm{cl}}\| + 2\lambda R_{\max}).$$

Alternatively, one may pivot through $M_{\lambda,S_{\mathrm{cl}}}(\hat{c}, c_{\mathrm{ref}})$ instead, yielding the analogous bound with D_{cl}, $\|c_{\mathrm{ref}} - 2\hat{c}\|$, and $\|c_{\mathrm{ref}}\|$. Taking the minimum of both pivots gives $T(\hat{c})$.

Step 2: Bounding $B(\hat{c})$. Write $v_i := v^\star_{\lambda,S_i}(c_i) = \Pi_{S_i}(-c_i/\lambda)$ for $i \in \{\mathrm{ref}, \mathrm{cl}\}$. The decision-term gap factors as:

$$\begin{aligned} |B_{\lambda,S_{\mathrm{ref}}}(\hat{c}, c_{\mathrm{ref}}) - B_{\lambda,S_{\mathrm{cl}}}(\hat{c}, c_{\mathrm{cl}})| &= |2\hat{c}^T v_{\mathrm{ref}} + \lambda\|v_{\mathrm{ref}}\|^2 - 2\hat{c}^T v_{\mathrm{cl}} - \lambda\|v_{\mathrm{cl}}\|^2| \\ &= |2\hat{c}^T(v_{\mathrm{ref}} - v_{\mathrm{cl}}) + \lambda\langle v_{\mathrm{ref}} + v_{\mathrm{cl}},\ v_{\mathrm{ref}} - v_{\mathrm{cl}}\rangle| \\ &= |\langle 2\hat{c} + \lambda(v_{\mathrm{ref}} + v_{\mathrm{cl}}),\ v_{\mathrm{ref}} - v_{\mathrm{cl}}\rangle| \\ &\leq (2\|\hat{c}\| + \lambda\|v_{\mathrm{ref}} + v_{\mathrm{cl}}\|)\, \|v_{\mathrm{ref}} - v_{\mathrm{cl}}\| \\ &\leq (2\|\hat{c}\| + 2\lambda R_{\max})\, \|v_{\mathrm{ref}} - v_{\mathrm{cl}}\|, \end{aligned}$$

where the last line uses $\|v_{\mathrm{ref}} + v_{\mathrm{cl}}\| \leq \|v_{\mathrm{ref}}\| + \|v_{\mathrm{cl}}\| \leq 2R_{\max}$ since each $v_i \in S_i$.

It remains to bound $\|v_{\mathrm{ref}} - v_{\mathrm{cl}}\|$. We decompose via the triangle inequality, pivoting through the intermediate projection $\Pi_{S_{\mathrm{ref}}}(-c_{\mathrm{cl}}/\lambda)$:

$$\begin{aligned} \|v_{\mathrm{ref}} - v_{\mathrm{cl}}\| &= \|\Pi_{S_{\mathrm{ref}}}(-c_{\mathrm{ref}}/\lambda) - \Pi_{S_{\mathrm{cl}}}(-c_{\mathrm{cl}}/\lambda)\| \\ &\leq \underbrace{\|\Pi_{S_{\mathrm{ref}}}(-c_{\mathrm{ref}}/\lambda) - \Pi_{S_{\mathrm{ref}}}(-c_{\mathrm{cl}}/\lambda)\|}_{\text{objective shift, fixed set}} + \underbrace{\|\Pi_{S_{\mathrm{ref}}}(-c_{\mathrm{cl}}/\lambda) - \Pi_{S_{\mathrm{cl}}}(-c_{\mathrm{cl}}/\lambda)\|}_{\text{set shift, fixed objective}}. \end{aligned}$$

The first term is bounded by $\|c_{\mathrm{ref}} - c_{\mathrm{cl}}\|/\lambda = c_d/\lambda$ since Euclidean projection onto a convex set is 1-Lipschitz (firmly non-expansive). The second term is bounded by Lemma 4 evaluated at $p = -c_{\mathrm{cl}}/\lambda$:

$$\|\Pi_{S_{\mathrm{ref}}}(-c_{\mathrm{cl}}/\lambda) - \Pi_{S_{\mathrm{cl}}}(-c_{\mathrm{cl}}/\lambda)\| \leq \sqrt{2\delta\Big(\frac{\|c_{\mathrm{cl}}\|}{\lambda} + R_{\max}\Big)}.$$

Combining yields the decision bound:

$$\|v_{\mathrm{ref}} - v_{\mathrm{cl}}\| \leq \frac{c_d}{\lambda} + \sqrt{2\delta\Big(\frac{\|c_{\mathrm{cl}}\|}{\lambda} + R_{\max}\Big)}.$$

Alternatively, pivoting through $\Pi_{S_{\mathrm{cl}}}(-c_{\mathrm{ref}}/\lambda)$ gives the bound with $\|c_{\mathrm{ref}}\|$ replacing $\|c_{\mathrm{cl}}\|$. Taking the minimum of both pivots and multiplying by the prefactor gives $B(\hat{c})$.

Combining. Summing the bounds from Steps 1 and 2 completes the proof:

$$|\ell_{S_{\mathrm{ref}}}(\hat{c}, c_{\mathrm{ref}}) - \ell_{S_{\mathrm{cl}}}(\hat{c}, c_{\mathrm{cl}})| \;\le\; T(\hat{c}) + B(\hat{c}).$$

□

Asymptotic comparison with SPO+. We verify the claims in Table 2 by examining the limiting behavior. As $c_d \to 0$ and $\delta \to 0$:

- **RSPO+:** Both $T(\hat{c})$ and $B(\hat{c})$ vanish since $T = \mathcal{O}(c_d + \delta)$ and $B = \mathcal{O}(c_d/\lambda + \sqrt{\delta})$, giving a combined rate of $\mathcal{O}(c_d + \sqrt{\delta})$. No geometric assumptions on S are required.
- **SPO+:** The heterogeneity bound takes the form $H(\hat{c}) = \mathcal{O}(c_d + \delta_N + D_{\min})$, where $D_{\min} := \min\{D(S_{\mathrm{ref}}), D(S_{\mathrm{cl}})\}$ persists as an irreducible bias even when $c_d = \delta_N = 0$, unless the feasible sets are strongly convex (in which case $D_{\min}$ can be replaced by $\mathcal{O}(\sqrt{\delta_N})$).

