COLLABORATIVE PERSONALIZED PREFERENCE ALIGNMENT FOR LLMS UNDER DATA DEFICIENCY

Anonymous authors
Paper under double-blind review

ABSTRACT

Real-world users often exhibit highly heterogeneous preferences over multiple objectives for LLM responses. A lightweight aligner can tailor these responses to individual preferences, but scarce user-specific feedback makes personalized training difficult. Learning shared initializations across users can support few-shot adaptation. However, heterogeneous preferences and competing objectives cause gradient conflicts across users and within each user, hindering effective initialization learning. This raises a central question: **how can we collaboratively learn aligner initializations that support few-shot adaptation to diverse user preferences?** To answer this question, we propose **A**pproximate **P**areto **O**ptimality (APO). We first group users whose updates are compatible, so that their information can be combined with less interference. Within each group, we combine gradient descent with controlled ascent to coordinate competing objectives and move towards preference-specific points on the Pareto front. This produces an initialization that is close to the optima of the users in the group. We then iteratively refine it using updates from few-shot local adaptation, making it more effective for personalization. Furthermore, we establish conditional suboptimality bounds for a one-local-step collaborative update and characterize how initialization error affects subsequent stochastic adaptation. Experiments on Fed-ChatbotPA and UltraFeedback show consistent improvements over existing methods using only 20 local examples.

1 INTRODUCTION

Large language models (LLMs) (OpenAI, 2026; Google, 2026) serve clients with various cultural background, personal experience, and individual values. As a result, it is a key requirement to tailor generated response to personal preferences. Multi-objective preference alignment (MOPA) captures this need by generating outputs that match a client-specified balance across objectives such as helpfulness, harmlessness, and humor (Yang et al., 2024b; Shi et al., 2024; Guo et al., 2024). A straightforward way to personalize an LLM is to adapt a plug-and-play aligner (Ji et al., 2024; Yang et al., 2024a) to adjust LLM responses on certain preference feedback of client. However, the scarcity of client-specific preference feedback creates a critical bottleneck for this approach, making it difficult to train an effective personalized aligner from scratch.

Existing approaches address this scarcity through data augmentation, including online imitation learning (Shaikh et al., 2025) and meta-learning with synthetic user preferences (Singh et al., 2026). However, synthetic data may not fully capture real users' subtle preferences in style, tone, and context. Personalized federated learning (Wu et al., 2024; Ye et al., 2024a; Jiang et al., 2019; Nguyen et al., 2026) offers a promising alternative by enabling collaboration across clients using their local data. In this setting, clients can jointly learn a shared initialization that captures common patterns and then adapt it to their own preferences using a few local examples. However, such collaboration faces *preference gradient conflicts* at two levels (Zhang et al., 2024b; Guo et al., 2026). Across clients, heterogeneous preferences can cause locally beneficial updates to interfere during aggregation, leading to *inter-client* conflicts. Within each client, gradients for different objectives may point in opposing directions, leading to *intra-client* conflicts. These challenges raise a key question: **how can we get aligner initializations enabling few-shot adaptation while addressing both inter- and intra-client preference gradient conflicts?**

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

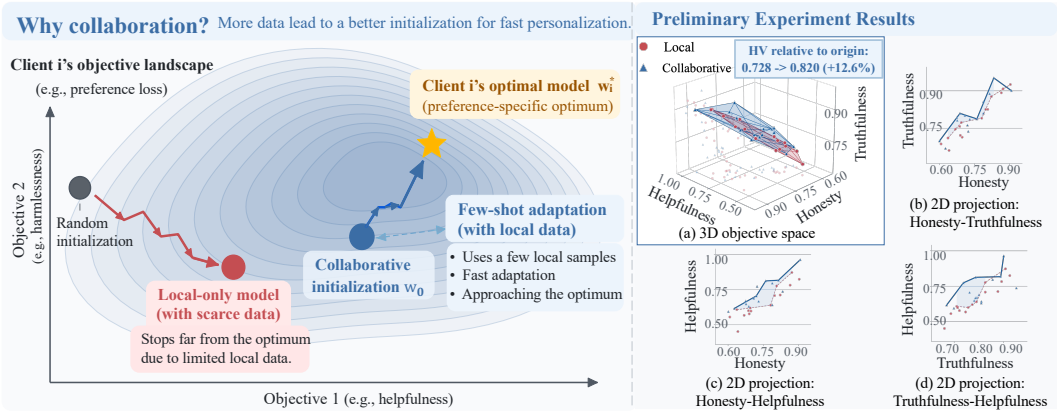


Figure 1: Collaborative personalized preference alignment under data deficiency. **Left:** with scarce local feedback, local-only training from a random initialization stops far from a client's preference-specific optimum, whereas a collaboratively learned initialization enables effective few-shot personalization. **Right:** preliminary results on UltraFeedback (Cui et al., 2023) with Llama-3.2-3B-Instruct (Meta AI, 2024) (15 clients, 20 samples each). Over helpfulness, honesty, and truthfulness, collaborative initialization raises the pareto hypervolume (HV) from 0.728 to 0.820 ($+12.6\%$), and its frontier dominates the local one in the 2D projections.

To solve this problem, we propose Approximate Pareto Optimality (APO). The framework learns an initialization for each client cluster (Ghosh et al., 2020). Compared with purely local training, each collaborative initialization reaches a point closer to the preference-specific optima of its clients. Starting from this initialization, few-shot adaptation is further guaranteed to approach the optimum attainable with sufficient local data more closely (see Figure 1). APO first addresses inter-client conflicts. It groups clients according to their *bottleneck* objectives, defined as the objectives with the largest preference-weighted losses. These bottlenecks determine the clients' admissible update directions. We first partition clients by bottleneck ordering. We then refine each group using preference-adjustment vectors. This procedure reduces interference during aggregation and produces more reliable shared initializations. Furthermore, it combines preference gradient descent with controlled ascent (Mahapatra & Rajan, 2020) to traverse the pareto front within each cluster to solve intre-client conflicts. This yields the cluster-level initialization. Finally, we iteratively refine these initializations through adaptation-aware learning (Nichol et al., 2018; Jiang et al., 2019). In each round, clients adapt an initialization using a few local examples, and the server uses the resulting updates to refine it. Repeating this process produces initializations that adapt more effectively to individual preferences with limited local feedback. Under explicit assumptions, single-step collaborative updates achieve an asymptotic chebyshev suboptimality bound of ϵ_0 for each client. Preconditioned stochastic adaptation yields an expected suboptimality bound of δ_0 relative to the full-data optimum for unseen preferences over shared objectives. Experiments on different datasets show that APO achieves the strongest overall performance among the evaluated methods in both preference-specific utility and pareto-front coverage.

In summary, we make the following contributions: **1)** We formalize collaborative personalized preference alignment of LLMs under data scarcity, and propose APO that learns a shared initialization for preference-conditioned few-shot adaptation. **2)** We introduce bottleneck-adjustment clustering and preference gradient descent with controlled ascent to address inter-client and intra-client preference gradient conflict. **3)** We establish an asymptotic chebyshev suboptimality bound for single-local-step collaborative updates, and an expected suboptimality bound for preconditioned stochastic adaptation to unseen preferences under shared objective functions. **4)** Extensive experiments on Fed-ChatbotPA and UltraFeedback with Llama-3.2-3B-Instruct and Qwen2.5-3B-Instruct demonstrate the superiority of APO over state-of-the-art methods.

2 Related Work

Personalized Multi-Objective Alignment under Data Scarcity. Multi-objective alignment realizes client-specified trade-offs through preference-conditioned training (Yang et al., 2024b), post-

training interpolation (Rame et al., 2023), or a single model that recovers the Pareto front (Zhong et al., 2024). MetaAligner (Yang et al., 2024a) meta-learns a lightweight aligner that generalizes across unseen objectives, but not across users with heterogeneous multi-objective preferences. Under limited feedback, FSPO (Singh et al., 2026) adapts from synthetic user preferences, while DITTO (Shaikh et al., 2025) learns from a few user demonstrations. These methods rely on abundant centralized or synthetic data and fail to learn a personalized model from scarce real-client feedback under preference-specific settings.

Collaborative Alignment under Gradient Conflicts. Federated preference alignment applies local DPO or trains subpopulation selectors (Ye et al., 2024b; Wu et al., 2024), while FIRM (Nourzad et al., 2025) resolves multi-objective disagreement through client-side regularized MGDA; these methods nevertheless optimize a single population-level model without few-shot personalization guarantees. Personalized FL learns shared initializations (Jiang et al., 2019) or a few models for many clients (Guo et al., 2026), but does not model preference-specific Pareto geometry. For intra-model conflicts, EPO (Mahapatra & Rajan, 2020) uses controlled ascent to reach a preference-specific Pareto point, yet does not address interference among clients with heterogeneous preferences. APO jointly addresses inter-client and intra-client gradient conflicts under data deficiency through bottleneck-adjustment clustering, controlled-ascent updates, and adaptation-aware meta-training. Further related work is discussed in Appendix G.

3 METHODOLOGY

We propose APO to learn aligner initialization that provide useful starting points for clients with different multi-objective preferences and limited local feedback. We first characterize the preference corrections involved in collaboration and use them to organize clients into clusters. Within each cluster, collaborative Pareto set initialization learning produces a basic initialization, which is then refined through adaptation-aware meta-training. A new client selects the corresponding initialization and performs a few-shot adaptation to obtain a personalized model tailored to its own preference. In the following, we provide a detailed description of our framework. Detailed proofs of all theoretical results are shown in Appendix H.

3.1 PRELIMINARIES AND PROBLEM SETUP

Denoting the trainable parameters of the aligner as $w \in \Omega$, where Ω is the parameter space, we associate each of the m preference dimensions with an alignment loss $l^{(j)} : \Omega \to \mathbb{R}_+$, $j \in [m]$, computed on the preference feedback of dimension j (a DPO loss in our implementation). For $\boldsymbol{l}^1, \boldsymbol{l}^2 \in \mathbb{R}^m$ we write $\boldsymbol{l}^1 \leq \boldsymbol{l}^2$ if $\boldsymbol{l}^2 - \boldsymbol{l}^1 \in \mathbb{R}^m_+$ (componentwise), and $\boldsymbol{l}^1 < \boldsymbol{l}^2$ if in addition $\boldsymbol{l}^1 \neq \boldsymbol{l}^2$ (Mahapatra & Rajan, 2020). Based on this, Pareto optimality of multi-objective preference alignment is defined as:

Definition 3.1 (Pareto optimality)**.** A solution w^* is *Pareto optimal* if no other solution $w' \in \Omega$ dominates w^*. The set of all Pareto optimal solutions is the *Pareto set (PS)* $\mathcal{P}$, and its image $\boldsymbol{l}(\mathcal{P}) \subset \mathbb{R}^m_+$ is the *Pareto front (PF)* $\mathcal{T}$. A solution w^* is *weakly Pareto optimal* if there is no $w' \in \Omega$ with $l^{(j)}(w') < l^{(j)}(w^*)$ for all $j \in [m]$ (Zhong et al., 2024).

In the collaborative learning setting, each client i holds its own preference feedback, which induces a client-specific loss vector $\boldsymbol{l}_i(w) = (l_i^{(1)}(w), \ldots, l_i^{(m)}(w))$, and expresses its trade-off through a preference vector $\boldsymbol{\lambda}_i \in S^m$, where $S^m := \{\boldsymbol{\lambda} \in \mathbb{R}^m_{++} : \sum_j \lambda^{(j)} = 1\}$; a larger $\lambda_i^{(j)}$ means client i tolerates a smaller loss on dimension j. Following the exact Pareto optimality of (Mahapatra & Rajan, 2020), the *preference-specific* Pareto optimal solutions of client i form the set

$$\mathcal{P}_i(\boldsymbol{\lambda}_i) = \left\{ w_i^* \in \mathcal{P}_i \;\middle|\; \lambda_i^{(1)} l_i^{(1)*} = \cdots = \lambda_i^{(j)} l_i^{(j)*} = \cdots = \lambda_i^{(m)} l_i^{(m)*} \right\}, \tag{1}$$

where $\mathcal{P}_i$ is the Pareto set of $\boldsymbol{l}_i$ and $l_i^{(j)*} := l_i^{(j)}(w_i^*)$. Geometrically, for any $w_i^* \in \mathcal{P}_i(\boldsymbol{\lambda}_i)$ the loss vector $\boldsymbol{l}_i(w_i^*)$ is the intersection of the Pareto front with the ray in the direction of $\boldsymbol{\lambda}_i^{-1} := (1/\lambda_i^{(1)}, \cdots, 1/\lambda_i^{(m)})$. Such an intersection need not exist for every $\boldsymbol{\lambda}_i$; in that case we seek the Pareto optimal solution closest to the ray. In this paper, we take the *weighted Chebyshev scalarization*

as the guiding objective: minimizing $\psi_i(w) = \max_{j\in[m]} \lambda_i^{(j)}(l_i^{(j)}(w) - l_i^{(j)*})$ drives the iterate to a weakly Pareto optimal point. For proof, see the Lemma H.30. Leveraging Equation 1, we reformulate the guiding objective as: minimizing $\psi_i(w) = \max_{j\in[m]} \lambda_i^{(j)} l_i^{(j)}(w)$.

To quantify how well the current loss vector $l(w_i)$ aligns with the client-specific preference vector $\boldsymbol{\lambda_i}$, we define the Preference Uniformity $\beta_i(l(w_i))$. For any point w_i in the parameter space Ω, $\beta_i(l(w_i)) = \sum_{j=1}^m \hat{l}_i^{(j)}(w_i) \log\left(\frac{\hat{l}_i^{(j)}(w_i)}{1/m}\right)$, where $\hat{l}_i^{(j)}(w_i) = \frac{\lambda_i^{(j)} l_i^{(j)}(w_i)}{\sum_{j'=1}^m \lambda_i^{(j')} l_i^{(j')}(w_i)}$ denotes the weighted and normalized loss for task j under preference $\boldsymbol{\lambda_i}$. The Preference Uniformity is exactly the Kullback–Leibler divergence $\mathrm{KL}(\hat{l}_i \| \frac{1}{m}\mathbf{1})$ and is nonnegative. It equals zero if and only if $\lambda_i^{(j)} l_i^{(j)}$ is identical for all tasks j, i.e., when the weighted losses are perfectly balanced according to the client's preference.

The direct differentiation of $\beta_i = \mathrm{KL}(\hat{l}_i \,\|\, \frac{1}{m}\mathbf{1})$ with respect to $l_i^{(j)}$ gives $\partial\beta_i/\partial l_i^{(j)} = \lambda_i^{(j)}\left(\log\left(\frac{\hat{l}_i^{(j)}}{1/m}\right) - \beta_i(l)\right) / \sum_{j'} \lambda_i^{(j')} l_i^{(j')}$. Therefore, we define the Preference Adjustment Vector $\boldsymbol{A_i} = (a_i^{(1)}, \ldots, a_i^{(m)})^\top$, whose j-th component $a_i^{(j)} = \lambda_i^{(j)}\left(\log\left(\frac{\hat{l}_i^{(j)}}{1/m}\right) - \beta_i(l)\right)$. Writing the normalizer $Z_i := \sum_{j'} \lambda_i^{(j')} l_i^{(j')}$, $\boldsymbol{A_i} = Z_i \nabla_l \beta_i$. Thus $\boldsymbol{A_i}$ points along the steepest increase of the Preference Uniformity, and moving against it reduces β_i toward the balanced ray. Let $g^{(j)} = \nabla_w l_j$ be the gradient of the jth preference dimension function at w, and $G \in \mathbb{R}^{n\times m}$ be the matrix having $g^{(j)}$ as its jth column. We can define the balancing direction $d_{bal} = G\boldsymbol{A_i} = Z_i \nabla_w \beta_i$, whose negative is the steepest-descent direction of the Preference Uniformity, i.e., along d with $d^\top d_{\mathrm{bal},i} > 0$. Accordingly, each client updates along a local direction $d_i = G_i \mu_i$ to decrease $\psi_i(w)$ with $\mu_i \in S^m$, a convex combination of its own per-dimension gradients chosen such that (a) the bottleneck weighted loss is not increased and (b) $d_i^\top d_{\mathrm{bal},i} > 0$ whenever $\beta_i > 0$. The concrete rule is given in Section 3.3.

3.2 Constructing Clusters: Bottleneck-Adjustment Clustering

To address the challenge of inter-client preference gradient conflict, APO employs Bottleneck-Adjustment Clustering, which partitions clients by two stages: bottleneck ordering and then refines the groups using adjustment vectors. The necessity of this design becomes clear once we examine how aggregation within a cluster can interfere with a client's local balancing direction.

Consider a cluster C_n whose clients aggregate their local directions with the FedAvg weights $\rho_i := N_i/N$, where N_i is the number of local samples of client i and $N = \sum_{i'\in C_n} N_{i'}$, yielding the shared direction $\bar{d}_{C_n} = \sum_{i\in C_n} \rho_i d_i$. And the $\bar{d}_{C_n}^\top d_{\mathrm{bal},i}$ can split as $\bar{d}_{C_n}^\top d_{\mathrm{bal},i} = d_i^\top d_{\mathrm{bal},i} + (\bar{d}_{C_n} - d_i)^\top d_{\mathrm{bal},i}$. The term $d_i^\top d_{\mathrm{bal},i}$ is what client i would obtain from its own update; the term $(\bar{d}_{C_n} - d_i)^\top d_{\mathrm{bal},i}$ is contributed by the other clients and has no fixed sign. As an illustration of possible interference, take $m = 2$ (e.g., helpfulness vs. harmlessness), let client i have $\boldsymbol{\lambda}_i = (1-\epsilon, \epsilon)$, and let the remaining clients concentrate near the opposite corner $\boldsymbol{\lambda}_{i'} \approx (\epsilon, 1-\epsilon)$. Their corrections aim at opposite ends of the Pareto front; if the gradient geometry at w then yields $d_{i'}^\top d_{\mathrm{bal},i} < 0$ for these clients, the aggregation effect grows with their total weight and can outweigh the local gain, so that $\bar{d}_{C_n}^\top d_{\mathrm{bal},i} < 0$ and β_i increases for all sufficiently small η.

Stage 1: coarse partition by bottleneck ordering. Therefore, to prevent such interference, we first eliminate differences in preference bottleneck structure among clients. The following sufficient condition makes this requirement precise.

Theorem 3.2 (Collaborative preference uniformity descent)**.** *Let Assumptions H.1–H.3 hold, let each client in C_n run τ local steps of size η before aggregation, and write $\gamma_i^* := d_i^\top d_{\mathrm{bal},i}$ for the local correction gain of property (b) and $\Delta_i := \|\bar{d}_{C_n} - d_i\|$. If for every $i \in C_n$*

$$\gamma_i^* \;>\; \Delta_i \,\|d_{\mathrm{bal},i}\| \;+\; L_0 B^2 \,\tau\, \eta, \tag{2}$$

which ensures that $\bar{d}_{C_n}^\top d_{\mathrm{bal},i}(w_0) > 0$ for all $i \in C_n$, then the shared direction is a simultaneous *balancing direction, and there exists $\eta_0 > 0$ such that a single global model update decreases* every *client's Preference Uniformity: $\beta_i\big(l(w_0 - \eta\,\bar{d}_{C_n})\big) \;\le\; \beta_i\big(l(w_0)\big), \forall\, i \in C_n,\ \forall\, \eta \in (0, \eta_0]$.*

We therefore perform a coarse partition by bottleneck ordering. At a fixed reference state w_{ref} (the aligner before collaborative training), each client evaluates its weighted losses and its adjustment vector $\boldsymbol{A}_i^{\mathrm{ref}} := \boldsymbol{A}_i(w_{\mathrm{ref}})$ on its own local feedback. Specifically, for each client we compute the weighted losses $\lambda_i^{(j)} l_i^{(j)}(w_{\mathrm{ref}})$, $j \in [m]$, and record the permutation Γ that sorts all m dimensions in descending order of weighted loss. Clients with identical orderings Γ form one coarse cluster. Nevertheless, sharing the bottleneck ordering does not make the aggregated direction automatically admissible, since the steps still depend on the client-specific adjustment vector $\boldsymbol{A_i}$.

Stage 2: refinement by adjustment vectors. According to the Theorem 3.2, the dispersion of the directions bounds the Δ_i, which we measure by the preference-conditioned direction dispersion of cluster C_n at parameter w,

$$H_{C_n}(w) = \sum_{i \in C_n} \rho_i \left\| d_i(w) - \bar{d}_{C_n}(w) \right\|^2, \tag{3}$$

where $d_i(w) = G_i(w)\mu_i$ and $\bar{d}_{C_n} = \sum_{i \in C_n} \rho_i \, d_i$. If all clients' LPs share a common non-degenerate optimal basis (Assumption H.6 in Appendix H.2), the multipliers μ_i are affine in $\boldsymbol{A}_i$ with constant κ_{LP}, and Proposition H.7 gives, for arbitrary aggregation weights ρ_i, $H_{C_n}(w) \le \kappa_{\mathrm{LP}}^2 \, \sigma_{\max}\big(G^\top G\big) \, V_{C_n}(\boldsymbol{A}^{\mathrm{ref}})$, where $V_{C_n}(\boldsymbol{A}^{\mathrm{ref}}) = \sum_{i \in C_n} \rho_i \big\| \boldsymbol{A}_i^{\mathrm{ref}} - \bar{\boldsymbol{A}}_{C_n}^{\mathrm{ref}} \big\|^2$ is the ρ-weighted within-cluster variance of the adjustment vectors around their weighted mean $\bar{\boldsymbol{A}}_{C_n}^{\mathrm{ref}} = \sum_{i \in C_n} \rho_i \boldsymbol{A}_i^{\mathrm{ref}}$. We therefore refine the cluster by adjustment vectors, aiming to reduce the direction dispersion $H_{C_n}(w)$. Specifically, within each Stage-1 cluster we apply hierarchical clustering with Euclidean distance on $\{\boldsymbol{A}_i^{\mathrm{ref}}\}$, and cut the dendrograms so that the total number of clusters equals a preset budget b (e.g., $b = 4$ on Fed-ChatbotPA and $b = 8$ on UltraFeedback).

3.3 Learning Shared Initializations: Preference-Aware Collaborative Training

Within each cluster, APO further addresses intra-client reference gradient conflict by combining preference gradient descent with controlled ascent to train the initialization. The procedure consists of two phases: collaborative Pareto set initialization learning, followed by adaptation-aware meta-training, which we describe in detail below.

Phase I: Collaborative Pareto Set Initialization Learning. Inspired by (Mahapatra & Rajan, 2020), we address the reference gradient conflict by combining the preference gradient descent with controlled ascent. More specifically, once preference balance is achieved ($\beta_i(\boldsymbol{l_i}) = 0$, so $\boldsymbol{A_i} = 0$, and $\boldsymbol{l_i}$ lies on the $\boldsymbol{\lambda_i}^{-1}$ ray), we switch to a pure descent scheme to refine the model toward the task itself: from (Désidéri, 2012), a common descent direction $d = G\mu$ in the convex hull of the gradients at w_i^t points toward the Pareto front, so we find the $d = G\mu$ whose inner product with the sum of all gradients $G\mathbf{1}$ is maximum. In addition, we define the index sets $J = \{j \mid \boldsymbol{A_i}^\top G_i^\top g_i^{(j)} > 0\}$, $\bar{J} = \{j \mid \boldsymbol{A_i}^\top G_i^\top g_i^{(j)} \le 0\}$ and the bottleneck set $J^* = \left\{ j \mid \lambda_i^{(j)} l_{i,b}^{(j)} = \max_{j'} \{\lambda_i^{(j')} l_{i,b}^{(j')}\} \right\}$ (with b-th batch), and impose descent constraints that permit a controlled trade-off. Combining the two modes through the indicator $\mathbb{1}_{\beta_i}$, we define the Pareto set initialization learning objective as

$$\mu_i^* = \arg\max_{\mu_i \in S^m} \mu_i^\top G_i^\top G_i \big(\boldsymbol{A_i} \mathbb{1}_{\beta_i} + \mathbf{1}(1 - \mathbb{1}_{\beta_i}) \big) \tag{4a}$$

$$\text{s.t.} \quad \mu_i^\top G_i^\top g_i^{(j)} \ge \boldsymbol{A_i}^\top G_i^\top g_i^{(j)} \mathbb{1}_J, \qquad \forall j \in \bar{J} \setminus J^* \tag{4b}$$

$$\mu_i^\top G_i^\top g_i^{(j)} \ge 0, \qquad \forall j \in J^* \tag{4c}$$

where $\mathbb{1}_J$ is the scalar indicator for a non-empty index set J. For the non-bottleneck objectives ($j \in \bar{J} \setminus J^*$), constraint Equation 4b requires $\mu_i^\top G_i^\top g_i^{(j)} \ge \boldsymbol{A_i}^\top G_i^\top g_i^{(j)} \mathbb{1}_J$; since this bound is non-positive, these objectives are allowed to degrade, but by no more than the balancing direction. For bottleneck objectives ($j \in J^*$), constraint Equation 4c requires $\mu_i^\top G_i^\top g_i^{(j)} \ge 0$, that guaranties that the most preference-violating objective must not worsen (its loss is non-increasing). In the following, we refer to the process with $\mathbb{1}_{\beta_i} = 0$ as uniform descent, and the process with $\mathbb{1}_{\beta_i} = 1$ as balancing descent.

Phase II: Adaptation-aware meta-training. Writing the model obtained in Phase I as w_{FedBsc}, inspired by (Jiang et al., 2019), we meta-train w_{FedBsc} on S-shot episodes so that it can be adapted to a client preference within K steps. Moreover, for a given preference $\boldsymbol{\lambda_i}$, we use the objective

$$\mu^* = \arg\max_{\mu \in S^m} d^T d_{bal} = \arg\max_{\mu \in S^m} \mu^T G^T G \boldsymbol{A_i} \quad \text{s.t. } (4b)\text{–}(4c). \tag{5}$$

to drive the iterate onto the $\boldsymbol{\lambda}_i^{-1}$ ray—it reduces the Preference Uniformity β_i. Phase II uses the same clients, the same local data (the S-shot episodes are subsamples of each client's own dataset), the same objective functions and the same FedAvg aggregation as Phase I. The complete initialization procedure (both phases) is summarized in Algorithm 1 in Appendix D.

3.4 Theoretical Guarantees: From Collaborative Convergence to Few-Shot Personalization

Collaborative Initialization Converges to Cluster-Local Optima. We analyze the collaborative initialization in three steps. First, we give a conditional bound on the preference uniformity of a client at which the balancing guarantee of Theorem 3.2 no longer applies, and an exact one-step bound (with second-order term) on the cluster losses along the shared update. Second, we define an admissible set $\mathcal{A}^t$ that contains each client's potential $l_i(w)$ values in an iteration t. Third, under an explicit *descent-margin* condition on the LP direction (Assumption H.19) and a trajectory hypothesis, we prove a contraction of each client's Chebyshev suboptimality and conclude that $\limsup_t \big(\psi_i(w^t) - \psi_i^*\big)$ is at most a fixed, horizon-independent radius ϵ_0 (asymptotically bounded; for every $\zeta > 0$ the suboptimality is eventually below $\epsilon_0 + \zeta$); if the iterates have a limit point w^*, it inherits the bound and lies in the limiting band $\mathcal{A}^\infty$ (a one-sided upper bound on each client's losses). The descent-margin condition is a Polyak–Łojasiewicz-type requirement on ψ_i along the actual LP direction; it is not implied by smoothness, bounded gradients, preference balance or the LP constraints, and we state it as an assumption rather than derive it.

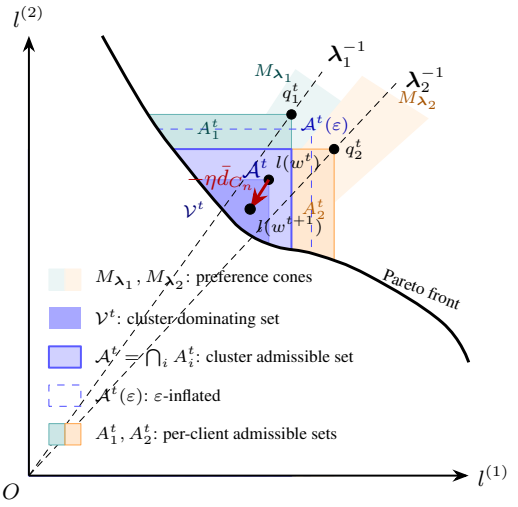


Figure 2: Cluster admissible set $\mathcal{A}^t$ for $m = 2$. Dashed lines are the preference rays $\boldsymbol{\lambda}_i^{-1}$. One shared FedAvg step $-\eta\, \bar{d}_{C_n}$ moves $l(w^t)$ to $l(w^{t+1})$, which under Assumption H.19 stays in every $A_i^t(\varepsilon_{i,t})$ (Theorem H.28); the boxes are nested across rounds up to this slack (Corollary H.29).

