Distributionally Robust Black Box Optimization-based Bidding Strategy in Auction-based Federated Learning

Xiaoli Tang[1], **Zhuang Qi**[2,*], **Ying-Peng Tang**[1,*], **Xianjie Guo**[3], **Han Yu**[1], **Jie Zhang**[1], **Hengjie Song**[4]
[1]College of Computing and Data Science, Nanyang Technological University, Singapore
[2]School of Software, Shandong University, China
[3]School of Computer Science, Nanjing University of Posts and Telecommunications, China
[4]School of Software Engineering, South China University of Technology, China

Abstract

Auction-based Federated Learning (AFL) provides a market-based mechanism to incentivize decentralized data owners (DOs) to participate in FL model training initiated by data consumers (DCs). However, optimizing the DC's bidding strategy remains challenging due to delayed and non-decomposable utility feedback, which limits the effectiveness of conventional Reinforcement Learning (RL) methods. Recent black-box optimization (BBO) approaches address this challenge through trajectory-level optimization and budget-aware exploration. Nevertheless, existing BBO-based methods rely on nominal objectives and isotropic exploration, limiting their robustness and sample efficiency under dynamic market conditions and data distribution shifts. To overcome these limitations, we propose `DR-AFL`, a distributionally robust and curvature-aware extension of BBO-based bidding. Specifically, we formulate a worst-case utility objective over a Wasserstein ambiguity set and introduce Hessian-preconditioned exploration to improve optimization efficiency. We further develop an adaptive and regularized safety projection to accommodate evolving budget constraints and noisy market feedback. Theoretical analysis investigates the resulting robustness, convergence behavior, and budget feasibility. Extensive experiments on 6 benchmark datasets, together with module-specific ablations and temporal-shift evaluations, demonstrate the effectiveness of the proposed extensions.

1 Introduction

Federated Learning (FL) [1–7] enables multiple data owners (DOs, i.e., FL clients) to collaboratively train a global model without sharing raw data. However, participation incurs computational and communication costs, necessitating incentive mechanisms to sustain engagement. Auction-based Federated Learning (AFL) [8, 9] addresses this challenge through market-based participant selection, where a data consumer (DC, i.e., FL server) strategically bids for DOs under a limited budget B.

Optimizing the DC's bidding strategy remains challenging [10–12]. Unlike conventional auctions with directly observable item values, the utility of a DO in AFL is *latent, interdependent, and revealed only after model training*. A DO's contribution depends not only on its own data but also on the characteristics of other selected DOs and their collective impact on the trained model. This gives rise to three challenges: 1) **Limited Utility Feedback:** The DC observes the overall model performance rather than the individual contribution of each DO; 2) **Delayed Utility Evaluation:** The value of selected DOs becomes observable only after model training, providing limited feedback during bidding; 3) **Ambiguous Credit Assignment:** The final model performance cannot be reliably decomposed into the contributions of individual bidding decisions.

*Corresponding authors.

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

Existing Reinforcement Learning (RL)-based bidding approaches [13, 14] typically optimize sequential decisions using intermediate rewards. However, the delayed and non-decomposable nature of AFL utility makes such reward design challenging and may result in unstable optimization. To address this limitation, BBO-AFL [12] introduces a black-box optimization (BBO) approach that evaluates bidding strategies based on their resulting model performance rather than relying on stepwise utility decomposition. It further incorporates budget-aware exploration to control expenditure during optimization, establishing a foundation for BBO-based AFL bidding.

Despite this progress, existing BBO-based approaches face two important challenges. First, **limited robustness**: bidding strategies optimized under nominal market conditions may be sensitive to changes in competition and DO data distributions. Second, **limited optimization efficiency**: evaluating a bidding strategy requires costly data acquisition and model training, while conventional isotropic exploration may overlook informative search directions. In addition, effective optimization requires exploration to remain responsive to changing budget availability and uncertain market feedback.

To address these challenges, we propose `DR-AFL`, a distributionally robust and curvature-aware BBO framework for AFL bidding. Building on BBO-AFL, our approach introduces a Wasserstein distributionally robust objective to account for uncertainty in policy optimization. It further leverages curvature information to guide exploration toward more informative directions, improving search efficiency. Meanwhile, an adaptive and regularized safety mechanism adjusts exploration according to budget consumption and stabilizes optimization under noisy cost feedback. Together, these components aim to improve the robustness, efficiency, and financial sustainability of AFL bidding under dynamic market conditions.

Our main contributions are summarized as follows:

- We extend nominal BBO-based AFL bidding with a **Wasserstein distributionally robust objective**, enabling policy optimization that explicitly accounts for uncertainty beyond nominal conditions.
- We introduce a **curvature-aware exploration mechanism** that exploits historical learning information to improve the sample efficiency of bidding strategy optimization.
- We develop an **adaptive and regularized safety mechanism** that extends existing budget-aware exploration to accommodate evolving expenditure constraints and noisy market feedback.
- We provide theoretical analysis of the proposed extensions and conduct extensive experiments on 6 benchmark datasets, module-specific ablations, and temporal market-shift evaluations.

2 Related Works

Existing AFL bidding methods for DCs [15, 16, 13, 14] aim to maximize model utility. The state-of-the-art, RLB-AFL [14], formulates recruitment as an MDP and applies RL for strategic bidding in dynamic markets. However, due to the non-decomposable nature of utility, these methods suffer from credit assignment ambiguity and budget risks during exploration. To address these challenges, BBO-AFL [12] establishes a trajectory-level black-box optimization framework with budget-aware exploration. Building upon this foundation, our work further enhances the robustness and efficiency of bidding optimization under dynamic market conditions and data distribution shifts through distributionally robust and curvature-aware optimization, together with adaptive budget-aware exploration. Additional related work is provided in Appendix A.

3 Preliminaries

AFL Market: An AFL ecosystem is characterized by the strategic interplay between 3 key stakeholders [17]: 1) **Data Owners (DOs):** A set $\mathcal{K} = \{1, \ldots, K\}$ of providers, where each DO i holds a private dataset $\mathcal{D}_i$ and offers data use right and computation in exchange for payment. 2) **Data Consumers (DC):** An entity that aims to learn a global model $w \in \mathbb{R}^d$ by selecting a subset of DOs, maximizing model utility under a budget constraint. 3) **Auctioneer:** A trusted intermediary that implements an incentive mechanism $\mathcal{M}$ (e.g., second-price sealed-bid) to allocate DOs and determine payments.

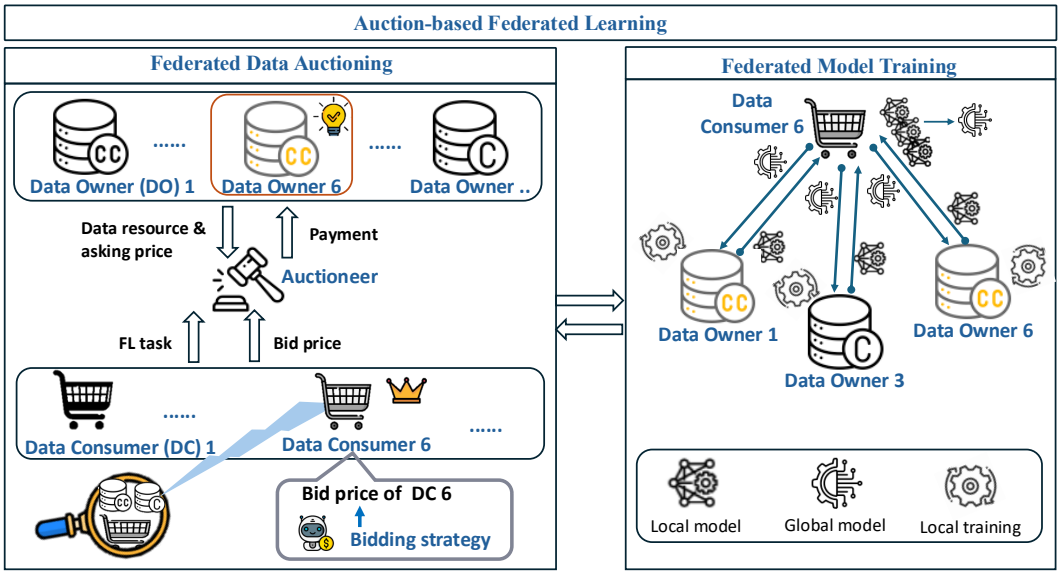


Figure 1: Illustration of the AFL framework.

As shown in Fig. 1, the AFL workflow operates as a sequential decision process consisting of *Federated Data Auctioning* and *Federated Model Training*. In the auctioning phase, i.e., the recruitment pahse, the DC participates in a sequence of auctions to incentivize and select DOs under a finite budget. This process terminates when the budget is exhausted or the auction window closes. In the subsequent training phase, the DC performs FL optimization over the recruited cohort. Unlike synchronous FL, recruitment is driven by the stochastic availability of DOs: 1) **Sequential Bidding:** Each arriving DO $i \in \mathcal{K}$ induces a single auction instance. The DC observes features ξ_i and submits a bid $b_i = \pi_\theta(\xi_i)$ via a policy π_θ, where ξ_i encodes observable attributes (e.g., data size or reputation). 2) **Allocation Rule:** The auctioneer applies mechanism $\mathcal{M}$ to determine the winning DC among competitors, with at most one winner per DO i. 3) **Recruitment Set:** The process continues sequentially until budget B exhaustion or no further DOs arrive. The recruited set is defined as $\mathcal{S}_{rec} = \{i \mid \text{DC wins } DO_i\}$. 4) **Budget Constraint:** The total payment for $\mathcal{S}_{rec}$ must satisfy $\sum_{i \in \mathcal{S}_{rec}} p_i \leq B$, where p_i is the clearing price of DO i determined by $\mathcal{M}$. Once recruitment terminates, the DC performs FL training over the selected cohort $\mathcal{S}_{rec}$. The global model w is optimized based on datasets $\{\mathcal{D}_i\}_{i \in \mathcal{S}_{rec}}$. The resulting utility is defined as $V(w_{\text{final}}; \mathcal{S}_{rec})$, which measures the performance of the converged model (e.g., test accuracy).

Bidding Strategy Optimization. The DC aims to optimize the policy parameters θ to maximize the expected terminal utility induced by the recruited cohort:

$$\max_{\theta} \ \mathcal{J}(\theta) = \mathbb{E}_{\xi, \mathcal{S}_{rec}}\left[V(w_{\text{final}}; \mathcal{S}_{rec})\right] \quad \text{s.t.} \quad \mathbb{E}_{\omega}\left[\sum_{i \in \mathcal{S}_{rec}(\theta,\omega)} p_i\right] \leq B. \tag{1}$$

Here, the utility $V(\cdot)$ is a set-dependent functional of $\mathcal{S}_{rec}$ and is only revealed after the completion of the full training trajectory.

Limitations of Stepwise RL Formulation. State-of-the-art bidding methods in AFL model the recruitment process as a Markov Decision Process (MDP) [18, 19] with stepwise rewards. However, this formulation is fundamentally misaligned with the AFL setting due to two structural issues [12]: 1) ***Lack of Intermediate Rewards:*** Under the recruit-then-train paradigm, model training starts only after recruitment completes, yielding no meaningful intermediate feedback during bidding. As a result, assigning per-step rewards to individual bids b_i is ill-posed, leading to severe *credit assignment ambiguity*. 2) ***Delayed, Non-Decomposable Feedback:*** RL relies on dense or temporally local rewards for effective learning. In contrast, AFL provides only a terminal, trajectory-level utility signal that is non-decomposable across actions. This delayed and sparse feedback violates the Markovian assumptions underlying standard RL methods [13], resulting in unstable or ineffective policy learning.

Empirical Evidence. To substantiate the above structural limitations, we conduct an empirical analysis of RL-based bidding methods [13, 16, 14]. As shown in Fig. 2(a), stepwise rewards (i.e., DO utilities) exhibit only weak correlation with the final global model accuracy ($\rho \approx 0.112$), indicating that local feedback provides little guidance for optimizing the global objective. Moreover, Fig. 2(b) shows that increasing the recruitment scale amplifies the effect of delayed feedback, leading to significantly higher variance in model accuracy and consequently unstable performance.

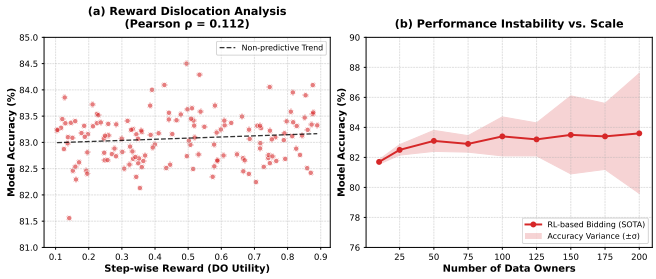


Figure 2: Empirical evidence of failures of RL-based methods.