C Experimental Details

This section provides the full experimental details of the setup described in Section 4.

Experiments were conducted on AMD EPYC Milan CPU nodes with 128 cores per node and 528 GB RAM. The total wall-clock runtime for the reported experiments was approximately 250 hours. For the Warcraft case study, each seed-method run took approximately 4 hours for MSE and SPO+ and approximately 6 hours for RSPO+ and the other DFL methods. For the synthetic knapsack experiments, one full heterogeneity-design run covering the three main methods (MSE, SPO+, RSPO+) and three heterogeneity levels across three axes and five seeds took approximately 90 minutes. The PJM experiments took approximately 2 minutes for five seeds.

C.1 Shared Configuration

For all experiments and baselines, we use a federated warm-start then local personalization. We use a shallow neural network with 64 neurons for the fractional knapsack and PJM experiments. For the Warcraft experiment, we use a truncated ResNet-18 [62], as per the literature [24]. We implement our method and baselines using the `InferOpt.jl` framework [56] in `Julia` [63] with the `Lux.jl` [64] deep learning package. We use the `Gurobi` [65] through the `JuMP` framework [66].

While FedRSPO+ uses the per-client λ and learning rates defined in Eq. (11), we use a learning rate of 10^{-3} for the baselines. We use $\kappa^{\mathrm{LR}} = 0.1, \kappa^{\lambda} = 1$, and $\alpha = 0.01$ for the fractional knapsack experiments, and $\alpha = 10^{-3}$ for the PJM and Warcraft experiments, selected to set initial learning rates within the same magnitude as the baseline learning rate 10^{-3}. We use a stopping threshold $\tau = 1\%$. We report absolute regret $R = z_{\mathrm{LP}}(c, \hat{w}) - z_{\mathrm{LP}}(c, w^{\star})$ and relative regret $R/z_{\mathrm{LP}}^{\mathrm{Avg}}(c, w^{\star})$ as metrics, where $z_{\mathrm{LP}}^{\mathrm{Avg}}$ is the per-client average ground-truth objective functions.

Optimizer. All experiments use Adam [67] with gradient clipping at norm 1.0 and no weight decay. Tabular experiments (knapsack, PJM) use `Float64` precision; the Warcraft experiment uses `Float32`.

Freeze mechanism. The annealing schedules for λ and η in Eq. (11) use a scheduler: once the stopping criterion (Eq. (6)) is satisfied, the schedule freezes and training continues. For SPO+ and MSE baselines, all schedules are held constant ($\kappa^{\lambda} = \kappa^{\mathrm{LR}} = 0$). We also tested early stopping a few rounds after freezing, and found that while it did help mitigate computational overhead, it did not significantly influence training performance, due to the already annealed learning rate slowing training in later rounds. We implement an early stopping condition 3 rounds after the schedule freezing criterion is met.

C.2 Synthetic Knapsack

Optimization problem. We employ a fractional knapsack with $d = 50$ items. Each client j solves

$$\max_{w \in S_j} \; c_j^T w, \qquad S_j = \{w \in [0,1]^d : \alpha_j^T w \le B_j\},$$

where c_j is the predicted cost (objective) vector, α_j the item weights, and B_j is the knapsack capacity. Since the problem is a maximization problem, we use a sign transform in our loss function implementation, $\max_{w \in S_j} \; c_j^T w = \min_{w \in S_j} \; (-c_j)^T w$.

Ground-truth objective. A single shared basis $W_{\text{true}} \in \mathbb{R}^{d \times p}$ is drawn once. Each client receives a rotated copy $W_j = W_{\text{true}} \, R_j$, where

$$R_j = \exp\Big(\eta_{\text{obj}} \, \frac{A_j}{\|A_j\|}\Big), \qquad A_j = G_j - G_j^T, \quad G_j \sim \mathcal{N}(0, I).$$

Because A_j is skew-symmetric, $\exp(\cdot)$ is an orthogonal matrix, so R_j is a rotation in feature space. When $\eta_{\text{obj}} = 0$ all clients share the same basis. Larger η_{obj} increases the spread of the per-client bases. Given a feature vector $x \in \mathbb{R}^p$, the true cost vector is

$$\theta_{\text{true}}(x) = -(1 + (1 + W_j \, x/\sqrt{p}\,)^4), \qquad \text{normalized by } \; \text{mean}(|\theta_{\text{true}}|). \tag{26}$$

Constraint heterogeneity (η_{constr}). A base capacity is set as $\bar{B} = 0.6 \times d = 30$. Each client draws a multiplicative factor $f_j \sim 2\,\text{Beta}(1/\eta_{\text{constr}}, \, 1/\eta_{\text{constr}})$ (mean 1; for $\eta_{\text{constr}} = 0$ the factor degenerates to 1) and receives capacity $B_j = \bar{B} \cdot f_j$. When the flag `eta_constr_affects_weights` is set, each item weight also receives an independent per-item Beta factor drawn from the same family, otherwise item weights are shared across clients.

Covariate-shift heterogeneity (η_{data}). Each client draws a mean shift $\mu_j = \eta_{\text{data}} \, z_j$, $z_j \sim \mathcal{N}(0, I)$, and per-feature positive scales σ_j constructed with the same Beta-factor mechanism used for constraints. Features are then sampled as

$$x \mid j \; \sim \; \mathcal{N}(\mu_j, \; \text{diag}(\sigma_j^2)).$$

Experimental protocol. Each of the 20 clients receives 100 training, 10 validation, and 100 test samples. Each heterogeneity axis $\eta \in \{0.0, 0.5, 1.0\}$ is varied independently while the others are held at zero. Results are averaged over 5 seeds $\in \{42, 69, 1993, 67, 2026\}$ and reported in Table 3.

Hyperparameters. Table 7 shows the full training configuration for both phases.