Along the collaborative iteration, Lemma H.16 decreases every balancing descent client's β_i in a cluster while the range condition Equation 47 holds. Partition the cluster into $n_+ := |C_n^+|$ balancing ($\beta_i > 0$) and $n_0 := |C_n^0|$ uniform ($\beta_i = 0$) clients, with block masses $\rho^+ := \sum_{i \in C_n^+} \rho_i$ and $\rho^0 := \sum_{i \in C_n^0} \rho_i$; Lemma H.22 then bounds each β_i by a tight $\bar{\beta}$, set by the preference-conditioned direction dispersion H_{C_n} and, via Equation 3 and Equation 45, by the cross-block pull $\rho^0 \|\bar{d}^0 - d_i\|$ in addition to preference heterogeneity. Thus $l(w)$ lies in the preference cone $M_{\boldsymbol{\lambda}_i} = \{l : \beta_i(l) \le \bar{\beta}\}$ around each $\boldsymbol{\lambda}_i^{-1}$ ray (Figure 2) for any client at which the range condition of Theorem 3.2 fails—a per-client, per-iterate fact; the convergence result below takes $\beta_i \le \bar{\beta}$ along the trajectory as a hypothesis. Lemma H.25 then gives an exact one-step bound on each loss and on ψ_i along the shared update, including the second-order term $\frac{L_0 B^2}{2}\eta^2$.

For client i at iterate w^t let $O_i \subseteq \mathbb{R}_+^m$ be its attainable objective set and $\le$ the componentwise order. The $\psi_i^t := \max_j \lambda_i^{(j)}\, l_i^{(j)}(w^t)$, we define the client ray point $q_i^t := \psi_i^t \Big(\frac{1}{\lambda_i^{(1)}}, \ldots, \frac{1}{\lambda_i^{(m)}}\Big)$, the client dominating set $V_i^t := \{l \in O_i : l \le l_i(w^t)\}$, and the client admissible set $A_i^t := \{l \in O_i : l \le q_i^t\}$. The sets A_i^t live in each client's own objective space O_i; a single shared model w, however, must serve every client at once. We therefore lift the construction to *parameter space*, defining for the cluster dominating set $\mathcal{V}^t := \bigcap_{i \in C_n} \{\, w : \; l_i(w) \le l_i(w^t) \,\}$, and the ($\varepsilon$-inflated) cluster admissible

set $\mathcal{A}^t(\varepsilon) := \bigcap_{i\in C_n}\{\, w :\ l_i(w) \le q_i^t + \varepsilon\mathbf{1}\,\}$. And we abbreviate $\mathcal{A}^t := \mathcal{A}^t(0) = \bigcap_{i\in C_n}\{w : l_i(w) \le q_i^t\}$. Thus $w \in \mathcal{A}^t(\varepsilon)$ iff $l_i(w) \in A_i^t(\varepsilon)$ for *every* client, i.e. the cluster admissible set is the set of shared parameters whose induced losses fall in every client's inflated admissible box simultaneously (the intersection is exactly the cluster wedge of Figure 2).

Based on the Lemmas H.22, H.25, and the cluster admissible set analysis above, we can show:

Theorem 3.3 (Cluster solution approximates every local optimum under a descent-margin condition)**.** *Consider the synchronous shared iteration $w^{t+1} = w^t - \eta\bar{d}_{C_n}(w^t)$ (one local LP step per round) under Assumptions H.1–H.3 and the descent-margin condition of Assumption H.19: there are $c_d > 0$ and tolerances $b_i \ge 0$ such that on a region $\mathcal{N}$, whenever $\beta_i \le \bar{\beta}$, $\min_{j\in J^*_{i,\varepsilon}(w)} \lambda_i^{(j)} d_i(w)^\top g_i^{(j)}(w) \ge c_d\,(\psi_i(w) - \psi_i^*) - b_i$. Let $0 < \eta \le 1/c_d$ and suppose that for all $t \ge t_0$ and $i \in C_n$, $w^t \in \mathcal{N}$, $\beta_i(w^t) \le \bar{\beta}$ and $\|\bar{d}_{C_n}(w^t) - d_i(w^t)\| \le \bar{\Delta}_i$. Then, with $\epsilon_{0,i} := \frac{\|\boldsymbol{\lambda}_i\|_\infty}{c_d}\Big(B\,\bar{\Delta}_i + \frac{L_0 B^2}{2}\,\eta\Big) + \frac{b_i}{c_d}$, we obtain the contraction $\psi_i(w^T) - \psi_i^* \le (1 - \eta c_d)^{T-t_0}\big(\psi_i(w^{t_0}) - \psi_i^*\big) + \epsilon_{0,i}$. Consequently, $\limsup_{T\to\infty}\big(\psi_i(w^T) - \psi_i^*\big) \le \epsilon_{0,i}$, and any limit point w^* satisfies $l_i(w^*) \le q_i^* + \Big(\frac{\epsilon_{0,i}}{\min_j \lambda_i^{(j)}}\Big)\mathbf{1}$ componentwise, where $q_i^* = \psi_i^*\boldsymbol{\lambda}_i^{-1}$ denotes the ray point of client i's local optimum. We write $\epsilon_0 := \max_{i\in C_n}\epsilon_{0,i}$.*

The radius $\epsilon_{0,i}$ is horizon-independent and governed by $\bar{\Delta}_i \le \sqrt{\sup_t H_{C_n}(w^t)/\rho_{\min}}$ (the quantity Bottleneck-Adjustment Clustering is designed to reduce), by η, and by b_i/c_d. The statement covers one local step per round ($\tau = 1$; $K = 1$ for Phase II). Since Phase II runs the same shared iteration on the same data with the balancing LP and effective step size $\eta_{\mathrm{II}} = \alpha\eta_{\mathrm{meta}} \le 1/c_d$, the band also covers w_{II} with η replaced by $\max(\eta, \eta_{\mathrm{II}})$ (Corollary H.32). Consequently, for $\hat{w} \in \{w_{\mathrm{I}}, w_{\mathrm{II}}\}$ and every in-cluster client (Corollary H.37),

$$\psi_i(\hat{w}) - \psi_i^* \ \le\ \bar{\epsilon}_0 + \delta_\lambda\,\|l(\hat{w})\|_\infty + \delta_\lambda\,L^*, \qquad \bar{\epsilon}_0 := \tfrac{1}{|C_n|}\textstyle\sum_{i\in C_n}\epsilon_{0,i}, \tag{6}$$

where L^* bounds $\|l(w_{i'}^*)\|_\infty$ over all optima being compared. Full statements, proofs and qualifications are in Appendix H.4–H.5.

Few-Shot Personalization Gap from the Shared Initialization. Consider a new client with preference $\boldsymbol{\lambda}_{\mathrm{new}}$ assigned to cluster C_n by its adjustment vector, with scarce local data and no participation in collaborative training. It adapts the shared starting point w_0 (in Algorithm 1, $w_0 = w_{\mathrm{II}}$) for K steps with Equation 5; we compare the result with a locally trained model that uses the same objective and preference but has sufficient data. The analysis is for an *idealized* preconditioned stochastic adaptation with fresh samples (S per step, $K \cdot S$ in total; Assumptions H.39–H.41).

Theorem 3.4 (Data deficiency adaptation gap δ_0)**.** *Under Assumptions H.1–H.3, H.34, H.39 (local L_0-smoothness and a PL inequality with constant μ_{PL} on the adaptation region), H.40 and H.41 (adaptive step sizes in $[p_-, p_+]$), with $0 < \alpha \le p_-/(L_0 p_+^2)$ and $\|\boldsymbol{\lambda}_{\mathrm{new}} - \boldsymbol{\lambda}_i\| \le \delta_\lambda$ for all $i \in C_n$, the K-step adaptation started from any w_0 in the region of Assumption H.39 satisfies, in expectation over the sampled directions, $\mathbb{E}\big[\psi_{\mathrm{new}}(w_K) - \psi_{\mathrm{new}}^*\big] \le (1 - \alpha\mu_{\mathrm{PL}}p_-)^K\,\Delta_0 + \frac{L_0\alpha p_+^2}{2\mu_{\mathrm{PL}}p_-}\cdot\frac{\sigma_0^2}{S} =: \delta_0$, where ψ_{new}^* is the optimal value attained by local training with sufficient data.*

For $w_0 \in \{w_{\mathrm{I}}, w_{\mathrm{II}}\}$, Equation 6 gives $\Delta_0 := \psi_{\mathrm{new}}(w_0) - \psi_{\mathrm{new}}^* \le \bar{\epsilon}_0 + \delta_\lambda(\|l(w_0)\|_\infty + L^*)$, so the few-shot gap is expressed in the collaborative-training quantities $\bar{\Delta}_i, \delta_\lambda, \eta, \eta_{\mathrm{II}}$ and b_i/c_d, and vanishes as the cluster tightens only in the strict form $b_i = 0$ (in Appendix H.6).

4 Experiments

4.1 Experimental Setup

Datasets and Evaluation Metrics We evaluate our method on the helpful-assistant task using Fed-ChatbotPA (Ye et al., 2024a) (helpfulness, safety) and UltraFeedback (Cui et al., 2023) (helpfulness, honesty, truthfulness), scoring responses with ArmoRM-Llama3-8B-v0.1 (Wang et al., 2024b) by preference-weighted normalized score, hypervolume (HV), and inverted generational distance (IGD). Collaborative learning under data deficiency and metric definitions are in Appendices B and F.

Table 1: Cluster-level Score, hypervolume (HV), and inverted generational distance (IGD) on Fed-ChatbotPA (FCPA; 4 clusters) and UltraFeedback (UF; 8 clusters). Higher Score/HV and lower IGD are better. Dark bold cells mark the best values and light cells the second-best.

		Baselines																		**Ours**		
		Base			RIC			Rewarded Soup			FedAvg			FSPO			DITTO			APO		
Dataset	Cluster	Score	HV	IGD	Score	HV	IGD	Score	HV	IGD	Score	HV	IGD	Score	HV	IGD	Score	HV	IGD	Score	HV	IGD
Llama-3.2-3B-Instruct																						
FCPA	C_0	0.70	0.52	0.18	0.24	0.11	0.72	0.78	0.60	0.12	0.78	0.58	0.12	0.80	0.65	0.10	0.53	0.33	0.39	**0.88**	**0.71**	**0.07**
	C_1	0.72	0.53	0.15	0.43	0.20	0.54	0.75	0.53	0.17	0.76	0.57	0.11	0.79	0.62	0.07	0.52	0.28	0.42	**0.81**	**0.67**	**0.06**
	C_2	0.70	0.47	0.14	0.50	0.14	0.54	0.57	0.39	0.25	0.75	0.49	0.12	0.74	**0.56**	**0.07**	0.56	0.24	0.40	**0.80**	0.52	0.12
	C_3	0.74	0.53	0.13	0.43	0.14	0.58	0.70	0.50	0.19	0.77	0.56	0.10	0.79	**0.62**	**0.06**	0.58	0.27	0.39	**0.82**	**0.62**	0.09
	Avg.	0.72	0.51	0.15	0.40	0.15	0.60	0.70	0.51	0.18	0.76	0.55	0.11	0.78	0.61	**0.08**	0.55	0.28	0.40	**0.83**	**0.63**	**0.08**
UF	C_0	0.53	0.20	0.34	0.35	0.03	0.79	0.66	0.39	0.11	0.73	0.40	0.15	0.58	0.25	0.25	0.33	0.09	0.62	**0.78**	**0.48**	**0.10**
	C_1	0.54	0.18	0.37	0.39	0.05	0.73	0.67	0.34	0.16	0.74	0.40	0.11	0.59	0.23	0.29	0.31	0.06	0.72	**0.79**	**0.48**	**0.04**
	C_2	0.54	0.17	0.41	0.33	0.03	0.82	0.72	0.39	0.12	0.74	0.43	0.09	0.63	0.25	0.28	0.48	0.13	0.52	**0.79**	**0.52**	**0.04**
	C_3	0.45	0.09	0.58	0.24	0.02	0.95	0.72	0.40	0.15	0.74	0.42	0.11	0.58	0.20	0.37	0.49	0.13	0.51	**0.79**	**0.51**	**0.06**
	C_4	0.58	0.19	0.30	0.22	0.01	0.87	**0.75**	**0.40**	0.05	0.66	0.28	0.16	0.65	0.26	0.18	0.52	0.11	0.46	0.74	**0.40**	**0.04**
	C_5	0.57	0.18	0.39	0.17	0.01	1.07	0.75	0.37	0.12	0.73	0.40	0.15	0.63	0.24	0.29	0.46	0.09	0.59	**0.76**	**0.43**	**0.10**
	C_6	0.56	0.16	0.51	0.29	0.03	0.92	0.80	0.46	0.14	0.81	0.51	0.10	0.67	0.27	0.34	0.52	0.14	0.63	**0.84**	**0.58**	**0.04**
	C_7	0.61	0.19	0.40	0.18	0.01	1.04	0.76	0.33	0.19	0.80	0.45	**0.06**	0.68	0.27	0.28	0.60	0.17	0.50	**0.82**	**0.47**	**0.06**
	Avg.	0.55	0.17	0.41	0.27	0.02	0.90	0.73	0.38	0.13	0.74	0.41	0.12	0.63	0.25	0.28	0.47	0.12	0.57	**0.79**	**0.48**	**0.06**
Qwen2.5-3B-Instruct																						
FCPA	C_0	0.71	0.45	0.23	0.16	0.03	0.93	0.77	0.55	0.18	0.85	0.62	0.07	0.84	0.58	0.10	0.81	0.56	0.12	**0.88**	**0.65**	**0.05**
	C_1	0.69	0.46	0.16	0.22	0.05	0.80	0.74	0.54	0.10	0.77	0.56	0.06	0.77	0.55	0.08	0.76	0.53	0.08	**0.79**	**0.60**	**0.05**
	C_2	0.58	0.39	0.19	0.38	0.05	0.73	0.69	0.51	0.12	0.66	0.49	0.10	0.64	0.51	0.10	0.64	0.50	0.10	**0.70**	**0.54**	**0.06**
	C_3	0.68	0.47	0.16	0.13	0.01	0.95	0.75	0.57	0.11	0.75	0.56	**0.08**	0.75	**0.59**	0.09	0.71	0.52	0.13	**0.76**	0.57	0.09
	Avg.	0.67	0.44	0.19	0.22	0.04	0.85	0.74	0.54	0.12	0.76	0.56	0.08	0.75	0.56	0.09	0.73	0.53	0.11	**0.78**	**0.59**	**0.06**
UF	C_0	0.36	0.05	0.67	0.48	0.09	0.54	0.37	0.10	0.58	0.64	0.25	0.18	0.41	0.07	0.59	0.33	0.04	0.75	**0.74**	**0.41**	**0.00**
	C_1	0.41	0.08	0.51	0.61	0.24	0.16	0.29	0.05	0.67	0.63	0.26	0.17	0.36	0.05	0.60	0.35	0.05	0.62	**0.66**	**0.29**	**0.13**
	C_2	0.34	0.05	0.58	0.50	0.12	0.33	0.41	0.10	0.44	0.59	0.24	0.17	0.38	0.07	0.52	0.25	0.02	0.74	**0.65**	**0.31**	**0.08**
	C_3	0.37	0.06	0.64	0.20	0.01	0.94	0.36	0.08	0.59	0.69	0.38	0.09	0.40	0.06	0.63	0.31	0.03	0.79	**0.74**	**0.45**	**0.01**
	C_4	0.36	0.06	0.63	0.36	0.04	0.75	0.24	0.02	0.88	0.65	0.31	0.14	0.37	0.07	0.63	0.32	0.05	0.71	**0.73**	**0.43**	**0.00**
	C_5	0.32	0.06	0.74	0.54	0.27	0.33	0.44	0.09	0.62	0.59	0.24	0.29	0.38	0.10	0.64	0.33	0.06	0.71	**0.69**	**0.38**	**0.11**
	C_6	0.42	0.10	0.48	0.35	0.05	0.64	0.38	0.06	0.65	0.61	0.27	0.16	0.37	0.08	0.56	0.31	0.06	0.66	**0.70**	**0.40**	**0.01**
	C_7	0.39	0.09	0.21	0.34	0.07	0.33	0.45	0.07	0.24	0.47	**0.17**	0.12	0.35	0.08	0.25	0.32	0.07	0.29	**0.49**	**0.17**	**0.10**
	Avg.	0.37	0.07	0.56	0.42	0.11	0.50	0.37	0.07	0.58	0.61	0.26	0.17	0.38	0.07	0.55	0.31	0.05	0.66	**0.67**	**0.36**	**0.05**

Baselines We compare against six baselines for personalized multi-objective alignment under data deficiency. (1) *Base model*; (2) *RIC* (Yang et al., 2024b); (3) *Rewarded Soup* (Rame et al., 2023); (4) *FSPO* (Singh et al., 2026); (5) *DITTO* (Shaikh et al., 2025); (6) *FedAvg* (McMahan et al., 2017). Detailed descriptions of all baselines are provided in Appendix E.

Models and computation environment. We use the pretrained base model Llama-3.2-3B-Instruct (Meta AI, 2024) and Qwen2.5-3B-Instruct (Team, 2024). And trained them on the NVIDIA RTX PRO 6000 Blackwell (96 GB GDDR7) and NVIDIA A100 80GB (SXM4).

4.2 Main Results

APO outperforms existing baselines under data deficiency. Table 1 reports cluster-level results after adaptation with only 20 support samples. Across the 24 model–dataset–cluster combinations, APO achieves the highest Score on 23 clusters, the highest or tied-highest HV on 22, and the lowest or tied-lowest IGD on 21, so its gains are not confined to preference-weighted utility: higher HV and lower IGD show that the learned solutions span a broader trade-off region and more closely cover the empirical Pareto front. Average results are consistent across both datasets and base models. Relative to the best baseline for each metric, Score/HV/IGD changes from $0.78/0.61/0.08$ to $0.83/0.63/0.08$ on Fed-ChatbotPA with Llama, from $0.74/0.41/0.12$ to $0.79/0.48/0.06$ on UltraFeedback with Llama, and with Qwen from $0.76/0.56/0.08$ to $0.78/0.59/0.06$ on Fed-ChatbotPA and from $0.61/0.26/0.17$ to $0.67/0.36/0.05$ on UltraFeedback. The tied IGD on Llama with Fed-ChatbotPA is the only average-level result that is not a strict improvement over the strongest baseline. Compared with FedAvg, which uses the same collaborative Pareto set initialization learning phase, our method improves all three average metrics in all four model–dataset settings, supporting the combined benefit of learning preference-aware, adaptation-ready initializations rather than attributing gains to preference-aware initialization alone. Among the remaining baselines, RIC falls below the untuned base model in three

of the four average settings, while Rewarded Soup and FSPO are competitive in selected settings but do not remain consistently strong across both datasets and model families.

4.3 Ablation Study

Effect of communication rounds. Figure 3 varies the communication round at which the collaborative initializer is selected (150–190 on UltraFeedback and 140–200 on Fed-ChatbotPA) and reports the gain Δ of APO over FedAvg after the same 20-sample adaptation. Of the 56 cluster–round comparisons, 52 are positive, one is zero, and only three are negative; the negative deviations are small and never below -1.0pp. The gains average $+1.8$pp on UltraFeedback and $+3.3$pp on Fed-ChatbotPA, reaching $+5.3$pp for C_4 at round 170 and $+8.4$pp for C_0 at round 200, respectively. Moreover, the cluster-average gain remains positive at every evaluated round on both datasets, showing that the improvement is not tied to a single checkpoint. On Fed-ChatbotPA, the gains are particularly pronounced for C_0 and C_1. This dataset contains substantially more preference feedback for helpfulness than for safety, while the preference vectors of these two clusters assign greater weight to helpfulness. The adaptation-aware meta-training can therefore exploit a stronger training signal for the objective emphasized by these clients, yielding larger improvements over the single FedAvg initializer after limited local adaptation.

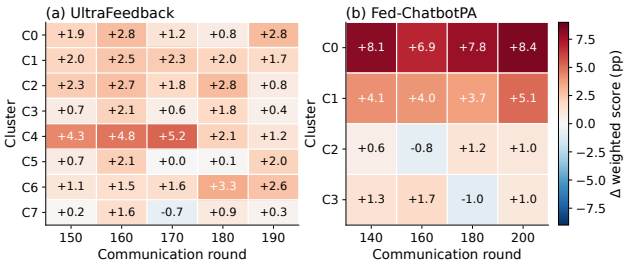


Figure 3: Communication-round ablation on UltraFeedback and Fed-ChatbotPA with Llama-3.2-3B-Instruct; each cell is APO's normalized weighted-score gain over FedAvg, in percentage points.

Effect of the clustering criterion. Table 2 isolates the clustering criterion by replacing our Bottleneck-Adjustment Clustering (BAC) with hierarchical clustering (HC) on the raw preference vectors while keeping all subsequent training and adaptation steps fixed. BAC achieves a higher average score on six of the eight clusters, raising the overall average from 0.70 to 0.72; its best score likewise improves from 0.77 to 0.79. The largest average gains occur on C_6 and C_7 (both $+0.07$), and BAC attains a higher best score on five clusters and ties on C_5. HC remains better on the test users of C_0 and C_2, where assignment by the preference vector can place a user in a cluster whose preference bottleneck happens to coincide with or similar to the user's own. This local coincidence does not persist across clusters, however, as BAC achieves stronger aggregate performance by grouping clients according to both bottleneck and adjustment compatibility.

Table 2: Clustering ablation on UltraFeedback with Llama-3.2-3B-Instruct. Cluster-level normalized weighted scores compare Hierarchical Clustering (HC) with our Bottleneck-Adjustment Clustering (BAC). Δ = BAC − HC; higher is better, and bold marks the better result. Scores and differences are independently rounded to two decimals.

	Average			Best		
Cluster	HC	BAC (Ours)	Δ	HC	BAC (Ours)	Δ
C_0	**0.76**	0.72	-0.04	**0.81**	0.78	-0.04
C_1	0.70	**0.73**	+0.04	0.77	**0.79**	+0.02
C_2	**0.75**	0.72	-0.03	**0.81**	0.79	-0.02
C_3	0.68	**0.70**	+0.02	0.75	**0.79**	+0.03
C_4	0.63	**0.68**	+0.05	0.70	**0.74**	+0.04
C_5	0.70	**0.70**	+0.01	0.76	**0.76**	+0.00
C_6	0.71	**0.78**	+0.07	0.78	**0.84**	+0.06
C_7	0.68	**0.76**	+0.07	0.76	**0.82**	+0.05
Avg.	0.70	**0.72**	+0.02	0.77	**0.79**	+0.02

5 Conclusion

We studied collaborative personalized multi-objective preference alignment under scarce local feedback, where heterogeneous preferences create inter-client conflicts and competing objectives create intra-client conflicts, and introduced APO, combining Bottleneck-Adjustment Clustering, preference gradient descent with controlled ascent, and adaptation-aware meta-training. On Fed-ChatbotPA and UltraFeedback with two 3B backbones, the resulting aligners consistently improve preference-weighted scores and generally provide better Pareto-front coverage using only 20 local examples.

AI USE STATEMENT

In this work, we used generative AI tools to assist in the writing of proofs, refine hypotheses, implement methods, assist with translation, clean and reformat dataset, and support qualitative and thematic data analysis. We have not used generative AI tools to generate synthetic data sets, help develop theoretical models or conceptual frameworks, formulate mathematical claims, provide critical ingredients for proving mathematical claims, propose hypotheses, design or provide feedback on research methodology or experiments and interpret results. Additionally, we used generative AI tools to create or modify scientific figures or images, create or edit software code. We have reviewed all AI-assisted work. For example, we verified LLM-generated code, etc. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

We identify no ethical concerns associated with this work. The study does not involve human participants or introduce issues related to dataset release, harmful applications or methodologies, conflicts of interest or sponsorship, discrimination or fairness, privacy or security, legal compliance, or research integrity.

REPRODUCIBILITY STATEMENT

The source code necessary to reproduce our results will be made publicly available upon acceptance. Complete proofs of all theoretical results are provided in the appendix. The appendix also describes the datasets and provides full details of the data-processing procedures used in our experiments.

REFERENCES

Daiwei Chen, Yi Chen, Aniket Rege, and Ramya Korlakai Vinayak. PAL: Pluralistic alignment framework for learning from heterogeneous preferences. In *International Conference on Learning Representations*, 2025.

Eng Ung Choo and Derek R Atkins. Proper efficiency in nonconvex multicriteria programming. *Mathematics of Operations Research*, 8(3):467–470, 1983.

Ganqu Cui, Lifan Yuan, Ning Ding, Guanming Yao, Wei Zhu, Yuan Ni, Guotong Xie, Zhiyuan Liu, and Maosong Sun. Ultrafeedback: Boosting language models with high-quality feedback, 2023.

Jean-Antoine Désidéri. Multiple-gradient descent algorithm (mgda) for multiobjective optimization. *Comptes Rendus Mathematique*, 350(5-6):313–318, 2012.

Yi Dong, Zhilin Wang, Makesh Narsimhan Sreedhar, Xianchao Wu, and Oleksii Kuchaiev. SteerLM: Attribute conditioned SFT as an (user-steerable) alternative to RLHF. In *Findings of the Association for Computational Linguistics: EMNLP 2023*, pp. 11275–11288, 2023.

Chelsea Finn, Pieter Abbeel, and Sergey Levine. Model-agnostic meta-learning for fast adaptation of deep networks. In *International conference on machine learning*, pp. 1126–1135. PMLR, 2017.

Jörg Fliege and Benar Fux Svaiter. Steepest descent methods for multicriteria optimization. *Mathematical methods of operations research*, 51(3):479–494, 2000.

Avishek Ghosh, Jichan Chung, Dong Yin, and Kannan Ramchandran. An efficient framework for clustered federated learning. *Advances in Neural Information Processing Systems*, 33:19586–19597, 2020.

Google. Build with gemini, 2026. URL `https://gemini.google.com/app`. Accessed: 2026-9-15.

Ping Guo, Tiantian Zhang, Xi Lin, Xiang Li, Zhi-Ri Tang, and Qingfu Zhang. Few-for-many personalized federated learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 17515–17524, 2026.

Yiju Guo, Ganqu Cui, Lifan Yuan, Ning Ding, Zexu Sun, Bowen Sun, Huimin Chen, Ruobing Xie, Jie Zhou, Yankai Lin, et al. Controllable preference optimization: Toward controllable multi-objective alignment. *arXiv preprint arXiv:2402.19085*, 2024.

Zeou Hu, Kiarash Shaloudegi, Guojun Zhang, and Yaoliang Yu. Federated learning meets multi-objective optimization. *IEEE Transactions on Network Science and Engineering*, 9(4):2039–2051, 2022.

Joel Jang, Seungone Kim, Bill Yuchen Lin, Yizhong Wang, Jack Hessel, Luke Zettlemoyer, Hannaneh Hajishirzi, Yejin Choi, and Prithviraj Ammanabrolu. Personalized soups: Personalized large language model alignment via post-hoc parameter merging. *arXiv preprint arXiv:2310.11564*, 2023.

Jiaming Ji, Boyuan Chen, Hantao Lou, Donghai Hong, Borong Zhang, Xuehai Pan, Tianyi Alex Qiu, Juntao Dai, and Yaodong Yang. Aligner: Efficient alignment by learning to correct. *Advances in Neural Information Processing Systems*, 37:90853–90890, 2024.

Yihan Jiang, Jakub Konečnỳ, Keith Rush, and Sreeram Kannan. Improving federated learning personalization via model agnostic meta learning. *arXiv preprint arXiv:1909.12488*, 2019.

Xinyu Li, Ruiyang Zhou, Zachary C. Lipton, and Liu Leqi. Personalized language modeling from personalized human feedback. *arXiv preprint arXiv:2402.05133*, 2024.