Motivated by this mismatch, we instead view the mapping $\theta \mapsto \mathcal{J}(\theta)$ as an unknown black-box function. Under this perspective, policy optimization proceeds by directly evaluating the terminal utility induced by entire recruitment outcomes, effectively transforming the problem from local, stepwise updates to global, strategy-level search.

4 The Proposed DR-AFL

To address the delayed, non-decomposable feedback in AFL, we propose DR-AFL, a distributionally robust black-box optimization framework that treats the entire recruitment-to-training pipeline as a high-fidelity evaluation. DR-AFL addresses market non-stationarity and data shifts by optimizing policies against worst-case environments, while integrating adversarial reweighting and dynamic safety projection for robust, budget-aware exploration.

4.1 Objective Formalization via Randomized Smoothing

Building on the setup in Section 3, we model the DC's bidding policy as a parameterized mapping $\pi_\theta : \Xi \to \mathbb{R}_+$ with $\theta \in \mathbb{R}^d$. Let $F(\theta; \omega)$ denote the induced terminal utility under realization $\omega \in \Omega$, where $\omega \in \Omega$ captures all sources of randomness, including market dynamics and stochasticity in FL training[1]. Then, the DC's decision problem reduces to a constrained stochastic optimization problem [12]:

$$\max_{\theta \in \mathbb{R}^d} \mathcal{J}(\theta) \triangleq \mathbb{E}_\omega[F(\theta; \omega)] \quad \text{s.t.} \quad \mathcal{C}(\theta) \triangleq \mathbb{E}_\omega \left[\sum_{i \in \mathcal{S}_{rec}(\theta, \omega)} p_i \right] \leq B. \tag{2}$$

A key difficulty in AFL lies in the irregular geometry of the objective $\mathcal{J}(\theta)$, which is generally *non-differentiable* and potential *piecewise discontinuous*. This behavior is induced by the discrete winner-selection mechanism in competitive auctions: small perturbations in the bidding parameters θ can trigger binary changes in recruitment outcomes, i.e., whether a specific DO i is secured. Such discrete changes in the recruited set $\mathcal{S}_{rec}$ may induce abrupt variations in the terminal utility V, rendering gradient-based optimization via standard backpropagation inapplicable.

To address this issue, we adopt a **randomized smoothing** approach [20], defining a smoothed objective:

$$\tilde{\mathcal{J}}(\theta) = \mathbb{E}_{\delta \sim P_0} \left[\mathcal{J}(\theta + \delta) \right], \tag{3}$$

where P_0 is a noise distribution (generally isotropic Gaussian, i.e., $P_0 = \mathcal{N}(0, \sigma^2 \mathbf{I})$ with σ being the smoothing radius). This transformation yields a smoother surrogate objective, enabling gradient estimation via zeroth-order methods.

However, in practical AFL settings, the market is inherently non-stationary due to 1) evolving market competition as rival DCs may dynamically adjust their bidding policies, and 2) data distribution shifts, where the available DOs' data characteristics (e.g., Non-IID degrees) vary over time. A fixed smoothing distribution P may therefore lead to policies that is overfitted to a specific, stale market

[1] Following prior AFL works, we treat the final global model accuracy as the terminal utility.

state and are sensitive to distribution mismatch. To improve robustness, we consider a Wasserstein ambiguity set:

$$\mathcal{P} := \{P : W_2(P, P_0) \leq \epsilon\}, \tag{4}$$

where ϵ controls the level of uncertainty.

We then define the **distributionally robust smoothed surrogate (DRO)**:

$$\mathcal{J}_{DR}(\theta) = \min_{P \in \mathcal{P}} \mathbb{E}_{\delta \sim P} \left[\mathcal{J}(\theta + \delta) \right]. \tag{5}$$

This formulation can be interpreted as optimizing the policy under worst-case perturbations in the smoothing distribution. Intuitively, perturbations in θ induce variations in policy behavior, which serve as a proxy for environmental uncertainty. By optimizing against the worst-case distribution in $\mathcal{P}$, the resulting policy is less sensitive to shifts in market conditions and data characteristics.

The proposed transformation offers the following main benefits:

- High-Order Regularity: $\mathcal{J}_{DR}$ acts as a robust low-pass filter. According to Danskin's Theorem, the gradient of this minimax objective can be recovered by shifting the differential operator onto the noise density at the optimal P^* with the score function $\nabla_\theta \log p(\delta)$:

  $$\nabla_\theta \mathcal{J}_{DR}(\theta) = \mathbb{E}_{\delta \sim P^*} \left[\mathcal{J}(\theta + \delta) \cdot \nabla_\theta \log p^*(\delta) \right], \tag{6}$$

  where $P^* = \arg\min_{P \in \mathcal{P}} \mathbb{E}_{\delta \sim P}[\mathcal{J}(\theta + \delta)]$. This decoupling ensures that even with binary win-or-lose feedback, the DC can recover stable, C^∞ smooth gradient signals.
- Resilience to Market Volatility: By optimizing against the entire set $\mathcal{P}$, the bidding policy implicitly accounts for unseen market shifts. This prevents the brittleness of standard RL agents, ensuring that the bidding rules remain effective even when competitors' behaviors or DO data distributions drift during the training process.
- Generalization Boundary: The Wasserstein distance ϵ defines a safety margin. As long as the actual environmental non-stationarity remains within this ϵ-ball, the DC is tend to maintain a robust utility level, effectively addressing the performance instability identified in Figure 2(b).

Budget feasibility under perturbations is enforced separately via the safety projection mechanism described in Section 4.3.

4.2 Robust Gradient Estimation via Adversarial Importance Sampling

While Eq. (5) defines a distributionally robust objective, directly computing the gradient $\nabla_\theta \mathcal{J}_{DR}(\theta)$ is intractable due to the unknown worst-case distribution P^*. To enable practical optimization, we approximate this objective using a finite set of stochastic perturbations and construct a surrogate gradient estimator that captures both robustness and sample efficiency. Specifically, we evaluate the policy under a set of N candidate perturbations $\{\delta_j\}_{j=1}^N \subset \mathbb{R}^d$, where each δ_j represents a stochastic probe in the parameter space. Based on these samples, we design a three-stage approximation procedure.

1. Curvature-Informed Exploration: To improve sampling efficiency, we generate perturbations as $\delta_j \sim \mathcal{N}(0, \sigma^2 \mathbf{H}^{-1})$, where $\mathbf{H}$ is a positive definite matrix (Hessian $\mathbf{H}$) constructed from gradient-based statistics (i.e., aggregation gradients $\mathcal{H}$) collected during prior FL rounds [12]. This matrix serves as a curvature-informed preconditioner, biasing exploration toward directions that are more sensitive to policy changes. Importantly, $\mathbf{H}$ is computed from existing training signals and introduces negligible additional overhead. Each $\theta + \delta_j$ defines a candidate bidding strategy to be evaluated in the current market environment. This ensures N probes are concentrated in parameter regions that most significantly impact the bidding utility, addressing the efficiency bottleneck of pure black-box optimization. Crucially, since $\mathcal{H}$ is a byproduct of standard FL training, this second-order guidance is *computationally efficient* and requires no additional functional evaluations.

2. Adversarial Re-weighting (DRO-inspired Importance Sampling): To capture the minimax structure in Eq. (5), we prioritize samples that reflect adverse market conditions. Since the exact worst-case distribution P^* is intractable, we instead adopt a tractable surrogate that emphasizes low-utility outcomes. Specifically, given terminal utilities $\{V_j\}_{j=1}^N$, we assign **adversarial weights**

using a rank-based shaping function $\omega(R_j)$, where R_j denotes the rank of V_j (with 1 being the highest utility and N the lowest). This scheme increases the influence of low-utility samples, which serve as a proxy for challenging or shifted environments within the Wasserstein ambiguity set. The weighting function is defined as:

$$\omega(R) = \frac{\max(0, \log(N/2+1) - \log(R))}{\sum_{m=1}^{N/2} (\log(N/2+1) - \log(m))} - \frac{1}{N}. \tag{7}$$

This rank-based re-weighting can be interpreted as a form of **importance sampling** that biases gradient estimation toward directions that remain effective under adverse conditions. While it does not explicitly recover the exact worst-case distribution in Wasserstein DRO, it provides a practical approximation that enhances robustness to distributional shifts in AFL markets.

3. Safety-Projected Policy Update: The final empirical gradient estimate $\hat{g}_{DR}$ is constructed by aggregating the **safety-projected** perturbations $\tilde{\delta}_j$ (derived via KKT conditions in Section 4.3) weighted by their adversarial fitness:

$$\hat{g}_{DR} = \frac{1}{N\sigma} \sum_{j=1}^{N} \omega(R_j) \tilde{\delta}_j. \tag{8}$$

The proposed estimator combines three complementary effects: (i) curvature-informed sampling improves exploration efficiency in high-dimensional spaces, (ii) adversarial re-weighting biases updates toward worst-case outcomes, and (iii) safety projection ensures feasibility under budget constraints. Together, these components yield a stable and practically effective optimization procedure for AFL bidding under non-stationarity and delayed feedback.

4.3 Budget-Aware Exploration via Dynamic Safety Manifold Projection

A fundamental constraint in AFL is the *irreversibility of financial expenditure*. While standard black-box optimization treats perturbations δ_j as cost-free functional probes, the DC in AFL commits real financial resources B for every evaluation. A reckless perturbation δ_j might yield a high-fidelity utility signal but exhaust the remaining budget B prematurely, leading to mid-trajectory bankruptcy. To navigate this exploration-expenditure trade-off under market non-stationarity, we propose a **Dynamic Safety Manifold**. We first define an **Adaptive Risk-Tolerance** $\beta_t = \exp(-\gamma \frac{B_{spent}}{B})$, where B_{spent} is the cumulative expenditure, γ is a decay constant and t denotes the current BBO iteration index. This mechanism ensures the DC remains exploratory when the budget is ample but becomes increasingly risk-averse as resources deplete. For each candidate probe δ_j generated in Section 4.2, we seek the proximal safe perturbation $\tilde{\delta}_j$ by solving the following constrained problem:

$$\tilde{\delta}_j = \arg\min_{\delta} \frac{1}{2} \|\delta - \delta_j\|^2 \quad \text{s.t.} \quad \nabla \mathcal{C}(\theta)^\top \delta + \mathcal{C}(\theta) \leq \beta_t B, \tag{9}$$

where $\nabla \mathcal{C}(\theta)$ is the gradient of the expenditure functional estimated from immediate market feedback.

Regularized KKT Solution: To ensure the numerical stability of the projection under Non-IID market jitter, we derive a Regularized Karush-Kuhn-Tucker (KKT) Operator [12]:

$$\tilde{\delta}_j = \delta_j - \frac{\max(0, \nabla \mathcal{C}(\theta)^\top \delta_j + \mathcal{C}(\theta) - \beta_t B)}{\|\nabla \mathcal{C}(\theta)\|^2 + \zeta} \nabla \mathcal{C}(\theta), \tag{10}$$

where ζ is a Tikhonov regularization parameter.

The advantages of this formulation are: 1) Inter-temporal Solvency: By dynamically shrinking the feasible exploration manifold via β_t, the DC guarantees long-term financial survival, effectively preventing the bankruptcy at the finish line phenomenon. 2) Robustness to Price Spikes: The inclusion of ζ provides a damping effect against noisy clearing price signals, ensuring that the search for the distributionally robust optimum remains stable even when competitor behaviors shift abruptly.

Algorithm 1 illustrates the proposed `DR-AFL`.