Additional sensitivity analysis. We perform additional sensitivity analysis across more of the hyperparameters of FedRSPO+ and demonstrate the results in Fig. 4. Overall FedRSPO+ is fairly insensitive to changes in annealing parameter κ, especially the learning rate parameter κ_{LR}. Reducing τ improves the model slightly, but it reduces the fraction of clients that end up frozen in the 10 rounds of personalization. To achieve even better results, a smaller τ with more annealing rounds until the freezing criterion is met should be considered. We also varied the number of calibration samples per client used to compute the freeze statistic $\tilde{U}_t$. Performance was largely insensitive to this: as $|V_c|$ increased from 2 to 32, the relative-regret improvement of RSPO+ over SPO+ remained nearly constant, varying only from 3.61% to 3.72%. This suggests that the freezing rule is robust to the exact calibration subset size, so we use $|V_c| = 8$ in the main experiments.

Additional ablation. We perform an additional ablation with the fractional knapsack experiment to complement the results of Table 5 across different heterogeneity settings. The results are shown in Table 9.

C.3 Warcraft Shortest Path

Projection model While we use Dijkstra's for the downstream LP oracle, we use a Gurobi QLP for the RSPO+ projections. For a Warcraft grid with vertex set $\mathcal{V}$ and canonical source–sink pair (s, t), let $\mathcal{A} \subseteq \mathcal{V} \times \mathcal{V}$ denote the directed 8-neighbor adjacency graph. For each vertex v, define incoming and outgoing arc sets

$$\delta^-(v) = \{a = (u, v) \in \mathcal{A}\}, \qquad \delta^+(v) = \{a = (v, u) \in \mathcal{A}\}.$$

Table 7: Synthetic knapsack hyperparameters. RSPO+-only entries are marked; SPO+ and MSE use the same configuration with constant schedules ($\kappa = 0$).

Hyperparameter	**Warm-start**	**Personalization**
Communication rounds	10	10
Local epochs per round	3	1
Batch size	64	64
Client fraction	0.4	1.0
κ^{λ} (RSPO+ only)	0.0	1.0
η_0 (baseline)	10^{-3}	10^{-3}
κ^{LR} (RSPO+ only)	0.0	0.1
Per-client λ/η (RSPO+ only)	✓	✓
α (Eq. (11))	0.01	0.01
Freeze threshold τ	—	0.01
Validation samples/client	—	8

Given transformed costs $c \in \mathbb{R}^{|\mathcal{V}|}$ and regularization parameter $\lambda > 0$, the RSPO+ projection oracle solves

$$\Pi_S(c) \in \arg\min_{x,e} \quad \sum_{v\in\mathcal{V}} c_v w_v + \frac{\lambda}{2}\sum_{v\in\mathcal{V}} w_v^2 \tag{27a}$$

$$\text{s.t.} \quad w_s = 1, \tag{27b}$$

$$w_v = \sum_{a\in\delta^-(v)} e_a, \qquad \forall v \in \mathcal{V}\setminus\{s\}, \tag{27c}$$

$$w_v = \sum_{a\in\delta^+(v)} e_a, \qquad \forall v \in \mathcal{V}\setminus\{t\}, \tag{27d}$$

$$\sum_{a\in\delta^-(s)} e_a = 0, \tag{27e}$$

$$\sum_{a\in\delta^+(t)} e_a = 0, \tag{27f}$$

$$0 \le w_v \le 1, \quad 0 \le e_a \le 1, \qquad \forall v \in \mathcal{V}, \quad \forall a \in \mathcal{A}. \tag{27g}$$

The returned relaxed Warcraft decision is $w^\star$. Equivalently, $\Pi_S(c) = \Pi_S(-c/\lambda)$, where S is the relaxed Warcraft path-flow feasible set defined by the constraints above. For a true cost vector c and prediction $\hat{c}$, RSPO+ calls this oracle $w_u = \Pi_S(2\hat{c} - c)$, and $w_v = \Pi_S(c)$, so that $w_u = \Pi_S((c - 2\hat{c})/\lambda)$ and $w_v = \Pi_S(-c/\lambda)$.

Dataset and federated partition. The Warcraft Shortest Path dataset [24] consists of 10,000 training, 1,000 validation, and 1,000 test 96×96 RGB terrain maps on a 12×12 grid with 8-connectivity. In the benchmark-aligned rerun, we form the federated benchmark from the original training split: we reserve 80% of the training set (8,000 samples) for federated training/validation and hold out the remaining 2,000 as a benchmark test set. The 8,000 federated samples are distributed round-robin (after shuffling) across 44 clients, yielding 182 samples per client before the $80/20$ intra-client split, i.e., roughly 146 training and 36 validation samples per client. Each client is evaluated on all 2,000 benchmark test maps. We report mean optimality gap: $(\text{predicted cost} - \text{optimal cost})/\text{optimal cost}$), and its standard deviation across seeds.

Client heterogeneity. Rather than 4 corner-defined clients, we use a border-antipodes construction on the 12×12 grid. Since the grid border contains $4 \cdot 12 - 4 = 44$ cells, this yields 44 clients. Each client's source is one border cell and its sink is the antipodal border cell under a 180° rotation of the grid, inducing all four diagonal endpoint orientations. We additionally enable objective heterogeneity: each client receives its own multiplicative reweighting of the 5 terrain values $c = c_0 \exp(0.6 \cdot x)$, where c_0 is the initial traversal cost of the terrain type and $x \sim \mathcal{N}(0, 1)$. Therefore, clients differ in both endpoints and path costs.

Architecture. The model uses a truncated ResNet-18 backbone, followed by adaptive max-pooling to a 12×12 grid and channel averaging to produce a 12×12 real-valued weight matrix per sample, following the methodology of [1]. Predicted cost grids are decoded into hard paths using the benchmark shortest-path solver (Dijkstra). Float32 precision is used throughout.

Additional baselines. Beyond RSPO+, SPO+, and MSE, the Warcraft experiment includes perturbed Fenchel–Young loss (PFYL), decision-focused perturbed MSE (DPO). PFYL and DPO both use 10 perturbation samples with perturbation scale $\epsilon = 1.0$.