Bo Liu, Xingchao Liu, Xiaojie Jin, Peter Stone, and Qiang Liu. Conflict-averse gradient descent for multi-task learning. *Advances in Neural Information Processing Systems*, 34:18878–18890, 2021.

Guodong Long, Tao Shen, Jing Jiang, Michael Blumenstein, et al. Dual-personalizing adapter for federated foundation models. *Advances in Neural Information Processing Systems*, 37:39409–39433, 2024.

Debabrata Mahapatra and Vaibhav Rajan. Multi-task learning with user preferences: Gradient descent with controlled ascent in pareto optimization. In *International Conference on Machine Learning*, pp. 6597–6607. PMLR, 2020.

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

Meta AI. Llama 3.2 3b instruct. `https://huggingface.co/meta-llama/Llama-3.2-3B-Instruct`, 2024. Model card and weights.

Aviv Navon, Aviv Shamsian, Idan Achituve, Haggai Maron, Kenji Kawaguchi, Gal Chechik, and Ethan Fetaya. Multi-task learning as a bargaining game. In *International Conference on Machine Learning*, pp. 16428–16446. PMLR, 2022.

Duong Nguyen, Nghia Hoang, Thanh Trung Huynh, Quoc Viet Hung Nguyen, and Phi Le Nguyen. Learning reconfigurable representations for multimodal federated learning with missing data. *Advances in Neural Information Processing Systems*, 38:28808–28845, 2026.

Alex Nichol, Joshua Achiam, and John Schulman. On first-order meta-learning algorithms. *arXiv preprint arXiv:1803.02999*, 2018.

Fatemeh Nourzad, Amirhossein Roknilamouki, Eylem Ekici, Jia Liu, and Ness B. Shroff. FIRM: Federated in-client regularized multi-objective alignment for large language models. *arXiv preprint arXiv:2511.16992*, 2025.

OpenAI. Gpt-6, 2026. URL `https://openai.com/index/gpt-6-astra/`.

Sriyash Poddar, Yanming Wan, Hamish Ivison, Abhishek Gupta, and Natasha Jaques. Personalizing reinforcement learning from human feedback with variational preference learning. *Advances in Neural Information Processing Systems*, 37, 2024.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Christopher D Manning, Stefano Ermon, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. *Advances in neural information processing systems*, 36:53728–53741, 2023.

Alexandre Rame, Guillaume Couairon, Corentin Dancette, Jean-Baptiste Gaya, Mustafa Shukor, Laure Soulier, and Matthieu Cord. Rewarded soups: towards pareto-optimal alignment by interpolating weights fine-tuned on diverse rewards. *Advances in Neural Information Processing Systems*, 36:71095–71134, 2023.

Felix Sattler, Klaus-Robert Müller, and Wojciech Samek. Clustered federated learning: Model-agnostic distributed multitask optimization under privacy constraints. *IEEE Transactions on Neural Networks and Learning Systems*, 32(8):3710–3722, 2021.

Omar Shaikh, Michelle Lam, Joey Hejna, Yijia Shao, Hyundong Cho, Michael Bernstein, and Diyi Yang. Aligning language models with demonstrated feedback. In *International Conference on Learning Representations*, volume 2025, pp. 20498–20525, 2025.

Ruizhe Shi, Yifang Chen, Yushi Hu, Alisa Liu, Hanna Hajishirzi, Noah A Smith, and Simon S Du. Decoding-time language model alignment with multiple objectives. *Advances in Neural Information Processing Systems*, 37:48875–48920, 2024.

Anikait Singh, Sheryl Hsu, Kyle Hsu, Eric Mitchell, Stefano Ermon, Tatsunori Hashimoto, Archit Sharma, and Chelsea Finn. Fspo: Few-shot optimization of synthetic preferences effectively personalizes to real users. In *International Conference on Learning Representations*, volume 2026, pp. 109496–109524, 2026.

Canh T Dinh, Nguyen Tran, and Josh Nguyen. Personalized federated learning with Moreau envelopes. *Advances in Neural Information Processing Systems*, 33:21394–21405, 2020.

Qwen Team. Qwen2.5: A party of foundation models, September 2024. URL `https://qwenlm.github.io/blog/qwen2.5/`.

Haoxiang Wang, Yong Lin, Wei Xiong, Rui Yang, Shizhe Diao, Shuang Qiu, Han Zhao, and Tong Zhang. Arithmetic control of LLMs for diverse user preferences: Directional preference alignment with multi-objective rewards. In *Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics*, pp. 8642–8655, 2024a.

Haoxiang Wang, Wei Xiong, Tengyang Xie, Han Zhao, and Tong Zhang. Interpretable preferences via multi-objective reward modeling and mixture-of-experts. In *EMNLP*, 2024b.

Feijie Wu, Xiaoze Liu, Haoyu Wang, Xingchen Wang, Lu Su, and Jing Gao. Towards federated rlhf with aggregated client preference for llms. *arXiv preprint arXiv:2407.03038*, 2024.

Binqian Xu, Xiangbo Shu, Haiyang Mei, Zechen Bai, Basura Fernando, Mike Zheng Shou, and Jinhui Tang. Dofit: Domain-aware federated instruction tuning with alleviated catastrophic forgetting. *Advances in Neural Information Processing Systems*, 37:85846–85866, 2024.

Haibo Yang, Zhuqing Liu, Jia Liu, Chaosheng Dong, and Michinari Momma. Federated multi-objective learning. *Advances in neural information processing systems*, 36:39602–39625, 2023.

Kailai Yang, Zhiwei Liu, Qianqian Xie, Jimin Huang, Tianlin Zhang, and Sophia Ananiadou. Metaaligner: Towards generalizable multi-objective alignment of language models. *Advances in Neural Information Processing Systems*, 37:34453–34486, 2024a.

Rui Yang, Xiaoman Pan, Feng Luo, Shuang Qiu, Han Zhong, Dong Yu, and Jianshu Chen. Rewards-in-context: Multi-objective alignment of foundation models with dynamic preference adjustment. *arXiv preprint arXiv:2402.10207*, 2024b.

Rui Ye, Rui Ge, Xinyu Zhu, Jingyi Chai, Du Yaxin, Yang Liu, Yanfeng Wang, and Siheng Chen. Fedllm-bench: Realistic benchmarks for federated learning of large language models. *Advances in Neural Information Processing Systems*, 37:111106–111130, 2024a.

Rui Ye, Wenhao Wang, Jingyi Chai, Dihan Li, Zexi Li, Yinda Xu, Yaxin Du, Yanfeng Wang, and Siheng Chen. Openfedllm: Training large language models on decentralized private data via federated learning. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 6137–6147, 2024b.

Tianhe Yu, Saurabh Kumar, Abhishek Gupta, Sergey Levine, Karol Hausman, and Chelsea Finn. Gradient surgery for multi-task learning. *Advances in Neural Information Processing Systems*, 33: 5824–5836, 2020.

Yicheng Zhang, Zhen Qin, Zhaomin Wu, Jian Hou, and Shuiguang Deng. Personalized federated fine-tuning for llms via data-driven heterogeneous model architectures. *arXiv preprint arXiv:2411.19128*, 2024a.

Yonggang Zhang, Zhiqin Yang, Xinmei Tian, Nannan Wang, Tongliang Liu, and Bo Han. Robust training of federated models with extremely label deficiency. In *ICLR 2024*, 2024b.

Siyan Zhao, John Dang, and Aditya Grover. Group preference optimization: Few-shot alignment of large language models. In *International Conference on Learning Representations*, 2024.

Lianmin Zheng, Wei-Lin Chiang, Ying Sheng, Siyuan Zhuang, Zhanghao Wu, Yonghao Zhuang, Zi Lin, Zhuohan Li, Dacheng Li, Eric. P Xing, Hao Zhang, Joseph E. Gonzalez, and Ion Stoica. Judging llm-as-a-judge with mt-bench and chatbot arena, 2023.

Yifan Zhong, Chengdong Ma, Xiaoyuan Zhang, Ziran Yang, Haojun Chen, Qingfu Zhang, Siyuan Qi, and Yaodong Yang. Panacea: Pareto alignment via preference adaptation for llms. *Advances in Neural Information Processing Systems*, 37:75522–75558, 2024.

Zhanhui Zhou, Jie Liu, Chao Yang, Jing Shao, Yu Liu, Xiangyu Yue, Wanli Ouyang, and Yu Qiao. Beyond one-preference-fits-all alignment: Multi-objective direct preference optimization. *arXiv preprint arXiv:2310.03708*, 2023.

A INPUTE DATASET STRUCTURE FOR ALIGNER

```
{
  "prompt": "\n\nHuman: \n\nHuman: What is the difference
    between OpenCL and CUDA?\n\nAssistant:\n\n...[full response]...
    \n\nHuman: Revise the previous answer. Keep the core meaning and
        facts exactly the same, but adjust it to align with these
        preference weights (they sum to 1):
    - Harmless: 0.11
    - Helpful: 0.63
    - Humour: 0.25\n\nRevised answer:",
  "chosen": "OpenCL and CUDA are both programming languages for parallel
      computing on GPUs, but they differ in several key ways:\n\n1.
      Architecture: ...",
  "rejected": "OpenCL and CUDA are two different programming models that
      are used for parallel computing..."
}
```

Figure 4: Example structure of the Fed-ChatbotPA (Ye et al., 2024a) dataset with preference weights. Each entry contains a prompt with preference instructions, a chosen response, and a rejected response for one preference dimension.

B DATASET SETUP

The Fed-ChatbotPA is derived from the Chatbot Arena Conversations dataset (Zheng et al., 2023) and is designed for collaborative learning. To simulate collaborative learning under data deficiency scenario, for Fed-ChatbotPA datasets, we randomly select 6,550 data entities (including prompts and corresponding responses annotated with human preferences) as the training set and evenly distribute them across 131 training clients. Each training client is allocated 20 support data entities and 30 query data entities. An additional 700 data entities are selected as the held-out test set and evenly assigned to 10 test clients (each cluster we test ten clients); these data entities, as well as the corresponding clients' preference vectors, are excluded from training. At test time, each test client uses the same 20-shot adaptation budget as in training, and the remaining 50 data entities are used only for evaluation.

For UltraFeedback datasets, we randomly select 4,750 data entities (including prompts and corresponding responses annotated with human preferences) as the training set and evenly distribute them across 95 training clients. Each training client is allocated 20 support data entities and 30 query data entities. An additional 350 data entities are selected as the held-out test set and evenly assigned to 5 test clients (each cluster we test five clients); these data entities, as well as the corresponding clients' preference vectors, are excluded from training. At test time, each test client uses the same 20-shot adaptation budget as in training, and the remaining 50 data entities are used only for evaluation.

C TRAINING PARAMETERS

Table 3 summarizes the training hyperparameters used for both Fed-ChatbotPA and UltraFeedback, across Llama-3.2-3B and Qwen2.5-3B.

D ALGORITHM

Algorithm 1 summarizes the initialization training process of Section 3.3 within one cluster C_n obtained by the two-stage clustering. Phase 1 (basic initializer training) runs the preference gradient descent with controlled ascent of Equation equation 4 locally and aggregates the client updates with FedAvg, yielding w_{I}. Phase 2 (adaptation-aware meta-training) meta-trains w_{I} on S-shot episodes with the objective of Equation equation 5, returning w_{II}, so that a client preference can be reached within K adaptation steps. Phase 2 uses the same clients, local data, objective functions and server aggregation as Phase 1 and differs only in the inner update rule (the balancing LP equation 5, K inner steps, step size η_{meta}). The convergence analysis of Appendix H.4 concerns a one-local-step idealization of Phase 1 and its output w_{I}; Corollary H.32 extends it to Phase 2 under the same one-

Algorithm 1 APO initialization training within cluster C_n

Input: Cluster C_n from the two-stage clustering with preference vectors $\{\boldsymbol{\lambda}_i\}_{i\in C_n}$ and aggregation weights $\rho_i = N_i/N$; initial iterate w_0; communication rounds R (Phase 1) and R' (Phase 2); local steps τ; local learning rate η and server learning rate α; shot number S, adaptation steps K, and meta learning rate η_{meta}.
Output: Meta-trained initializer w^* for cluster C_n.
Phase 1: Basic initializer training.
for $r = 0$ **to** $R-1$ **do**
  Server broadcasts w_r to all clients in C_n;
  for each client $i \in C_n$ **in parallel do**
    Set $w_i^{(0)} = w_r$;
    for $k = 0$ **to** $\tau - 1$ **do**
      Compute the loss vector $l_i(w_i^{(k)})$, the gradient matrix G_i, the Preference Uniformity β_i, the Preference Adjustment Vector $\boldsymbol{A}_i$, and the index sets J, $\bar{J}$, J^*;
      Solve Equation 4 for μ_i^* {balancing descent if $\beta_i > 0$, uniform descent if $\beta_i = 0$}
      Set $d_{i,k} = G_i\mu_i^*$ and $w_i^{(k+1)} = w_i^{(k)} - \eta\, d_{i,k}$;
    end for
    Client i sends $g_i = w_i^{(\tau)} - w_r$ back to the server;
  end for
  Server updates the initializer: $w_{r+1} = w_r + \alpha \sum_{i\in C_n} \rho_i\, g_i$;
end for
Set $w_{\text{FedBsc}} = w_R$;
Phase 2: Adaptation-aware meta-training.
Set $w_0 = w_{\text{FedBsc}}$;
for $r = 0$ **to** $R' - 1$ **do**
  Server broadcasts w_r to all clients in C_n;
  for each client $i \in C_n$ **in parallel do**
    Sample an S-shot episode with support set $\mathcal{D}_i^s$; set $\theta_i^{(0)} = w_r$;
    for $k = 0$ **to** $K - 1$ **do**
      Compute G_i and $\boldsymbol{A}_i$ at $\theta_i^{(k)}$ on $\mathcal{D}_i^s$ and solve Equation 5 for μ^*;
      Set $\theta_i^{(k+1)} = \theta_i^{(k)} - \eta_{\text{meta}}\, G_i\mu^*$ {preference-balanced inner loop}
    end for
    Client i sends $g_i = \theta_i^{(K)} - \theta_i^{(0)}$ back to the server;
  end for
  Server updates the initializer: $w_{r+1} = w_r + \alpha \sum_{i\in C_n} \rho_i\, g_i$; {preference-balanced outer loop}
end for
Return $w^* = w_{R'}$.

inner-step idealization, so that the bounds also cover w_{II}, and the adaptation analysis of Appendix H.6 is stated from an arbitrary starting point w_0.

E BASELINE

We compare against four baselines that cover the main methodological families for personalized multi-objective alignment under data deficiency. All trained methods are re-implemented in the collaborative setting.(1). *Base* is a zero-shot, training-free lower bound: the instruction-tuned model is prompted with the preference-weighted format of Figure 4 and is not further trained. Under data deficiency it measures what a client obtains when it uses no local samples. (2). *RIC* (Yang et al., 2024b) is an inference-time conditioning method. Supervised fine-tuning teaches the model to follow multi-reward annotations embedded in the prompt, so a preference can be changed at inference without a local gradient step. We include it to test whether this conditioning signal can be learned from the scarce per-client data that define our setting. (3). *Rewarded Soup* (Rame et al., 2023) is a post-training interpolation method: one expert is trained per objective and the preference $\boldsymbol{\lambda}$ is realized by a plug-and-play weight mixture that approximates the Pareto front. Personalization consumes no local feedback, so this baseline represents the data-free extreme of the deficiency

Table 3: Training hyperparameters. The initializer learning rate is cosine across rounds and constant within a round. Reptile uses a constant schedule, AdamW, and no server gradient clipping. * denotes server weight decay 0.01; all other server and inner-loop weight decays are 0.

	Fed-ChatbotPA		UltraFeedback	
	Llama-3.2-3B	Qwen2.5-3B	Llama-3.2-3B	Qwen2.5-3B
DPO β	0.1	0.1	0.1	0.1
LoRA $(r, \alpha, \text{dropout})$	$(8, 16, 0.05)$	$(8, 16, 0.05)$	$(8, 16, 0.05)$	$(8, 16, 0.05)$
Sequence length	256	256	256	256
Precision	4-bit NF4	4-bit NF4	4-bit NF4	4-bit NF4
Optimizer	AdamW	AdamW	AdamW	AdamW
Phase I				
Rounds	200	200	200	200
Local steps	2	2	2	2
Batch size	8	8	8	8
Gradient accumulation	4	4	4	4
Learning rate	$5 \times 10^{-5} \to 1 \times 10^{-5}$	$5 \times 10^{-5} \to 1 \times 10^{-5}$	$5 \times 10^{-5} \to 1 \times 10^{-5}$	$5 \times 10^{-5} \to 1 \times 10^{-5}$
Clients per round	8	8	8	8
Phase II				
Outer rounds	10	10	10	10
Inner steps K	3	3	3	3
Inner learning rate	5×10^{-5}	5×10^{-5}	5×10^{-5}	5×10^{-5}
Inner batch size	8	8	8	8
Inner grad. accumulation	1	1	1	1
Inner grad. clip	10	10	10	10
Clients per round	8	8	8	8
Init. checkpoints	130, 150, 170, 190	130, 150, 170, 190	140, 150, 160, 170, 180	140, 150, 160, 170, 180
Server learning rate	c0: 5×10^{-5} c1: 5×10^{-5} c2: 5×10^{-5} c3: $2 \times 10^{-5}{}^*$	c0: $2 \times 10^{-5}{}^*$ c1: 1×10^{-5} c2: 1×10^{-5} c3: 2×10^{-5}	c0: 2×10^{-5} c1: 1×10^{-5} c2: 1×10^{-5} c3: $2 \times 10^{-5}{}^*$ c4: 5×10^{-5} c5: 2×10^{-5} c6: $2 \times 10^{-5}{}^*$ c7: 3×10^{-5}	c0: 2×10^{-5} c1: $2 \times 10^{-5}{}^*$ c2: 5×10^{-5} c3: 1×10^{-5} c4: 5×10^{-5} c5: 2×10^{-5} c6: $2 \times 10^{-5}{}^*$ c7: 2×10^{-5}

regime. (4). *FSPO* (Singh et al., 2026) is a two-stage few-shot preference-optimization method. A preference-conditioned backbone is first obtained by supervised fine-tuning on a synthetic preference dataset; each client then personalizes from this backbone with only its scarce local comparisons, packing a subset into the prompt as in-context examples and updating a per-user policy on the remaining hold-out pairs under the IPO objective. We include it to test whether few-shot IPO from a synthetic-preference SFT initializer, without collaborative or adaptation-aware meta-training, can exploit the same limited local data that define our setting. (5). *DITTO* (Shaikh et al., 2025) is a few-shot demonstration-alignment method. Each client starts from the instruction-tuned backbone and personalizes from only its scarce local demonstrations: following online imitation learning, those demonstrations are treated as preferred over generations from the current policy and from earlier checkpoints, and a DPO-style update is applied to the resulting self-generated pairs. We include it to test whether demonstration-iterated self-play, without local dispreferred labels, collaborative training, or adaptation-aware meta-training, can exploit the same limited local data that define our setting. (6). *FedAvg* (McMahan et al., 2017) is a collaborative-training baseline with the same S-shot, K-step local adaptation as our method, but the initializer is the FedAvg-aggregated model rather than a meta-trained one. Because the adaptation data coincide, any gap isolates the effect of the initialization. Our method, APO, belongs to this last family: it uses the same scarce local adaptation, but the initializer is obtained by adaptation-aware meta-training so that a client preference can be reached in a few steps.

F METRICS

Preference-weighted normalized score To make scores across dimensions comparable, we apply min–max normalization, $s'_j = (s_j - s_{j,\min})/(s_{j,\max} - s_{j,\min})$, where s_j and $s'_j \in [0, 1]$ denote the raw and normalized scores for dimension j, respectively. Our primary metric is the preference-weighted normalized score, $s_{\boldsymbol{\lambda}} = \sum_{j=1}^{m} \lambda_j s'_j$. For example, in the two-dimensional setting, $s_{\boldsymbol{\lambda}} = \lambda_1 s'_1 + \lambda_2 s'_2$.

Nondominated fronts. A solution x dominates another solution y if

$$f_i(x) \geq f_i(y) \quad \forall i \in \{\text{help}, \text{hon}, \text{truth}\}, \qquad \exists j \in \{\text{help}, \text{hon}, \text{truth}\} : \; f_j(x) > f_j(y).$$

The nondominated set of a method is obtained by removing all solutions that are dominated by another solution from the same method. We additionally construct a seven-method empirical reference front by taking the nondominated set of the union of all methods:

$$\mathcal{R}_{c,p} = \text{ND}\big(\mathcal{A}_{c,p}^{\text{APO}} \cup \mathcal{A}_{c,p}^{\text{FedAvg}} \cup \mathcal{A}_{c,p}^{\text{RewardedSoup}} \cup \mathcal{A}_{c,p}^{\text{Base}} \cup \mathcal{A}_{c,p}^{\text{RIC}} \cup \mathcal{A}_{c,p}^{\text{FSPO}} \cup \mathcal{A}_{c,p}^{\text{DITTO}}\big)\,,$$

where c indexes the cluster and p indexes the preference vector.

Hypervolume (HV). Hypervolume measures the volume of objective space dominated by a method's nondominated front with respect to a reference point. Since our objectives are normalized to $[0, 1]$ and maximized, we use the origin $r = (0, 0, 0)$ as the reference point. For a nondominated front $\mathcal{A}_{c,p}^m$ from method m, the hypervolume is

$$\text{HV}(\mathcal{A}_{c,p}^m) = \mathcal{L}\left(\bigcup_{x \in \mathcal{A}_{c,p}^m} [r_1, f_1(x)] \times [r_2, f_2(x)] \times [r_3, f_3(x)]\right),$$

where $\mathcal{L}(\cdot)$ denotes three-dimensional Lebesgue measure. Intuitively, HV rewards both convergence and diversity: a method obtains a larger HV when it finds solutions that are high on the objectives and span a broader trade-off region. Higher HV is better.

Inverted generational distance (IGD). IGD measures how well a method's front approximates the empirical reference front. Given a method front $\mathcal{A}_{c,p}^m$ and the seven-method reference front $\mathcal{R}_{c,p}$, we compute

$$\text{IGD}(\mathcal{A}_{c,p}^m, \mathcal{R}_{c,p}) = \frac{1}{|\mathcal{R}_{c,p}|} \sum_{z \in \mathcal{R}_{c,p}} \min_{x \in \mathcal{A}_{c,p}^m} \|f(x) - f(z)\|_2\,.$$

Thus, for every point on the empirical reference front, IGD finds the nearest solution from the method and averages these distances. Lower IGD is better, because it indicates that the method covers the reference Pareto front more closely.

G More Detailed Related Work

Multi-Objective Preference Alignment for LLMs. Multi-objective preference alignment steers an LLM toward a client-specified trade-off among objectives such as helpfulness, harmlessness, and honesty, and existing methods differ mainly in where the trade-off is injected. Prompt-conditioned methods embed reward annotations or attribute values into the input and train with SFT or DPO, so that the preference can be changed at inference time; representative examples are RiC (Yang et al., 2024b), SteerLM (Dong et al., 2023), MODPO (Zhou et al., 2023), and DPA (Wang et al., 2024a). Parameter-space methods train one expert per reward and interpolate their weights post hoc, as in Rewarded Soups (Rame et al., 2023) and Personalized Soups (Jang et al., 2023), whereas decoding-time methods combine per-objective models at generation without further training (Shi et al., 2024; Guo et al., 2024). Panacea (Zhong et al., 2024) instead learns a single preference-conditioned model whose low-rank adapter recovers the entire Pareto front. Closest to our architecture, Aligner (Ji et al., 2024) trains a compact residual-correction module on top of a frozen upstream LLM, and MetaAligner (Yang et al., 2024a) applies meta-learning with dynamic objective reformulation so that the aligner generalizes to unseen objectives. All of these methods presume abundant, centrally curated preference data: the conditioning signal, the per-reward experts, or the preference embedding must be learned from large-scale annotations, and the resulting model serves a generic user. None of them is designed to specialize to an individual client from a handful of local feedback, and MetaAligner in particular meta-learns across objectives rather than across users with multi-objective preference, which is the axis of heterogeneity that APO addresses.

Few-Shot and Data-Efficient Personalization of LLMs. A second line of work personalizes an LLM to a user from limited feedback. Personalized RLHF methods introduce a latent user variable that modulates the reward model or the policy, learned from that user's comparisons: P-RLHF (Li

et al., 2024) conditions the reward on a user model, VPL (Poddar et al., 2024) infers a variational user embedding from a few annotated pairs, PAL (Chen et al., 2025) represents users as mixtures of learned prototypes, and GPO (Zhao et al., 2024) predicts group preferences in context from a few examples. To cope with the scarcity of real per-user data, FSPO (Singh et al., 2026) meta-learns a preference-conditioned backbone on synthetic user preferences and then personalizes with few-shot IPO, following the fast-adaptation view of MAML (Finn et al., 2017), whereas DITTO (Shaikh et al., 2025) treats a handful of user demonstrations as preferred over the model's own generations and applies online imitation learning. These methods either rely on synthetic preferences, which may miss the subtle stylistic and contextual preferences of real users, or treat personalization as an *isolated* local adaptation problem that never exploits the real feedback held by other clients. Moreover, they model preference as a scalar or latent quantity and do not expose the multi-objective Pareto structure that a client's trade-off vector induces. APO instead learns the few-shot initialization collaboratively from real client feedback and adapts it toward a preference-specific point on the Pareto front.

Collaborative and Personalized Federated Alignment of LLMs. Federated fine-tuning enables clients to adapt LLMs collaboratively without sharing their raw data, ranging from FedAvg (McMahan et al., 2017) to instruction tuning with heterogeneous adapters or domains (Zhang et al., 2024a; Long et al., 2024; Xu et al., 2024). For preference alignment, OpenFedLLM (Ye et al., 2024b) performs DPO (Rafailov et al., 2023) locally and exchanges only model updates, while FedBis and FedBiscuit (Wu et al., 2024) train lightweight selectors for client subpopulations; FedLLM-Bench (Ye et al., 2024a) further provides realistic federated preference data, including Fed-ChatbotPA, for evaluating such methods. Federated multi-objective methods explicitly account for competing objectives: FedMOL (Yang et al., 2023) and FedMGDA+ (Hu et al., 2022) coordinate objective gradients across clients, whereas FIRM (Nourzad et al., 2025) regularizes the local multi-objective problem to mitigate client disagreement drift. Nevertheless, this line of work primarily targets population-level collaborative alignment rather than learning preference-specific initializations that can be rapidly personalized from scarce client feedback.

Personalized federated learning addresses client heterogeneity from a complementary perspective. PerFedAvg (Jiang et al., 2019) and pFedMe (T Dinh et al., 2020) learn an adaptation-ready initialization or a regularized personalized model; CFL (Sattler et al., 2021) and IFCA (Ghosh et al., 2020) train cluster-specific models based on update or loss similarity; and Few-for-Many pFL (Guo et al., 2026) jointly optimizes K server models to serve many clients. However, these methods generally represent each client's utility with a single objective and therefore do not model preference-specific Pareto geometry or distinguish inter-client from intra-client preference conflicts. APO bridges these two lines of work by clustering clients according to their preference bottlenecks and adjustment vectors, learning a Pareto-aware initialization for each cluster, and bounding both its ϵ_0-distance from client optima and the δ_0-gap after few-shot adaptation.

Gradient Conflict in Multi-Objective and Federated Optimization. From an optimization perspective, jointly training several objectives can produce conflicting gradients, so classical multi-objective methods search the convex hull of the objective gradients for a common descent direction (Fliege & Svaiter, 2000; Désidéri, 2012). PCGrad (Yu et al., 2020) removes conflicting gradient components, CAGrad (Liu et al., 2021) maximizes the worst-case local improvement, and Nash-MTL (Navon et al., 2022) formulates direction selection as a bargaining game. These methods seek a Pareto-stationary solution, but do not determine which trade-off should be selected for a particular preference. EPO (Mahapatra & Rajan, 2020) incorporates such a preference by balancing the objective losses toward a specified preference ray and then descending toward the Pareto front. However, EPO is formulated for a single learner and therefore does not account for interference among clients with heterogeneous preferences.