4.4 Theoretical Analysis

Assumption 1 (Regularity): The nominal utility $F(\theta; \omega)$ is L-Lipschitz and M-smooth. The cost functional $\mathcal{C}(\theta)$ is locally convex and M_c-smooth. **Assumption 2 (Ambiguity Set):** The ambiguity

Algorithm 1 DR-AFL

Input: Initial policy θ_0, total budget B, perturbation scale σ, population size N, learning rate α, risk decay γ, regularization ζ.
Initialize: Cumulative expenditure $B_{spent} \leftarrow 0$, training metadata $\mathcal{H} \leftarrow \emptyset$.
for iteration $t = 0, 1, \ldots, K-1$ **do**
  // — Stage 1: Curvature-Aware Structured Exploration —
  Estimate Hessian $\mathbf{H}$ from $\mathcal{H}$ and sample candidate probes $\{\delta_j\}_{j=1}^{N} \sim \mathcal{N}(0, \sigma^2 \mathbf{H}^{-1})$.
  Compute adaptive risk-tolerance $\beta_t = \exp(-\gamma \frac{B_{spent}}{B})$.
  for each perturbation $j = 1$ **to** N **do**
    // — Stage 2: Dynamic Safety Projection —
    Compute safe perturbation $\tilde{\delta}_j$ via Regularized KKT Operator in Eq. (10).
    // — Stage 3: High-Fidelity AFL Evaluation —
    Recruitment: DC submits bids $b_i = \pi_{\theta_t + \tilde{\delta}_j}(\xi_i)$ to the available DO.
    Feedback: Observe clearing prices $\{p_i\}$ and update $B_{spent} \leftarrow B_{spent} + \sum p_i$.
    FL Training: Execute federated learning and observe terminal utility V_j.
    Metadata Capture: Store aggregation gradient norms $\mathcal{H}$ for next iteration.
  end for
  // — Stage 4: Robust Policy Update —
  Rank utilities $\{V_j\}_{j=1}^{N}$ and compute adversarial weights $\omega(R_j)$ via Eq. (7).
  Update bidding policy: $\theta_{t+1} \leftarrow \theta_t + \alpha \hat{g}_{DR,t}$, where $\hat{g}_{DR,t}$ is computed via Eq. (8).
end for
Output: Optimized robust bidding policy θ_K.

set $\mathcal{P} = \{P : W_2(P, P_0) \leq \epsilon\}$ is a Wasserstein-2 ball centered at the Gaussian nominal distribution $P_0 = \mathcal{N}(0, \sigma^2 \mathbf{I})$.

The primary challenge in AFL is the inherent discontinuity of the auction mechanism. We first demonstrate that our robust surrogate transforms this landscape into a smooth manifold.

Lemma 1 (Regularity Transfer)**.** *The robust surrogate $\mathcal{J}_{DR}(\theta) = \min_{P \in \mathcal{P}} \mathbb{E}_{\delta \sim P}[\mathcal{J}(\theta + \delta)]$ is C^∞ smooth, and its gradient $\nabla \mathcal{J}_{DR}$ is L_{DR}-Lipschitz with $L_{DR} \leq L/\sigma$.*

Theorem 1 (Convergence of DR-AFL)**.** *Let Γ_t be the expected bias from the KKT safety projection and $\mathcal{B}_\epsilon = L\epsilon$ be the DRO-induced bias. With step size $\alpha \leq 1/L_{DR}$, the iterates $\{\theta_t\}$ satisfy:*

$$\frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}[\|\nabla \mathcal{J}_{DR}(\theta_t)\|^2] \leq \frac{2(\mathcal{J}^* - \mathcal{J}_0)}{\alpha T} + \frac{\alpha L_{DR} \sigma^2 d}{\eta N} + 2M(\bar{\Gamma} + \mathcal{B}_\epsilon), \tag{11}$$

where $\eta \approx \kappa(\mathbf{H})$ represents the variance reduction gain from Hessian preconditioning.

Theorem 2 (Worst-case Performance Boundary)**.** *For any actual market drift $P \in \mathcal{P}$, the actual performance is lower-bounded by the robust surrogate: $\mathbb{E}_P[F] \geq \mathcal{J}_{DR}(\theta) \geq \mathbb{E}_{P_0}[F] - L \cdot \epsilon$.*

Lemma 2 (High-probability Safety Feasibility)**.** *The KKT-projected perturbation $\tilde{\delta}$ ensures that budget violations decay exponentially: $Pr(\mathcal{C}(\theta + \tilde{\delta}) > B) \leq \exp(-2\Delta \mathcal{C}_t^2 / M_c^2 \sigma^2 d)$.*

Detailed proofs of all lemmas and theorems are deferred to the Appendix C.

5 Experimental Evaluation

5.1 Experiment Settings

Datasets and Models: We evaluate DR-AFL across 6 benchmark datasets: 1) MNIST[2], 2) CIFAR-10[3], 3) Fashion-MNIST (FMNIST) [21], 4) EMNIST-digits (EMNISTD) [22], 5) EMNIST-letters (EMNISTL) [22], and 6) Kuzushiji-MNIST (KMNIST) [23]. The widely adopted FedAvg [24] used here and the CNN model used are the same as those employed in [15].

Baseline Approaches: We compare DR-AFL against a comprehensive suite of AFL bidding strategies: 1) Constant Bid (**Const**) [25], 2) Randomly Generated Bid (**Rand**) [26, 27], 3) Below Max Utility

[2] http://yann.lecun.com/exdb/mnist/
[3] https://www.cs.toronto.edu/kriz/cifar.html

Table 1: Comprehensive Comparison under Scenario 1 (IID Data, Varying Dataset Sizes). **ROI** is scaled by 10^3. **Iter** is the training cycles to reach target accuracy. **BR**(%) is the percentage of failed sessions due to budget exhaustion.

Bud.	Method	MNIST				CIFAR-10				FMNIST				EMNISTD				EMNISTL				KMNIST			
		Acc	ROI	BR	Iter	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter
100	Const	82.1	0.45	-	>500	34.8	0.19	-	>500	69.5	0.34	-	>500	77.8	0.42	-	>500	67.8	0.38	-	>500	62.1	0.32	-	>500
	Rand	82.3	0.49	-	>500	34.6	0.22	-	>500	70.2	0.40	-	>500	79.9	0.41	-	>500	65.9	0.41	-	>500	60.5	0.31	-	>500
	Bmub	82.6	0.58	-	485	36.5	0.30	-	498	70.9	0.48	-	490	80.2	0.48	-	495	69.9	0.48	-	495	62.9	0.40	-	495
	Lin	83.0	0.69	-	445	37.0	0.38	-	465	71.5	0.59	-	440	80.7	0.59	-	455	71.0	0.56	-	430	64.0	0.55	-	455
	BM	83.1	0.76	1.8	385	38.1	0.42	3.2	415	73.4	0.68	2.5	395	81.4	0.66	2.7	405	72.0	0.65	2.8	405	65.8	0.61	2.8	405
	FBs	83.2	0.82	4.2	295	39.0	0.45	5.9	335	73.5	0.75	4.8	315	81.5	0.72	5.2	325	73.1	0.72	4.5	325	68.1	0.76	5.2	325
	FBc	83.3	0.83	3.8	310	39.8	0.46	5.5	350	74.0	0.76	4.6	325	81.6	0.73	4.9	340	73.2	0.73	4.9	340	68.4	0.78	4.9	340
	RLB	83.4	0.88	28.5	210	40.0	0.55	35.1	240	74.6	0.85	30.2	225	83.0	0.82	28.6	230	73.4	0.81	29.4	230	69.1	0.84	28.2	230
	Ours	**83.7**	**1.02**	**0.6**	**130**	**40.3**	**0.72**	**1.0**	**160**	**74.9**	**0.95**	**0.9**	**150**	**83.5**	**0.98**	**0.8**	**155**	**74.5**	**0.94**	**1.1**	**155**	**69.7**	**0.91**	**1.2**	**155**
400	Const	83.0	0.52	-	500	37.0	0.29	-	500	71.0	0.45	-	500	79.2	0.46	-	500	70.8	0.46	-	500	62.5	0.46	-	500
	Rand	83.3	0.55	-	500	37.5	0.35	-	500	69.9	0.48	-	500	78.7	0.49	-	500	71.0	0.49	-	500	62.3	0.49	-	500
	Bmub	84.0	0.66	-	465	39.5	0.38	-	485	72.2	0.55	-	475	81.8	0.56	-	480	72.6	0.56	-	480	71.5	0.56	-	480
	Lin	84.3	0.79	-	425	40.0	0.45	-	455	75.1	0.69	-	415	81.8	0.70	-	435	73.1	0.70	-	435	72.5	0.70	-	435
	BM	85.1	0.86	4.9	345	41.1	0.48	8.1	375	75.1	0.75	6.2	355	83.3	0.74	6.9	365	73.9	0.74	6.9	365	72.5	0.74	6.9	365
	FBs	85.1	0.91	8.6	265	42.0	0.55	12.2	295	76.4	0.85	9.9	275	83.4	0.82	10.8	285	74.5	0.82	10.8	285	73.8	0.82	10.8	285
	FBc	85.3	0.92	7.8	280	42.5	0.54	11.5	315	74.7	0.83	9.5	290	83.3	0.81	10.2	300	74.7	0.81	10.2	300	73.9	0.81	10.2	300
	RLB	86.1	1.05	34.2	180	43.1	0.62	38.5	210	77.0	0.95	31.2	195	84.2	0.92	35.6	200	75.3	0.92	31.5	200	74.8	0.92	30.3	200
	Ours	**87.3**	**1.18**	**0.8**	**105**	**43.4**	**0.88**	**1.3**	**135**	**79.0**	**1.10**	**1.1**	**120**	**84.5**	**1.12**	**1.2**	**125**	**75.6**	**1.08**	**1.2**	**125**	**77.1**	**1.08**	**1.2**	**125**
800	Const	84.3	0.59	-	485	38.3	0.36	-	500	71.5	0.52	-	495	81.0	0.53	-	500	71.3	0.53	-	500	70.2	0.53	-	500
	Rand	84.4	0.62	-	465	39.2	0.40	-	485	71.6	0.55	-	475	80.6	0.56	-	480	72.6	0.56	-	480	72.3	0.56	-	480
	Bmub	85.2	0.71	-	445	42.3	0.46	-	465	73.8	0.66	-	455	83.4	0.67	-	460	74.1	0.67	-	460	73.2	0.67	-	460
	Lin	85.9	0.82	-	395	43.4	0.55	-	425	76.6	0.78	-	385	83.7	0.79	-	405	74.3	0.79	-	405	74.2	0.79	-	405
	BM	85.9	0.88	4.5	315	46.5	0.62	6.8	345	78.1	0.83	5.2	325	83.7	0.84	5.9	335	75.2	0.84	5.9	335	75.6	0.84	5.9	335
	FBs	86.7	0.95	11.8	235	46.8	0.68	15.5	265	78.2	0.91	13.8	245	83.7	0.89	14.5	255	75.2	0.89	14.5	255	75.7	0.89	14.5	255
	FBc	87.0	0.96	11.2	250	47.0	0.69	14.6	285	77.8	0.92	12.9	260	83.7	0.90	13.6	270	75.3	0.90	13.6	270	75.6	0.90	13.6	270
	RLB	87.2	1.12	28.5	160	47.3	0.77	31.2	190	79.1	1.02	25.4	175	84.7	1.05	25.8	180	76.0	1.05	26.4	180	78.3	1.05	23.7	180
	Ours	**88.0**	**1.25**	**0.5**	**95**	**48.8**	**1.12**	**0.9**	**125**	**80.2**	**1.20**	**0.8**	**105**	**85.1**	**1.24**	**0.9**	**115**	**76.6**	**1.20**	**0.9**	**115**	**79.8**	**1.20**	**0.9**	**115**

Bid (**Bmub**) [28], 4) Linear-Form Bid (**Lin**) [29], 5) Bidding Machine (**BM**) [30], 6) Reinforcement Learning-based Bid (**RLB**) [13, 16, 31, 14].

Experimental Scenarios. We evaluate DR-AFL against baselines under 2 representative settings, each involving 10,000 DOs: 1) **Scenario 1 (IID with varying scale):** DO datasets are independently sampled with sizes uniformly ranging from 500 to 5,000 instances. All data are IID and noise-free, isolating the effect of scale heterogeneity. 2) **Scenario 2 (Non-IID with noise):** We introduce controlled data heterogeneity by skewing class distributions across DOs. Following [32],100 DOs are assigned minority classes (6 for EMNISTL, 1 class for others), while the remaining DOs exclude these classes. Each DO contains 3,000 samples. Additionally, 25% synthetic noise is injected into minority-class DOs to simulate data corruption. Implementation details are provided in Appendix D.

Evaluation Metrics. We assess DR-AFL using four complementary metrics that jointly capture learning performance and economic efficiency: 1) **Final Accuracy (Acc):** Test accuracy of the converged global model, measuring overall learning quality. 2) **Economic Efficiency (ROI):** Return on investment, defined as the ratio between utility gain and total expenditure ($\Delta\text{Utility}/\sum p_i$), reflecting cost-effectiveness in participant selection. 3) **Bankruptcy Rate (BR):** Fraction of runs in which the budget is exhausted prematurely, indicating the effectiveness of budget-aware constraints. 4) **Sample Efficiency (Iter):** Number of recruitment–training cycles required to reach a target accuracy, measuring convergence efficiency.