Hyperparameters. Table 8 reports the main configuration. The Warcraft rerun uses five seeds rather than a single seed; reported variability therefore reflects both seed-to-seed and client-level variation.

Table 8: Warcraft shortest path hyperparameters.

Hyperparameter	**Warm-start**	**Personalization**
Communication rounds	10	10
Local epochs per round	3	2
Batch size	64	64
Client fraction	0.5	1.0
λ_0 (RSPO+ only)	2.0	2.0
κ^{λ} (RSPO+ only)	0.0	1.0
η_0	10^{-3}	10^{-3}
κ^{LR} (RSPO+ only)	0.0	0.1
Per-client λ/η (RSPO+ only)	✓	✓
α (Eq. (11), RSPO+ only)	0.001	0.001
Perturbation samples (PFYL/DPO)	10	10
Perturbation scale ϵ (PFYL/DPO)	1.0	1.0
Freeze threshold τ (RSPO+ only)	0.03	0.01

Table 9: Fractional-knapsack component ablation. Values are mean test relative regret (%) $\pm$ sample standard deviation across five seeds, after averaging clients within each seed; lower is better.

Method	$(0,0,0)\downarrow$	$(1,0,0)\downarrow$	$(0,1,0)\downarrow$
Local-only RSPO+	9.74 ± 2.73	9.80 ± 2.78	6.99 ± 1.89
FedSPO+	9.69 ± 2.08	10.93 ± 2.19	12.89 ± 1.99
RSPO+ *(surrogate only)*	2.26 ± 0.40	3.05 ± 0.25	4.71 ± 0.49
RSPO+ *w/* λ_j/η_j	2.26 ± 0.40	3.05 ± 0.25	4.63 ± 0.46
RSPO+ *w/* λ_j/η_j + *P*	2.16 ± 0.36	2.41 ± 0.38	4.08 ± 0.41
RSPO+ (Full)	$\mathbf{1.86} \pm 0.26$	$\mathbf{2.02} \pm 0.26$	$\mathbf{2.58} \pm 0.40$

C.4 PJM Energy Pricing

We use hourly day-ahead locational marginal pricing (LMP) data from the PJM Interconnection, a regional transmission organization spanning 13 U.S. states [25]. We use the following nodes as clients: DOM, PEPCO, PPL, DUQ, BGE, AECO, ATSI, RECO, PSEG, DPL, DAY, METED, JCPL, APS, DEOK, AEP, COMED, PECO, and PENELEC. Prices are retrieved for the year 2025 and assigned to clients, with variation arising from localized congestion effects and are in dollars per MWh [68]. Energy markets provide a natural setting for decision-focused federated learning, as operational data are often siloed across entities and subject to regulatory protections such as NERC CIP-011 [69].

Data processing. The 8 weather covariates are: mean/min/max temperature (°C), mean humidity, mean wind speed, mean/max solar irradiance, and day of week. The 24 target dimensions are $c_h = \mathrm{LMP}_h + \mathrm{congestion}_h$ for hours $h = 0, \ldots, 23$. The demand satisfaction constraint for client k is $\sum_h w_h = \mathrm{demand}_k$ with $w_h \in [0, 1]$, solved as a fractional knapsack with equality constraint.

Each client corresponds to a fixed daily demand customer with demand normalized to a range of 4 to 12 MWh. This setup captures heterogeneity across participants while maintaining realistic operational constraints.

Optimization problem. The downstream task is to minimize daily electricity cost subject to daily demand. This can be formulated as a capacity-constrained fractional knapsack problem, as described in Section 4.1. Specifically, the feasible set is defined as $S = \{w \in [0,1]^{24} : \mathbf{1}^T w = D_j\}$, where D_j is the demand at client j, and the problem is defined as $\min_{w \in S} c^T w$. Demand is randomly assigned from $\{4, 6, 8, 10, 12\}$ to the PJM clients.

Data splits. The data is split chronologically with no shuffling: 50% train, 10% validation, 40% test, allocated by date index, with 19 total clients.

Ablation Study. For our ablation study, we first evaluate a federated-only training run with standard MSE loss and evaluate the mean and standard deviation across 5 seeds. We then replace the MSE loss with the DFL SPO+ loss and further incorporate RSPO+. We use the average of the would-be per-client calibrated parameters, or $\frac{1}{|C|}\sum_{j=1}^{|C|}\lambda_j$ and $\frac{1}{|C|}\sum_{j=1}^{|C|}\eta_j$ from Eq. (11), where $|C|$ is the number of clients. We set $\kappa^\lambda = \kappa^{\mathrm{LR}} = 0$ for the first 2 RSPO+ ablations. We then gradually introduce more aspects of the FedRSPO+ algorithm and evaluate similarly. To showcase the modularity of the algorithm, we incorporate two popular FL methods: FedProx [14] and Ditto [16]. During the federated stage, FedProx adds a regularization term $\mu/2\|\theta_t - \theta\|^2$ to each client's local objective, where θ_t is the global model at the start of the round and θ is the client model. That means that FedProx only affects the federated training stage, not the personalization stage. We use $\mu = 1$ for our experiment, and keep all other parameters the same as in standard FedRSPO+. For Ditto, the federated training phase is identical to standard FedRSPO+. Similarly to FedProx, each client solves a regularized version of their local objective, adding the term $\mu/2\|\theta_t - \theta\|^2$. As opposed to FedProx, Ditto is introduced in the personalization phase only. We use $\mu = 1$ for the regularization parameter, and keep all other settings the same as before.

Table 10: PJM hyperparameters.

Hyperparameter	**Warm-start**	**Personalization**
Communication rounds	10	10
Local epochs per round	3	1
Batch size	64	64
Client fraction	0.4	1.0
κ^λ (RSPO+ only)	0.0	1.0
η_0 (baseline)	10^{-3}	10^{-3}
κ^{LR} (RSPO+ only)	0.0	0.1
Per-client λ/η (RSPO+ only)	✓	✓
α (Eq. (11))	0.01	10^{-3}
Freeze threshold τ	—	0.03
Validation samples/client	—	8

Leave-one-client-out PJM. In the original experiment, each client has roughly six months of training data, and local-only RSPO+ already achieves 61.53 mean regret, compared with 60.65 for full FedRSPO+. The large local datasets leave little estimation error for federation to reduce. To investigate a case where federation would be very helpful for PJM, we design a leave-one-client-out experiment, where we train 18/19 clients normally in a federation, and leave one client out, who has access to only [3, 7, 14, 30, 60] days of data. The client then compares their local model with the federated model, fine-tuned on their data. We repeat this procedure for all 19 client nodes. In both cases we use RSPO+ to isolate the federated effect. The results are shown in Tables 11 and 12.