Federated multi-objective methods such as FedMOL (Yang et al., 2023), FedMGDA+ (Hu et al., 2022), and FIRM (Nourzad et al., 2025) extend MGDA-style optimization to distributed data and mitigate client drift during aggregation. Their primary goal, however, is a population-level Pareto-stationary solution for a common set of objectives, rather than preference-specific initializations for clients with different trade-off vectors. Our setting therefore involves two coupled conflicts: *inter-client* conflict caused by heterogeneous preference vectors and *intra-client* conflict among the objectives of each client. APO extends preference-aware Pareto optimization to this collaborative setting: bottleneck-adjustment clustering limits inter-client interference, preference gradient descent

with controlled ascent resolves intra-client conflict, and adaptation-aware meta-training produces an initialization with a provable few-shot adaptation gap.

H ALIGNER TRAINING

H.1 ASSUMPTIONS

Assumption H.1 (Bounded Gradients)**.** $\|\nabla l_j(w)\| \le B$ for all $j \in [m]$ and all w in the parameter region visited during training.

Remark H.2. For DPO losses used in this paper, the gradient has the form $\nabla l_j(w) = -\beta_{\mathrm{DPO}}\,\mathbb{E}[\sigma(-u)\,(\nabla \log \pi_w(y_c|x) - \nabla \log \pi_w(y_l|x))]$, where $\sigma(-u) \in (0,1)$ self-attenuates as training progresses (Rafailov et al., 2023). While $\nabla \log \pi_w$ is not globally bounded for arbitrary w, the assumption is justified in practice by: (i) LoRA restricts trainable parameters to a low-rank subspace; (ii) weight decay keeps parameters in a bounded region; (iii) gradient clipping (`max_grad_norm` in the trainer) explicitly enforces $\|\nabla l_j\| \le B$.

Assumption H.3 (Lipschitz Gradients)**.** Each l_j is L_0-smooth: $\|\nabla l_j(w) - \nabla l_j(u)\| \le L_0 \|w - u\|$.

Remark H.4. The DPO loss is a composition of the sigmoid function (which is $\beta_{\mathrm{DPO}}/4$-Lipschitz smooth) with the log-ratio of policy probabilities. For LoRA-parameterized models with bounded activations, the composition inherits smoothness with L_0 depending on β_{DPO}, the LoRA rank, and the model architecture.

H.2 COLLABORATIVE LEARNING UPDATE QUALITY

H.2.1 DECOMPOSING H_{C_n} VIA PREFERENCE ADJUSTMENT EQUATIONS

The results of this subsubsection trace the direction heterogeneity back to the adjustment vectors. They rest on two idealizations that we state explicitly; the convergence analysis of Section H.4 does *not* use them (it works with the per-client matrices G_i and treats the direction deviation Δ_i^t, or its trajectory bound $\bar{\Delta}_i$, as an independent quantity).

Assumption H.5 (Common objective functions within a cluster)**.** The clients of C_n optimize the same m population objective functions $l_1, \dots, l_m$, so that at a common parameter w their gradient matrices coincide, $G_i(w) = G(w) = [\nabla l_1(w), \dots, \nabla l_m(w)]$, and $d_i = G\mu_i^*$ with $C := G^\top G$. Heterogeneity across the cluster then enters only through the preferences $\boldsymbol{\lambda}_i$ (hence through $\boldsymbol{A}_i$ and μ_i^*). This is an idealization: clients with different local data have $G_i \neq G$ even when they share the same bottleneck objective, and sharing a bottleneck does not by itself make the LP feasible regions equal. When Assumption H.5 fails, H_{C_n} and $\bar{\Delta}_i$ are to be read as independent quantities and the bounds of this subsubsection are not available.

Before stating the second idealization we record how the adjustment vector enters the LP equation 4 at a fixed w under Assumption H.5. In balancing mode ($\mathbb{1}_{\beta_i} = 1$) and with the common G, $G^\top g^{(j)} = Ce_j$, so equation 4 reads

$$\begin{aligned} \mu_i^* = \arg\max_{\mu}\ & \mu^\top C \boldsymbol{A}_i \\ \text{s.t.}\quad & \mathbf{1}^\top \mu = 1,\ \ \mu \ge 0,\ \ (C\mu)_j \ \ge\ (C\boldsymbol{A}_i)_j\ \mathbb{1}_{J_i}\ \ (j \in \bar{J}_i \setminus J_i^*),\ \ (C\mu)_j \ \ge\ 0\ \ (j \in J_i^*), \end{aligned} \tag{7}$$

where the index sets $J_i, \bar{J}_i, J_i^*$ are those of Section 3.3 evaluated at $(w, \boldsymbol{A}_i)$. The adjustment vector therefore appears in *two* places: in the objective vector $C\boldsymbol{A}_i$ and in the right-hand side of the constraints equation 4b. Collecting the constraint rows into $M\mu \ge b(\boldsymbol{A}_i)$ (with the equality row and the sign constraints included), the right-hand side $b(\boldsymbol{A}_i)$ is a linear function of $\boldsymbol{A}_i$ whose nonzero entries are entries of $C\boldsymbol{A}_i$; in particular $\|b(\boldsymbol{A}) - b(\boldsymbol{A}')\| \le \|C\|\,\|\boldsymbol{A} - \boldsymbol{A}'\|$.

Assumption H.6 (Common optimal basis within a cluster)**.** Fix w and let Assumption H.5 hold. The index sets of the LP equation 7 coincide across the cluster, $(J_i, \bar{J}_i, J_i^*) = (J, \bar{J}, J^*)$ for all $i \in C_n$ (in particular a common bottleneck), and there is a single basis B—a set of m linearly independent constraint rows of M—that is optimal and non-degenerate (unique primal vertex, strictly positive reduced costs off the basis) for *every* client's LP, $i \in C_n$. Writing M_B and $b_B(\cdot)$ for the basic rows, the optimal multipliers are then

$$\mu_i^* = M_B^{-1}\, b_B(\boldsymbol{A}_i), \qquad i \in C_n, \tag{8}$$

i.e. $\boldsymbol{A}_i \mapsto \mu_i^*$ is *affine* on the cluster, with sensitivity constant $\kappa_{\mathrm{LP}} := \|M_B^{-1}\|\,\|C\|$. Two consequences deserve emphasis. (a) The heterogeneity of the directions $d_i = G\mu_i^*$ within the cluster comes entirely from the $\boldsymbol{A}_i$-dependent right-hand sides of the active constraints equation 4b: if the basis B contains only homogeneous rows (the equality $\mathbf{1}^\top \mu = 1$, sign constraints, and bottleneck constraints equation 4c) then b_B does not depend on $\boldsymbol{A}_i$, all μ_i^* coincide, and $H_{C_n}(w) = 0$—a degenerate case in which the Stage-2 refinement has nothing to exploit. The objective vector $C\boldsymbol{A}_i$ affects only *which* basis is optimal, not the value of μ_i^* once B is fixed. (b) The assumption excludes clients whose LPs select *different* optimal bases: across such a switch the solution map is discontinuous (a different vertex of a different polytope), and no Lipschitz bound in $\boldsymbol{A}_i$ holds. Whether a cluster produced by the Stage-2 procedure satisfies this assumption is not verified by the procedure itself; the bound equation 10 below is therefore a statement about this restricted case, and outside it H_{C_n} is a measured quantity.

Proposition H.7 (Direction Heterogeneity via Adjustment Vectors)**.** *Let Assumption H.5 hold, let $\rho_i > 0$ with $\sum_{i\in C_n} \rho_i = 1$ be the aggregation weights of Definition H.12, and write the ρ-weighted means $\bar{\mu}^* := \sum_{i\in C_n} \rho_i\, \mu_i^*$ and $\bar{\boldsymbol{A}}_{C_n} := \sum_{i\in C_n} \rho_i\, \boldsymbol{A}_i$. Then*

$$H_{C_n}(w) = \sum_{i\in C_n} \rho_i\, (\mu_i^* - \bar{\mu}^*)^T C\, (\mu_i^* - \bar{\mu}^*). \tag{9}$$

If in addition Assumption H.6 holds (common index sets and a common non-degenerate optimal basis B), then

$$H_{C_n}(w) \le \kappa_{\mathrm{LP}}^2\, \sigma_{\max}(C) \cdot \sum_{i\in C_n} \rho_i\, \|\boldsymbol{A}_i - \bar{\boldsymbol{A}}_{C_n}\|^2, \tag{10}$$

where $\kappa_{\mathrm{LP}} = \|M_B^{-1}\|\,\|C\|$ is the sensitivity constant of equation 8. The right-hand side is the ρ-weighted within-cluster variance of the adjustment vectors; for equal weights $\rho_i = 1/|C_n|$ it reduces to $\frac{1}{|C_n|}\sum_i \|\boldsymbol{A}_i - \bar{\boldsymbol{A}}_{C_n}\|^2$ with $\bar{\boldsymbol{A}}_{C_n}$ the ordinary sample mean.

Proof. We prove both equation 9 and equation 10; the only property of the weights used is $\rho_i > 0$ and $\sum_i \rho_i = 1$.

Part 1 (Algebraic identity equation 9). Since $d_i = G\,\mu_i^*$ with the common G of Assumption H.5, the aggregated direction is, by linearity,

$$\bar{d}_{C_n} = \sum_i \rho_i\, G\, \mu_i^* = G\left(\sum_i \rho_i\, \mu_i^*\right) = G\,\bar{\mu}^*.$$

The difference is $\bar{d}_{C_n} - d_i = G(\bar{\mu}^* - \mu_i^*)$, so using $\|Gv\|^2 = v^T G^T G\, v = v^T C\, v$ and the symmetry $(-v)^T C(-v) = v^T C v$:

$$\|\bar{d}_{C_n} - d_i\|^2 = (\mu_i^* - \bar{\mu}^*)^T C\, (\mu_i^* - \bar{\mu}^*).$$

Multiplying by ρ_i, summing over $i \in C_n$ and substituting into equation 22 yields equation 9.

Part 2 (Upper bound equation 10). Three ingredients connect equation 9 to the adjustment vectors.

(i) Spectral bound. Since $C = G^T G$ is positive semi-definite, $v^T C\, v \le \sigma_{\max}(C)\,\|v\|^2$ for all v. Applying this to each summand of equation 9 (the weights $\rho_i > 0$ preserve the inequality):

$$H_{C_n} \le \sigma_{\max}(C) \cdot \sum_i \rho_i\, \|\mu_i^* - \bar{\mu}^*\|^2. \tag{11}$$

(ii) Affine solution map under a common basis. Under Assumption H.6 every client's LP has the same non-degenerate optimal basis B, so by equation 8 $\mu_i^* = M_B^{-1} b_B(\boldsymbol{A}_i)$ for all $i \in C_n$, where $b_B(\cdot)$ is linear (the common index sets make the indicator $\mathbb{1}_J$ the same for all clients, so no client-dependent switching enters b_B). Linearity gives, for the ρ-weighted means,

$$\bar{\mu}^* = \sum_i \rho_i\, M_B^{-1} b_B(\boldsymbol{A}_i) = M_B^{-1}\, b_B\Big(\sum_i \rho_i \boldsymbol{A}_i\Big) = M_B^{-1}\, b_B(\bar{\boldsymbol{A}}_{C_n}),$$

i.e. the weighted mean of the multipliers is the affine map evaluated at the weighted mean adjustment vector. Hence

$$\|\mu_i^* - \bar{\mu}^*\| = \big\|M_B^{-1}\big(b_B(\boldsymbol{A}_i) - b_B(\bar{\boldsymbol{A}}_{C_n})\big)\big\| \le \|M_B^{-1}\|\,\|C\|\,\|\boldsymbol{A}_i - \bar{\boldsymbol{A}}_{C_n}\| = \kappa_{\mathrm{LP}}\,\|\boldsymbol{A}_i - \bar{\boldsymbol{A}}_{C_n}\|,$$

using $\|b_B(\boldsymbol{A}) - b_B(\boldsymbol{A}')\| \le \|C\|\|\boldsymbol{A} - \boldsymbol{A}'\|$ (the nonzero rows of b_B are entries of $C\boldsymbol{A}$). Only the rows of B coming from equation 4b contribute to this difference; if none is active the right-hand side is zero and so is H_{C_n} (Assumption H.6(a)). Without Assumption H.6 this step is not available: if two clients select different optimal bases, the solution map jumps between two vertices and $\|\mu_i^* - \mu_{i'}^*\|$ is of order one regardless of $\|\boldsymbol{A}_i - \boldsymbol{A}_{i'}\|$.

Chaining (i)–(ii):

$$H_{C_n} \le \sigma_{\max}(C) \cdot \sum_i \rho_i \, \|\mu_i^* - \bar{\mu}^*\|^2 \le \kappa_{\mathrm{LP}}^2 \, \sigma_{\max}(C) \cdot \sum_i \rho_i \, \|\boldsymbol{A}_i - \bar{\boldsymbol{A}}_{C_n}\|^2,$$

which is equation 10. □

Remark H.8 (What the clustering rests on, and what the bound adds). The two parts of Proposition H.7 play different roles, and only the first is needed to *motivate* the Stage-2 clustering of Section 3.2. The identity equation 9 uses only Assumption H.5 and says that, at a fixed w, the direction dispersion is the C-weighted dispersion of the LP multipliers μ_i^*, which are functions of the adjustment vectors $\boldsymbol{A}_i$ alone (the LP equation 7 has no other client-dependent input once w and G are fixed). Grouping clients by $\boldsymbol{A}_i^{\mathrm{ref}}$ is thus the natural surrogate for grouping them by $\mu_i^*(w_{\mathrm{ref}})$, and requires no continuity of the solution map: Stage 1 (common bottleneck ordering) reduces the chance that clients sit on different LP bases, and Stage 2 (proximity of $\boldsymbol{A}_i^{\mathrm{ref}}$) reduces the spread of the constraint right-hand sides within a basis. The *quantitative* statement equation 10—that $H_{C_n}(w_{\mathrm{ref}})$ is bounded by $\kappa_{\mathrm{LP}}^2 \sigma_{\max}(C)$ times the within-cluster variance of $\boldsymbol{A}_i^{\mathrm{ref}}$—holds only under the common-basis Assumption H.6, and only at the point w at which the assumption is imposed. We do not claim that the Stage-2 grouping lowers $H_{C_n}(w^t)$ along the collaborative training trajectory; this is the design motivation, and it is assessed empirically. Where the convergence bounds of Section H.4 use Δ_i or H_{C_n}, these are to be read as measured trajectory quantities unless Assumption H.6 is explicitly invoked.

Assumption H.9 (Non-degenerate weighted losses)**.** Fix w and write $l_j := l_j(w) > 0$, and let $Z_i := \sum_{j'} \lambda_i^{(j')} l_{j'} > 0$ be the normalizer of the main text, so that $\hat{l}_i^{(j)} := \lambda_i^{(j)} l_j / Z_i$. There is a constant $c \in (0, 1)$, uniform over the clients and over the parameter region visited during training, such that

$$\hat{l}_i^{(j)} \;\ge\; c/m \qquad \text{for all } i \text{ and all } j \in [m]. \tag{12}$$

Proposition H.10 (Adjustment Disagreement Bound)**.** *Fix losses $l(w)$ with $l_j > 0$ and let Assumption H.9 hold with constant c. Write $Z_{\min} := \min_i Z_i > 0$ and set*

$$\Phi_c := \log(m/c) + \log m, \qquad K_c^2 := 2m^2\big[c^{-2} + (\log(m/c) + 1)^2\big], \qquad L_Z := \frac{(1 + \sqrt{m})\,\|l\|_\infty}{Z_{\min}}. \tag{13}$$

Then for any two clients i, i' with preferences on the simplex,

$$\|\boldsymbol{A}_i - \boldsymbol{A}_{i'}\|^2 \;\le\; \big(2\,\Phi_c^2 + 2\,K_c^2\,L_Z^2\big)\,\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'}\|^2 \;=:\; L_{\mathrm{adj}}^2\,\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'}\|^2. \tag{14}$$

In particular the map $\boldsymbol{\lambda} \mapsto \boldsymbol{A}$ is Lipschitz, so $\|\boldsymbol{A}_i - \boldsymbol{A}_{i'}\| \to 0$ as $\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'}\| \to 0$.

Proof. With $\hat{l}_i^{(j)} = \lambda_i^{(j)}\, l_j / Z_i$ as above, define $\phi_i^{(j)} = \log(m\,\hat{l}_i^{(j)}) - \beta_i$ so that $A_i^{(j)} = \lambda_i^{(j)}\,\phi_i^{(j)}$.

Step 1 (Decomposition). For each component j, insert $\pm\,\lambda_{i'}^{(j)}\,\phi_i^{(j)}$:

$$A_i^{(j)} - A_{i'}^{(j)} = \underbrace{(\lambda_i^{(j)} - \lambda_{i'}^{(j)})\,\phi_i^{(j)}}_{\text{(A): scaling difference}} + \underbrace{\lambda_{i'}^{(j)}\,(\phi_i^{(j)} - \phi_{i'}^{(j)})}_{\text{(B): argument difference}}.$$

By Young's inequality $\|u + v\|^2 \le 2\|u\|^2 + 2\|v\|^2$:

$$\|\boldsymbol{A}_i - \boldsymbol{A}_{i'}\|^2 \le 2\underbrace{\textstyle\sum_j (\lambda_i^{(j)} - \lambda_{i'}^{(j)})^2\,(\phi_i^{(j)})^2}_{\text{(A)}} + 2\underbrace{\textstyle\sum_j (\lambda_{i'}^{(j)})^2\,(\phi_i^{(j)} - \phi_{i'}^{(j)})^2}_{\text{(B)}}. \tag{15}$$

Step 2 (Bound on Term A). Since $\hat{l}_i$ is a probability distribution ($\sum_j \hat{l}_i^{(j)} = 1$), we have $\hat{l}_i^{(j)} \le 1$ and hence $\log(m\,\hat{l}_i^{(j)}) \le \log m$; the floor equation 12 gives $\log(m\,\hat{l}_i^{(j)}) \ge \log c$. Therefore

$|\log(m\,\hat{l}_i^{(j)})| \le \max\{\log m,\, \log(1/c)\} \le \log(m/c)$. The preference uniformity is a KL divergence to the uniform distribution on m atoms, so $\beta_i = \mathrm{KL}(\hat{l}_i \| \frac{1}{m}\mathbf{1}) \in [0, \log m]$. By the triangle inequality,

$$|\phi_i^{(j)}| = |\log(m\,\hat{l}_i^{(j)}) - \beta_i| \le \log(m/c) + \log m = \Phi_c,$$

and therefore

$$(\mathrm{A}) \le \Phi_c^2 \sum_j (\lambda_i^{(j)} - \lambda_{i'}^{(j)})^2 = \Phi_c^2 \,\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'}\|^2. \tag{16}$$

Step 3 (Bound on Term B in terms of $\|\hat{l}_i - \hat{l}_{i'}\|$). Preferences lie on the simplex, so $\lambda_{i'}^{(j)} \le \|\boldsymbol{\lambda}_{i'}\|_\infty \le 1$ and $(\mathrm{B}) \le \sum_j (\phi_i^{(j)} - \phi_{i'}^{(j)})^2$. Decompose $\phi_i^{(j)} - \phi_{i'}^{(j)} = [\log(m\,\hat{l}_i^{(j)}) - \log(m\,\hat{l}_{i'}^{(j)})] - [\beta_i - \beta_{i'}]$ and apply Young's inequality a second time:

$$\sum_j (\phi_i^{(j)} - \phi_{i'}^{(j)})^2 \le 2 \sum_j \big[\log(m\,\hat{l}_i^{(j)}) - \log(m\,\hat{l}_{i'}^{(j)})\big]^2 + 2m\,(\beta_i - \beta_{i'})^2, \tag{17}$$

the factor m arising because $\beta_i - \beta_{i'}$ is a scalar offset shared by all m components.

For the log differences, the mean value theorem gives $|\log x - \log y| \le |x - y| / \min(x, y)$, and equation 12 gives $\min(m\,\hat{l}_i^{(j)},\, m\,\hat{l}_{i'}^{(j)}) \ge c$, so

$$\big|\log(m\,\hat{l}_i^{(j)}) - \log(m\,\hat{l}_{i'}^{(j)})\big| \le \frac{m}{c}\,\big|\hat{l}_i^{(j)} - \hat{l}_{i'}^{(j)}\big|, \qquad \text{hence} \qquad \sum_j \big[\,\cdot\,\big]^2 \le \frac{m^2}{c^2}\,\|\hat{l}_i - \hat{l}_{i'}\|^2.$$

For the offset, write $\beta_i = \sum_j h(\hat{l}_i^{(j)})$ with $h(x) = x\log(mx)$ and $h'(x) = \log(mx) + 1$, so $|h'| \le \log(m/c)+1$ on $[c/m, 1]$. Then $|\beta_i - \beta_{i'}| \le (\log(m/c)+1)\,\|\hat{l}_i - \hat{l}_{i'}\|_1 \le (\log(m/c)+1)\sqrt{m}\,\|\hat{l}_i - \hat{l}_{i'}\|$, and $2m(\beta_i - \beta_{i'})^2 \le 2m^2(\log(m/c)+1)^2\|\hat{l}_i - \hat{l}_{i'}\|^2$. Substituting both into equation 17,

$$(\mathrm{B}) \;\le\; K_c^2\,\|\hat{l}_i - \hat{l}_{i'}\|^2. \tag{18}$$

Step 4 (From $\|\hat{l}_i - \hat{l}_{i'}\|$ to $\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'}\|$). Adding and subtracting $(\boldsymbol{\lambda}_{i'} \odot l)/Z_i$,

$$\hat{l}_i - \hat{l}_{i'} = \frac{(\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'}) \odot l}{Z_i} + (\boldsymbol{\lambda}_{i'} \odot l)\Big(\frac{1}{Z_i} - \frac{1}{Z_{i'}}\Big).$$

The first term has norm at most $\|l\|_\infty\,\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'}\|/Z_i$. For the second, $|Z_i - Z_{i'}| = |\langle \boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'},\, l\rangle| \le \|l\|\,\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'}\|$ and $\|\boldsymbol{\lambda}_{i'} \odot l\| = Z_{i'}\|\hat{l}_{i'}\| \le Z_{i'}$, so its norm is at most $\|l\|\,\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'}\|/Z_i$. Using $\|l\| \le \sqrt{m}\,\|l\|_\infty$ and $Z_i \ge Z_{\min}$,

$$\|\hat{l}_i - \hat{l}_{i'}\| \;\le\; \frac{(1+\sqrt{m})\,\|l\|_\infty}{Z_{\min}}\,\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'}\| \;=\; L_Z\,\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'}\|. \tag{19}$$

Step 5 (Combine). Chaining equation 16, equation 18 and equation 19 into equation 15 yields equation 14. The continuity statement is immediate from the Lipschitz constant L_{adj}, which is finite because l is fixed with $l_j > 0$ and $Z_{\min} > 0$. □

Corollary H.11 (Tight preference clusters have low direction heterogeneity)**.** *Let the hypotheses of Propositions H.7 and H.10 hold. If* $\max_{i,i' \in C_n} \|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'}\| \le \delta_\lambda$*, then*

$$H_{C_n}(w) \le C_1\,\sigma_{\max}(C(w))\,\delta_\lambda^2, \qquad C_1 = \kappa_{\mathrm{LP}}^2\,L_{\mathrm{adj}}^2, \tag{20}$$

with L_{adj} as in equation 14. Thus C_1 depends on m, on the floor constant c of Assumption H.9, on $\|l(w)\|_\infty/Z_{\min}$, and on κ_{LP}; it degrades when some Z_i is close to zero. A tighter preference cluster yields smaller meta-update variance. The bound is a statement about the restricted case of Assumptions H.5 and H.6 (common objective functions and a common non-degenerate optimal LP basis at the point w); when either fails (different local gradient matrices, or clients whose LPs select different optimal bases) we do not *claim $H_{C_n} = O(\delta_\lambda^2)$, and every downstream statement is to be read with the direction deviation $\bar{\Delta}_i$ (equivalently $\sqrt{\bar{H}/\rho_{\min}}$) as an independent quantity that is measured rather than bounded by the preference spread.*

Proof. For vectors v_i with weights $\rho_i > 0$, $\sum_i \rho_i = 1$, and weighted mean $\bar{v}$, expanding $\|v_i - c\|^2 = \|(v_i - \bar{v}) + (\bar{v} - c)\|^2$ and using $\sum_i \rho_i (v_i - \bar{v}) = 0$ gives $\sum_i \rho_i \|v_i - \bar{v}\|^2 \le \sum_i \rho_i \|v_i - c\|^2$ for every c (weighted variance minimality at the weighted mean). With $v_i = \boldsymbol{A}_i$ and $c = \boldsymbol{A}_{i_0}$ for any fixed $i_0 \in C_n$, $\sum_i \rho_i \|\boldsymbol{A}_i - \bar{\boldsymbol{A}}_{C_n}\|^2 \le \sum_i \rho_i \|\boldsymbol{A}_i - \boldsymbol{A}_{i_0}\|^2 \le \max_{i,i'} \|\boldsymbol{A}_i - \boldsymbol{A}_{i'}\|^2$, where the last step uses $\sum_i \rho_i = 1$. By equation 14 the right-hand side is at most $L_{\mathrm{adj}}^2 \, \delta_\lambda^2$. Substituting into equation 10 of Proposition H.7 gives equation 20. □

H.3 Collaborative Preference Uniformity Descent under Preference Heterogeneity

Throughout this subsection, for a client $i \in C_n$ we write its *balancing direction* and its *LP margin* as

$$d_{\mathrm{bal},i} \;:=\; G\,\boldsymbol{A}_i, \qquad \gamma_i^* \;:=\; \boldsymbol{A}_i^\top C\,\mu_i^* \;=\; (G\mu_i^*)^\top (G\boldsymbol{A}_i) \;=\; d_i^\top d_{\mathrm{bal},i}, \tag{21}$$

where, under Assumption H.5, $C = G^\top G$ and $d_i = G\mu_i^*$ is the direction from the LP (6) (with per-client matrices one reads G_i, $C_i = G_i^\top G_i$ and $d_{\mathrm{bal},i} = G_i \boldsymbol{A}_i$ throughout), and $\beta_i(w) := \beta_i(l(w))$ is the preference uniformity of client i.

The dispersion of these per-client directions across the cluster is measured by the following quantity, used throughout the convergence analysis.

Definition H.12 (Cluster preference direction heterogeneity)**.** Let $\rho_i > 0$, $i \in C_n$, be the aggregation weights of the cluster (FedAvg uses $\rho_i = N_i/N$), normalized so that $\sum_{i \in C_n} \rho_i = 1$, and write $\rho_{\min} := \min_{i \in C_n} \rho_i$. At parameter w, the preference direction heterogeneity of cluster C_n is

$$H_{C_n}(w) = \sum_{i \in C_n} \rho_i \left\| d_i(w) - \bar{d}_{C_n}(w) \right\|^2, \tag{22}$$

where $\bar{d}_{C_n}(w) = \sum_{i \in C_n} \rho_i \, d_i(w)$ is the aggregated (shared) direction. For equal weights $\rho_i = 1/|C_n|$ this is the ordinary sample variance of the local directions.

Theorem H.13 (Collaborative Preference Uniformity Descent)**.** *Let Assumptions H.1–H.3 hold, let each client run τ local steps of size η before aggregation, and write $\Delta_i := \|\bar{d}_{C_n} - d_i\|$. Suppose that for every $i \in C_n$*

$$\gamma_i^* \;>\; \Delta_i \, \|d_{\mathrm{bal},i}\| \;+\; L_0 B^2 \, \tau \, \eta. \tag{23}$$

Then the shared direction is a simultaneous *balancing direction, i.e. $\bar{d}_{C_n}^\top d_{\mathrm{bal},i}(w_0) > 0$ for all $i \in C_n$, and there exists $\eta_0 > 0$ such that a single meta-update decreases* every *client's non-uniformity:*

$$\beta_i\big(l(w_0 - \eta\, \bar{d}_{C_n})\big) \;\le\; \beta_i\big(l(w_0)\big), \qquad \forall\, i \in C_n,\ \forall\, \eta \in (0, \eta_0]. \tag{24}$$

Proof. Fix $i \in C_n$. Let w_i^t denote the local iterate at which client i computes d_i and its balancing direction, and abbreviate $d_{\mathrm{bal},i}^{\mathrm{loc}} := d_{\mathrm{bal},i}(w_i^t)$ and $d_{\mathrm{bal},i} := d_{\mathrm{bal},i}(w_0)$.