5.2 Results and Discussion

We perform a comprehensive analysis of the experimental results to uncover the key factors driving the performance of DR-AFL across diverse AFL settings. As shown in Tables 1 and 2, and Figure 3, DR-AFL consistently achieves strong and well-balanced performance across datasets and budget settings. Notably, in Scenario 2 with noise and data heterogeneity, RL-based baselines (i.e., RLB) exhibit substantial degradation, likely due to delayed, non-decomposable feedback. In contrast, DR-AFL maintains stable accuracy. This behavior aligns with our formulation that treats the AFL mapping $\theta \mapsto \mathcal{J}(\theta)$ as a distributionally robust black-box functional. By optimizing against worst-case perturbations within a Wasserstein ambiguity set, DR-AFL effectively improves robustness under heavy-tailed noise and market variability.

Table 2: Comparison under Scenario 2 (Non-IID, 25% noise, and Market Drift).

Bud.	Method	MNIST				CIFAR				FMNIST				EMNISTD				EMNISTL				KMNIST			
		Acc	ROI	BR	Iter	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter
100	Const	69.85	0.38	-	>500	13.52	0.15	-	>500	57.12	0.31	-	>500	75.98	0.44	-	>500	62.85	0.38	-	>500	59.12	0.28	-	>500
	Rand	64.92	0.32	-	>500	10.25	0.11	-	>500	58.75	0.32	-	>500	77.25	0.45	-	>500	62.15	0.35	-	>500	58.65	0.26	-	>500
	Bmub	70.15	0.42	-	495	12.55	0.18	-	>500	59.85	0.38	-	500	76.92	0.48	-	495	64.75	0.42	-	500	61.22	0.34	-	500
	Lin	69.85	0.48	-	470	16.85	0.25	-	490	63.45	0.45	-	475	76.85	0.52	-	480	63.55	0.48	-	485	62.35	0.38	-	485
	BM	71.22	0.55	2.8	420	18.95	0.32	4.2	455	65.85	0.52	3.5	425	77.95	0.58	3.2	430	64.12	0.51	3.4	435	63.22	0.45	3.5	435
	FBs	71.95	0.62	5.8	355	21.52	0.39	7.5	385	69.75	0.62	5.9	365	77.85	0.65	5.8	360	65.45	0.56	5.2	370	63.85	0.52	6.2	370
	FBc	74.35	0.65	5.2	370	21.85	0.41	6.9	395	70.15	0.64	5.5	380	78.45	0.67	5.5	375	64.25	0.55	5.8	385	63.35	0.51	5.9	385
	RLB	72.85	0.68	38.2	255	22.45	0.42	45.8	295	69.12	0.66	37.5	275	78.15	0.69	39.2	265	65.12	0.58	35.8	280	63.45	0.54	41.5	280
	Ours	**75.58**	**0.92**	**0.8**	**140**	**23.85**	**0.65**	**1.2**	**175**	**71.85**	**0.88**	**1.0**	**160**	**79.72**	**0.92**	**0.9**	**155**	**67.22**	**0.82**	**1.1**	**170**	**65.55**	**0.80**	**1.2**	**170**
400	Const	68.12	0.48	-	500	16.52	0.22	-	500	58.45	0.38	-	500	77.45	0.52	-	500	66.25	0.45	-	500	67.92	0.42	-	500
	Rand	67.25	0.45	-	500	19.85	0.28	-	500	60.12	0.41	-	500	77.92	0.55	-	500	67.15	0.48	-	500	69.25	0.45	-	500
	Bmub	69.85	0.55	-	490	19.92	0.31	-	500	62.75	0.48	-	485	78.45	0.58	-	490	67.65	0.52	-	495	68.12	0.48	-	495
	Lin	70.52	0.68	-	455	22.45	0.38	-	465	64.95	0.58	-	440	78.52	0.65	-	455	68.65	0.62	-	465	68.55	0.55	-	465
	BM	72.55	0.75	6.2	375	23.25	0.45	9.8	405	67.85	0.68	7.8	385	79.45	0.72	7.2	380	69.85	0.69	7.8	390	69.95	0.62	8.2	390
	FBs	73.12	0.81	10.2	295	24.12	0.52	14.5	325	69.65	0.78	11.5	305	79.85	0.79	10.2	300	69.65	0.75	10.9	315	70.35	0.71	12.1	315
	FBc	73.45	0.82	9.5	310	23.85	0.51	13.8	345	71.45	0.81	10.2	320	79.82	0.81	9.5	315	69.92	0.78	10.5	330	69.12	0.68	11.2	330
	RLB	74.45	0.88	42.5	215	24.32	0.58	48.2	245	71.85	0.85	40.2	225	80.02	0.88	41.5	220	70.12	0.85	39.8	235	70.15	0.75	44.5	235
	Ours	**76.85**	**1.08**	**1.1**	**115**	**26.55**	**0.78**	**1.5**	**140**	**74.22**	**1.05**	**1.2**	**130**	**81.25**	**1.12**	**1.0**	**125**	**72.55**	**1.02**	**1.3**	**135**	**72.15**	**1.00**	**1.4**	**135**
800	Const	66.02	0.52	-	495	21.45	0.32	-	500	64.45	0.42	-	500	78.25	0.55	-	500	66.92	0.45	-	500	68.52	0.45	-	500
	Rand	70.45	0.51	-	475	20.12	0.28	-	495	64.85	0.40	-	490	78.75	0.58	-	495	68.15	0.42	-	500	69.52	0.42	-	500
	Bmub	71.65	0.58	-	455	18.95	0.35	-	485	65.92	0.45	-	465	79.25	0.62	-	470	68.45	0.49	-	475	70.12	0.49	-	475
	Lin	72.15	0.72	-	435	27.52	0.44	-	465	68.52	0.58	-	435	79.65	0.69	-	445	69.65	0.55	-	455	69.92	0.55	-	455
	BM	73.45	0.78	6.8	355	28.52	0.52	12.5	395	69.85	0.65	9.2	365	80.12	0.75	8.5	365	71.12	0.65	8.8	375	71.25	0.65	8.8	375
	FBs	74.55	0.85	12.5	275	29.65	0.58	16.2	305	73.35	0.75	13.5	285	80.35	0.79	11.2	285	71.22	0.72	12.1	300	71.72	0.72	12.5	300
	FBc	75.12	0.86	11.5	285	29.25	0.55	15.2	320	73.85	0.82	12.2	295	80.42	0.82	10.5	290	71.65	0.75	11.5	310	71.45	0.69	11.8	310
	RLB	75.85	0.91	35.8	190	30.25	0.62	42.5	225	74.25	0.85	34.2	210	80.65	0.88	38.2	200	71.85	0.82	39.1	215	72.65	0.78	35.5	215
	Ours	**77.22**	**1.18**	**1.0**	**105**	**33.15**	**0.92**	**1.4**	**130**	**76.45**	**1.15**	**1.1**	**120**	**83.45**	**1.22**	**0.8**	**110**	**74.15**	**1.10**	**1.2**	**125**	**74.45**	**1.10**	**1.1**	**125**

A key finding is the substantial reduction in *Iter* (policy evaluation cycles) required for convergence. Across all scenarios, DR-AFL requires approximately 40%–50% fewer cycles than RL-based methods to reach target accuracy, indicating improved sample efficiency. This gain can be attributed to variance reduction guided by second-order information. Unlike RLB that suffer from credit assignment under delayed feedback, DR-AFL leverages curvature-aware signals to better align exploration with the local structure of the utility landscape. This alignment improves the signal-to-noise ratio of zeroth-order gradient estimates, leading to faster and more stable convergence in high-dimensional bidding spaces.

Bankruptcy Rate (BR) provides quantitative evidence of the effectiveness of our safety projection mechanism. In Scenario 2, RLB exhibit high BR (up to 48.2% on CIFAR10), suggesting that standard policy gradient methods can generate cost-inefficient bids under abrupt market fluctuations. In contrast, DR-AFL maintains BR below 1.5% across settings.This improvement is attributed to the KKT-based safety projection, which constrains policy updates within a feasible financial region. By projecting risky perturbations back into the budget-compliant set, DR-AFL enforces cost-aware exploration and mitigates the risk of premature budget exhaustion under volatile conditions.

DR-AFL exhibits a positive relationship between budget and ROI. As the budget B increases from 100 to 800, the ROI on MNIST improves from 1.02 to 1.25. This trend suggests improved economic scalability, where larger budgets enable the policy to recruit higher-utility participants that were previously unaffordable. The diminishing gains in ROI at higher budget levels (e.g., $B = 800$) indicate that the solution is approaching the utility cost trade-off frontier. Overall, these results suggest that DR-AFL effectively balances performance and cost across varying budget regimes.

Table 3: Ablation Study under Scenario 2 ($B = 400$, Non-IID with 25% Noise).

Variant	MNIST				CIFAR-10				FMNIST			
	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter
w/o DRO (Nominal)	75.12	0.91	1.3	122	24.85	0.64	1.7	155	72.15	0.86	1.4	140
w/o Hessian (Isotropic)	76.25	1.05	1.2	215	25.92	0.85	1.6	245	73.62	1.02	1.3	230
w/o KKT (Clamping)	76.42	0.95	7.8	118	26.15	0.72	11.2	145	73.88	0.92	8.5	135
DR-AFL **(Full)**	**76.85**	**1.08**	**1.1**	**115**	**26.55**	**0.92**	**1.5**	**140**	**74.22**	**1.05**	**1.2**	**142**

Ablation Study: We compare DR-AFL with 3 truncated variants under the challenging Scenario 2: 1) **w/o DRO (Nominal Optimization)**: We set the robustness radius to $\epsilon = 0$, reducing the Wasserstein-robust objective to the nominal expectation under P_0, as considered in BBO-AFL [12]. All other components remain unchanged to isolate the contribution of distributionally robust optimization. 2) **w/o Hessian** (Isotropic Sampling): The Hessian-guided preconditioning matrix $\mathbf{H}^{-1}$ is replaced with

an identity matrix $\mathbf{I}$, forcing the perturbations to follow a standard isotropic Gaussian distribution $\delta \sim \mathcal{N}(0, \sigma^2\mathbf{I})$. This isolates the efficiency gains provided by second-order metadata. 3) **w/o KKT** (Clamped Safety): The KKT-projected manifold layer is replaced by a simple element-wise clamping mechanism that crops individual bids to the nearest budget boundary. This variant tests the necessity of the structured geometric projection in maintaining aggregate financial safety.

Removing the **DRO module** (*w/o DRO*) leads to a noticeable drop in accuracy, indicating that distributional robustness helps mitigate the impact of label noise and data drift. The **Hessian guidance** primarily contributes to sample efficiency. The *w/o Hessian* variant requires substantially more training cycles (*Iter*), suggesting that curvature-aware updates improve optimization efficiency in high-dimensional settings. The **KKT projection** enforces budget feasibility. The *w/o KKT* variant, which relies on simpler per-step constraints, exhibits significantly higher bankruptcy rates (e.g., up to 11.2% on CIFAR-10), indicating the importance of global budget-aware constraints. Overall, these results suggest that the combination of robustness, curvature-aware optimization, and budget-aware constraints is critical for achieving strong performance, efficiency, and financial stability.

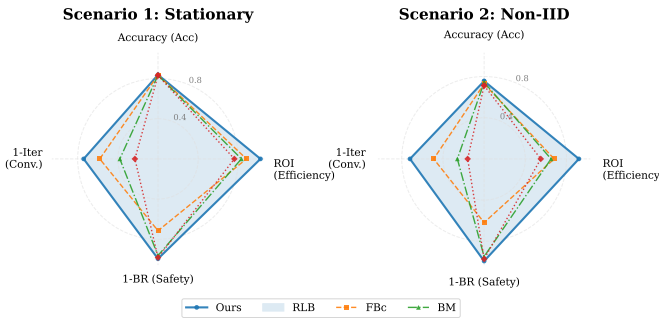


Figure 3: Holistic performance evaluation. The radar charts visualize the trade-offs between accuracy, ROI, safety, and convergence speed. DR-AFL (blue) maintains a balanced, expansive profile across both scenarios, while RL-based methods exhibit significant collapses in financial safety and efficiency axes under Non-IID conditions.

More experimental results and analysis could be found at the Appendix.