The federated-only stays constant since it does not train on the held-out target's dataset, so is unaffected by a larger target dataset. In the data-poor regimes the federated model provides the largest benefit, with the personalization providing marginal or no benefit (even 0.06 worse than federated-only at 7 days, which is negligible), due to the small target dataset. Personalization performance improves noticeably at 60 target days. In terms of per-client differences, 19/19 clients prefer the

Table 11: Leave-one-client-out experiment for PJM. Results are reported in absolute decision regret; lower is better.

Target training days	Local RSPO$^+$	Federated only	FedRSPO$^+$ personalized	Gap vs. local (Fed+P)	Relative reduction	
					Fed-only	Fed+P
3	168.48 ± 10.29	73.58 ± 1.86	73.06 ± 2.19	95.43 ± 9.30	56.3%	56.6%
7	167.49 ± 9.49	73.58 ± 1.86	73.64 ± 2.14	93.85 ± 8.71	56.1%	56.0%
14	164.44 ± 10.91	73.58 ± 1.86	73.31 ± 2.15	91.13 ± 10.46	55.3%	55.4%
30	164.37 ± 7.91	73.58 ± 1.86	73.12 ± 2.11	91.26 ± 7.62	55.2%	55.5%
60	153.86 ± 8.17	73.58 ± 1.86	72.20 ± 1.84	81.67 ± 6.51	52.2%	53.1%

Table 12: Client-level benefit of federated-only training over the original local RSPO$^+$ baseline with the PJM experiment.

Target training days	Federated-only beats local	Median client reduction [IQR]
3	19/19	51.1% [37.2%, 66.3%]
7	19/19	47.9% [33.9%, 65.4%]
14	19/19	48.6% [27.4%, 65.6%]
30	19/19	45.1% [28.1%, 64.8%]
60	18/19	41.0% [27.4%, 57.2%]

federated-only model over the local model for target days = 3,7,14, and 30, and 18/19 for target days = 60. We provide a second table in Table 12 with inter-client median reduction and the interquartile range (IQR) for federated only versus local.

At 60 days, PENELEC is the only client with negative transfer from local to federated. These values are lower than those in Table 11 because the Table 11 computes a single ratio of aggregate means, which weights high-regret clients more heavily, whereas in Table 12 each client contributes one percentage equally. We report only federated-only gains, since the benefits of personalization do not meaningfully change the results in this experiment. These IQRs are across the 19 client-level percentage reductions. For each client, regret is first averaged over five seeds and then converted to a percentage relative to that client's local mean. Each experiment rotates all 19 PJM clients as the held-out target and uses five training seeds, giving $19 \times 5 = 95$ observations per target-data budget. For each seed, we first average the metric over the 19 held-out targets.

Fixed demand in PJM. To isolate the effect of feasible set heterogeneity in the PJM experiment and to show that optimization heterogeneity plays a role and FedRSPO+ performs better than federated SPO+ under heterogeneity, we reran the PJM experiment but set each client's demand to be equal. We repeat this for all 5 demand values [4,6,8,10,12] for 5 seeds. At each demand level, the gap between FedRSPO+ and FedSPO+ decreased on average. For seed s, let $G_{\mathrm{het},s}$ and $G_{d,s}$ denote the regret gaps under heterogeneous demands and common demand d, respectively. Each gap is defined as federated-only SPO$^+$ regret minus personalized FedRSPO$^+$ regret. We compute,

$$\Delta_{d,s} = 100\,\frac{G_{d,s} - G_{\mathrm{het},s}}{G_{\mathrm{het},s}}.$$

In Table 13 we report the percentage change, where a negative gap corresponds to a narrower gap. In other words, a negative gap means FedRSPO+ and federated SPO+'s performance was closer with homogeneous feasible sets. The mean percentage change is negative at every demand level. This is consistent with feasible-set heterogeneity contributing to the observed performance difference.

C.5 Centralized and Local-only Baselines

To contextualize the federated results in Table 3 and Table 5, we additionally evaluate two non-federated training regimes. **Centralized** pools all client training data into a single dataset and trains one global model, evaluated on each client's downstream objective at test time. It is a centralized proxy for the value of data sharing when no privacy constraint applies. **Local-only** trains a separate

Table 13: Effect of equal client demands on the regret gap between federated-only SPO$^+$ and personalized FedRSPO$^+$ on the PJM-derived constructed-cost benchmark.

Common demand	Percentage change in gap
4	$-18.50\% \pm 9.73\%$
6	$-8.37\% \pm 10.94\%$
8	$-4.97\% \pm 10.03\%$
10	$-24.33\% \pm 8.71\%$
12	$-20.13\% \pm 8.16\%$

model per client on that client's data alone, with no cross-client communication. It is a lower-bound proxy for federation that quantifies how much each client gains from collaboration. Both regimes use the same architectures, optimizer, and number of gradient updates per client as the federated runs to keep compute comparable, and we report mean $\pm$ std over the same 5 seeds.

Table 14: Synthetic knapsack: relative regret (%) of centralized and local-only baselines on representative heterogeneity configurations, alongside the federated results from Table 3 for reference. Mean $\pm$ std over 5 seeds. Lower is better (↓).