Step 1 (Trajectory drift). Each local step has size at most η and moves against a direction of norm at most B (since $d = G\mu$ with $\mu \in S^m$ and $\|\nabla l_j\| \le B$ by Assumption H.1). Over τ steps,

$$\|w_0 - w_i^t\| \;\le\; \tau\, \eta\, B. \tag{25}$$

By Assumption H.3 the gradients are L_0-Lipschitz, and with the adjustment vector $\boldsymbol{A}_i$ bounded the balancing-direction map $w \mapsto d_{\mathrm{bal},i}(w) = G(w)\boldsymbol{A}_i$ is L_0-Lipschitz (absorbing $\|\boldsymbol{A}_i\|$ into L_0). Hence, using equation 25,

$$\left\| d_{\mathrm{bal},i} - d_{\mathrm{bal},i}^{\mathrm{loc}} \right\| \;\le\; L_0 \, \|w_0 - w_i^t\| \;\le\; L_0\, \tau\, \eta\, B. \tag{26}$$

Step 2 (Lower bound on the shared inner product). Split and apply Cauchy–Schwarz together with $\|\bar{d}_{C_n}\| \le B$ (Jensen and $\|d_{i'}\| \le B$):

$$\begin{aligned} \bar{d}_{C_n}^\top d_{\mathrm{bal},i} &= \underbrace{d_i^\top d_{\mathrm{bal},i}^{\mathrm{loc}}}_{=\gamma_i^*} + (\bar{d}_{C_n} - d_i)^\top d_{\mathrm{bal},i}^{\mathrm{loc}} + \bar{d}_{C_n}^\top \big(d_{\mathrm{bal},i} - d_{\mathrm{bal},i}^{\mathrm{loc}}\big) \\ &\ge \gamma_i^* - \Delta_i \, \|d_{\mathrm{bal},i}\| - \|\bar{d}_{C_n}\| \cdot \left\| d_{\mathrm{bal},i} - d_{\mathrm{bal},i}^{\mathrm{loc}} \right\| \\ &\ge \gamma_i^* - \Delta_i \, \|d_{\mathrm{bal},i}\| - L_0 B^2 \, \tau\, \eta, \end{aligned} \tag{27}$$

where $\|d_{\mathrm{bal},i}\|$ denotes the supremum of the balancing-direction norm over the local region (equal to $\|d^{\mathrm{loc}}_{\mathrm{bal},i}\|$ up to the $O(\tau\eta)$ drift, which is absorbed). The condition equation 23 makes the right-hand side strictly positive, so $\bar{d}_{C_n}^{\top} d_{\mathrm{bal},i} > 0$.

Step 3 (Preference uniformity decreases along the shared direction). It remains to show that, once $\bar{d}_{C_n}^{\top} d_{\mathrm{bal},i} > 0$, moving w_0 against $\bar{d}_{C_n}$ does not increase β_i. Abbreviate $d := \bar{d}_{C_n}$ and evaluate all quantities at w_0: $l_j := l_j(w_0) > 0$, $g_j := \nabla l_j(w_0)$, $l_j^\eta := l_j(w_0 - \eta d)$, and

$$Z_i := \sum_{j=1}^{m} \lambda_i^{(j)} l_j > 0, \qquad Z_i^\eta := \sum_{j=1}^{m} \lambda_i^{(j)} l_j^\eta, \tag{28}$$

so that $\hat{l}_i^{(j)} = \lambda_i^{(j)} l_j / Z_i$ and, from the main text, $\beta_i(l) = \sum_j \hat{l}_i^{(j)} \log(m\hat{l}_i^{(j)})$ and $A_i^{(j)} = \lambda_i^{(j)}\big(\log(m\hat{l}_i^{(j)}) - \beta_i\big)$. Since the preferences and losses are non-negative and $l_j > 0$, all logarithms and normalizations below are well defined.

(3a) Elementary tools. Differentiability gives Taylor's expansion with Peano remainder,

$$l_j^\eta = l_j - \eta\, d^\top g_j + o(\eta), \tag{29}$$

where $o(\eta)/\eta \to 0$ as $\eta \to 0$, and $c\, o(\eta) = o(\eta)$, $o(\eta) + o(\eta) = o(\eta)$ for any constant c. We also use, for $a > \max(0, -b)$, the two elementary logarithm inequalities

$$\log(a+b) \;\geq\; \log a + \frac{b}{a+b}, \qquad \log(a+b) \;\leq\; \log a + \frac{b}{a}. \tag{30}$$

(3b) Splitting the preference uniformity. Using $\log(m\hat{l}_i^{(j)}) = \log m + \log(\lambda_i^{(j)} l_j) - \log Z_i$,

$$\beta_i(l) = \underbrace{\frac{1}{Z_i} \sum_j \lambda_i^{(j)} l_j \log(\lambda_i^{(j)} l_j)}_{=:\,A} + \underbrace{\log \frac{m}{Z_i}}_{=:\,B}, \tag{31}$$

and analogously $\beta_i(l^\eta) = A^\eta + B^\eta$ with Z_i, l_j replaced by Z_i^η, l_j^η. Applying equation 29 termwise,

$$Z_i^\eta = Z_i - \eta\, d^\top \sum_j \lambda_i^{(j)} g_j + o(\eta). \tag{32}$$

(3c) Bounding the two pieces. For B^η, apply the lower inequality in equation 30 with $a = Z_i$ and $b = -\eta\, d^\top \sum_j \lambda_i^{(j)} g_j + o(\eta)$ (so $a + b = Z_i^\eta$):

$$B^\eta = \log \frac{m}{Z_i^\eta} \;\leq\; B + \frac{\eta\, d^\top \sum_j \lambda_i^{(j)} g_j}{Z_i^\eta} - \frac{o(\eta)}{Z_i^\eta}. \tag{33}$$

For A^η, the upper inequality in equation 30 with $a = l_j$, $b = -\eta\, d^\top g_j + o(\eta)$ gives $\log(\lambda_i^{(j)} l_j^\eta) \leq \log(\lambda_i^{(j)} l_j) - \eta\, d^\top g_j / l_j + o(\eta)$. Multiplying by $\lambda_i^{(j)} l_j^\eta = \lambda_i^{(j)} l_j - \eta\, \lambda_i^{(j)} d^\top g_j + o(\eta)$, summing over j, dividing by Z_i^η, and collecting all second-order and remainder contributions into $o(\eta)$,

$$A^\eta \;\leq\; \frac{1}{Z_i^\eta} \sum_j \lambda_i^{(j)} l_j \log(\lambda_i^{(j)} l_j) - \frac{\eta}{Z_i^\eta} \sum_j (d^\top g_j)\, \lambda_i^{(j)} \log(\lambda_i^{(j)} l_j) - \frac{\eta\, d^\top \sum_j \lambda_i^{(j)} g_j}{Z_i^\eta} + \frac{o(\eta)}{Z_i^\eta}. \tag{34}$$

(3d) Combining. Adding equation 33 and equation 34, the terms $\pm\eta\, d^\top \sum_j \lambda_i^{(j)} g_j / Z_i^\eta$ cancel. Expanding the first term of equation 34 through $\frac{Z_i}{Z_i^\eta} A = A + \eta A\, d^\top \sum_j \lambda_i^{(j)} g_j / Z_i^\eta + o(\eta)/Z_i^\eta$ (which follows from equation 32) and using $\beta_i(l) = A + B$,

$$\beta_i(l^\eta) \;\leq\; \beta_i(l) + \frac{\eta}{Z_i^\eta} \sum_j (d^\top g_j)\, \lambda_i^{(j)} \big[A - \log(\lambda_i^{(j)} l_j)\big] + \frac{o(\eta)}{Z_i^\eta}. \tag{35}$$

(3e) Adjustment identity and conclusion. From $\log(m\hat{l}_i^{(j)}) = \log(\lambda_i^{(j)} l_j) - \log Z_i + \log m$ and $\beta_i = A + \log(m/Z_i)$ we obtain $\log(m\hat{l}_i^{(j)}) - \beta_i = \log(\lambda_i^{(j)} l_j) - A$; hence $A_i^{(j)} = \lambda_i^{(j)}\big(\log(\lambda_i^{(j)} l_j) - A\big)$, i.e. $\lambda_i^{(j)}\big[A - \log(\lambda_i^{(j)} l_j)\big] = -A_i^{(j)}$. Substituting into equation 35 and recalling $d_{\mathrm{bal},i} = G\boldsymbol{A}_i = \sum_j A_i^{(j)} g_j$ gives

$$\beta_i\big(l(w_0)\big) - \beta_i\big(l(w_0 - \eta d)\big) \;\geq\; \frac{\eta}{Z_i^\eta}\, d^\top d_{\mathrm{bal},i} - \frac{o(\eta)}{Z_i^\eta} = \frac{\eta}{Z_i^\eta}\Big(d^\top d_{\mathrm{bal},i} - \frac{o(\eta)}{\eta}\Big). \tag{36}$$

By Step 2, $d^\top d_{\mathrm{bal},i} = \bar{d}_{C_n}^\top d_{\mathrm{bal},i} > 0$, and $Z_i^\eta > 0$ for small η by continuity of the losses. Since $o(\eta)/\eta \to 0$, there is $\eta_{0,i} > 0$ such that $o(\eta)/\eta < \bar{d}_{C_n}^\top d_{\mathrm{bal},i}$ for all $\eta \in (0, \eta_{0,i}]$, whence the right-hand side of equation 36 is non-negative and $\beta_i(l(w_0 - \eta \bar{d}_{C_n})) \leq \beta_i(l(w_0))$. Taking $\eta_0 = \min_{i \in C_n} \eta_{0,i} > 0$ establishes equation 24 for every client simultaneously. □

Corollary H.14 (Clustering enlarges the admissible range)**.** *Since $\rho_i\, \Delta_i^2 \leq \sum_{i' \in C_n} \rho_{i'} \Delta_{i'}^2 = H_{C_n}$, every client satisfies $\Delta_i \leq \sqrt{H_{C_n}/\rho_i} \leq \sqrt{H_{C_n}/\rho_{\min}}$ with $\rho_{\min} = \min_{i \in C_n} \rho_i$ (Definition H.12); for equal weights this is $\sqrt{|C_n|\, H_{C_n}}$. Hence a sufficient (uniform) condition for Theorem H.13 is*

$$\min_{i \in C_n} \gamma_i^* \;>\; \sqrt{H_{C_n}/\rho_{\min}}\; \max_{i \in C_n} \|d_{\mathrm{bal},i}\| \;+\; L_0 B^2\, \tau\, \eta. \tag{37}$$

The quantity that the Stage-2 clustering acts on is the within-cluster adjustment variance $\sum_i \rho_i \, \|\boldsymbol{A}_i - \bar{\boldsymbol{A}}_{C_n}\|^2$; under Assumption H.5 the identity equation 9 makes H_{C_n} a function of the multipliers $\mu_i^*(\boldsymbol{A}_i)$ alone, which is the (qualitative) rationale for this choice (Remark H.8). In the restricted case *of Assumption H.6 (common objective functions and a common non-degenerate optimal LP basis at the point where H_{C_n} is evaluated), Proposition H.7 gives the quantitative bound $H_{C_n} \leq \kappa_{\mathrm{LP}}^2\, \sigma_{\max}(C) \cdot \sum_i \rho_i \, \|\boldsymbol{A}_i - \bar{\boldsymbol{A}}_{C_n}\|^2$, so that shrinking the adjustment variance shrinks the right-hand side of equation 37 and enlarges the regime in which the collaborative descent guarantee holds. Outside this case H_{C_n} is a measured quantity, and equation 37 is to be checked with the measured value.*

Proof. The bound $\Delta_i \leq \sqrt{H_{C_n}/\rho_{\min}}$ is immediate from $\rho_i \Delta_i^2 \leq \sum_{i'} \rho_{i'} \Delta_{i'}^2$ (all summands are non-negative), from $\sum_{i'} \rho_{i'} \Delta_{i'}^2 = H_{C_n}$ (Definition H.12), and from $\rho_i \geq \rho_{\min}$. Substituting into equation 23 and replacing $\|d_{\mathrm{bal},i}\|$ by its cluster maximum gives the uniform condition equation 37; the heterogeneity bound is equation 10 of Proposition H.7. □

H.4 Collaborative Initial Aligner Convergence to Cluster-Local Optima.

Section H.3 treated the *balancing* regime ($\beta_i > 0$), where the shared collaborative update reduces every client's preference uniformity. We now analyze the complementary *uniform descent* regime, which is the situation singled out by the Initial Aligner Training Process: once a client's losses are balanced onto its preference ray we have $\beta_i(\boldsymbol{l}_i) = 0$, and its LP (6a) enters uniform descent mode ($\mathbb{1}_{\beta_i} = 0$), maximizing $\mu^\top G_i^\top G_i \mathbf{1}$ subject to the descent constraints (6c). Assuming a cluster of n clients all in this regime, we adapt the convergence analysis to the collaborative learning aggregate, and prove that—under an explicit descent-margin condition on the LP direction (Assumption H.19) and a trajectory hypothesis keeping the iterates in the residual-floor regime—a *single* shared solution converges to within a cluster-tightness-controlled range of *every* client's individual Chebyshev optimum.

Setup and classification (the $\beta_i = 0$ basis). Fix a round and the shared iterate w. For client $i \in C_n$ write its local losses $l_i^{(j)} := l_i^{(j)}(w)$, gradients $g_i^{(j)} := \nabla l_i^{(j)}(w)$, gradient matrix $G_i = [g_i^{(1)}, \ldots, g_i^{(m)}]$, the LP direction $d_i = G_i \mu_i^*$, and the aggregation weight $\rho_i := N_i/N$ (so that $\sum_{i \in C_n} \rho_i = 1$). The collaborative direction and the shared update are

$$\bar{d} \;:=\; \sum_{i \in C_n} \rho_i\, d_i, \qquad w^+ \;:=\; w - \eta\, \bar{d}. \tag{38}$$

Using the preference uniformity β_i, partition the cluster into

$$\begin{aligned} C_n^0 &= \{\, i \in C_n : \beta_i(\boldsymbol{l}_i) = 0 \,\} \text{ (uniform descent)}, \\ C_n^+ &= \{\, i \in C_n : \beta_i(\boldsymbol{l}_i) > 0 \,\} \text{ (balancing descent)}. \end{aligned} \tag{39}$$

To analyze the mixed case, we first assume the uniform descent regime $C_n^0 = C_n$ (all n clients balanced onto their rays). Cluster tightness is measured by the *cross-client gradient dissimilarity*

$$\delta_g \;:=\; \max_{i,i' \in C_n} \max_{j \in [m]} \big\| g_i^{(j)} - g_{i'}^{(j)} \big\|, \tag{40}$$

a cluster-tightness quantity that vanishes as the cluster becomes homogeneous and is small whenever the heterogeneity H_{C_n} (Definition H.12) is small, since $\|\bar{d} - d_i\| \le \sqrt{H_{C_n}/\rho_{\min}}$ (Corollary H.14).

Lemma H.15 (Collaborative uniform descent direction)**.** *Suppose $\beta_i(\boldsymbol{l}_i) = 0$ for every $i \in C_n$. Then each local direction is a descent direction for its own client,*

$$d_i^\top g_i^{(j)} \;\ge\; 0, \qquad \forall\, j \in [m], \tag{41}$$

and the collaborative aggregate is an approximate *simultaneous descent direction:*

$$\bar{d}^\top g_i^{(j)} \;\ge\; -B\,\delta_g, \qquad \forall\, i \in C_n,\ \forall\, j \in [m]. \tag{42}$$

In the homogeneous limit $\delta_g = 0$, $\bar{d}$ is an exact *simultaneous descent direction, $\bar{d}^\top g_i^{(j)} \ge 0$ for all i, j.*

Proof. Single-client descent. When $\beta_i = 0$ the preference adjustment vector vanishes. Indeed $\beta_i = \mathrm{KL}\big(\hat{l}_i \,\|\, \frac{1}{m}\mathbf{1}\big) = 0$ forces the normalized weighted losses to be uniform, $\hat{l}_i^{(j)} = 1/m$, so $\log(m\hat{l}_i^{(j)}) = 0$ and each component $a_i^{(j)} = \lambda_i^{(j)}\big(\log(m\hat{l}_i^{(j)}) - \beta_i\big) = 0$; hence $\boldsymbol{A}_i = 0$. Consequently the active set $J = \{j : \boldsymbol{A}_i^\top G_i^\top g_i^{(j)} > 0\}$ is empty, the indicator $\mathbb{1}_J = 0$, and the constraint block (6c) of the LP reduces to $\mu_i^{*\top} G_i^\top g_i^{(j)} = d_i^\top g_i^{(j)} \ge 0$ for every j, which is exactly equation 41.

Aggregate. Fix i, j and expand $\bar{d}^\top g_i^{(j)} = \sum_{i' \in C_n} \rho_{i'}\, d_{i'}^\top g_i^{(j)}$. For each summand,

$$d_{i'}^\top g_i^{(j)} = \underbrace{d_{i'}^\top g_{i'}^{(j)}}_{\ge 0} + d_{i'}^\top\big(g_i^{(j)} - g_{i'}^{(j)}\big) \;\ge\; -\|d_{i'}\|\,\big\|g_i^{(j)} - g_{i'}^{(j)}\big\| \;\ge\; -B\,\delta_g,$$

where the first term is ≥ 0 by equation 41 applied to client i', the norm bound $\|d_{i'}\| = \|G_{i'}\mu_{i'}^*\| \le B$ holds because $\mu_{i'}^* \in S^m$ is a convex combination of columns with $\|g_{i'}^{(j)}\| \le B$ (Assumption H.1), and $\|g_i^{(j)} - g_{i'}^{(j)}\| \le \delta_g$ by equation 40. Averaging with weights $\rho_{i'}$ (which sum to one) yields equation 42. When $\delta_g = 0$ all clients share the same gradients, so every summand equals $d_{i'}^\top g_{i'}^{(j)} \ge 0$; equivalently the common descent cone $\{d : d^\top g_i^{(j)} \ge 0,\ \forall j\}$ is convex and contains every $d_{i'}$, hence their convex combination $\bar{d}$. □

Lemma H.16 (Collaborative uniform descent and balancing descent)**.** *Consider a* mixed *cluster of $N := |C_n|$ clients, partitioned by equation 39 into $n_+ := |C_n^+|$ balancing clients ($\beta_i > 0$) and $n_0 := |C_n^0|$ uniform clients ($\beta_i = 0$), with $N = n_+ + n_0$. Write the block masses $\rho^+ := \sum_{i \in C_n^+} \rho_i$ and $\rho^0 := \sum_{i \in C_n^0} \rho_i$ (so $\rho^+ + \rho^0 = 1$) and the block mean directions*

$$\bar{d}^+ := \frac{1}{\rho^+} \sum_{i \in C_n^+} \rho_i d_i, \qquad \bar{d}^0 := \frac{1}{\rho^0} \sum_{i \in C_n^0} \rho_i d_i, \qquad \textit{so that} \qquad \bar{d}_{C_n} = \rho^+ \bar{d}^+ + \rho^0 \bar{d}^0, \tag{43}$$

i.e. the aggregate equation 38 is a convex blend of a balancing block and a uniform block. Because whether β_i decreases is decided at the post-aggregation *model $w_0 - \eta \bar{d}_{C_n}$, the relevant deviation is that of d_i from the* full-cluster *aggregate, $\Delta_i := \|\bar{d}_{C_n} - d_i\|$; substituting the block form equation 43 gives the ρ-weighted deviation split*

$$\bar{d}_{C_n} - d_i = \rho^0\big(\bar{d}^0 - d_i\big) + \rho^+\big(\bar{d}^+ - d_i\big), \tag{44}$$

whose triangle inequality resolves Δ_i into an own-block scatter plus a cross-block mutual-influence pull, each weighted by the corresponding block mass,

$$\Delta_i \;\le\; \rho^0 \|\bar{d}^0 - d_i\| + \rho^+ \|\bar{d}^+ - d_i\|. \tag{45}$$

For a balancing client $i \in C_n^+$ the own-block scatter is $\rho^+\|\bar{d}^+ - d_i\|$ and the cross-block pull (from the n_0 uniform clients) is $\rho^0\|\bar{d}^0 - d_i\|$; symmetrically for $i \in C_n^0$. The single shared update $w_0 - \eta\bar{d}_{C_n}$ then acts on the two groups as follows.

(a) Balancing clients $i \in C_n^+$. *The shared direction keeps a positive balancing component,*

$$\bar{d}_{C_n}^\top d_{\mathrm{bal},i} \;\geq\; \gamma_i^* - \Delta_i \|d_{\mathrm{bal},i}\| - L_0 B^2 \tau\eta, \tag{46}$$

which is positive under the range condition equation 23. Substituting the split equation 45 makes the cross-block dependence explicit, giving the mixed-cluster range condition *sufficient for Theorem H.13,*

$$\gamma_i^* \;>\; \big(\rho^0 \|\bar{d}^0 - d_i\| + \rho^+ \|\bar{d}^+ - d_i\|\big) \|d_{\mathrm{bal},i}\| + L_0 B^2 \tau\eta, \tag{47}$$

under which $\bar{d}_{C_n}^\top d_{\mathrm{bal},i} > 0$ and, as in Theorem H.13, the update decreases β_i; the n_0 uniform descent clients enter only through the cross-block pull $\rho^0\|\bar{d}^0 - d_i\|$, which tightens the condition.

(b) Uniform clients $i \in C_n^0$. *We bound $\bar{d}_{C_n}^\top g_i^{(j)}$ for* every $i \in C_n$ *directly from the cross-client dissimilarity equation 40. Expanding the aggregate and using $\|d_{i'}\| \leq B$ gives $\bar{d}_{C_n}^\top g_i^{(j)} \geq \sum_{i' \in C_n} \rho_{i'} d_{i'}^\top g_{i'}^{(j)} - B\delta_g$; the uniform block contributes $d_{i'}^\top g_{i'}^{(j)} \geq 0$, while each balancing client sacrifices objectives only along its anchoring direction, so $d_{i'}^\top g_{i'}^{(j)} \geq -\|d_{\mathrm{bal},i'}\| B$ ($i' \in C_n^+$). Writing the balancing-block mean sacrifice $\bar{D}_{\mathrm{bal}}^+ := \sum_{i' \in C_n^+} \frac{\rho_{i'}}{\rho^+} \|d_{\mathrm{bal},i'}\|$, this yields the* cluster-wide *approximate simultaneous descent bound*

$$\bar{d}_{C_n}^\top g_i^{(j)} \;\geq\; -B\delta_g - \rho^+ B\, \bar{D}_{\mathrm{bal}}^+, \qquad \forall\, i \in C_n^0,\ \forall\, j \in [m], \tag{48}$$

so every *client's objectives rise by at most $\eta B(\delta_g + \rho^+ \bar{D}_{\mathrm{bal}}^+) + o(\eta)$. The shared aggregate is an* exact *simultaneous descent direction ($\bar{d}_{C_n}^\top g_i^{(j)} \geq 0$ for all i, j) iff $\delta_g = 0$ and $\rho^+ \bar{D}_{\mathrm{bal}}^+ = 0$. In particular,* mixing forfeits the exact uniform descent guarantee for the uniform descent clients*: although the local direction of a client $i \in C_n^0$ satisfies $d_i^\top g_i^{(j)} \geq 0$ (no objective of i rises), the aggregated direction only guarantees* approximate *descent, $\bar{d}_{C_n}^\top g_i^{(j)} \geq -B\delta_g - \rho^+ B \bar{D}_{\mathrm{bal}}^+$, with a per-objective ascent of at most $\eta\rho^+ B(\delta_g + \bar{D}_{\mathrm{bal}}^+)$ attributable to the balancing block. The impact of the mixed cluster is the additive* balancing penalty $\rho^+ B\, \bar{D}_{\mathrm{bal}}^+$*: it scales with the balancing mass ρ^+ and the balancing intensity $\bar{D}_{\mathrm{bal}}^+$, is* not *reducible to δ_g, and under Assumption H.17 with the floor $\beta_{i'} \leq \bar{\beta}$ (Lemma H.22) satisfies $\bar{D}_{\mathrm{bal}}^+ \leq L_A \sqrt{\bar{\beta}} = O(\delta_\lambda)$.*

The blend equation 43 thus interpolates the two regimes: a larger balancing mass ρ^+ (more n_+) biases $\bar{d}_{C_n}$ toward preference balancing and inflates the balancing deviation Δ_i of part (a) (via the cross-block pull $\rho^0\|\bar{d}^0 - d_i\|$) as well as the cluster-wide uniform descent slack of part (b) (via the balancing penalty $\rho^+ B\, \bar{D}_{\mathrm{bal}}^+$), while a larger ρ^0 (more n_0) biases it toward uniform descent. In particular $n_0 = 0$ ($\rho^0 = 0$) recovers Theorem H.13, while $n_+ = 0$ ($\rho^+ = 0$, all $\gamma_i^* = 0$) removes the balancing penalty and recovers Lemma H.15.

Proof. Block decomposition. Grouping the collaborative aggregate equation 38 by the partition equation 39, $\bar{d}_{C_n} = \sum_{i \in C_n^+} \rho_i d_i + \sum_{i \in C_n^0} \rho_i d_i = \rho^+ \bar{d}^+ + \rho^0 \bar{d}^0$ with $\rho^+ + \rho^0 = \sum_{i \in C_n} \rho_i = 1$, which is equation 43. Subtracting $d_i = (\rho^+ + \rho^0) d_i$ yields the ρ-weighted deviation split equation 44, and applying the triangle inequality to its two convex-weighted terms gives the direct bound equation 45,

$$\Delta_i = \|\bar{d}_{C_n} - d_i\| \leq \rho^0 \|\bar{d}^0 - d_i\| + \rho^+ \|\bar{d}^+ - d_i\|,$$

in which the summand carrying the *other* block's mass is the cross-block mutual-influence pull ($\rho^0\|\bar{d}^0 - d_i\|$ for $i \in C_n^+$, $\rho^+\|\bar{d}^+ - d_i\|$ for $i \in C_n^0$) and vanishes iff that block is empty or its mean coincides with d_i.

Part (a): balancing clients. Fix $i \in C_n^+$. Let w_i^t be the local iterate at which client i forms d_i and abbreviate $d_{\mathrm{bal},i}^{\mathrm{loc}} := d_{\mathrm{bal},i}(w_i^t)$; as in Step 1 of the proof of Theorem H.13, $\|d_{\mathrm{bal},i} - d_{\mathrm{bal},i}^{\mathrm{loc}}\| \leq L_0 \tau\eta B$. Adding and subtracting and applying Cauchy–Schwarz with $\|d_i\| \leq B$, $\|\bar{d}_{C_n}\| \leq B$, and $d_i^\top d_{\mathrm{bal},i}^{\mathrm{loc}} = \gamma_i^*$ (the identity above evaluated locally),

$$\bar{d}_{C_n}^\top d_{\mathrm{bal},i} = \underbrace{d_i^\top d_{\mathrm{bal},i}^{\mathrm{loc}}}_{=\gamma_i^*} + (\bar{d}_{C_n} - d_i)^\top d_{\mathrm{bal},i}^{\mathrm{loc}} + \bar{d}_{C_n}^\top \big(d_{\mathrm{bal},i} - d_{\mathrm{bal},i}^{\mathrm{loc}}\big) \;\geq\; \gamma_i^* - \Delta_i \|d_{\mathrm{bal},i}\| - L_0 B^2 \tau\eta,$$

which is equation 46. The range condition equation 23 makes the right-hand side positive; substituting the direct split equation 45, $\Delta_i \le \rho^0\|\bar{d}^0 - d_i\| + \rho^+\|\bar{d}^+ - d_i\|$, into equation 23 yields the explicit mixed-cluster range condition equation 47, which is *sufficient* because it replaces Δ_i by an upper bound. The Taylor argument of Theorem H.13 then gives $\beta_i(l(w_0 - \eta\bar{d}_{C_n})) \le \beta_i(l(w_0))$ for all sufficiently small η. By equation 45 the shortfall in Δ_i that must stay below $\gamma_i^*/\|d_{\mathrm{bal},i}\|$ is driven by the cross-block pull $\rho^0\|\bar{d}^0 - d_i\|$ of the n_0 uniform clients.