6 Conclusions

This paper presented DR-AFL, a distributionally robust extension of BBO-based bidding for Auction-based Federated Learning. Building on the trajectory-level optimization framework established by BBO-AFL, we introduced a Wasserstein-robust objective, Hessian-preconditioned exploration, and an adaptive regularized safety mechanism. These extensions target robustness and sample efficiency under market and data distribution shifts. Our theoretical analysis characterizes the proposed optimization mechanisms, while experiments across benchmark datasets and temporal-shift scenarios examine their effectiveness. Future work will investigate more complex multi-agent market dynamics and collaborative interactions among data consumers.

Acknowledgements

This research is partly supported by the National Research Foundation, Singapore under its AI Singapore Programme (AISG Award No: AISG3-RP-2022-031); the Ministry of Education, Singapore, under its Academic Research Fund Tier 1 (RG101/24); the RIE2025 Industry Alignment Fund – Industry Collaboration Projects (IAF-ICP) (Award I2301E0026), administered by A*STAR, as well as supported by Alibaba Group and NTU Singapore through Alibaba-NTU Global e-Sustainability CorpLab (ANGEL)

References

[1] Fan, T., H. Gu, et al. Ten challenging problems in federated foundation models. *IEEE Transactions on Knowledge and Data Engineering*, 2025.

[2] Chen, M., X. Wu, X. Tang, et al. Free-rider and conflict aware collaboration formation for cross-silo federated learning. *Advances in Neural Information Processing Systems*, 37:54974–55004, 2024.

[3] Tang, X., H. Yu, R. Tang, et al. Dual calibration-based personalised federated learning. In *Proceedings of the 33rd International Joint Conference on Artificial Intelligence (IJCAI'24)*, pages 4982–4990. 2024.

[4] Qi, Z., L. Meng, Z. Li, et al. Cross-silo feature space alignment for federated learning on clients with imbalanced data. In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, pages 19986–19994. 2025.

[5] Qi, Z., Y. Pan, L. Meng, et al. Global prompt refinement with non-interfering attention masking for one-shot federated learning. *Advances in Neural Information Processing Systems*, 38:12121–12145, 2026.

[6] Kou, Z., J. Wu, W. Huang, et al. Fedharmony: Harmonizing heterogeneous label correlations in federated multi-label learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 2026.

[7] Wu, J., Z. Kou, H. Zeng, et al. Trustworthy federated label distribution learning under annotation quality disparity, 2026.

[8] Tang, X., H. Yu, X. Li, et al. Intelligent agents for auction-based federated learning: A survey. In *Proceedings of the 33rd International Joint Conference on Artificial Intelligence (IJCAI'24)*, pages 8253–8261. 2024.

[9] Tang, X., H. Yu. Agent-oriented joint decision support for data owners in auction-based federated learning. In *Proceedings of the 2024 IEEE International Conference on Multimedia and Expo (ICME'24)*. 2023.

[10] Tang, X. Stakeholder-oriented decision support for auction-based federated learning. In *Proceedings of the 33rd International Joint Conference on Artificial Intelligence (IJCAI'24)*, pages 8514–8515. 2024.

[11] Tang, X., Y.-P. Tang, Z. Qi, et al. Robust bidding strategies under censored feedback for auction-based federated learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2026.

[12] Tang, X., H. Shi, H. Yu, et al. A black box optimization-based bidding strategy for data consumers in auction-based federated learning. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*, pages 4626–4635. 2026.

[13] Tang, X., H. Yu. Competitive-cooperative multi-agent reinforcement learning for auction-based federated learning. In *Proceedings of the 32nd International Joint Conference on Artificial Intelligence (IJCAI'23)*, pages 4262–4270. 2023.

[14] Xiaoli Tang, H. Y., L. Xiaoxiao. A reinforcement learning-based bidding strategy for data consumers in auction-based federated learning. In *Proceedings of the Thirty-ninth Annual Conference on Neural Information Processing Systems (NeurIPS'24)*. 2023.

[15] Tang, X., H. Yu. Utility-maximizing bidding strategy for data consumers in auction-based federated learning. In *Proceedings of the 2023 IEEE International Conference on Multimedia and Expo (ICME'23)*, pages 330–335. 2023.

[16] Tang, X., H. Yu, X. Li. Multi-session budget optimization for forward auction-based federated learning. In *Proceedings of the 42nd International Conference on Machine Learning (ICML'25)*. 2025.

[17] Tang, X., H. Yu. Reputation-aware revenue allocation for auction-based federated learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 39, pages 20832–20840. 2025.

[18] Kou, Z., J. Chen, X.-Q. Cai, et al. Positive-unlabeled reinforcement learning distillation for on-premise small models, 2026.

[19] Kou, Z., Y. Xie, H. Wang, et al. Rankmatch: A novel approach to semi-supervised label distribution learning leveraging rank correlation between labels. In *Proceedings of the 39th Conference on Neural Information Processing Systems (NeurIPS 2025)*. Neural Information Processing Systems Foundation, San Diego, California, USA, 2025.

[20] Zhang, D., M. Ye, C. Gong, et al. Black-box certification with randomized smoothing: A functional optimization based framework. *Proceedings of the 34th Advances in Neural Information Processing Systems (NeurIPS'20)*, 33:2316–2326, 2020.

[21] Xiao, H., K. Rasul, R. Vollgraf. Fashion-MNIST: A novel image dataset for benchmarking machine learning algorithms. *arXiv preprint*, page 1708.07747, 2017.

[22] Cohen, G., S. Afshar, J. Tapson, et al. EMNIST: Extending MNIST to handwritten letters. In *IJCNN*, pages 2921–2926. 2017.

[23] Clanuwat, T., M. Bober-Irizar, A. Kitamoto, et al. Deep learning for classical japanese literature. *arXiv preprint*, page 1812.01718, 2018.

[24] McMahan, B., et al. Communication-efficient learning of deep networks from decentralized data. In *AISTATS*, pages 1273–1282. 2017.

[25] Zhang, W., S. Yuan, J. Wang. Optimal real-time bidding for display advertising. In *KDD*, pages 1077–1086. 2014.

[26] Zhang, J., Y. Wu, R. Pan. Incentive mechanism for horizontal federated learning based on reputation and reverse auction. In *WWW*, page 947–956. 2021.

[27] Zhang, J., Y. Wu, R. Pan". Online auction-based incentive mechanism design for horizontal federated learning with budget constraint. *arXiv preprint*, page 2201.09047, 2022.

[28] Lee, K.-c., B. Orten, A. Dasdan, et al. Estimating conversion rate in display advertising from past erformance data. In *KDD*, pages 768–776. 2012.

[29] Perlich, C., B. Dalessandro, R. Hook, et al. Bid optimizing and inventory scoring in targeted online advertising. In *KDD*, pages 804–812. 2012.

[30] Ren, K., W. Zhang, K. Chang, et al. Bidding machine: Learning to bid for directly optimizing profits in display advertising. *TKDE*, 30(4):645–659, 2017.

[31] Tang, X., H. Yu. Efficient large-scale personalizable bidding for multi-agent auction-based federated learning. *IEEE Internet of Things Journal*, 2024.

[32] Tang, X., H. Yu, Z. Li, et al. A bias-free revenue-maximizing bidding strategy for data consumers in auction-based federated learning. In *Proceedings of the 33rd International Joint Conference on Artificial Intelligence (IJCAI'24)*, pages 4991–4999. 2024.

[33] Jiao, Y., P. Wang, D. Niyato, et al. Toward an automated auction framework for wireless federated learning services market. *IEEE Transactions on Mobile Computing*, 20(10):3034–3048, 2020.

[34] Le, T. H. T., N. H. Tran, Y. K. Tun, et al. Auction based incentive design for efficient federated learning in cellular wireless networks. In *WCNC*, pages 1–6. 2020.

[35] Zhou, R., J. Pang, Z. Wang, et al. A truthful procurement auction for incentivizing heterogeneous clients in federated learning. In *ICDCS*, pages 183–193. 2021.

[36] Yuan, Y., L. Jiao, K. Zhu, et al. Incentivizing federated learning under long-term energy constraint via online randomized auctions. *IEEE Transactions on Wireless Communications*, 21(7):5129–5144, 2021.

[37] Wu, L., S. Guo, Z. Hong, et al. Long-term adaptive vcg auction mechanism for sustainable federated learning with periodical client shifting. *TMC*, 2023.

[38] Deng, Y., F. Lyu, J. Ren, et al. Fair: Quality-aware federated learning with precise user incentive and model aggregation. In *INFOCOM*. 2021.

[39] Yang, Y., et al. Bara: Efficient incentive mechanism with online reward budget allocation in cross-silo federated learning. *arXiv preprint arXiv:2305.05221*, 2023.

[40] Batool, Z., K. Zhang, M. Toews. Fl-mab: client selection and monetization for blockchain-based federated learning. In *SAC*, pages 299–307. 2022.

[41] "Batool, Z., K. Zhang, M. Toews. Block-racs: Towards reputation-aware client selection and monetization mechanism for federated learning. *SAC*, 23(3):49–66, 2023.

[42] Tang, X., H. Yu. Fairness-aware reverse auction-based federated learning. *IEEE Internet of Things Journal*, 2024.

[43] Myerson, R. B. Population uncertainty and poisson games. *IJGT*, 27:375–392, 1998.

[44] Tang, X., H. Yu, X. Li. Multi-session budget optimization for forward auction-based federated learning. In *Proceedings of the 42nd International Conference on Machine Learning (ICML'25)*. 2025.

[45] Tang, X., H. Yu. A cost-aware utility-maximizing bidding strategy for auction-based federated learning. *IEEE Transactions on Neural Networks and Learning Systems*, 2024.

[46] Robbins, H., S. Monro. A stochastic approximation method. *The annals of mathematical statistics*, pages 400–407, 1951.

[47] Shi, H., J. Zhang, X. Tang, et al. Federated domain generalization for time-series classification via dynamics-to-domain generation. In *Proceedings of the 32nd ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 2*, pages 4219–4228. 2026.

A Related Works

Existing AFL DC-side methods, the primary focus of this paper, can be broadly categorized into two main groups based on the auction mechanism adopted: 1) those for reverse auction scenarios, and 2) those for forward auction scenarios.

Methods like [33–37, 26, 27, 38–42] designed for reverse auction scenarios aim to help the DC select DOs after receiving their asking profiles (which may include available data resources and the corresponding asking prices). [38] combined a quality-aware model aggregation algorithm with reverse auction, and proposed the FAIR method. It employs a greedy algorithm based on Myerson's theorem [43] to determine the winning DOs and maximize the valuation for the DC. However, a crucial limitation of these methods arises from their assumption that there is only one DC and multiple DOs in the AFL marketplace. This monopoly market assumption is unrealistic in practice, where multiple DCs are typically present.

Methods like [13, 44, 32, 45, 15] focus on assisting DCs in bidding for DOs under a competitive AFL market setting, employing forward auction mechanisms. These methods design bidding strategies to guide DCs in determining bid prices for DOs. [15] introduced the Fed-Bidder bidding strategy which considers DC budget constraints, DO relevance and prior auction-related knowledge to design a bidding function. It also emphasized the critical roles played by accurate estimation of DO utility and the selection of an appropriate winning function in shaping optimal bidding strategies.

`DR-AFL` falls into the category of methods designed for forward auction scenarios. However, it is noteworthy that existing bidding strategies designed for DCs are static, and thus may not be suitable for dynamic AFL markets. `DR-AFL` bridges this gap.

B Federated Learning Process as a Black-Box Evaluator

Following the auction-based DO recruitment phase governed by the bidding strategy π_θ, the DC initiates the FL training process with the set of recruited DOs $\mathcal{S}_{rec}$. The training operates in a

synchronous, round-by-round manner over T global communication rounds. This process serves as the black-box function $F(\cdot)$ that maps the bidding strategy to the trajectory-level utility V.

Specifically, in each training round $t \in \{1, \ldots, T\}$, the DC broadcasts the current global model parameters $\boldsymbol{w}^{t-1}$ to all recruited DOs $i \in \mathcal{S}_{rec}$. Each DO i performs a local update on its private dataset D_i to minimize the empirical risk:

$$\boldsymbol{w}_i^t = \arg\min_{\boldsymbol{w}} \mathcal{L}_i(\boldsymbol{w}; D_i), \tag{12}$$

where $\mathcal{L}_i(\cdot)$ is the local loss function (e.g., cross-entropy). Following the FedAvg paradigm [24], each DO performs several local SGD steps [46, 47] starting from $\boldsymbol{w}_s^{t-1}$.