(η_o, η_c, η_d)	Centralized	Local-only	Fed	Fed + P
SPO+				
$(0, 0, 0)$	0.32 ± 0.06	2.42 ± 0.30	9.7 ± 2.1	5.6 ± 1.4
$(0, 0, 0.5)$	3.49 ± 0.62	2.93 ± 0.54	14.1 ± 2.0	6.9 ± 1.3
$(0, 0, 1)$	6.97 ± 0.89	2.47 ± 0.39	17.7 ± 2.1	6.9 ± 0.8
RSPO+ (Ours)				
$(0, 0, 0)$	0.75 ± 0.08	9.74 ± 2.73	2.3 ± 0.4	1.9 ± 0.3
$(0, 0, 0.5)$	4.12 ± 0.40	6.93 ± 1.15	6.5 ± 0.8	3.8 ± 0.2
$(0, 0, 1)$	7.42 ± 0.72	4.15 ± 0.69	9.3 ± 1.2	3.4 ± 0.2

Synthetic knapsack. Table 14 reports relative regret for different heterogeneity configurations. We can see two patterns. First, centralized training is the strongest baseline when clients share a data distribution (rows with $\eta_d = 0$): centralized SPO+ achieves 0.32% relative regret at $(0, 0, 0)$, far below any federated or local-only number. But centralized performance degrades sharply once clients differ in their feature distributions: at $(0, 0, 1)$, centralized SPO+ rises to 6.97%, and centralized RSPO+ to 7.42%, because pooling distributionally distinct clients creates a misspecified training problem. Second, local-only training stays relatively stable under heterogeneity: local-only SPO+ averages $2.96 \pm 0.73\%$ across all 27 heterogeneity configurations and beats centralized SPO+ in every configuration with $\eta_d > 0$. Local-only RSPO+, by contrast, is much weaker than its federated counterpart ($9.74 \pm 2.73\%$ at $(0, 0, 0)$ versus 1.9% for FedRSPO+ with personalization in Table 3), supporting the interpretation that RSPO+'s advantage lies in federation and mitigating heterogeneity.

Table 15: PJM energy pricing: mean absolute regret across training regimes. Mean $\pm$ std over 5 seeds. Federated rows are reproduced from Table 5 for reference. Lower is better (↓).

Regime	MSE	SPO+	RSPO+ (Ours)
Centralized	70.26 ± 0.61	69.37 ± 0.44	69.32 ± 1.18
Local-only	163.76 ± 10.12	69.57 ± 1.72	61.53 ± 0.66
Federated (naive)	175.45 ± 7.26	81.68 ± 1.77	72.89 ± 4.23
FedRSPO+ (Full)	—	—	$\mathbf{60.65 \pm 0.80}$

PJM energy pricing. Table 15 reports the same comparison on the PJM case study, using mean absolute regret to match Table 5. Three observations stand out. First, centralized training close the gap between the three methods: MSE, SPO+, and RSPO+ all sit between 69 and 71, indicating that the PJM clients share enough structure that pooled fitting is a strong strategy regardless of the loss. Second, local-only MSE remains poor (163.76 ± 10.12, close to the federated MSE result of 175.45 ± 7.26), confirming that the improvements in the federated pipeline require decision-aware training to achieve competitive results. Third, local-only RSPO+ achieves 61.53 ± 0.66,

Table 16: QP–LP discrepancies after personalization. Values are mean normalized gaps (%; lower is better), averaged over clients and five seeds, and additionally over 27 heterogeneity configurations for knapsack. The PJM diagnostic uses 10 personalization rounds.

Dataset	Decision gap (%)	Objective gap (%)
Knapsack	21.44	3.26
PJM	42.57	10.43

nearly matching the full FedRSPO+ result of 60.65 ± 0.80 from Table 5. This reflects the relatively homogeneous downstream structure of PJM (all clients face the same top-k selection problem on overlapping price signals), where personalization captures most of the available gain. On synthetic instances with structurally heterogeneous feasible sets, the analogous local-only RSPO+ baseline is much weaker, indicating that federation contributes most when clients differ in their optimization structure rather than only in scale.

Discussion. Together, these baselines reinforce the interpretation of the main results. Centralized training is a useful upper-bound comparator but is not uniformly dominant: it excels when clients are statistically aligned and degrades sharply under data heterogeneity. Local-only training is a strong personalization baseline in two specific regimes (SPO+ in synthetic data-heterogeneous settings and RSPO+ on PJM) but offers no shared model and is inconsistent across heterogeneity types, due to the addition of the downstream objective. The main FedRSPO+ result is a robust middle ground that retains performance across the heterogeneity axes while supporting data privacy. The PJM observation that local-only RSPO+ is competitive with FedRSPO+ also identifies a regime where the federated machinery is nearly redundant. We view formalizing this regime, where shared training adds little beyond the surrogate itself, as a useful direction for follow-up work.

C.6 QP–LP decision alignment

To complement the oracle-bias freeze statistic $\tilde{U}_t$, we directly measure the decision-space discrepancy between the regularized predicted decision and the LP predicted decision,

$$\Delta_{\mathrm{QP-LP}} = \mathbb{E}\left[\left\|\Pi_S(-\hat{c}/\lambda_t) - \arg\min_{w \in S} \hat{c}^T w\right\|\right].$$

On synthetic knapsack, personalization consistently reduces this gap across all configurations: the mean gap decreases from 3.715 after the federated warm start to 1.166 after personalization, a 69% reduction, and this decrease occurs in all $135/135$ configurations. The freeze statistic aligns with this result. Clients that satisfy the freeze criterion have lower average decision gaps (1.112) than clients that do not freeze (1.338), and lower final oracle-bias values $\tilde{U}_t$ (0.00858 versus 0.01735). The hardest cases are primarily driven by data heterogeneity, where increasing η_d from 0 to 1 raises the personalized decision gap from 1.036 to 1.335 and reduces the frozen-client fraction from 0.917 to 0.578.

We observe the same qualitative effect in the PJM case study, where personalization reduces the predicted QP–LP gap from 1.955 to 1.149, a 41% reduction. This residual gap is not negligible, about 42.6% of the average predicted LP decision norm, but it confirms that personalization moves the regularized predicted decision substantially closer to the LP decision used at evaluation. Finally, the final freeze statistic is strongly correlated with the predicted QP–LP decision gap on knapsack at the configuration level ($r = 0.807$), suggesting that $\tilde{U}_t$ is a useful proxy for QP–LP alignment. The direct decision-gap diagnostic captures predicted-decision alignment more explicitly, but adds computational overhead if tracked in situ ($|V_c|$ additional projections per client per round).