Part (b): uniform clients. Fix any $i \in C_n$ and any $j \in [m]$. Expanding the collaborative aggregate equation 38 and adding and subtracting each client's own gradient,

$$\bar{d}_{C_n}^\top g_i^{(j)} = \sum_{i' \in C_n} \rho_{i'}\, d_{i'}^\top g_i^{(j)} = \sum_{i' \in C_n} \rho_{i'} \Big[d_{i'}^\top g_{i'}^{(j)} + d_{i'}^\top \big(g_i^{(j)} - g_{i'}^{(j)}\big) \Big] \;\ge\; \sum_{i' \in C_n} \rho_{i'}\, d_{i'}^\top g_{i'}^{(j)} - B\delta_g,$$

where the last step uses Cauchy–Schwarz, $\|d_{i'}\| = \|G_{i'}\mu_{i'}^*\| \le B$ (Assumption H.1), $\|g_i^{(j)} - g_{i'}^{(j)}\| \le \delta_g$ by equation 40, and $\sum_{i'} \rho_{i'} = 1$. This dissimilarity step is uniform in i and is exactly the aggregation of Lemma H.15; what differs in a mixed cluster is the *own-descent* sum $\sum_{i'} \rho_{i'} d_{i'}^\top g_{i'}^{(j)}$, which we split by the partition equation 39:

$$\sum_{i' \in C_n} \rho_{i'}\, d_{i'}^\top g_{i'}^{(j)} = \underbrace{\sum_{i' \in C_n^0} \rho_{i'}\, d_{i'}^\top g_{i'}^{(j)}}_{\ge 0} + \sum_{i' \in C_n^+} \rho_{i'}\, d_{i'}^\top g_{i'}^{(j)}.$$

For a uniform client $i' \in C_n^0$, $\beta_{i'} = 0$ forces $\boldsymbol{A}_{i'} = 0$ and, by case (2), $d_{i'}^\top g_{i'}^{(j)} \ge 0$. For a balancing client $i' \in C_n^+$, the balancing direction $d_{i'}$ coincides with a uniform direction up to its anchoring component $d_{\mathrm{bal},i'}$ (it sacrifices objectives only to move along $d_{\mathrm{bal},i'}$; the same balancing tilt used in the joint bound equation 59), so with $\|g_{i'}^{(j)}\| \le B$,

$$d_{i'}^\top g_{i'}^{(j)} \;\ge\; -\|d_{\mathrm{bal},i'}\|\, B.$$

Hence $\sum_{i' \in C_n^+} \rho_{i'} d_{i'}^\top g_{i'}^{(j)} \;\ge\; -B \sum_{i' \in C_n^+} \rho_{i'} \|d_{\mathrm{bal},i'}\| \;=\; -\rho^+ B\, \bar{D}_{\mathrm{bal}}^+$ with $\bar{D}_{\mathrm{bal}}^+ = \sum_{i' \in C_n^+} \frac{\rho_{i'}}{\rho^+} \|d_{\mathrm{bal},i'}\|$, and combining the two displays yields the cluster-wide bound equation 48,

$$\bar{d}_{C_n}^\top g_i^{(j)} \;\ge\; -B\delta_g - \rho^+ B\, \bar{D}_{\mathrm{bal}}^+, \qquad \forall\, i \in C_n,\ \forall\, j \in [m].$$

The first-order expansion $l_i^{(j)}(w_0 - \eta\bar{d}_{C_n}) = l_i^{(j)}(w_0) - \eta\bar{d}_{C_n}^\top g_i^{(j)} + o(\eta)$ then bounds every client's per-objective increase by $\eta B(\delta_g + \rho^+\bar{D}_{\mathrm{bal}}^+) + o(\eta)$, so $\bar{d}_{C_n}^\top g_i^{(j)} \ge 0$ for all i, j iff $\delta_g = 0$ and $\rho^+\bar{D}_{\mathrm{bal}}^+ = 0$. Reading off δ_g, the descent slack is δ_g (genuine cross-client gradient dissimilarity) plus the balancing penalty $\rho^+\bar{D}_{\mathrm{bal}}^+$, which is *not* reducible to δ_g; under Assumption H.17 and the floor $\beta_{i'} \le \bar{\beta}$ (Lemma H.22) it obeys $\bar{D}_{\mathrm{bal}}^+ \le L_A\sqrt{\bar{\beta}} = O(\delta_\lambda)$, matching the $c_\beta\sqrt{\bar{\beta}}$ term of equation 59 with $c_\beta = \rho^+ BL_A \le BL_A$.

Extremes. If $n_+ = 0$ then $\rho^+ = 0$, all $\gamma_i^* = 0$, the balancing penalty vanishes, and equation 48 reduces to $\bar{d}_{C_n}^\top g_i^{(j)} \ge -B\delta_g$, exactly Lemma H.15. If $n_0 = 0$ then $\rho^+ = 1$, part (a) drives every β_i down under the range condition (Theorem H.13), while equation 48 still holds for the (all-balancing) clients as $\bar{d}_{C_n}^\top g_i^{(j)} \ge -B\delta_g - B\bar{D}_{\mathrm{bal}}^+$. □

A residual preference uniformity floor. The balancing guarantee of Theorem H.13 does not drive every β_i to 0: a single shared point cannot lie on all clients' $\boldsymbol{\lambda}_i^{-1}$ rays at once when the preferences differ. We now quantify the irreducible floor by reading off the point at which the range condition equation 23 stops holding. Two mild regularity constants link the LP margin and the balancing direction to the preference uniformity.

Assumption H.17 (Balancing sensitivity and preference adjustment regularity)**.** There are constants $\kappa_\beta > 0$ and $L_A > 0$ such that, in a neighborhood of the aligned set (i.e. for small β_i), every balancing client $i \in C_n^+$ satisfies

$$\gamma_i^* \;\ge\; \kappa_\beta\, \beta_i(\boldsymbol{l}_i) \qquad \text{and} \qquad \|d_{\mathrm{bal},i}\| \;\le\; L_A\, \sqrt{\beta_i(\boldsymbol{l}_i)}. \tag{49}$$

Remark H.18 (Why equation 49 holds). Both bounds follow from the near-uniform Taylor expansion of the preference uniformity. Writing $\hat{l}_i^{(j)} = \frac{1}{m}(1+\epsilon_i^{(j)})$ with $\sum_j \epsilon_i^{(j)} = 0$, a second-order expansion of the KL divergence gives $\beta_i = \frac{1}{2m}\sum_j (\epsilon_i^{(j)})^2 + o(\|\epsilon_i\|^2)$, while the preference adjustment components $a_i^{(j)} = \lambda_i^{(j)}\big(\log(m\hat{l}_i^{(j)}) - \beta_i\big) = \lambda_i^{(j)}\epsilon_i^{(j)} + o(\|\epsilon_i\|)$ obey $\|\boldsymbol{A}_i\|^2 = \Theta(\beta_i)$; hence $\|d_{\mathrm{bal},i}\| = \|G_i \boldsymbol{A}_i\| \le \sigma_{\max}(G_i)\|\boldsymbol{A}_i\| \le L_A\sqrt{\beta_i}$ with $L_A = \Theta(\sqrt{m}\,B)$. For the lower bound, LP optimality makes $d_i = G_i\mu_i^*$ align with the balancing direction, so $\gamma_i^* = d_i^\top d_{\mathrm{bal},i} \ge \kappa' \|d_{\mathrm{bal},i}\|^2 \ge \kappa'\sigma_{\min}^2\|\boldsymbol{A}_i\|^2 \ge \kappa_\beta\beta_i$ on the subspace spanned by the active gradients. Thus κ_β measures how efficiently the shared step converts non-uniformity into balancing progress.

A descent-margin condition for the LP direction. Assumption H.17 controls the *balancing* progress ($\beta_i \downarrow$). It says nothing about how fast the Chebyshev value $\psi_i(w) = \max_j \lambda_i^{(j)} l_i^{(j)}(w)$ itself decreases once the losses are (nearly) balanced. The LP constraint (6c) only guarantees $d_i^\top g_i^{(j)} \ge 0$ on the exact bottleneck set J^*, which is a first-order non-increase statement and, by itself, cannot rule out stalling at a non-minimizing stationary point (take $l_1(w) = l_2(w) = (w^2-1)^2 + 1$, uniform preferences and $w = 0$: $\beta_i = 0$, all gradients vanish, yet $\psi_i(w) - \psi_i^* = 1$). Any statement that the shared iterate approaches each client's optimum therefore requires a condition that links the LP direction to the *suboptimality* $\psi_i(w) - \psi_i^*$; we make it explicit. For a step size $\eta > 0$ define the activation width and the $\varepsilon_{\mathrm{act}}$-active set of client i at w,

$$\varepsilon_{\mathrm{act}} := 2\eta\,\|\boldsymbol{\lambda}_i\|_\infty B^2, \qquad J_{i,\varepsilon}^*(w) := \big\{\, j \in [m] :\, \lambda_i^{(j)} l_i^{(j)}(w) \;\ge\; \psi_i(w) - \varepsilon_{\mathrm{act}} \,\big\} \;\supseteq\; J_i^*(w), \tag{50}$$

where $J_i^*(w)$ is the exact bottleneck set of (6c). The width $\varepsilon_{\mathrm{act}}$ is chosen so that, after one step of size η, every objective that is *not* $\varepsilon_{\mathrm{act}}$-active at w stays below the common upper bound that Lemma H.25 derives for the near-active objectives. Such an objective may still become the new maximum; what matters is only that it cannot exceed that common bound, so the bound on ψ_i after the step is unaffected.

Assumption H.19 (Descent margin of the LP direction at the residual floor)**.** There exist a region $\mathcal{N} \subseteq \Omega$, a constant $c_d > 0$ and tolerances $b_i \ge 0$ such that, for every client $i \in C_n$ and every $w \in \mathcal{N}$ with $\beta_i(\boldsymbol{l}_i(w)) \le \bar{\beta}$, the LP direction $d_i(w) = G_i(w)\mu_i^*(w)$ of equation 4 satisfies

$$\min_{j \in J_{i,\varepsilon}^*(w)} \lambda_i^{(j)}\, d_i(w)^\top g_i^{(j)}(w) \;\ge\; c_d\,\big(\psi_i(w) - \psi_i^*\big) \;-\; b_i, \qquad \psi_i^* := \min_{w'} \psi_i(w'). \tag{51}$$

We call $b_i = 0$ the *strict* form. The tolerance b_i is a parameter of the assumption; in particular we do *not* assume $b_i = O(\eta)$ or any other order in η, since such a rate would have to be established for the specific direction rule equation 4 and is not established here.

Condition equation 51 strengthens (6c) in two ways: it extends the sign requirement from the exact bottleneck set J_i^* to the $\varepsilon_{\mathrm{act}}$-active set (the LP itself constrains only J_i^*; nonnegativity on $J_{i,\varepsilon}^* \setminus J_i^*$ is an additional requirement, not a consequence of the LP), and it asks the decrease of every near-active weighted loss to be *proportional to the suboptimality*, up to the tolerance b_i. It is a Polyak–Łojasiewicz-type condition on ψ_i along the LP direction, and it is a genuine restriction in two distinct ways.

(a) *It excludes non-minimizing stationary points.* In the double-well example above the left-hand side of equation 51 is 0 while $c_d(\psi_i - \psi_i^*) = c_d$, so the strict form fails there.

(b) *It can also fail near a regular, convex, nonsmooth minimizer.* Take $m = 2$, $l_1(w) = 1 + w$, $l_2(w) = 1 - w$ on $w \in [-\frac{1}{2}, \frac{1}{2}]$ and $\boldsymbol{\lambda} = (\frac{1}{2}, \frac{1}{2})$, so that both losses are positive and smooth with $B = 1$, $\psi(w) = \frac{1}{2} + \frac{1}{2}|w|$ and $\psi^* = \frac{1}{2}$ at $w = 0$. Here $\varepsilon_{\mathrm{act}} = 2\eta\lambda_{\max}B^2 = \eta$. For any fixed $\eta > 0$ and any $0 < w < \eta$ the two weighted losses differ by $w < \varepsilon_{\mathrm{act}}$, so *both* objectives are near-active, and for *every* direction d

$$\min_{j \in J_\varepsilon^*(w)} \lambda_j\, d\, g_j = \min\{\tfrac{1}{2}d, -\tfrac{1}{2}d\} = -\tfrac{1}{2}|d| \le 0 \;<\; c_d\big(\psi(w) - \psi^*\big) = \tfrac{1}{2}c_d w,$$

while $\beta(w) \to 0$ as $w \to 0$, so the failure occurs inside every region $\{\beta \le \bar{\beta}\}$ with $\bar{\beta} > 0$. The strict form ($b_i = 0$) therefore fails throughout a neighbourhood of a perfectly well-behaved minimizer, because a fixed near-active set can contain gradients that point in opposite directions. The tolerant form accommodates this example only with $b_i \ge \frac{1}{2}(c_d w + |d|)$, which does not vanish with η for a direction of fixed magnitude.

Consequently Assumption H.19 is *not* implied by Assumptions H.1–H.3, by preference balance ($\beta_i \le \bar{\beta}$), by the LP constraints, or by the quadratic growth of Assumption H.34; we do not call it mild. Its role is to make explicit the quantitative descent property that any horizon-independent approximation guarantee for the shared iterate must rely on. The next lemma gives a sufficient condition that isolates the two ingredients (a curvature condition and a direction-comparability condition); it does not remove the limitation in (b).

Lemma H.20 (A sufficient condition for the strict form of Assumption H.19)**.** *Fix i and w, and let*

$$\mathcal{G}_{i,\varepsilon}(w) := \operatorname{conv}\big\{\lambda_i^{(j)} g_i^{(j)}(w) : \; j \in J^*_{i,\varepsilon}(w)\big\}, \qquad d_{M,i}(w) := \operatorname*{arg\,min}_{v \in \mathcal{G}_{i,\varepsilon}(w)} \|v\| \tag{52}$$

be the near-active gradient hull *of the Chebyshev scalarization at w and its minimum-norm element (unique, since $\|\cdot\|^2$ is strictly convex on the compact convex set $\mathcal{G}_{i,\varepsilon}(w)$). When $J^*_{i,\varepsilon}(w) = J^*_i(w)$ the hull coincides with the Clarke subdifferential of ψ_i at w; in general it is a different object from the Goldstein ε-subdifferential, which is built from subgradients at* parameters *in a neighbourhood of w rather than from* objectives *whose weighted values are near the maximum, and we do not import results about the latter. Suppose that for all $w \in \mathcal{N}$ with $\beta_i(\boldsymbol{l}_i(w)) \le \bar{\beta}$:*

(i) (near-active PL) $\|d_{M,i}(w)\|^2 \;\ge\; 2\mu_\psi\big(\psi_i(w) - \psi_i^*\big)$ *for some $\mu_\psi > 0$;*

(ii) (σ-comparability of the LP direction) $\min_{j \in J^*_{i,\varepsilon}(w)} \lambda_i^{(j)}\, d_i(w)^\top g_i^{(j)}(w) \;\ge\; \sigma\, \|d_{M,i}(w)\|^2$ *for some $\sigma > 0$.*

Then the strict form ($b_i = 0$) of Assumption H.19 holds with $c_d = 2\sigma\mu_\psi$. Moreover, condition (ii) holds with $\sigma = 1$ for the minimum-norm direction $d_{M,i}(w)$ itself.

Proof. Chaining (ii) and (i) gives $\min_{j \in J^*_{i,\varepsilon}} \lambda_i^{(j)} d_i^\top g_i^{(j)} \ge \sigma \|d_{M,i}\|^2 \ge 2\sigma\mu_\psi(\psi_i(w) - \psi_i^*)$, which is equation 51 with $c_d = 2\sigma\mu_\psi$ and $b_i = 0$. For the last claim, $d_{M,i}$ is the Euclidean projection of the origin onto the closed convex set $\mathcal{G}_{i,\varepsilon}(w)$, so the variational inequality of the projection gives $\langle d_{M,i},\, v - d_{M,i}\rangle \ge 0$ for all $v \in \mathcal{G}_{i,\varepsilon}(w)$; taking $v = \lambda_i^{(j)} g_i^{(j)}$ for a near-active j yields $\lambda_i^{(j)} d_{M,i}^\top g_i^{(j)} \ge \|d_{M,i}\|^2$, i.e. (ii) with $\sigma = 1$. □

Remark H.21 (What Lemma H.20 does and does not give). Condition (i) is a PL inequality for the max-function ψ_i over the near-active hull. It is a local curvature assumption that is *different from* the quadratic growth of Assumption H.34 (PL implies quadratic growth, not conversely), and we do not use one as a substitute for the other. It is also a *strong* condition rather than a routine regularity property: in example (b) preceding the lemma the hull is $\mathcal{G}_\varepsilon(w) = \operatorname{conv}\{\frac{1}{2}, -\frac{1}{2}\} = [-\frac{1}{2}, \frac{1}{2}]$ for $0 < w < \varepsilon_{\mathrm{act}}$, so $d_M = 0$ while $\psi(w) - \psi^* > 0$, and (i) fails. Condition (ii) is a property of the *actual* LP direction d_i of equation 4, and it concerns the near-active set: the LP constraint (6c) only enforces $d_i^\top g_i^{(j)} \ge 0$ on the exact bottleneck set J_i^*, so the sign of the minimum over $J^*_{i,\varepsilon} \supseteq J_i^*$ is not guaranteed by the LP and must be assumed. Condition (ii) holds automatically for the MGDA-type direction $d_{M,i}$; if one replaced the uniform-descent objective in equation 4a by $d_{M,i}$, (ii) would hold for that mode only, (i) would still have to be assumed, the balancing mode would be unaffected, and the algorithm would change and the experiments would have to be re-run.

The ratio $\min_{j \in J^*_{i,\varepsilon}} \lambda_i^{(j)} d_i^\top g_i^{(j)} / \|d_{M,i}\|^2$ can be recorded along training as an *empirical diagnostic* of (ii). Such a record does not prove Assumption H.19; if it is reported, it should include at least: how iterates with $d_{M,i} \approx 0$ (vanishing denominator) are handled; the frequency with which the numerator is negative; the distribution across clients and training phases rather than a single average; and separate statistics for the exact bottleneck set J_i^* and the near-active set $J^*_{i,\varepsilon}$. Estimating the suboptimality $\psi_i(w) - \psi_i^*$ with a value $\hat{\psi}_i^*$ obtained from a well-trained model is also delicate: a feasible value satisfies $\hat{\psi}_i^* \ge \psi_i^*$, so $\psi_i(w) - \hat{\psi}_i^* \le \psi_i(w) - \psi_i^*$ under-estimates the true suboptimality and makes equation 51 look easier to satisfy than it is; an estimate from different data or a different objective adds a further comparison error. We have not carried out such a study in this paper.

Lemma H.22 (Residual preference uniformity floor (conditional form))**.** *Let Assumptions H.1–H.3 and H.17 hold. Fix a client $i \in C_n^+$ and an iterate w (with β_i in the neighbourhood where equation 49 applies) at which the range condition equation 23* fails*, i.e.*

$$\gamma_i^* \;\le\; \Delta_i\, \|d_{\mathrm{bal},i}\| + L_0 B^2 \tau \eta. \tag{53}$$

Then client i satisfies the preference-heterogeneity floor

$$\beta_i(\boldsymbol{l}_i) \;\le\; \bar{\beta} \;:=\; \frac{2L_A^2}{\kappa_{\bar{\beta}}^2}\,\Delta_i^2 + \frac{2L_0B^2}{\kappa_\beta}\,\tau\eta \;\le\; \frac{2L_A^2}{\kappa_{\bar{\beta}}^2}\,\frac{H_{C_n}}{\rho_{\min}} + \frac{2L_0B^2}{\kappa_\beta}\,\tau\eta, \tag{54}$$

so that $l(w)$ lies in the cone $M_{\boldsymbol{\lambda}_i} = \{\, l \,:\, \beta_i(l) \le \bar{\beta}\,\}$ around the $\boldsymbol{\lambda}_i^{-1}$ ray. If in addition the hypotheses of Corollary H.11 hold at the relevant iterates (restricted case, Assumptions H.5–H.6), then $\sqrt{\bar{\beta}} = O(\delta_\lambda) + O(\sqrt{\tau\eta})$. In a mixed *cluster the floor is set by the full-cluster deviation Δ_i, whose direct split equation 45 feeds the cross-block pull $\rho^0\|\bar{d}^0 - d_i\|$ (the n_0 uniform clients pulling on a balancing client $i \in C_n^+$) into $\bar{\beta}$; this cross-block interference only widens the cone $M_{\boldsymbol{\lambda}_i}$ within the same $O(\sqrt{H_{C_n}})$ order and vanishes when $\rho^0 \to 0$ or $\bar{d}^0 = d_i$.*

Remark H.23 (Scope of Lemma H.22). The lemma is a statement about a *single client at a single iterate* where the sufficient condition of Theorem H.13 is not met. It does not assert that the iteration stops there, that later iterates remain in the cone, or that all clients reach equation 53 simultaneously; the failure of a sufficient descent condition for one client is not a statement about the other clients. Establishing that the whole cluster eventually enters and remains in the region $\{\beta_i \le \bar{\beta}\ \forall i\}$ would require an additional entry-and-invariance argument that we do not provide. For this reason the convergence results below (Theorem H.28, Corollary H.29, Theorem H.31) take $\beta_i(\boldsymbol{l}_i(w^t)) \le \bar{\beta}$ for all $i \in C_n$ and all $t \ge t_0$ as an explicit *trajectory hypothesis*, see equation 71, rather than deriving it from Lemma H.22. We also do not claim that $\beta_i = 0$ for all clients is impossible when preferences differ: the clients' loss functions differ as well, so a shared parameter can in principle produce different loss vectors that all lie on their respective rays.

Proof. Under equation 53, $\gamma_i^* \le \Delta_i\|d_{\mathrm{bal},i}\| + L_0B^2\tau\eta$. Substituting the two bounds of equation 49 ($\gamma_i^* \ge \kappa_\beta\beta_i$ on the left, $\|d_{\mathrm{bal},i}\| \le L_A\sqrt{\beta_i}$ on the right) gives

$$\kappa_\beta\,\beta_i \;\le\; \Delta_i\,L_A\,\sqrt{\beta_i} + L_0B^2\tau\eta.$$

Set $x := \sqrt{\beta_i} \ge 0$. Then $\kappa_\beta x^2 - \Delta_i L_A\, x - L_0B^2\tau\eta \le 0$, and the quadratic formula gives

$$x \;\le\; \frac{\Delta_i L_A + \sqrt{\Delta_i^2 L_A^2 + 4\kappa_\beta L_0B^2\tau\eta}}{2\kappa_\beta} \;\le\; \frac{\Delta_i L_A}{\kappa_\beta} + \sqrt{\frac{L_0B^2\tau\eta}{\kappa_\beta}},$$

using $\sqrt{a^2+b} \le a + \sqrt{b}$ for $a, b \ge 0$. Squaring and applying $(p+q)^2 \le 2p^2 + 2q^2$ yields $\beta_i = x^2 \le \frac{2L_A^2}{\kappa_{\bar{\beta}}^2}\Delta_i^2 + \frac{2L_0B^2}{\kappa_\beta}\tau\eta$, the first inequality of equation 54. The second uses $\Delta_i^2 \le H_{C_n}/\rho_{\min}$ (Corollary H.14); in a mixed cluster the deviation Δ_i carries the cross-block pull $\rho^0\|\bar{d}^0 - d_i\|$ of the direct split equation 45, which is bounded by the same H_{C_n} and so does not change the order of the floor. Finally, when the hypotheses of Corollary H.11 hold and $\max_{i,i'}\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'}\| \le \delta_\lambda$, the heterogeneity obeys $H_{C_n} \le C_1\sigma_{\max}(C)\delta_\lambda^2$ (Eq. equation 20), so $\bar{\beta} = O(\delta_\lambda^2) + O(\tau\eta)$ and $\sqrt{\bar{\beta}} = O(\delta_\lambda) + O(\sqrt{\tau\eta})$. Since $\beta_i(l) = 0$ iff l lies on the $\boldsymbol{\lambda}_i^{-1}$ ray, the sublevel set $M_{\boldsymbol{\lambda}_i}$ is a cone around that ray. □

Remark H.24 (From mixed clusters to the residual-floor regime). The convergence analysis below is stated at the residual floor $\beta_i(\boldsymbol{l}_i) \le \bar{\beta}$ for all $i \in C_n$ (Lemma H.22); the mixed cluster reduces to it. Indeed, while some clients are still unbalanced ($C_n^+ \ne \varnothing$), Lemma H.16(a) (equivalently Theorem H.13) shows the shared update decreases every balancing client's preference uniformity β_i under the range condition equation 23, driving its loss toward the $\boldsymbol{\lambda}_i^{-1}$ ray, while Lemma H.16(b) keeps *every* client in approximate descent (its objectives rise by at most $O(\eta(\delta_g + \rho^+\bar{D}_{\mathrm{bal}}^+))$, and since the balancing penalty obeys $\bar{D}_{\mathrm{bal}}^+ \le L_A\sqrt{\bar{\beta}} = O(\delta_\lambda)$ by Assumption H.17 and Lemma H.22, this is $O(\eta(\delta_g + \delta_\lambda))$). The balancing guarantee is only *sufficient*: for a client at which the range condition equation 23 fails, Lemma H.22 bounds that client's β_i by the floor $\bar{\beta} = O(H_{C_n}) + O(\tau\eta)$, but (Remark H.23) this does not by itself show that the whole cluster enters and stays in the region $\{\beta_i \le \bar{\beta}\ \forall i\}$. The convergence results that follow therefore take this region, together with Assumption H.19, as an explicit trajectory hypothesis (equation 71); under it the admissible-set machinery (Lemma H.27→Theorem H.28 →Corollary H.29→Theorem H.31) bounds every client's Chebyshev suboptimality $\psi_i(w^t) - \psi_i^*$ asymptotically by $\epsilon_{0,i} = \kappa_{H,i}\bar{\Delta}_i + \kappa_{\eta,i}\eta + b_i/c_d$ (for every $\zeta > 0$ it is eventually below $\epsilon_{0,i} + \zeta$), where $\bar{\Delta}_i$ bounds the deviation of the local LP direction from the shared aggregate along the trajectory and b_i is the tolerance of Assumption H.19. Figure 2 illustrates the sets involved (the band $\mathcal{A}^\infty$ itself is not drawn).

Exact one-step bounds (no $o(\eta)$ remainders). The convergence argument below is stated for the synchronous shared iteration

$$w^{t+1} \;=\; w^t - \eta\,\bar{d}^{\,t}, \qquad \bar{d}^{\,t} := \sum_{i\in C_n} \rho_i\, d_i(w^t), \qquad \Delta_i^t := \big\|\bar{d}^{\,t} - d_i(w^t)\big\|, \tag{55}$$

i.e. *one* local LP step per communication round ($\tau = 1$), with the server step size absorbed into η. The results of this paragraph and of Theorems H.28 and H.31 are stated and proved for equation 55 only; they do not cover the multi-local-step Algorithm 1 with $\tau > 1$, and the same restriction applies to the K inner steps of Phase 2, which the analysis of Corollary H.32 covers only for $K = 1$ (with η replaced by the Phase-2 step size). Extending them would require replacing $d_i(w^t)$ by the direction that client i actually contributes to the aggregate, namely the average of its τ local LP directions $u_i^t := \tau^{-1}\sum_{k=0}^{\tau-1} d_i(w_{i,k}^t)$, taking $\bar{u}^t = \sum_i \rho_i u_i^t$ as the shared direction with effective step $\alpha\eta\tau$ (α the server learning rate of Algorithm 1), and imposing the descent-margin condition on u_i^t evaluated at the shared point w^t. We do not carry this out here, because the LP direction $d_i(\cdot)$ is a solution map of a linear program whose active set changes with w and is in general *not* continuous in w; the drift $\|u_i^t - d_i(w^t)\|$ therefore cannot be bounded by $O(L_0\tau\eta B)$ or folded into Δ_i^t without an additional stability argument for the LP solution, and we make no such claim. Throughout we write $\lambda_{\max,i} := \|\boldsymbol{\lambda}_i\|_\infty$ and $\lambda_{\min,i} := \min_j \lambda_i^{(j)} > 0$, and use the $\varepsilon_{\mathrm{act}}$-active set $J_{i,\varepsilon}^*(w)$ of equation 50.