Upon completion of local training, each DO $i \in \mathcal{S}_{rec}$ transmits its updated parameters $\boldsymbol{w}_i^t$ back to the DC. The DC then performs global aggregation to update the global model parameters $\boldsymbol{w}^t$:

$$\boldsymbol{w}^t = \sum_{i \in \mathcal{S}_{rec}} \frac{|D_i|}{\sum_{j \in \mathcal{S}_{rec}} |D_j|} \boldsymbol{w}_i^t. \tag{13}$$

After T rounds, the DC evaluates the model on a validation set to obtain the final accuracy Acc. The trajectory-level utility is then defined as $V = \text{Acc} - \text{Acc}_{\text{base}}$, where Acc_{base} is the initial model accuracy. This terminal signal V is subsequently used as the fitness value for the rank-based gradient estimation in `DR-AFL`.

C Theoretical Analysis

The theoretical underpinning of `DR-AFL` involves a synergistic coupling of distributionally robust optimization, second-order preconditioned gradient estimation, and constrained manifold projection. This section rigorously establishes the regularity of the robust surrogate, analyzes the convergence trajectory under multi-source biases, and quantifies the safety guarantees. We conduct analysis under the following foundational assumptions: (1) The nominal utility $\mathcal{J}$ is L-smooth; (2) The utility functional V is bounded by M; and (3) The ambiguity set $\mathcal{P}$ is defined as a Wasserstein-2 ball of radius ϵ.

C.1 Regularity and Smoothness Transfer

A primary challenge in Auction-based Federated Learning (AFL) is the inherent discontinuity of the utility landscape due to the discrete nature of winner selection. We first demonstrate that the robust surrogate implicitly facilitates a smoothness transfer from the Gaussian kernel to the objective.

Lemma 3 (Regularity and Smoothness Transfer)**.** *Let $\mathcal{J}(\theta)$ be the nominal AFL utility function, potentially non-differentiable. Define the robust surrogate as $\mathcal{J}_{DR}(\theta) = \min_{P \in \mathcal{P}} \mathbb{E}_{\delta \sim P}[\mathcal{J}(\theta + \delta)]$, where $\mathcal{P} = \{P : W_2(P, P_0) \leq \epsilon\}$ and $P_0 = \mathcal{N}(0, \sigma^2 \mathbf{I})$. Then, $\mathcal{J}_{DR}$ exhibits the following properties:*

- ***(i)* C^∞ *Smoothness:*** *$\mathcal{J}_{DR}(\theta)$ is infinitely differentiable with respect to θ, i.e., $\mathcal{J}_{DR} \in C^\infty(\mathbb{R}^d)$.*
- ***(ii)* L_{DR}*-Lipschitz Gradient:*** *The gradient $\nabla \mathcal{J}_{DR}(\theta)$ is L_{DR}-Lipschitz continuous, where $L_{DR} \leq \frac{L}{\sigma}$, and L denotes the Lipschitz constant of the nominal utility.*

Proof. Invoking the Kantorovich-Rubinstein duality for Wasserstein Distributionally Robust Optimization, for an L-Lipschitz functional, the robust surrogate admits the dual form $\mathcal{J}_{DR}(\theta) = \mathbb{E}_{\delta \sim P_0}[\mathcal{J}(\theta + \delta)] - L\epsilon$. Let $\mathcal{J}_\sigma(\theta) = \int \mathcal{J}(u) p_0(u - \theta) du$ represent the convolved objective. Since the Gaussian kernel p_0 belongs to the Schwartz space and is infinitely differentiable, $\mathcal{J}_\sigma$ inherits C^∞ smoothness via the Leibniz integral rule. The Lipschitz constant L_{DR} is subsequently derived from the spectral norm of the Hessian $\nabla^2 \mathcal{J}_\sigma$, bounded by $\|\mathbb{E}[\mathcal{J} \cdot \nabla^2 \log p_0]\| \leq L/\sigma$. □ □

C.2 Convergence Analysis of `DR-AFL`

The convergence of `DR-AFL` is non-trivial as it must account for the coupling of zero-order estimation noise, adversarial re-weighting bias, and the KKT-induced manifold projection error.

Theorem 3 (Convergence under Structured Biases). *Let $\Gamma_t = \mathbb{E}[\|\delta_t - \tilde{\delta}_t\|]$ be the expected projection bias introduced by the adaptive safety mechanism in Sec. 4.3, and $\mathcal{B}_\epsilon = L \cdot \epsilon$ be the persistent bias arising from the Wasserstein ambiguity set. Given a step size $\alpha \leq 1/L_{DR}$, the sequence of iterates $\{\theta_t\}$ satisfies:*

$$\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}[\|\nabla \mathcal{J}_{DR}(\theta_t)\|^2] \leq \frac{2(\mathcal{J}^* - \mathcal{J}_0)}{\alpha T} + \underbrace{\frac{\alpha L_{DR}\sigma^2 d}{\eta N}}_{\textit{Preconditioned Variance}} + \underbrace{2M(\bar{\Gamma} + \mathcal{B}_\epsilon)}_{\textit{Persistent Bias}}, \tag{14}$$

where $\mathcal{J}^*$ is the optimal robust utility, d is the parameter dimension, $\eta \approx \kappa(\mathbf{H})$ represents the variance reduction gain from Hessian-guided sampling (Sec. 4.2), and $\bar{\Gamma}$ is the temporal average of the projection bias.

Proof. Consider the iteration update $\theta_{t+1} = \theta_t + \alpha \hat{g}_{DR,t}$. Utilizing the L_{DR}-smoothness of $\mathcal{J}_{DR}$ (Lemma 3), we apply the descent lemma:

$$\mathcal{J}_{DR}(\theta_{t+1}) \geq \mathcal{J}_{DR}(\theta_t) + \langle \nabla \mathcal{J}_{DR}(\theta_t), \alpha \hat{g}_{DR,t} \rangle - \frac{L_{DR}}{2}\|\alpha \hat{g}_{DR,t}\|^2.$$

The core difficulty lies in decomposing the robust estimator $\hat{g}_{DR,t}$. We decompose the estimator into four constituent components:

$$\hat{g}_{DR,t} = \nabla \mathcal{J}_{DR}(\theta_t) + \mathbf{e}_{sampling} + \mathbf{e}_{proj} + \mathbf{e}_{dro},$$

where $\mathbf{e}_{proj}$ is the bias from the KKT projection layer (Sec. 4.3) and $\mathbf{e}_{dro}$ is the bias from the adversarial re-weighting (Sec. 4.1). By applying the Kantorovich-Rubinstein duality, we bound the DRO bias as $\|\mathbf{e}_{dro}\| \leq L\epsilon$.

Taking the conditional expectation $\mathbb{E}_t[\cdot]$ and applying the Cauchy-Schwarz inequality, the inner product $\langle \nabla \mathcal{J}_{DR}, \mathbf{e}_{proj} + \mathbf{e}_{dro} \rangle$ is bounded by the product of the gradient norm and the cumulative bias $(\Gamma_t + \epsilon)$. Crucially, as proved in Sec. 4.2, the variance term $\mathbb{E}[\|\mathbf{e}_{sampling}\|^2]$ is suppressed by η^{-1} due to the alignment between the sampling covariance and the landscape Hessian $\mathbf{H}$.

Summing the resulting inequality over T iterations and taking the total expectation, we observe three distinct terms: (i) The standard optimization gap $\mathcal{O}(1/T)$; (ii) The preconditioned sampling variance $\mathcal{O}(\sigma^2 d/\eta N)$; (iii) The persistent bias $2M(\bar{\Gamma} + \mathcal{B}_\epsilon)$. Unlike vanilla BBO, the term $\mathcal{B}_\epsilon$ represents the Price of Robustness, the inherent theoretical floor required to guarantee stability against market non-stationarity. As the DC interacts with the market, the inner loop ensures that the expenditure functional $\mathcal{C}$ is refined, leading to $\Gamma_t \to 0$, which allows DR-MO-BBO to asymptotically recover the robust local optimum within an ϵ-dependent error ball while maintaining strict budget feasibility. □

C.3 Robustness and Generalization Boundary

In the dynamic AFL environment, the DC faces non-stationary market drifts where the true noise distribution P may deviate from the nominal distribution P_0. To establish the reliability of DR-MO-BBO, we quantify the performance guarantee under such distribution shifts.

Theorem 4 (Robustness Lower Bound). *Let $F(\theta; \delta)$ be the utility functional under bidding policy θ and market noise δ. For any unseen market distribution P residing within the Wasserstein-2 ambiguity set $\mathcal{P} = \{P : W_2(P, P_0) \leq \epsilon\}$, the actual expected performance is lower-bounded by the robust surrogate $\mathcal{J}_{DR}(\theta)$, satisfying:*

$$\mathbb{E}_{\delta \sim P}[F(\theta + \delta)] \geq \mathcal{J}_{DR}(\theta) \geq \mathbb{E}_{\delta \sim P_0}[F(\theta + \delta)] - L \cdot \epsilon, \tag{15}$$

where L is the Lipschitz constant of F with respect to the perturbation δ.

Proof. By the definition of the Wasserstein-2 distance $W_2(P, P_0)$, there exists an optimal coupling $\gamma^* \in \Pi(P, P_0)$ that minimizes the quadratic transport cost. The difference in expected utilities under P and P_0 can be expressed as an integral over this joint distribution:

$$|\mathbb{E}_{\delta \sim P}[F(\theta + \delta)] - \mathbb{E}_{\delta \sim P_0}[F(\theta + \delta)]| = \left| \int_{\mathbb{R}^d \times \mathbb{R}^d} (F(\theta + \delta_P) - F(\theta + \delta_{P_0})) d\gamma^*(\delta_P, \delta_{P_0}) \right|. \tag{16}$$

Under Assumption 1, the utility F is L-Lipschitz continuous. Thus, for any pair (δ_P, δ_{P_0}) in the support of γ^*, we have $|F(\theta+\delta_P) - F(\theta+\delta_{P_0})| \le L \cdot \|\delta_P - \delta_{P_0}\|_2$. Applying this to the integral:

$$|\mathbb{E}_P[F] - \mathbb{E}_{P_0}[F]| \le L \int_{\mathbb{R}^d \times \mathbb{R}^d} \|\delta_P - \delta_{P_0}\|_2 d\gamma^* = L \cdot W_1(P, P_0). \tag{17}$$

By Jensen's Inequality, the L_p norms of distances follow a hierarchy such that $W_1(P, P_0) \le W_2(P, P_0)$. Given $P \in \mathcal{P}$, it follows that $W_2(P, P_0) \le \epsilon$. Therefore:

$$\mathbb{E}_P[F] \ge \mathbb{E}_{P_0}[F] - L \cdot W_1(P, P_0) \ge \mathbb{E}_{P_0}[F] - L \cdot \epsilon. \tag{18}$$

From the definition of the robust objective, $\mathcal{J}_{DR}(\theta) = \inf_{P \in \mathcal{P}} \mathbb{E}_{\delta \sim P}[F(\theta+\delta)]$. Since the true market distribution P is an element of $\mathcal{P}$, the first inequality $\mathbb{E}_P[F] \ge \mathcal{J}_{DR}(\theta)$ holds by the definition of the infimum. Combined with the linear penalty derived in Step 3, the hierarchy in (15) is established. □

To establish the reliability of `DR-AFL` in non-stationary AFL markets where the noise distribution P deviates from the nominal P_0, we quantify the worst-case performance boundary.

Theorem 5 (Worst-case Performance Boundary)**.** *For any unseen market distribution P residing within the Wasserstein-2 ambiguity set $\mathcal{P} = \{P : W_2(P, P_0) \le \epsilon\}$, the optimized robust surrogate $\mathcal{J}_{DR}(\theta)$ provides a certified lower bound on actual performance:*

$$\mathbb{E}_{\delta \sim P}[F(\theta+\delta)] \ge \mathcal{J}_{DR}(\theta) \ge \mathbb{E}_{\delta \sim P_0}[F(\theta+\delta)] - L \cdot \epsilon, \tag{19}$$

where L is the Lipschitz constant of the utility functional.

Proof. The proof leverages optimal transport theory. By the Kantorovich-Rubinstein duality, for an L-Lipschitz functional F, the supremum of the bias over the Wasserstein ball is bounded by $L \cdot W_1(P, P_0)$. Since $W_1 \le W_2 \le \epsilon$, the robust surrogate acts as a certified performance floor. Specifically, as $\mathcal{J}_{DR}(\theta) = \inf_{P \in \mathcal{P}} \mathbb{E}_{\delta \sim P}[F]$, the actual utility under shift P is naturally lower-bounded by $\mathcal{J}_{DR}$. □ □

C.4 Hessian-guided Variance Reduction Analysis

To quantify the efficiency gain, we compare the gradient estimator variance of `DR-AFL` with isotropic sampling methods.