We also evaluate the decision and objective gaps for the fractional knapsack and PJM experiments. For client j, let $\widehat{w}_{ji}^{\mathrm{QP}}$ and $\widehat{w}_{ji}^{\mathrm{LP}}$ denote the QP and LP decisions under the same predicted cost vector $\widehat{c}_{ji}$, with the QP using the client's final smoothing coefficient λ. Let $z_{ji}^{\star} = \min_{w \in S_j} c_{ji}^{\top} w$ be the LP optimum under the true cost vector. With $\epsilon = 10^{-8}$, the reported values in Table 16 use the following

Table 17: QP–LP decision alignment on synthetic knapsack. Personalization substantially reduces the predicted decision discrepancy between the regularized RSPO+ decision and the LP decision used at evaluation.

Metric	Mean	Median	Range
Global QP–LP gap	3.715	3.654	3.080–4.456
Personalized QP–LP gap	1.166	1.118	0.922–1.521
Personalized / global ratio	0.314	0.316	0.258–0.354
Final freeze statistic	0.01067	0.00959	0.00822–0.01761
Frozen-client fraction	0.761	0.800	0.300–1.000

Table 18: Bias-free linear predictor results on synthetic knapsack. Entries report mean client relative regret (%) as mean $\pm$ sample standard deviation across five seeds. The heterogeneity tuple represents objective, constraint, and data-distribution heterogeneity, respectively. "Fed" denotes the federated endpoint; "+P" includes subsequent personalization. Bold indicates the lowest mean regret within each training stage.

	MSE		PFYL		DPO		SPO+		RSPO+	
Heterogeneity	Fed	+P	Fed	+P	Fed	+P	Fed	+P	Fed	+P
$(0, 0, 0)$	37.76 ± 2.44	34.89 ± 2.16	41.08 ± 3.14	38.93 ± 3.22	44.99 ± 4.03	44.43 ± 4.32	39.57 ± 3.16	36.79 ± 3.14	$\mathbf{12.64 \pm 0.91}$	$\mathbf{11.04 \pm 0.70}$
$(0.5, 0, 0)$	37.88 ± 2.45	34.97 ± 2.20	41.17 ± 3.10	39.01 ± 3.19	45.02 ± 4.00	44.45 ± 4.27	39.68 ± 3.10	36.90 ± 3.16	$\mathbf{13.30 \pm 1.04}$	$\mathbf{11.54 \pm 0.82}$
$(1, 0, 0)$	38.42 ± 2.43	35.42 ± 2.04	41.57 ± 2.95	39.38 ± 3.06	45.08 ± 3.78	44.49 ± 4.05	40.20 ± 3.00	37.35 ± 3.02	$\mathbf{15.09 \pm 1.14}$	$\mathbf{13.00 \pm 0.97}$
$(0, 0.5, 0)$	42.26 ± 2.69	39.38 ± 2.68	47.22 ± 2.96	45.43 ± 2.89	49.87 ± 3.27	49.38 ± 3.31	45.00 ± 3.06	42.46 ± 2.95	$\mathbf{14.83 \pm 1.19}$	$\mathbf{12.98 \pm 0.93}$
$(0, 1, 0)$	44.03 ± 2.24	41.08 ± 2.22	49.45 ± 2.82	47.93 ± 2.75	51.97 ± 2.87	51.57 ± 2.97	47.11 ± 2.76	44.68 ± 2.87	$\mathbf{15.71 \pm 0.58}$	$\mathbf{13.73 \pm 0.52}$
$(0, 0, 0.5)$	40.93 ± 2.02	38.34 ± 1.67	43.42 ± 2.59	41.60 ± 2.66	46.07 ± 3.04	45.52 ± 3.32	42.43 ± 2.49	40.01 ± 2.54	$\mathbf{18.85 \pm 1.67}$	$\mathbf{16.62 \pm 1.48}$
$(0, 0, 1)$	44.53 ± 1.99	41.90 ± 1.51	45.96 ± 2.62	44.20 ± 2.56	47.39 ± 2.81	46.90 ± 2.88	45.35 ± 2.32	43.05 ± 2.26	$\mathbf{28.29 \pm 2.31}$	$\mathbf{25.19 \pm 2.35}$

normalizations,

$$D_j = 100 \, \frac{\frac{1}{n_j} \sum_{i=1}^{n_j} \left\| \widehat{w}_{ji}^{\mathrm{QP}} - \widehat{w}_{ji}^{\mathrm{LP}} \right\|_2}{\frac{1}{n_j} \sum_{i=1}^{n_j} \left\| \widehat{w}_{ji}^{\mathrm{LP}} \right\|_2 + \epsilon}, \tag{28}$$

$$O_j = \frac{100}{n_j} \sum_{i=1}^{n_j} \frac{\left| c_{ji}^{\top} \left(\widehat{w}_{ji}^{\mathrm{QP}} - \widehat{w}_{ji}^{\mathrm{LP}} \right) \right|}{|z_{ji}^{\star}| + \epsilon}. \tag{29}$$

Both objectives are evaluated using the true linear costs, without the quadratic penalty.

C.7 Linear predictor results

To strengthen the connection to Theorem 4, we reran all seven synthetic-knapsack settings using a bias-free linear predictor, consistent with the predictor class assumed in the convergence analysis. RSPO+ achieves the lowest regret in all 14 federated and personalized comparisons. Averaged across settings, RSPO+ obtains 16.96% federated regret, compared with 40.83% for the strongest baseline, and 14.87% personalized regret, compared with 38.00% for the strongest baseline. This corresponds to approximately $2.4\times$ and $2.6\times$ lower regret, respectively. Detailed results are shown in Table 18.

C.8 Hyperparameter Selection

Hyperparameters were not selected by automated search. The RSPO+-specific parameters (λ_j, η_j) were set using the principled per-client defaults derived in Eq. (11). The remaining parameters (rounds, epochs, batch size, client fraction) were fixed to standard values from the FL literature. A sensitivity analysis over personalization parameters for the knapsack experiment is available in the supplementary code and the results are shown in Fig. 4.