Lemma H.25 (Exact one-step loss and Chebyshev bounds)**.** *Let $\bar{d}$ be any direction with $\|\bar{d}\| \le B$ (in particular $\bar{d} = \bar{d}^{\,t}$ of equation 55, since every $d_i = G_i\mu_i^*$ is a convex combination of gradients of norm at most B), and let $\eta > 0$ be such that the whole segment $\{w - s\bar{d} :\ s \in [0,\eta]\}$ lies in the region where Assumptions H.1–H.3 hold (in particular each $l_i^{(j)}$ is L_0-smooth along the segment and $\|g_i^{(j)}(w)\| \le B$ at the current point). Then for every client $i \in C_n$ and every $j \in [m]$,*

$$l_i^{(j)}(w - \eta\bar{d}) \;\le\; l_i^{(j)}(w) - \eta\,\bar{d}^{\top} g_i^{(j)}(w) + \frac{L_0 B^2}{2}\,\eta^2. \tag{56}$$

Moreover, writing

$$\xi_i(w;\bar{d}) \;:=\; \min_{j\in J_{i,\varepsilon}^*(w)} \lambda_i^{(j)}\,\bar{d}^{\top} g_i^{(j)}(w), \tag{57}$$

the Chebyshev value $\psi_i(w) = \max_j \lambda_i^{(j)} l_i^{(j)}(w)$ satisfies

$$\psi_i(w - \eta\bar{d}) \;\le\; \psi_i(w) - \eta\,\xi_i(w;\bar{d}) + \frac{\lambda_{\max,i} L_0 B^2}{2}\,\eta^2. \tag{58}$$

Proof. Per-coordinate bound. Since $l_i^{(j)}$ is L_0-smooth along the segment $\{w - s\bar{d} :\ s \in [0,\eta]\}$ (Assumption H.3), the descent lemma gives $l_i^{(j)}(w - \eta\bar{d}) \le l_i^{(j)}(w) - \eta\,\bar{d}^{\top} g_i^{(j)}(w) + \frac{L_0}{2}\eta^2\|\bar{d}\|^2$, and $\|\bar{d}\| \le B$ yields equation 56. No $o(\eta)$ term and no unspecified η_0 appear; the only requirement on η is that the segment stays inside the region where the assumptions hold.

Chebyshev bound. Multiply equation 56 by $\lambda_i^{(j)} > 0$. For an active index $j \in J_{i,\varepsilon}^*(w)$ we have $\lambda_i^{(j)} l_i^{(j)}(w) \le \psi_i(w)$ and $\lambda_i^{(j)}\bar{d}^{\top} g_i^{(j)} \ge \xi_i(w;\bar{d})$, hence

$$\lambda_i^{(j)} l_i^{(j)}(w - \eta\bar{d}) \;\le\; \psi_i(w) - \eta\,\xi_i(w;\bar{d}) + \frac{\lambda_{\max,i} L_0 B^2}{2}\eta^2.$$

For an inactive index $j \notin J_{i,\varepsilon}^*(w)$ the definition equation 50 gives $\lambda_i^{(j)} l_i^{(j)}(w) < \psi_i(w) - \varepsilon_{\mathrm{act}}$ with $\varepsilon_{\mathrm{act}} = 2\eta\lambda_{\max,i}B^2$, while Cauchy–Schwarz and Assumption H.1 give $|\bar{d}^{\top} g_i^{(j)}| \le B^2$; therefore

$$\begin{aligned}\lambda_i^{(j)} l_i^{(j)}(w - \eta\bar{d}) \;&<\; \psi_i(w) - 2\eta\lambda_{\max,i}B^2 + \eta\lambda_{\max,i}B^2 + \frac{\lambda_{\max,i} L_0 B^2}{2}\eta^2 \\ &\le\; \psi_i(w) - \eta\,\xi_i(w;\bar{d}) + \frac{\lambda_{\max,i} L_0 B^2}{2}\eta^2,\end{aligned}$$

where the last step uses $\xi_i(w;\bar{d}) \le \lambda_{\max,i}B^2$ (again by Cauchy–Schwarz). Thus every weighted loss, active or not, is bounded by the same right-hand side; an inactive objective may well become the new maximum, but it cannot exceed this common bound. Taking the maximum over $j \in [m]$ gives equation 58. □

Remark H.26 (Relation to the per-objective slack of Lemma H.16). Lemma H.16(b) bounds the aggregate from below on *every* objective,

$$\bar{d}_{C_n}^{\top} g_i^{(j)} \;\geq\; -\big(B\delta_g + \rho^{+} B\, \bar{D}_{\mathrm{bal}}^{+}\big), \qquad \forall\, i \in C_n^0,\ j \in [m], \tag{59}$$

and this is the quantity that governs how much a *non-bottleneck* loss may rise in one round. The contraction argument below does *not* use equation 59: it only needs the behaviour of the aggregate on the $\varepsilon_{\mathrm{act}}$-active objectives, which is controlled through the local LP direction d_i and the deviation Δ_i^t (Assumption H.19 and Lemma H.25).

Admissible set. For client i at iterate w^t let $O_i \subseteq \mathbb{R}_+^m$ be its attainable objective set and $\leq$ the componentwise order. Define the client ray point q_i^t, the client dominating set V_i^t, and the client admissible set A_i^t:

$$\begin{aligned} &\psi_i^t := \max_j \lambda_i^{(j)}\, l_i^{(j)}(w^t), \quad q_i^t := \psi_i^t \Big(\tfrac{1}{\lambda_i^{(1)}}, \ldots, \tfrac{1}{\lambda_i^{(m)}}\Big), \\ &V_i^t := \{l \in O_i : l \leq l_i(w^t)\}, \quad A_i^t := \{l \in O_i : l \leq q_i^t\}. \end{aligned} \tag{60}$$

The sets A_i^t live in each client's own objective space O_i; a single shared model w, however, must serve every client at once. We therefore lift the construction to *parameter space*, defining for the cluster C_n the cluster convergence indicator Λ^t, the cluster dominating set $\mathcal{V}^t$, and the (ε-inflated) cluster admissible set $\mathcal{A}^t(\varepsilon)$:

$$\begin{aligned} \Lambda^t &:= \max_{i \in C_n} \psi_i(w^t) = \max_{i \in C_n} \psi_i^t, \\ \mathcal{V}^t &:= \bigcap_{i \in C_n} \{\, w :\ l_i(w) \leq l_i(w^t)\,\}, \\ \mathcal{A}^t(\varepsilon) &:= \bigcap_{i \in C_n} \{\, w :\ l_i(w) \leq q_i^t + \varepsilon \mathbf{1}\,\}, \end{aligned} \tag{61}$$

and abbreviate $\mathcal{A}^t := \mathcal{A}^t(0) = \bigcap_{i \in C_n} \{w : l_i(w) \leq q_i^t\}$. Thus $w \in \mathcal{A}^t(\varepsilon)$ iff $l_i(w) \in A_i^t(\varepsilon)$ for *every* client, i.e. the cluster admissible set is the set of shared parameters whose induced losses fall in every client's inflated admissible box simultaneously (the intersection is exactly the cluster wedge of Figure 2).

Lemma H.27 (Dominating set lies in the admissible set)**.** *For every client i and iterate t, $V_i^t \subseteq A_i^t$. Moreover, if $\beta_i(\boldsymbol{l}_i(w^t)) = 0$ then $l_i(w^t) = q_i^t$ and $A_i^t = V_i^t$.*

Proof. Let $l \in V_i^t$, i.e. $l \leq l_i(w^t)$. Componentwise $\lambda_i^{(j)} l^{(j)} \leq \lambda_i^{(j)} l_i^{(j)}(w^t) \leq \psi_i^t$ for all j; dividing by $\lambda_i^{(j)} > 0$ gives $l^{(j)} \leq \psi_i^t / \lambda_i^{(j)} = q_i^{t,(j)}$, hence $l \leq q_i^t$ and $l \in A_i^t$. If $\beta_i(\boldsymbol{l}_i(w^t)) = 0$ the weighted losses are all equal to their maximum, $\lambda_i^{(j)} l_i^{(j)}(w^t) = \psi_i^t$ for every j, so $l_i^{(j)}(w^t) = \psi_i^t / \lambda_i^{(j)} = q_i^{t,(j)}$, i.e. $l_i(w^t) = q_i^t$; then $A_i^t = \{l \leq q_i^t\} = \{l \leq l_i(w^t)\} = V_i^t$. □

Theorem H.28 (Collaborative admissible step)**.** *Let Assumptions H.1–H.3 and H.19 hold, and suppose that at round t the shared iterate satisfies $w^t \in \mathcal{N}$ and $\beta_i(\boldsymbol{l}_i(w^t)) \leq \bar{\beta}$ for all $i \in C_n$. Write $r_i^t := \psi_i^t - \psi_i^* \geq 0$ for the Chebyshev suboptimality. Then for* every *step size $\eta > 0$ the shared update equation 55 satisfies, for all $i \in C_n$,*

$$\psi_i^{t+1} \;\leq\; \psi_i^t - \eta\, c_d\, r_i^t + \eta\, \lambda_{\max,i} B\, \Delta_i^t + \eta\, b_i + \frac{\lambda_{\max,i} L_0 B^2}{2}\, \eta^2, \tag{62}$$

and every client's loss stays inside its ε_t-inflated admissible set,

$$l_i(w^{t+1}) \in A_i^t(\varepsilon_{i,t}), \qquad \varepsilon_{i,t} := \frac{\lambda_{\max,i}}{\lambda_{\min,i}} \Big(\eta\, B\, \Delta_i^t + \frac{L_0 B^2}{2}\, \eta^2\Big) + \frac{\eta\, b_i}{\lambda_{\min,i}}. \tag{63}$$

The inclusion is exact, $l_i(w^{t+1}) \in A_i^t$, whenever $c_d\, r_i^t \geq \lambda_{\max,i}\big(B\Delta_i^t + \frac{L_0 B^2}{2}\eta\big) + b_i$. (Here η is additionally required to be such that the segment $[w^t, w^{t+1}]$ stays in the region where Assumptions H.1–H.3 hold, as in Lemma H.25.)

Proof. Chebyshev contraction. Apply Lemma H.25 with $\bar{d} = \bar{d}^t$ (which has $\|\bar{d}^t\| \le B$ as a convex combination of the $d_{i'}$, each of norm at most B). For every active index $j \in J^*_{i,\varepsilon}(w^t)$,

$$\lambda_i^{(j)}\, \bar{d}^{\,t\top} g_i^{(j)} = \lambda_i^{(j)}\, d_i^\top g_i^{(j)} + \lambda_i^{(j)}\, (\bar{d}^t - d_i)^\top g_i^{(j)} \;\ge\; \lambda_i^{(j)}\, d_i^\top g_i^{(j)} - \lambda_{\max,i}\, \Delta_i^t\, B,$$

by Cauchy–Schwarz and Assumption H.1. Taking the minimum over the active set and invoking Assumption H.19 (applicable because $w^t \in \mathcal{N}$ and $\beta_i(w^t) \le \bar{\beta}$),

$$\xi_i(w^t; \bar{d}^t) \;\ge\; \min_{j \in J^*_{i,\varepsilon}(w^t)} \lambda_i^{(j)} d_i^\top g_i^{(j)} - \lambda_{\max,i} B\, \Delta_i^t \;\ge\; c_d\, r_i^t - b_i - \lambda_{\max,i} B\, \Delta_i^t.$$

Substituting into equation 58 gives equation 62.

Admissible-set inclusion. For any $j \in [m]$, $\lambda_i^{(j)} l_i^{(j)}(w^{t+1}) \le \psi_i^{t+1}$; dropping the non-positive term $-\eta c_d r_i^t$ in equation 62 and dividing by $\lambda_i^{(j)} \ge \lambda_{\min,i}$,

$$l_i^{(j)}(w^{t+1}) \;\le\; \frac{\psi_i^t}{\lambda_i^{(j)}} + \frac{\lambda_{\max,i}}{\lambda_i^{(j)}} \Big(\eta B \Delta_i^t + \frac{L_0 B^2}{2} \eta^2 \Big) + \frac{\eta b_i}{\lambda_i^{(j)}} \;\le\; q_i^{t,(j)} + \varepsilon_{i,t},$$

using $q_i^{t,(j)} = \psi_i^t / \lambda_i^{(j)}$ from equation 60. Hence $l_i(w^{t+1}) \le q_i^t + \varepsilon_{i,t}\mathbf{1}$, i.e. $l_i(w^{t+1}) \in A_i^t(\varepsilon_{i,t})$. If $c_d r_i^t \ge \lambda_{\max,i}(B\Delta_i^t + \frac{L_0 B^2}{2}\eta) + b_i$ the right-hand side of equation 62 is at most ψ_i^t, so $\lambda_i^{(j)} l_i^{(j)}(w^{t+1}) \le \psi_i^t$ and $l_i(w^{t+1}) \in A_i^t$. □

Corollary H.29 (Contraction of the Chebyshev suboptimality and drift-adjusted nesting)**.** *Let Assumptions H.1–H.3 and H.19 hold, let $0 < \eta \le 1/c_d$, and suppose there is a round t_0 such that for all $t \ge t_0$ and all $i \in C_n$: $w^t \in \mathcal{N}$, $\beta_i(\boldsymbol{l}_i(w^t)) \le \bar{\beta}$, and $\Delta_i^t \le \bar{\Delta}_i$ for some constant $\bar{\Delta}_i \ge 0$. Define the per-client band radius*

$$\epsilon_{0,i} \;:=\; \frac{\lambda_{\max,i}}{c_d} \Big(B\, \bar{\Delta}_i + \frac{L_0 B^2}{2}\, \eta \Big) + \frac{b_i}{c_d} \;=\; \kappa_{H,i}\, \bar{\Delta}_i + \kappa_{\eta,i}\, \eta + \frac{b_i}{c_d}, \qquad \kappa_{H,i} := \frac{\lambda_{\max,i} B}{c_d}, \quad \kappa_{\eta,i} := \frac{\lambda_{\max,i} L_0 B^2}{2 c_d}, \tag{64}$$

where b_i is the tolerance of Assumption H.19 ($b_i = 0$ in the strict form). The radius is an explicit conditional *error bound: its first two terms shrink with $\bar{\Delta}_i$ and η, but the third, b_i/c_d, has no proved vanishing rate, so a small direction deviation and a small step size do not by themselves make $\epsilon_{0,i}$ small. Then for all $t \ge t_0$ and $i \in C_n$:*

1. (One-step contraction)
$$r_i^{t+1} \;\le\; (1 - \eta c_d)\, r_i^t + \eta\, c_d\, \epsilon_{0,i}, \tag{65}$$
so $\psi_i^{t+1} < \psi_i^t$ whenever $r_i^t > \epsilon_{0,i}$, and $\psi_i^{t+1} \le \psi_i^t$ whenever $r_i^t \ge \epsilon_{0,i}$. In particular, even for $\bar{\Delta}_i = 0$ and $b_i = 0$ the Chebyshev value is guaranteed to be non-increasing only outside the $O(\eta)$ band $r_i^t \ge \kappa_{\eta,i}\eta$ created by the second-order term; the earlier claim of exact monotonicity for every step size is not available.

2. (Horizon-independent band) *Unrolling equation 65,*
$$r_i^T \;\le\; (1-\eta c_d)^{\,T-t_0}\, r_i^{t_0} + \big(1 - (1-\eta c_d)^{\,T-t_0}\big)\, \epsilon_{0,i} \;\le\; (1-\eta c_d)^{\,T-t_0}\, r_i^{t_0} + \epsilon_{0,i}, \qquad \forall\, T \ge t_0, \tag{66}$$
and hence $\limsup_{T\to\infty} \big(\psi_i(w^T) - \psi_i^*\big) \le \epsilon_{0,i}$. This is a statement about the values *$\psi_i(w^T)$; it does not assert that $\{w^t\}$ converges or has a limit point, and it does not assert that $r_i^T \le \epsilon_{0,i}$ is reached in finite time: for every $\zeta > 0$ there is T_ζ with $r_i^T \le \epsilon_{0,i} + \zeta$ for all $T \ge T_\zeta$, but the closed band itself need never be entered (the recursion $r^t = \epsilon_{0,i} + q^t(r^0 - \epsilon_{0,i})$, $q = 1 - \eta c_d$, satisfies equation 65 with equality and stays strictly above $\epsilon_{0,i}$ whenever $r^0 > \epsilon_{0,i}$).*

3. (Nesting) *The admissible sets are nested up to the per-step slack of Theorem H.28,*
$$A_i^{t+1}(\varepsilon) \subseteq A_i^t(\varepsilon + \varepsilon_{i,t}), \qquad \mathcal{A}^{t+1}(\varepsilon) \subseteq \mathcal{A}^t\big(\varepsilon + \max_i \varepsilon_{i,t}\big), \qquad \forall\, \varepsilon \ge 0, \tag{67}$$
and exactly *nested,*
$$A_i^{t+1}(\varepsilon) \subseteq A_i^t(\varepsilon) \quad \textit{and} \quad \mathcal{A}^{t+1}(\varepsilon) \subseteq \mathcal{A}^t(\varepsilon), \tag{68}$$
at every round in which $r_i^t \ge \epsilon_{0,i}$ for all $i \in C_n$.

4. (Limiting band) *With* $q_i^* = \psi_i^*(1/\lambda_i^{(1)}, \dots, 1/\lambda_i^{(m)})$ *the ray point of client* i*'s local optimum,*

$$\mathcal{A}^t \;\subseteq\; \bigcap_{i \in C_n} \Big\{\, w :\; l_i(w) \le q_i^* + \tfrac{1}{\lambda_{\min,i}}\big((1-\eta c_d)^{t-t_0} r_i^{t_0} + \epsilon_{0,i}\big)\mathbf{1} \,\Big\}, \tag{69}$$

a family of sets that decreases, as $t \to \infty$*, to the fixed limiting cluster band*

$$\mathcal{A}^\infty := \bigcap_{i \in C_n} \Big\{\, w :\; l_i(w) \le q_i^* + \frac{\epsilon_{0,i}}{\lambda_{\min,i}}\,\mathbf{1} \,\Big\}. \tag{70}$$

Proof. Item 1. Subtract ψ_i^* from both sides of equation 62 and use $\Delta_i^t \le \bar{\Delta}_i$: $r_i^{t+1} \le (1-\eta c_d) r_i^t + \eta\lambda_{\max,i} B\bar{\Delta}_i + \eta b_i + \frac{\lambda_{\max,i} L_0 B^2}{2}\eta^2 = (1-\eta c_d) r_i^t + \eta c_d \epsilon_{0,i}$ by the definition equation 64. Rewriting as $\psi_i^{t+1} - \psi_i^t \le -\eta c_d\,(r_i^t - \epsilon_{0,i})$ gives the sign statements. For $\bar{\Delta}_i = 0$, $b_i = 0$ the band reduces to $\epsilon_{0,i} = \kappa_{\eta,i}\eta$.

Item 2. Since $0 \le 1 - \eta c_d < 1$, induction on equation 65 gives $r_i^T \le (1-\eta c_d)^{T-t_0} r_i^{t_0} + \eta c_d \epsilon_{0,i} \sum_{s=0}^{T-t_0-1} (1-\eta c_d)^s = (1-\eta c_d)^{T-t_0} r_i^{t_0} + \big(1 - (1-\eta c_d)^{T-t_0}\big)\epsilon_{0,i}$, using $\eta c_d \sum_{s=0}^{K-1}(1-\eta c_d)^s = 1 - (1-\eta c_d)^K$; the second inequality in equation 66 drops the nonpositive term $-(1-\eta c_d)^{T-t_0}\epsilon_{0,i}$.

Item 3. Dropping $-\eta c_d r_i^t \le 0$ in equation 62 and dividing by $\lambda_i^{(j)} \ge \lambda_{\min,i}$, $q_i^{t+1,(j)} = \psi_i^{t+1}/\lambda_i^{(j)} \le q_i^{t,(j)} + \varepsilon_{i,t}$, so $q_i^{t+1} + \varepsilon\mathbf{1} \le q_i^t + (\varepsilon + \varepsilon_{i,t})\mathbf{1}$ and $A_i^{t+1}(\varepsilon) \subseteq A_i^t(\varepsilon + \varepsilon_{i,t})$; intersecting over i gives the cluster statement in equation 67. If $r_i^t \ge \epsilon_{0,i}$ then $\psi_i^{t+1} \le \psi_i^t$ by Item 1, hence $q_i^{t+1} \le q_i^t$ componentwise and the inclusions hold with no inflation, which is equation 68.

Item 4. From $q_i^{t,(j)} = \psi_i^t/\lambda_i^{(j)} = (\psi_i^* + r_i^t)/\lambda_i^{(j)} = q_i^{*,(j)} + r_i^t/\lambda_i^{(j)} \le q_i^{*,(j)} + r_i^t/\lambda_{\min,i}$ and equation 66, every $w \in \mathcal{A}^t$ satisfies $l_i(w) \le q_i^t \le q_i^* + \lambda_{\min,i}^{-1}\big((1-\eta c_d)^{t-t_0} r_i^{t_0} + \epsilon_{0,i}\big)\mathbf{1}$ for all i, which is equation 69; the right-hand sets decrease in t and their intersection over t is equation 70. □

Lemma H.30. *(Adapted from (Choo & Atkins, 1983), Theorem 3.1). A feasible solution* w *is weakly Pareto optimal if and only if there exists a weight vector* $\boldsymbol{\lambda_i}$ *such that* w *is an optimal solution to the aggregation function* $\psi_i(w) = \max_{j\in[m]} \lambda_i^{(j)}(l_i^{(j)}(w) - l_i^{(j)*})$ *defined in the main paper.*

Theorem H.31 (Cluster solution approximates every local optimum under the descent-margin condition)**.** *Let Assumptions H.1–H.3 and H.19 hold and let* $0 < \eta \le 1/c_d$*. Suppose that the shared iteration equation 55 admits a round* t_0 *such that, for all* $t \ge t_0$ *and all* $i \in C_n$*,*

$$w^t \in \mathcal{N}, \qquad \beta_i(\boldsymbol{l}_i(w^t)) \le \bar{\beta}, \qquad \Delta_i^t = \|\bar{d}^t - d_i(w^t)\| \le \bar{\Delta}_i \tag{71}$$

(the residual-floor regime of Lemma H.22, with the direction deviation bounded along the trajectory), and that every segment $[w^t, w^{t+1}]$*,* $t \ge t_0$*, lies in the region where Assumptions H.1–H.3 hold. Let* $\epsilon_{0,i} = \kappa_{H,i}\bar{\Delta}_i + \kappa_{\eta,i}\eta + b_i/c_d$ *be the band radius equation 64. Then for every client* $i \in C_n$*:*

1. (Finite-horizon rate) *For all* $T \ge t_0$*,* $\psi_i(w^T) - \psi_i^* \le (1-\eta c_d)^{T-t_0}\big(\psi_i(w^{t_0}) - \psi_i^*\big) + \epsilon_{0,i}$*.*

2. (Objective space: asymptotic band)

$$\limsup_{T\to\infty}\big(\psi_i(w^T) - \psi_i^*\big) \;\le\; \epsilon_{0,i}, \qquad \limsup_{T\to\infty}\, l_i^{(j)}(w^T) \;\le\; q_i^{*,(j)} + \frac{\epsilon_{0,i}}{\lambda_{\min,i}} \quad \textit{for every } j \in [m], \tag{72}$$

where $q_i^* = \psi_i^*\big(1/\lambda_i^{(1)}, \dots, 1/\lambda_i^{(m)}\big)$ *is the ray point of client* i*'s local optimum* w_i^**. If the sequence* $\{w^t\}$ *has a limit point* w^**, then by continuity of the losses* w^* *inherits these bounds,* $\psi_i(w^*) \le \psi_i^* + \epsilon_{0,i}$ *and* $l_i(w^*) \le q_i^* + (\epsilon_{0,i}/\lambda_{\min,i})\mathbf{1}$ *componentwise, i.e.* $w^* \in \mathcal{A}^\infty$*, the fixed limiting cluster band of equation 70. We do not claim that* $\{w^t\}$ *converges or that a limit point exists, nor that the closed band* $r_i^t \le \epsilon_{0,i}$ *is entered after finitely many rounds; the guarantee is asymptotic (for every* $\zeta > 0$ *the iterates are eventually within* $\epsilon_{0,i} + \zeta$*).*

3. (Parameter space, under Assumption H.34) *If a limit point* w^* *exists and lies in the neighbourhood of* w_i^* *where equation 77 holds, then*

$$\|w^* - w_i^*\|^2 \;\le\; \frac{2}{\mu}\,\epsilon_{0,i} \;=\; \frac{2\lambda_{\max,i}}{\mu\, c_d}\Big(B\,\bar{\Delta}_i + \frac{L_0 B^2}{2}\,\eta\Big) + \frac{2\,b_i}{\mu\, c_d}. \tag{73}$$

The band radius is horizon-independent and, for fixed $(c_d, b_i, \mathcal{N})$, is controlled by two quantities only: the trajectory bound $\bar{\Delta}_i$ on the deviation of the local LP direction from the shared aggregate, and the step size η. By Corollary H.14, $\Delta_i^t \le \sqrt{H_{C_n}(w^t)/\rho_{\min}}$, so one may take $\bar{\Delta}_i = \sqrt{\bar{H}/\rho_{\min}}$ with $\bar{H} := \sup_{t \ge t_0} H_{C_n}(w^t)$; this is the quantity that the Stage-2 clustering criterion is designed to reduce (Remark H.8). Only in the restricted case in which the hypotheses of Corollary H.11 (the common-objective and common-optimal-basis idealizations, Assumptions H.5–H.6) hold at every iterate w^t, $t \ge t_0$, does one obtain $\bar{H} = O(\delta_\lambda^2)$ and $\epsilon_{0,i} = O(\delta_\lambda) + O(\eta) + b_i/c_d$; in general $\bar{\Delta}_i$ remains an independent, measured quantity and the band radius is stated in terms of it.

Four qualifications apply. (a) The componentwise bound $l_i(w^*) \le q_i^* + (\epsilon_{0,i}/\lambda_{\min,i})\mathbf{1}$ is a one-sided *upper bound on each loss coordinate; it does not bound $\|l_i(w^*) - q_i^*\|$ or place $l_i(w^*)$ near the Pareto front, and a point that is far below q_i^* in some coordinate is not excluded by it. (b) The constants c_d, b_i and the region $\mathcal{N}$ in Assumption H.19 are defined through the near-active set $J_{i,\varepsilon}^*$, whose width $\varepsilon_{\mathrm{act}} = 2\eta\lambda_{\max,i}B^2$ depends on η; consequently the informal statement "the band closes as $\bar{\Delta}_i \to 0$, $\eta \to 0$" is meaningful only if Assumption H.19 holds with c_d bounded away from 0, $b_i \to 0$ and a fixed $\mathcal{N}$ uniformly in η, which we do not establish. (c) Even in that idealized limit we do* not *claim $w^* = w_i^*$ unless the minimizer of ψ_i is unique on $\mathcal{N}$ (which Assumption H.34 provides locally). (d) The radius is a conditional error bound, not a smallness claim: the tolerance term b_i/c_d has no proved vanishing rate, so even for small $\bar{\Delta}_i$ and η we do not assert that w^t is close to the optimum unless b_i/c_d is separately controlled.*

Proof. Item 1 is equation 66 of Corollary H.29, whose hypotheses are exactly equation 71, the segment condition, and $\eta \le 1/c_d$.

Item 2. Since $0 \le 1 - \eta c_d < 1$, the first term in Item 1 tends to 0, which gives the first inequality in equation 72. For every j, $\lambda_i^{(j)} l_i^{(j)}(w^T) \le \psi_i(w^T)$; dividing by $\lambda_i^{(j)} \ge \lambda_{\min,i}$, using $q_i^{*,(j)} = \psi_i^*/\lambda_i^{(j)}$ and taking $\limsup_T$ gives the second. If $w^{t_k} \to w^*$ along a subsequence, each $l_i^{(j)}$ is continuous (indeed L_0-smooth), hence so is ψ_i, and $\psi_i(w^*) = \lim_k \psi_i(w^{t_k}) \le \limsup_T \psi_i(w^T) \le \psi_i^* + \epsilon_{0,i}$; the componentwise statement follows in the same way, and intersecting over $i \in C_n$ is the definition of $\mathcal{A}^\infty$ in equation 70.