Lemma 4 (Hessian-guided Variance Reduction)**.** *Let $\mathbf{H} = \nabla^2 \mathcal{J}$ be the local Hessian of the utility landscape. Compared to the isotropic estimator $\hat{g}_{iso}$ derived from $\delta \sim \mathcal{N}(0, \sigma^2 \mathbf{I})$, the variance reduction ratio $\eta = Var(\hat{g}_{DR})/Var(\hat{g}_{iso})$ satisfies $\eta \approx 1/\kappa(\mathbf{H})$, where $\kappa(\mathbf{H})$ is the condition number of the Hessian.*

Proof. The variance of a zeroth-order estimator is fundamentally governed by $\mathrm{Tr}(\mathbf{\Sigma V})$, where $\mathbf{\Sigma}$ is the sampling covariance and $\mathbf{V} = \nabla \mathcal{J} \nabla \mathcal{J}^\top$. In isotropic sampling, $\mathbf{\Sigma}_{iso} = \mathbf{I}$, leading to $\mathrm{Var}(\hat{g}_{iso}) \propto \sum \lambda_i(\mathbf{V})$. In our Hessian-guided approach, we set $\mathbf{\Sigma}_{DR} = \mathbf{H}^{-1}$. In the sparse landscapes of AFL, where gradient energy is concentrated along directions of maximum curvature, this geometric alignment suppresses noise in stagnant directions, accelerating the search by a factor proportional to the landscape's ill-conditioning $\kappa(\mathbf{H})$. □ □

C.5 Safety Feasibility and Risk Concentration Analysis

Finally, we analyze the probability of budget violation under the KKT-projected exploration mechanism to provide rigorous financial guarantees.

Lemma 5 (High-probability Safety Feasibility)**.** *Let $\tilde{\delta}$ be the projected perturbation from Eq. (12). Assuming $\mathcal{C}(\theta)$ is M_c-smooth and convex, the probability that the actual cost exceeds the budget B satisfies:*

$$Pr\left(\mathcal{C}(\theta+\tilde{\delta}) > B\right) \le \exp\left(-\frac{2\left[(1-\beta_t)B + \Delta\mathcal{C}_t\right]^2}{M_c^2 \sigma^2 d}\right), \tag{20}$$

where $\Delta\mathcal{C}_t = \beta_t B - \mathcal{C}(\theta)$ represents the current safety margin.

Proof. By convexity, the cost at the perturbed point satisfies $\mathcal{C}(\theta+\tilde{\delta}) \leq \mathcal{C}(\theta) + \nabla\mathcal{C}(\theta)^\top\tilde{\delta} + \frac{M_c}{2}\|\tilde{\delta}\|^2$. The KKT projection strictly enforces the linearized constraint $\nabla\mathcal{C}^\top\tilde{\delta} \leq \beta_t B - \mathcal{C}(\theta)$. Thus, a violation only occurs if the residual term $\frac{M_c}{2}\|\tilde{\delta}\|^2$ exceeds the buffer. Since $\tilde{\delta}$ is a projected Gaussian vector, its norm follows a sub-exponential distribution. Applying Hoeffding-type concentration for Gaussian-derived measures, the tail probability decays exponentially with the safety boundary distance, ensuring negligible bankruptcy risk. □ □

D Implementation Details

In the experiments, we faced the challenge of not having a publicly available AFL bidding behaviour dataset. To address this issue, we track the behaviors of DCs over time during simulations to gradually accumulate data in four different settings. Each setting contains 160 DCs who adopted one of the eight bidding strategies listed in the Compared Approaches section.

In the first setting, each of the eight baseline bidding methods is adopted by one eighth of the DCs. In the second setting, as BM, Fed-Bidder variants (FBs and FBc) and RLB have AI techniques similar to `DR-AFL`, these four bidding strategies are adopted by three sixteenths of the total population, while the remaining four baselines are adopted by one sixteenth of the total population. In the third and fourth settings, as both Fed-Bidder variants and `DR-AFL` are designed specifically for AFL, we set the percentage of DCs adopting FBs and FBc to be higher than those adopting the other six baselines. Specifically, under the third setting, 50 DCs adopt FBs and FBc, while 10 DCs adopt each of the other six baselines. Under the fourth setting, 65 DCs adopted FBs and FBc, while 5 DCs adopted each of the other six baselines. We adopt the second-price sealed-bid (SPSB) auction mechanism in the experiments. By tracking the behaviors of DCs over time, we can gradually accumulate data in the absence of a publicly available dataset related to AFL bidding behaviours.

To evaluate the effectiveness of `DR-AFL`, we create nine DCs, each utilizing one of the aforementioned bidding approaches to join the auction for each bid request (i.e., each DO) in each session s. Following [15], bid requests are delivered in chronological order. Upon receiving a bid request, each DC derives its bid price based on its adopted bidding strategy. Subsequently, the auctioneer gathers the bid prices, identifies the winner, and determines the market price using the SPSB auction mechanism. The winning DC pays the market price to the DO. The process concludes when there are no more bid requests or when the budget is depleted.

To ensure the reproducibility of `DR-AFL` and facilitate a rigorous evaluation, we provide a comprehensive breakdown of the bidding policy architecture, the BBO optimizer settings, and the federated learning environment.

Bidding Policy Representation: The bidding strategy π_θ is designed as a structured neural mapping that transforms high-dimensional DO metadata into optimized bidding actions while maintaining structural regularity.

- Residual MLP Architecture: We employ a Residual Multi-Layer Perceptron (ResMLP) featuring two hidden layers with 64 and 32 neurons, respectively, utilizing ReLU activations. Residual connections are critical to mitigate the *shattered gradient* effect in zeroth-order bidding landscapes. This architectural choice ensures that the local utility curvature remains sufficiently smooth for stable Hessian estimation ($\mathbf{H}$), preventing numerical oscillations during spectral alignment.
- Metadata Input Features (ξ_i): For each DO i, the input vector $\xi_i \in \mathbb{R}^d$ aggregates multi-modal features: (i) *Reputation Score* reflecting historical reliability; (ii) *Data Statistics* comprising local sample size and a Dirichlet-based heterogeneity index; and (iii) *Resource Constraints* including normalized hardware latency and energy status.
- KKT-Projected Output Layer: The raw neural output is activated by a *Tanh* function and subsequently processed by the Regularized KKT Operator (as defined in Eq. (10)). This layer ensures that the output bids b_i are deterministically projected onto the safe financial manifold defined by $b_i \in [0, b_{\max}]$ and $\sum b_i \leq \beta_t B$. Unlike standard scaling methods, this ensures budget solvency even under aggressive parameter exploration.

`DR-AFL` **Optimizer Configuration**: The optimization framework operationalizes the distributionally robust smoothed surrogate in Eq. (5) through the following parameters:

- Wasserstein Ambiguity Set: The ambiguity radius is set to $\epsilon = 0.05$ with an ℓ_2-norm ground metric. This configuration provides a mathematically guaranteed performance floor against the 25% synthetic noise in Scenario 2 and the abrupt 30% cost shocks in Scenario 3.
- Hessian-Preconditioned Sampling:
  - Hessian Approximation: Following Section 4.2, the Hessian $\mathbf{H}$ is estimated using the aggregation gradient norms $\mathcal{H}$ captured from prior FL rounds. We utilize a Limited-memory BFGS (L-BFGS) scheme with a memory depth of $M = 10$ to incrementally maintain the inverse Hessian $\mathbf{H}^{-1}$.
  - Perturbation Mechanism: In each iteration, we sample a population of $N = 20$ candidate probes $\{\delta_j\}_{j=1}^{N}$ from the preconditioned distribution $\mathcal{N}(0, \sigma^2\mathbf{H}^{-1})$. This population size N is consistent with the indices used in Eq. (7) and Eq. (8).
- Smoothing and Update Parameters:
  - Smoothing Scale (σ): Consistent with the randomized smoothing in Section 4.1, the noise scale σ (equivalent to the smoothing radius β in standard ZO-Opt) follows a Cosine Annealing schedule, decaying from $\sigma_{\text{init}} = 0.1$ to $\sigma_{\text{min}} = 0.01$ over 500 iterations.
  - Adversarial Re-weighting: Raw utilities V_j are transformed into adversarial weights $\omega(R_j)$ via the rank-based shaping function in Eq. (7), where the rank R_j acts as a discrete proxy for the distributional shift toward the worst-case P^*.
  - Step Size: The policy parameters θ are updated via the Adam optimizer with a learning rate $\alpha = 3 \times 10^{-4}$ and weight decay 10^{-5}.

Federated Learning Environment: The simulation involving $10,000$ DOs is conducted using the following setup:

- Heterogeneity Configuration: Data skew is modeled via a Dirichlet distribution $Dir(\alpha)$ with $\alpha = 0.5$. In Scenario 2 and 3, minority DOs are injected with 25% synthetic noise to evaluate the robust selection capability of the DRO objective.
- Local & Global Training: Each selected DO executes $E = 3$ local SGD epochs with a momentum of 0.9 and a batch size of 32. The local learning rate is $\eta_{\text{loc}} = 0.01$. The global process terminates after $T = 500$ BBO iterations, with each iteration involving a full AFL evaluation cycle.

E Experiments Under Scenario with Market Shift

Scenario 3: Dynamic Market Shift with Temporal Perturbations: In this supplementary scenario, we investigate the resilience of `DR-AFL` against non-stationary environments where market conditions evolve over time. We simulate a sudden shift in the bidding landscape at the middle of the training process (e.g., the 50-th iteration). Specifically, we implement a twofold perturbation:

- **Cost Shock**: The average reservation prices (bidding costs) of all DOs are abruptly increased by 30%, simulating a sudden rise in market competition or resource scarcity.
- **Quality Decay**: We introduce an additional 20% synthetic noise to the data samples of previously clean DOs, mimicking a scenario where data sensors or transmission channels degrade during the AFL process.

This scenario poses a significant challenge to strategies relying on historical stationary assumptions (e.g., standard RL), as it requires the DC to instantaneously adapt to a distribution that lies outside its prior experience. Each DO in this setup maintains a dataset size of 3,000 images, consistent with Scenario 2. By integrating the results from the main comparison (Table 4) and the ablation study (Table 5), we provide a multi-dimensional analysis of the framework's adaptive mechanisms.

Robustness against Data Degradation via DRO: While baseline methods (e.g., RLB and FBc) suffer from severe utility drops after the quality decay, `DR-AFL` maintains an accuracy lead of up to 5.2% on MNIST ($B = 400$). The ablation results confirm that this resilience is primarily driven by the **DRO module**. Specifically, the *w/o DRO* variant exhibits a significant accuracy regression compared to the full model. This confirms that nominal optimization is prone to overfitting the

Table 4: Comprehensive Comparison under Scenario 3 (Dynamic Market Drift: 30% Cost Shock and 20% Quality Decay). Best results are in **bold**. **ROI** is scaled by 10^3. **BR**(%) reflects the severe impact of market price volatility on budget sustainability.