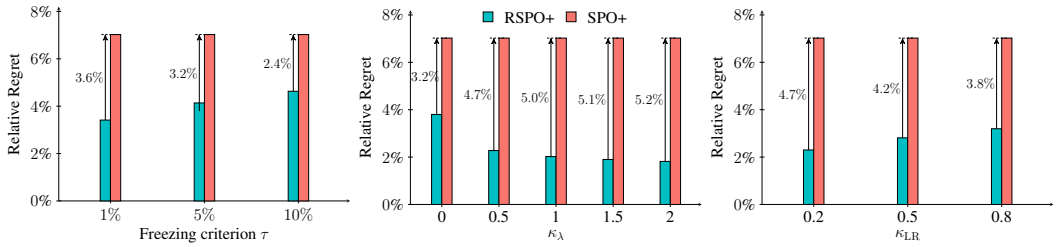


Figure 4: (a) Relative regret (%) of SPO+ and RSPO+ across different (a) values of the freezing threshold τ, (b) values of κ^{λ}, and (c) values of κ^{LR}. Lower is better (↓).

C.9 Algorithms

Input: client datasets $\{(\mathcal{D}_i^{\mathrm{tr}}, \mathcal{V}_i, n_i)\}_{i=1}^N$, rounds $R_{\mathrm{fed}}, R_{\mathrm{per}}$, local epochs E, participation rate ρ, freeze threshold τ, post-freeze budget K

initialize global model ϕ^0
for $i = 1$ **to** N **do**
  initialize per-client schedules $(\lambda_{i,0},\ \eta_{i,0})$
end
global freeze flag $g \leftarrow 0$

`/* Phase 1: Federated training */`
for $r = 0$ **to** $R_{\mathrm{fed}} - 1$ **do**
  if $g = 0$ **then** advance schedule counters $\lambda_i^{(r)}, \eta_i^{(r)}$ for all i
  sample $I_r \subseteq \{1, \dots, N\}$ with $|I_r| = \lceil \rho N \rceil$
  forall $i \in I_r$ ***in parallel*** **do**
    $\tilde{\phi}_i^{r+1} \leftarrow \texttt{LocalTrain}(\phi^r, \mathcal{D}_i^{\mathrm{tr}}, \lambda_i^{(r)}, \eta_i^{(r)}, E)$
    $\Delta_i^r \leftarrow \tilde{\phi}_i^{r+1} - \phi^r$
  end
  $\phi^{r+1} \leftarrow \phi^r + \sum_{i \in I_r} \frac{n_i}{\sum_{j \in I_r} n_j} \Delta_i^r$
  if $\texttt{StopMetric}(\phi^{r+1}, \{\mathcal{V}_i\}, \{\lambda_i^{(r)}\}) \le \tau$ **then** $g \leftarrow 1$ `// freeze annealing`
end

`/* Phase 2: Personalization (warm-started from` $\phi^{R_{\mathrm{fed}}}$`) */`
forall $i = 1, \dots, N$ ***in parallel*** **do**
  $\phi_i^{\mathrm{per}} \leftarrow \phi^{R_{\mathrm{fed}}}$
  initialize personal schedules $(\lambda_{i,\mathrm{per}}^{(0)},\ \eta_{i,\mathrm{per}}^{(0)})$
  client freeze flag $g_i \leftarrow 0$,
  freeze round $t_i \leftarrow +\infty$
  for $r = 0$ **to** $R_{\mathrm{per}} - 1$ **do**
    if $r > t_i + K$ **then break**
    if $g_i = 0$ **then** advance personal schedule counters $\lambda_{i,\mathrm{per}}^{(r)}, \eta_{i,\mathrm{per}}^{(r)}$
    $\phi_i^{\mathrm{per}} \leftarrow \texttt{LocalTrain}(\phi_i^{\mathrm{per}}, \mathcal{D}_i^{\mathrm{tr}}, \lambda_{i,\mathrm{per}}^{(r)}, \eta_{i,\mathrm{per}}^{(r)}, E)$
    if $\texttt{StopMetric}(\phi_i^{\mathrm{per}}, \mathcal{V}_i, \lambda_{i,\mathrm{per}}^{(r)}) \le \tau$ ***and*** $g_i = 0$ **then** $g_i \leftarrow 1$,
    $t_i \leftarrow r$
  end
end
return $\phi^{R_{\mathrm{fed}}}$ and $\{\phi_{i,\star}^{\mathrm{per}}\}_{i=1}^N$

Algorithm 1: FedRSPO+: federated warm start with co-annealing scheduling and personalization.

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 introduction clearly state our three contributions: the RSPO+ surrogate with regret/error bounds, FedRSPO+ with heterogeneity bounds that vanish at homogeneity, and empirical evaluation across three tasks.

   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: We discuss limitations of our approach in Section 5.

   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: All proofs are provided in the Appendix, Section B.

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: We have implementation details for each experiment, including hyperparameter choices and data splits used, to reproduce all our results. All data sources are provided and well-documented. We also provide the code for reproducibility.

   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: Our paper provides the anonymized code necessary to reproduce the experiments, and all data sources are open and well-documented.

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: We explicitly provide all the experimental details and setups for each experiment in the Appendix, Section C.

   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: We report mean $\pm 1\sigma$ (standard deviation) over 5 random seeds for all experiments and describe error bars and methodology in the Appendix, Section C.

   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 provide the computer resources in the implementation details in the Appendix.

   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 read carefully and align with the Ethics Guidelines shown.

   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: [Yes]

    Justification: We state the potential broader impacts in Section 5.

    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: Our datasets are widely used ones, such as PJM energy markets and the Warcraft DFL benchmark. Their safety has been proven by existing work.

    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 cite all works and tools used to properly credit their authors.

    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: [Yes]

    Justification: We have documented our models and methods used, and release them along with the anonymized code.

    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: This 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: This paper does not need to do IRB approvals or equivalent since it does not involve human subject research.

    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: We do not use LLMs for any non-standard uses.

    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.