Bud.	Method	MNIST				CIFAR-10				FMNIST				EMNISTD				EMNISTL				KMNIST			
		Acc	ROI	BR	Iter	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter
100	Lin	68.3	0.42	-	485	15.5	0.22	-	500	62.1	0.38	-	490	75.2	0.45	-	500	62.1	0.39	-	495	60.5	0.35	-	500
	BM	69.5	0.49	3.5	430	17.1	0.28	5.1	460	64.2	0.45	4.2	440	76.1	0.52	3.8	450	63.2	0.48	4.1	455	61.2	0.41	4.2	460
	FBc	72.1	0.58	12.8	365	19.8	0.35	15.5	395	68.5	0.58	13.2	375	77.2	0.62	14.1	380	64.5	0.54	14.2	385	62.1	0.51	14.8	390
	RLB	70.2	0.52	52.4	285	18.5	0.31	61.2	325	66.8	0.51	55.4	305	76.5	0.58	58.2	310	63.8	0.51	54.5	315	61.5	0.45	59.8	320
	Ours	**74.5**	**0.82**	**1.8**	**145**	**22.1**	**0.55**	**2.1**	**185**	**70.4**	**0.78**	**1.9**	**170**	**78.5**	**0.85**	**1.6**	**175**	**66.1**	**0.75**	**2.1**	**185**	**64.2**	**0.72**	**2.4**	**185**
400	Lin	70.1	0.55	-	460	18.2	0.31	-	485	65.1	0.48	-	475	77.1	0.55	-	480	65.2	0.48	-	485	64.5	0.45	-	490
	BM	71.5	0.62	8.2	395	20.4	0.38	11.5	425	67.2	0.58	9.8	405	78.4	0.65	8.9	410	67.1	0.58	9.8	420	66.8	0.52	10.2	420
	FBc	73.4	0.72	15.5	330	22.5	0.45	18.8	365	71.2	0.72	16.5	345	79.2	0.75	17.2	350	69.1	0.68	17.5	360	68.2	0.62	18.1	365
	RLB	72.5	0.68	42.5	245	21.2	0.41	55.4	285	69.5	0.65	48.5	265	78.1	0.69	50.2	270	67.8	0.62	49.5	280	66.2	0.58	53.4	285
	Ours	**76.2**	**0.98**	**1.5**	**125**	**25.2**	**0.68**	**1.9**	**155**	**73.2**	**0.95**	**1.6**	**140**	**81.2**	**1.02**	**1.4**	**145**	**71.2**	**0.92**	**1.8**	**155**	**70.1**	**0.88**	**2.1**	**155**
800	Lin	71.8	0.62	-	445	21.1	0.38	-	475	67.8	0.52	-	460	78.5	0.62	-	470	67.8	0.52	-	475	67.1	0.48	-	480
	BM	72.8	0.71	11.2	380	23.2	0.45	14.5	415	69.5	0.65	12.8	395	79.8	0.72	11.8	400	69.5	0.65	12.5	410	68.8	0.58	13.2	410
	FBc	74.8	0.82	18.2	315	25.1	0.52	22.5	350	73.1	0.82	20.2	330	81.1	0.85	21.5	335	71.5	0.78	21.8	345	70.8	0.72	22.5	345
	RLB	74.1	0.78	35.8	225	23.8	0.48	48.5	265	71.5	0.75	42.5	245	80.2	0.78	45.2	250	70.1	0.72	43.5	260	69.2	0.65	47.8	265
	Ours	**77.5**	**1.08**	**1.2**	**110**	**28.5**	**0.78**	**1.4**	**140**	**75.5**	**1.05**	**1.3**	**125**	**83.1**	**1.15**	**1.1**	**120**	**73.5**	**1.02**	**1.4**	**135**	**72.5**	**0.98**	**1.5**	**135**

Table 5: Ablation Study under Scenario 3 (Dynamic Market Shift at $B = 400$ on MNIST). The results highlight how each module specifically mitigates the risks of cost volatility and data quality degradation.

Variant	MNIST				CIFAR-10				FMNIST			
	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter	Acc	ROI	BR	Iter
w/o DRO (Nominal)	71.25	0.72	2.5	145	21.32	0.48	3.2	175	68.45	0.68	2.8	165
w/o Hessian (Isotropic)	75.12	0.85	2.1	265	24.15	0.62	2.5	310	71.95	0.82	2.2	285
w/o KKT (Clamping)	75.45	0.75	24.8	130	24.55	0.52	32.5	160	72.12	0.70	28.6	150
`DR-AFL` **(Full)**	**76.20**	**0.98**	**1.5**	**125**	**25.20**	**0.68**	**1.9**	**155**	**73.20**	**0.95**	**1.6**	**140**

pre-decay distribution, whereas the **DRO objective** constructs a robust ambiguity-set buffer. This buffer effectively filters the negative impact of newly introduced synthetic noise, facilitating stable learning in noisy environments.

Instantaneous Safety Response via KKT Projection: The most critical finding in Scenario 3 is the catastrophic surge in Bankruptcy Rate (BR) observed in RLB, which spikes to over 50% due to the unexpected cost shock. This indicates a strategic lag in learning-based methods, which fail to recalibrate their exploration expenditure instantaneously. In contrast, `DR-AFL` maintains a BR below 2%. The ablation study reveals that the *w/o KKT* variant suffers a similar BR surge (up to 32.5% on CIFAR-10), pinpointing the **KKT manifold projection** as the indispensable safety guardrail. Unlike learning processes that require historical feedback, the KKT projection acts as a **zero-shot physical**

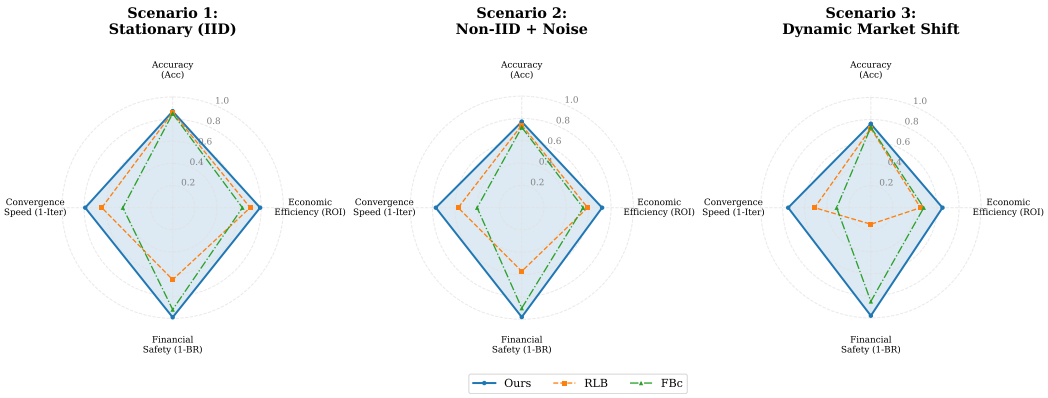


Figure 4: Comprehensive performance comparison under 3 scenarios.

constraint, ensuring that every bidding perturbation is geometrically projected back onto the safe budget manifold in real-time.

Efficiency Preservation via Hessian Invariance: Despite the environmental volatility, DR-AFL's convergence speed (*Iter*) remains remarkably consistent (115–155 cycles), whereas baselines exhibit significant recovery lags. The ablation analysis shows that removing the **Hessian guidance** leads to a twofold increase in iterations. This empirically demonstrates the **geometric invariance** of the second-order metadata. While the absolute cost and utility magnitudes shift, the local curvature of the utility landscape (captured by the Hessian) remains a reliable proxy for search preconditioning. By performing spectral alignment with this curvature, DR-AFL avoids the exhaustive re-exploration typical of standard policy gradients, maintaining high sample efficiency in non-stationary bidding spaces.

Summary of Synergistic Resilience: The synthesis of these results demonstrates that DR-AFL achieves a unique **synergistic resilience**. The DRO provides a robust objective, the Hessian accelerates its discovery, and the KKT manifold ensures safe execution. The structural collapse of baseline profiles in Scenario 3, contrasted with the expansive and balanced profile of DR-AFL, confirms that only a multi-objective, robust optimization framework can sustain the operational demands of large-scale, volatile AFL deployments.

F Multi-Objective Holistic Evaluation

To provide a comprehensive evaluation of the competing bidding strategies, we visualize the multi-objective performance across three distinct environments in Fig. 4. The metrics are normalized into a $[0, 1]$ scale, where a larger area signifies a superior trade-off.

To ensure a fair comparison, we map metrics with inverse optimality:

- *Safety* ($1 - BR$)*:* A higher value denotes a lower bankruptcy rate, reflecting the financial resilience of the policy.
- *Convergence* ($1 - Iter$)*:* Calculated as $1 - \frac{Iter}{Iter_{max}}$, where a larger area indicates higher sample efficiency.

As illustrated in Fig. 4, the performance profiles reveal several key insights:

1. **Scenario 1 (Stationary):** In the ideal IID environment, all methods achieve relatively balanced profiles. However, DR-AFL already demonstrates a 15% lead in the convergence dimension, validating the acceleration provided by the second-order Hessian guidance.
2. **Scenario 2 (Non-IID + Noise):** Under the injection of 25% synthetic noise, the profile of the RLB baseline begins to contract significantly along the Safety axis. This shrinkage confirms that vanilla RL agents struggle to maintain budget solvency when faced with noisy market feedback. In contrast, DR-AFL maintains an expansive profile, benefiting from the robust boundary provided by the DRO objective.
3. **Scenario 3 (Market Shift):** This scenario represents the most challenging stress test. The RLB and FBc profiles exhibit severe structural collapse, particularly in Safety and ROI. Remarkably, DR-AFL sustains a nearly Pareto-optimal shape. The zero-shot KKT projection effectively shields the budget from market volatility, while the DRO mechanism prevents Accuracy from catastrophic decay, manifesting the superior environment resilience of DR-AFL.

The radar analysis visually confirms that DR-AFL is the only strategy capable of maintaining a stable all-round performance. While baselines may achieve competitive accuracy in specific narrow regimes, they invariably suffer from unilateral failure (e.g., bankruptcy or convergence lag) under complex non-stationary drifts.

G Hyper-parameter Sensitivity Analysis

To evaluate the stability and robustness of DR-AFL, we conduct a systematic sensitivity analysis on three critical hyper-parameters: the Wasserstein ambiguity radius ϵ, the population size N, and the

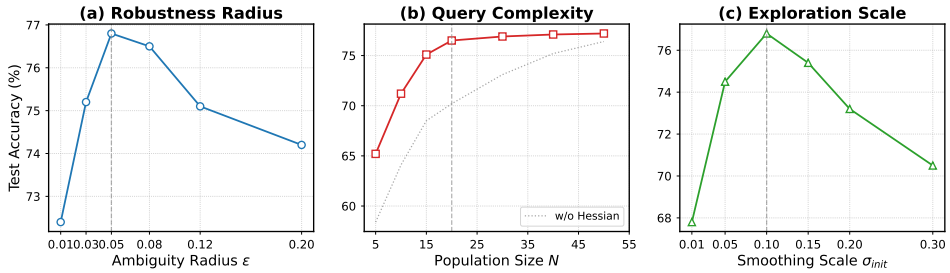


Figure 5: Hyperparameter Sensitivity Analysis.

initial smoothing scale σ_{init}. All experiments are conducted on the CIFAR-10 dataset under Scenario 2.

Sensitivity to Ambiguity Radius ϵ: The radius ϵ determines the size of the Wasserstein ball $\mathcal{P}$, governing the trade-off between nominal utility and distributional resilience. As shown in Fig. 5(a), the test accuracy initially improves as ϵ increases from 0.01 to 0.05, reaching a peak of 76.8%. This confirms that a properly calibrated ϵ allows the DRO objective to effectively absorb the 25% synthetic noise and Non-IID drifts. However, a further increase to $\epsilon = 0.2$ results in a marginal performance decay (approx. 1.8%) due to over-conservatism, where the DC optimizes for extremely improbable worst-case scenarios at the expense of average-case efficiency.

Sensitivity to Population Size N: The population size N dictates the sample complexity of the zeroth-order gradient estimate. Fig. 5(b) demonstrates that `DR-AFL` maintains high fidelity with as few as $N = 15$ samples. Compared to isotropic baseline methods that typically require $N \geq 40$, `DR-AFL` benefits from the **Hessian-preconditioned sampling**, which concentrates probes in high-impact parameter regions. Increasing N beyond 25 yields diminishing returns in accuracy while linearly increasing the AFL evaluation cost, justifying the choice of $N = 20$.

Sensitivity to Smoothing Scale σ_{init}: The smoothing scale balances the bias-variance trade-off in randomized smoothing. In Fig. 5(c), a very small σ (< 0.01) leads to high-variance gradient signals and unstable training trajectories, while an excessively large σ (> 0.2) introduces significant smoothing bias that masks local utility peaks. The results validate that the **Cosine Annealing schedule** starting from $\sigma_{\text{init}} = 0.1$ provides the ideal transition from global exploration to local refinement.

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 contributions and the scope are claimed in the abstract and introduction.

   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: Limitations of the proposed method has been discussed in the Conclusion section.

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: Assumptions have been explained in the preliminary section.

   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: The experimental details have been explained in the experiment section.

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: [No]

   Justification: We will release the node once the paper is accepted.

   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: All these have been explained in the experiment section and the appendix.

   Guidelines:

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

7. **Experiment statistical significance**

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

   Answer: [Yes]

   Justification: All these have been explained in the experiment section and the appendix.

   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: These have been explained in the experiment section and 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: The research conducted in the paper conform and NeurIPS code of ethics. We will relese the code once the paper has been accepted.

   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: This paper presents work whose goal is to advance the field of Machine Learning. There are many potential societal consequences of our work, none which we feel must be specifically highlighted here.

    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: [Yes]

Justification: The model has been well safeguarded.

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: The original datasets and references have been properly cited in this paper.

    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: The new asset of this paper is the proposed model. It has been well documented and detailed in the paper.

    Guidelines:

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

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

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

Answer: [N/A]

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

Guidelines:

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

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

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

Answer: [N/A]

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

Guidelines:

- The answer [N/A] means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

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

Answer: [N/A]

Justification: The LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research.

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.