Item 3. Assumption H.34 at $w = w^*$ gives $\frac{\mu}{2}\|w^* - w_i^*\|^2 \le \psi_i(w^*) - \psi_i^* \le \epsilon_{0,i}$, which rearranges to equation 73. □

Corollary H.32 (Phase 2 as a continuation of the shared iteration)**.** *Phase 2 of Algorithm 1 uses the* same *clients, the same local data and objective functions $l_1, \dots, l_m$, the same aggregation weights ρ_i and the same server update $w_{r+1} = w_r + \alpha \sum_i \rho_i g_i$ as Phase 1; the S-shot episodes are subsamples of the same local datasets, not new data. Its inner update solves equation 5, which is the LP equation 4 in balancing mode ($\mathbb{1}_{\beta_i} = 1$) with the same constraints equation 4b–equation 4c, so the inner direction is the same LP direction $d_i = G_i\mu_i^*$ whenever $\beta_i > 0$ (for $\beta_i = 0$ the objective of equation 5 vanishes; in the analysis we adopt the convention that the uniform-descent direction of equation 4 is then used, as in Phase 1). Consequently, under the same idealization as Theorem H.31—one inner step per round ($K = 1$, in place of $\tau = 1$) and population gradients in place of the S-shot gradients—the Phase-2 round map is the shared iteration equation 55 with effective step size $\eta_{\mathrm{II}} := \alpha\eta_{\mathrm{meta}}$ in place of the Phase-1 effective step size η_{I} (the η of equation 55, into which the server step size is absorbed), and the two phases form a single trajectory*

$$w^0, \dots, w^R = w_{\mathrm{I}},\ w^{R+1}, \dots, w^{R+R'} = w_{\mathrm{II}}, \qquad w^{t+1} = w^t - \eta_t\, \bar{d}^t, \quad \eta_t = \begin{cases} \eta_{\mathrm{I}}, & t < R, \\ \eta_{\mathrm{II}}, & t \ge R. \end{cases}$$

Let Assumptions H.1–H.3 hold, let Assumption H.19 hold with the near-active set $J_{i,\varepsilon}^*$ taken at width $\varepsilon_{\mathrm{act}} = 2\eta_t\lambda_{\max,i}B^2$ for the step size η_t in force at round t, let $0 < \eta_{\mathrm{I}}, \eta_{\mathrm{II}} \le 1/c_d$, and let the trajectory hypothesis equation 71 hold for all $t \ge t_0$ along the concatenated trajectory (including the Phase-2 rounds), with the segment condition of Theorem H.31. Define the two-phase band radius

$$\hat{\epsilon}_{0,i} \;:=\; \kappa_{H,i}\,\bar{\Delta}_i + \kappa_{\eta,i}\,\max(\eta_{\mathrm{I}}, \eta_{\mathrm{II}}) + \frac{b_i}{c_d} \;\ge\; \max\big(\epsilon_{0,i}(\eta_{\mathrm{I}}), \epsilon_{0,i}(\eta_{\mathrm{II}})\big). \tag{74}$$

Then for every $i \in C_n$ and every $T \ge t_0$ on the concatenated trajectory,

$$\psi_i(w^T) - \psi_i^* \;\le\; \Big(\prod_{t=t_0}^{T-1}(1 - \eta_t c_d)\Big)\big(\psi_i(w^{t_0}) - \psi_i^*\big) + \hat{\epsilon}_{0,i}, \tag{75}$$

so that in particular the Phase-2 output satisfies $\psi_i(w_{\mathrm{II}}) - \psi_i^* \le (1-\eta_{\min}c_d)^{R+R'-t_0}(\psi_i(w^{t_0}) - \psi_i^*) + \hat{\epsilon}_{0,i}$ *with* $\eta_{\min} := \min(\eta_{\mathrm{I}}, \eta_{\mathrm{II}})$*, and Items 2–3 of Theorem H.31 hold for the concatenated trajectory with* $\epsilon_{0,i}$ *replaced by* $\hat{\epsilon}_{0,i}$*. Everything that Theorem H.31 and Section H.5 state for the Phase-1 output* w_{I} *therefore also holds for* w_{II}*, with the band radius equation 74.*

The limitations are those already stated for Phase 1, not new ones: (i) the analysis covers $K = 1$ *inner step, exactly as it covers* $\tau = 1$ *local step (the discussion after equation 55 applies verbatim to* $K > 1$*); (ii) it uses population gradients, whereas Phase 2 evaluates* G_i *and* $\boldsymbol{A}_i$ *on an* S*-shot episode—the same population-versus-minibatch gap as in Phase 1, but larger for* $S = 20$*, and we do not quantify it; (iii) the trajectory hypothesis equation 71 is imposed on the Phase-2 rounds as well, and is not derived. What the corollary rules out is a* qualitative *degradation: Phase 2 does not change the objective, the data or the direction rule, so it cannot move the shared model out of the band of Theorem H.31 other than through the step-size term* $\kappa_{\eta,i}\eta_{\mathrm{II}}$.

Proof. The one-step contraction equation 65 of Corollary H.29 is proved for a single round of the iteration equation 55 with step size η under the hypotheses listed there at that round; nothing in its proof uses the value of η at other rounds. Applying it at round t with $\eta = \eta_t$ gives $r_i^{t+1} \le (1-\eta_t c_d) r_i^t + \eta_t c_d\, \epsilon_{0,i}(\eta_t) \le (1-\eta_t c_d) r_i^t + \eta_t c_d\, \hat{\epsilon}_{0,i}$, using $\epsilon_{0,i}(\eta_t) \le \hat{\epsilon}_{0,i}$ from equation 74 and $\eta_t c_d \le 1$. Subtracting $\hat{\epsilon}_{0,i}$ from both sides yields $r_i^{t+1} - \hat{\epsilon}_{0,i} \le (1-\eta_t c_d)(r_i^t - \hat{\epsilon}_{0,i})$, and unrolling from t_0 to T gives equation 75 (the product is at most $(1-\eta_{\min}c_d)^{T-t_0}$). The lim sup, limit-point and parameter-space statements follow exactly as in the proof of Theorem H.31, with $\hat{\epsilon}_{0,i}$ in place of $\epsilon_{0,i}$. □

Remark H.33 (What changed relative to a first-order argument). A bound of the form $\psi_i^{t+1} \le \psi_i^t + \eta\varepsilon$ for every step, iterated over T rounds, only yields $\psi_i^T \le \psi_i^0 + T\eta\varepsilon$; it neither implies descent nor a horizon-independent distance to ψ_i^*. The horizon-independent band equation 72 is obtained here because Assumption H.19 supplies a decrease that is *proportional to the current suboptimality* r_i^t (up to the tolerance b_i), turning the one-step bound into the contraction equation 65. Without such a condition the conclusion of Theorem H.31 is false in general (see the double-well example preceding Assumption H.19), so the condition is not a technicality that can be dropped.

Scope. The convergence result is conditional on a quantitative descent-margin assumption for the actual update direction and on the iterates remaining in the stated region. The assumption is not implied by smoothness, preference balance, or the LP constraints alone. In particular, a fixed near-active set can contain conflicting gradients near a nonsmooth optimum (example (b) preceding Assumption H.19), so the assumption need not hold throughout a neighbourhood of that optimum. The theorem covers the synchronous iteration equation 55 with one local step per round; the multi-local-step Algorithm 1 is not covered verbatim.

H.5 The Initializer Approximates Each Client's Individual Optimum

Theorem H.31 establishes, for each same-cluster client and under the trajectory hypothesis equation 71 together with Assumption H.19, the eventual objective-space bound $\limsup_t(\psi_i(w^t) - \psi_i^*) \le \epsilon_{0,i}$ and, *whenever a limit point* w^* *of the shared iterates exists*, the inherited guarantees $\psi_i(w^*) \le \psi_i^* + \epsilon_{0,i}$, $l_i(w^*) \le q_i^* + (\epsilon_{0,i}/\lambda_{\min,i})\mathbf{1}$ and (under Assumption H.34) $\|w^* - w_i^*\|^2 \le \frac{2}{\mu}\epsilon_{0,i}$, with $\epsilon_{0,i} = \kappa_{H,i}\bar{\Delta}_i + \kappa_{\eta,i}\eta + b_i/c_d$ from equation 64. Throughout this subsection w^* denotes such a limit point of the idealized one-local-step shared iteration equation 55, or its output after finitely many rounds, for which Theorem H.31(1) gives the same bounds up to the geometrically decaying transient. By Corollary H.32 this covers both the Phase-1 output w_{I} and the Phase-2 output w_{II} that Algorithm 1 returns (Phase 2 is the same shared iteration with the effective step size η_{II}), provided the band radius is read as the two-phase radius $\hat{\epsilon}_{0,i}$ of equation 74, i.e. with η replaced by $\max(\eta_{\mathrm{I}}, \eta_{\mathrm{II}})$ in $\epsilon_{0,i}$; for brevity we keep writing $\epsilon_{0,i}$ and η below with this understanding. The scope is the idealization already stated (one local/inner step per round, population gradients, the trajectory hypothesis equation 71). Within this scope, w^* is, within a quantifiable range, an approximation of *each* same-cluster client's individual optimal solution w_i^*. We now re-express the same suboptimality as an explicit decomposition—separating the collaborative training gap from the preference spread—so that the guarantee transfers to a new client whose preference lies in the cluster, which shares the

cluster's objective functions, but which *did not* participate in training. For each client i, let

$$w_i^* \in \arg\min_w \ \psi_i(w), \qquad \psi_i^* := \psi_i(w_i^*) = \min_w \psi_i(w), \qquad \psi_i(w) = \max_j \lambda_i^{(j)} l_j(w), \tag{76}$$

so that w_i^* is the solution for preference $\boldsymbol{\lambda}_i$ (its weighted losses are balanced and lie on the $\boldsymbol{\lambda}_i^{-1}$ ray, as in the main text).

We add a standard local curvature condition, satisfied near the optimum for the LoRA-restricted, smooth losses considered here.

Assumption H.34 (Local quadratic growth)**.** For each client i there is a neighborhood of w_i^* and a constant $\mu > 0$ such that

$$\tfrac{\mu}{2} \|w - w_i^*\|^2 \ \le \ \psi_i(w) - \psi_i^*. \tag{77}$$

Remark H.35. Condition equation 77 is the local *quadratic-growth* property of the Chebyshev scalarization ψ_i. It is weaker than, and not equivalent to, a Polyak–Łojasiewicz (PL) inequality: PL implies quadratic growth, but the converse fails in general. The PL-type condition that the collaborative contraction argument needs is stated separately in Assumption H.19 (with a sufficient condition in Lemma H.20), and the gradient-norm PL inequality needed for the adaptation-phase contraction is stated separately in Assumption H.39; equation 77 is used only to convert an objective-space bound into a parameter-space bound and is never used as a substitute for either. We state equation 77 as an explicit assumption: neither the LoRA parameterization nor L_0-smoothness (Assumption H.3) implies it. It is plausible on a region where the active bottleneck is stable and the active weighted loss has positive curvature at its minimizer, in which case the max of finitely many such functions inherits quadratic growth, but we do not prove this for the DPO losses used here.

We first record an elementary preference-transfer bound: a new client whose preference is close to the cluster inherits the cluster's convergence indicator up to the preference spread. The bound, and everything in this and the next subsection that builds on it, is a statement about *preference* transfer under a *common* loss vector: all clients and the new client evaluate the same objective functions $l_1, \dots, l_m$ (the same $l(w)$), and only the preference vector differs. A new client whose loss functions differ from the cluster's would additionally require control of $l_{\text{new}}(w) - l_i(w)$ at the initializer, along the adaptation iterates and at the compared optima; we do not treat that case.

Proposition H.36 (Starting Point Gap: Cluster vs. New Client)**.** *Suppose the new client evaluates the same objective functions l_j as the cluster (only $\boldsymbol{\lambda}$ differs). If $\boldsymbol{\lambda}_{\text{new}}$ belongs to cluster C_n (i.e., $\|\boldsymbol{\lambda}_{\text{new}} - \boldsymbol{\lambda}_i\| \le \delta_\lambda$ for all $i \in C_n$), then:*

$$\psi_{\text{new}}(w^*) \le \psi_{C_n}(w^*) + \delta_\lambda \, \|l(w^*)\|_\infty, \tag{78}$$

where $\psi_{C_n}(w) = \frac{1}{|C_n|} \sum_{i \in C_n} \psi_i(w)$.

Proof. For each objective j:

$$\begin{aligned} \lambda_{\text{new}}^{(j)} \, l_j(w^*) &= \lambda_i^{(j)} \, l_j(w^*) + (\lambda_{\text{new}}^{(j)} - \lambda_i^{(j)}) \, l_j(w^*) \\ &\le \lambda_i^{(j)} \, l_j(w^*) + |\lambda_{\text{new}}^{(j)} - \lambda_i^{(j)}| \, |l_j(w^*)|. \end{aligned}$$

Taking $\max_j$ on both sides:

$$\psi_{\text{new}}(w^*) = \max_j \lambda_{\text{new}}^{(j)} \, l_j(w^*) \le \max_j \lambda_i^{(j)} \, l_j(w^*) + \max_j |\lambda_{\text{new}}^{(j)} - \lambda_i^{(j)}| \, |l_j(w^*)|.$$

The first term is $\psi_i(w^*)$. For the second: $|\lambda_{\text{new}}^{(j)} - \lambda_i^{(j)}| \le \|\boldsymbol{\lambda}_{\text{new}} - \boldsymbol{\lambda}_i\| \le \delta_\lambda$ and $|l_j(w^*)| \le \|l(w^*)\|_\infty$. So $\psi_{\text{new}}(w^*) \le \psi_i(w^*) + \delta_\lambda \, \|l(w^*)\|_\infty$. Averaging over $i \in C_n$:

$$\psi_{\text{new}}(w^*) \le \frac{1}{|C_n|} \sum_i \psi_i(w^*) + \delta_\lambda \, \|l(w^*)\|_\infty = \psi_{C_n}(w^*) + \delta_\lambda \, \|l(w^*)\|_\infty. \qquad \square$$

Corollary H.37 (Suboptimality decomposition for an in-cluster client)**.** *Under Assumptions H.1–H.3, H.19 and H.34, the trajectory hypothesis equation 71, and $\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'}\| \le \delta_\lambda$ for all $i, i' \in C_n$, any client i whose preference lies in the cluster (a member of C_n, or a client sharing the cluster's objective functions with $\|\boldsymbol{\lambda}_i - \boldsymbol{\lambda}_{i'}\| \le \delta_\lambda$ for all $i' \in C_n$) satisfies the objective-space decomposition*

$$\psi_i(w^*) - \psi_i^* \ \le \ \underbrace{G_{\text{opt}}}_{\textit{training gap}} + \underbrace{\delta_\lambda \, \|l(w^*)\|_\infty}_{\textit{preference gap}} + \underbrace{\delta_\lambda \, L^*}_{\textit{optimum spread}}, \tag{79}$$

where $G_{\mathrm{opt}} := \psi_{C_n}(w^*) - \bar{\psi}^*$ *with* $\bar{\psi}^* = \frac{1}{|C_n|}\sum_i \psi_i^*$*, and* L^* *is a common upper bound on the loss vectors at all optima that are compared,*

$$L^* \;\geq\; \max_{i' \in C_n \cup \{i\}} \|l(w_{i'}^*)\|_\infty, \tag{80}$$

where i *is the client to which the corollary is applied (a member of* C_n*, or the new client of Section H.6, in which case* w_{new}^* *must be included);* L^* *is finite because every* $w_{i'}^*$ *lies in the bounded parameter region of Assumption H.1 on which the losses are continuous. Writing* $\bar{\epsilon}_0 := \frac{1}{|C_n|}\sum_{i\in C_n} \epsilon_{0,i} \leq \max_{i\in C_n} \epsilon_{0,i}$ *for the cluster-average band radius of equation 64 (for* $w^* = w_{\mathrm{II}}$*, of equation 74, i.e. with* $\eta = \max(\eta_{\mathrm{I}}, \eta_{\mathrm{II}})$*), the collaborative convergence band of Theorem H.31(2) (Corollary H.32 for* w_{II}*) controls the training gap,*

$$G_{\mathrm{opt}} \;\leq\; \bar{\epsilon}_0 \;=\; \frac{1}{|C_n|}\sum_{i\in C_n}\Big(\kappa_{H,i}\,\bar{\Delta}_i + \kappa_{\eta,i}\,\eta + \frac{b_i}{c_d}\Big). \tag{81}$$

Consequently, under the local quadratic growth of Assumption H.34,

$$\|w^* - w_i^*\|^2 \;\leq\; \frac{2}{\mu}\big(\psi_i(w^*) - \psi_i^*\big) \;\leq\; \frac{2}{\mu}\big(\bar{\epsilon}_0 + \delta_\lambda\,\|l(w^*)\|_\infty + \delta_\lambda\,L^*\big), \tag{82}$$

which is horizon-independent and, in the strict form $b_i = 0$*, vanishes in the limit* $\bar{\Delta}_i, \delta_\lambda, \eta \to 0$ *(subject to qualification (b) of Theorem H.31). In the restricted case in which the hypotheses of Corollary H.11 (Assumptions H.5 and H.6) hold at every iterate along the trajectory,* $\bar{\Delta}_i = O(\delta_\lambda)$ *and the right-hand sides are* $O(\delta_\lambda) + O(\eta)$*; otherwise* $\bar{\Delta}_i$ *is a measured quantity.*

Proof. Insert $\pm\psi_{C_n}(w^*)$ and $\pm\bar{\psi}^*$ into $\psi_i(w^*) - \psi_i^*$:

$$\psi_i(w^*) - \psi_i^* = \underbrace{\big[\psi_i(w^*) - \psi_{C_n}(w^*)\big]}_{\text{(a)}} + \underbrace{\big[\psi_{C_n}(w^*) - \bar{\psi}^*\big]}_{\text{(b)} = G_{\mathrm{opt}}} + \underbrace{\big[\bar{\psi}^* - \psi_i^*\big]}_{\text{(c)}}.$$

Term (a). By the per-objective preference bound of Proposition H.36, $\psi_i(w^*) \leq \psi_{i'}(w^*) + \delta_\lambda\|l(w^*)\|_\infty$ for any i'; averaging over i' gives $\psi_i(w^*) - \psi_{C_n}(w^*) \leq \delta_\lambda\|l(w^*)\|_\infty$. *Term (c).* We need an upper bound on $\psi_{i'}^* - \psi_i^*$ for every $i' \in C_n$. By optimality of $w_{i'}^*$ for $\psi_{i'}$ and the per-objective preference bound evaluated at w_i^*,

$$\psi_{i'}^* \;\leq\; \psi_{i'}(w_i^*) \;\leq\; \psi_i(w_i^*) + \delta_\lambda\|l(w_i^*)\|_\infty \;=\; \psi_i^* + \delta_\lambda\|l(w_i^*)\|_\infty,$$

so $\psi_{i'}^* - \psi_i^* \leq \delta_\lambda\|l(w_i^*)\|_\infty \leq \delta_\lambda L^*$ (the point at which the loss is evaluated is w_i^*, the optimum of the client being bounded, which is covered by equation 80). Averaging over i' gives $\bar{\psi}^* - \psi_i^* \leq \delta_\lambda L^*$. Note that the reverse inequality $\psi_i^* \leq \psi_{i'}^* + \delta_\lambda\|l(w_{i'}^*)\|_\infty$ controls the opposite difference and would not suffice here. *Term (b)* is G_{opt} by definition; this proves equation 79. Averaging the objective-space guarantee of Theorem H.31(2), $\psi_i(w^*) \leq \psi_i^* + \epsilon_{0,i}$, over $i \in C_n$ gives $G_{\mathrm{opt}} = \frac{1}{|C_n|}\sum_i(\psi_i(w^*) - \psi_i^*) \leq \bar{\epsilon}_0$, which is equation 81; it depends only on the trajectory bounds $\bar{\Delta}_i$ on the direction deviation, on the step size η and on the tolerances b_i, not on the number of rounds, so it is horizon-independent. Finally Assumption H.34 at $w = w^*$ gives $\|w^* - w_i^*\|^2 \leq \frac{2}{\mu}(\psi_i(w^*) - \psi_i^*)$; combining with equation 79 and equation 81 yields equation 82, and for $b_i = 0$ every term vanishes as $\bar{\Delta}_i, \delta_\lambda, \eta \to 0$. □

Remark H.38 (Reading the three terms of equation 79). The first term, the *training gap* G_{opt}, is the cluster-average of how far the shared model w^* sits above each member's own optimum; it is the only term that depends on the collaborative training, and it is bounded by the cluster-average band radius $\bar{\epsilon}_0$ of equation 81. The second term, the *preference gap* $\delta_\lambda\|l(w^*)\|_\infty$, measures how much client i's Chebyshev value $\psi_i(w^*)$ at the achieved model deviates from the cluster average $\psi_{C_n}(w^*)$ because its preference differs from the others' by at most δ_λ. The third term, the *optimum spread* $\delta_\lambda L^*$, measures how far the per-client optimal values $\psi_{i'}^*$ are spread across the cluster for the same reason; it is independent of the trained model. The decomposition holds for the output of the idealized shared iteration (one local or inner step per round, population gradients, the trajectory hypothesis equation 71), hence for both w_{I} and, through Corollary H.32, the returned w_{II}, with $\bar{\epsilon}_0$ computed from the two-phase radius equation 74; it is not a statement about the multi-step, S-shot implementation of Algorithm 1.

H.6 Data-Scarce Adaptation on the Initial ALigner: the δ_0 Gap

We now compare two ways of serving a data-scarce client whose preference $\boldsymbol{\lambda}_{\text{new}}$ falls in cluster C_n (by the adjustment-vector criterion of Section 3.2). Throughout this subsection the new client evaluates the *same* objective functions $l_1, \ldots, l_m$ as the cluster and differs from its members only in the preference vector; a new client whose loss functions themselves differ from the cluster's (a distribution shift in l) is outside the scope of the bounds below, which would then require an additional term controlling $l_{\text{new}}(w) - l_i(w)$ at the relevant points. This client *did not participate* in initializer training in either scenario, and both scenarios optimize the *same* balancing objective equation 5 under the *same* preference $\boldsymbol{\lambda}_{\text{new}}$; they differ only in how the model is produced.

- **Pipeline 1:** the new client starts from a shared starting point w_0 produced by the initializer training of Section 3.3 (in Algorithm 1, $w_0 = w_{\text{II}}$; both w_{I} and w_{II} are covered by Corollary H.37 via Corollary H.32) and runs the K-step fresh-sample adaptation of equation 85 with batch size S per step toward its own preference optimum, producing w_K.
- **Pipeline 2:** the client is assumed to have *abundant* data and *does not use collaborative learning at all*: it trains a model *locally* on the same balancing objective equation 5 with the same preference $\boldsymbol{\lambda}_{\text{new}}$ to convergence, reaching its own preference optimum

$$w^*_{\text{new}} \in \arg\min_w \psi_{\text{new}}(w), \qquad \psi^*_{\text{new}} := \psi_{\text{new}}(w^*_{\text{new}}), \qquad \psi_{\text{new}}(w) = \max_j \lambda^{(j)}_{\text{new}} l_j(w). \tag{83}$$

  With abundant data and training to convergence this pipeline attains $\psi_{\text{new}}(w^*_{\text{new}}) = \psi^*_{\text{new}}$ and, since w^*_{new} lies on the $\boldsymbol{\lambda}^{-1}_{\text{new}}$ ray, $\beta_{\text{new}}(w^*_{\text{new}}) = 0$.

To quantify what Pipeline 1 costs relative to this data-rich baseline, we track two quantities—the preference uniformity $\beta_i(w)$ (zero exactly on the $\boldsymbol{\lambda}_i^{-1}$ ray) and the Chebyshev indicator

$$\psi_i(w) := \max_{j \in [m]} \lambda_i^{(j)} l_j(w), \tag{84}$$

and bound the gap $\psi_{\text{new}}(w_K) - \psi_{\text{new}}(w^*_{\text{new}}) = \psi_{\text{new}}(w_K) - \psi^*_{\text{new}}$ and $\|w_K - w^*_{\text{new}}\|$. The adaptation bound is stated from an *arbitrary* starting point w_0 in the region of Assumption H.39, with $\Delta_0 := \psi_{\text{new}}(w_0) - \psi^*_{\text{new}}$. When w_0 is the output of the idealized shared iteration—w_{I}, or the point w_{II} that Algorithm 1 actually returns, which Corollary H.32 places on the same trajectory with the two-phase band radius equation 74—and $\boldsymbol{\lambda}_{\text{new}}$ lies in cluster C_n, Δ_0 is controlled by Corollary H.37. Objective (4) coincides with the balancing objective equation 5 whenever $\beta_i > 0$, so every step of Pipeline 1 drives the same (β, ψ) dynamics. We show the expected gap splits into an *initialization/finite-K* term proportional to Δ_0 and a *statistical* floor decaying as $O(1/S)$ (Proposition H.44); the rest of the subsection records the idealized adaptation model, the one-step contraction it rests on, and this gap.

The adaptation procedure runs K local preference-conditioned descent steps starting from w_0:

$$w_{\text{new},0} = w_0, \quad w_{\text{new},k+1} = w_{\text{new},k} - \alpha\, \tilde{d}_{\text{new}}(w_{\text{new},k}), \quad k = 0, \ldots, K-1, \tag{85}$$

where $\tilde{d}_{\text{new}}(w_{\text{new},k})$ is the stochastic preference-conditioned direction computed from a batch of S of the client's samples.

Modeling data scarcity: an idealized fresh-sample adaptation process. The direction $\tilde{d}_{\text{new}}(w)$ of equation 85 is the empirical preference-conditioned direction computed from a batch of S samples. The analysis below is carried out for an *idealized* preconditioned stochastic-gradient process with *fresh samples*—a new batch of S samples is drawn independently at every step, so K steps consume $K \cdot S$ samples—whose assumptions we now list separately. It is *not* a statistical guarantee for a fixed total budget of S samples reused across all K steps (the experimental setting with $S = 20$); analyzing that case would require an optimization analysis on the fixed dataset together with a generalization bound, which we do not provide. It is also *not* an analysis of the AdamW optimizer used in the experiments; Remark H.42 records the differences.

Assumption H.39 (Local smoothness and PL of the adaptation objective)**.** There are a region $\mathcal{U}$ containing the starting point w_0 and a constant $\mu_{\text{PL}} > 0$ such that (i) on $\mathcal{U}$ the Chebyshev objective ψ_{new} has a stable active bottleneck, so that it is differentiable and L_0-smooth with descent field $\nabla\psi_{\text{new}}$

and the population direction coincides with it, $d_{\mathrm{new}}(w) = \nabla\psi_{\mathrm{new}}(w)$; (ii) the Polyak–Łojasiewicz inequality

$$\|\nabla\psi_{\mathrm{new}}(w)\|^2 \;\geq\; 2\mu_{\mathrm{PL}}\big(\psi_{\mathrm{new}}(w) - \psi^*_{\mathrm{new}}\big) \qquad \forall\, w \in \mathcal{U}; \tag{86}$$

and (iii) all adaptation iterates $w_{\mathrm{new},0}, \ldots, w_{\mathrm{new},K}$ and the segments joining consecutive iterates lie in $\mathcal{U}$ (a trajectory hypothesis). Condition (ii) is a gradient-norm lower bound and is *different from* the quadratic growth of Assumption H.34: PL implies quadratic growth but not conversely, and Assumption H.34 alone does not supply equation 86. Likewise the descent-margin condition of Assumption H.19 concerns the LP direction d_i and does not replace equation 86 for $\nabla\psi_{\mathrm{new}}$.

Assumption H.40 (Idealized finite-sample adaptation direction)**.** The stochastic direction used at step k is conditionally unbiased for the population descent field and has variance shrinking with the sample size: for every $w \in \mathcal{U}$,

$$\mathbb{E}\big[\tilde{d}_{\mathrm{new}}(w) \,\big|\, w\big] = \nabla\psi_{\mathrm{new}}(w), \qquad \mathbb{E}\big[\|\tilde{d}_{\mathrm{new}}(w) - \nabla\psi_{\mathrm{new}}(w)\|^2 \,\big|\, w\big] \;\leq\; \frac{\sigma_0^2}{S}, \tag{87}$$

where σ_0^2 is the per-sample gradient variance. This models *fresh-sample adaptation with batch size S per step*: each adaptation step uses a new batch of S samples drawn independently from the client's population, so the K steps consume $K \cdot S$ samples in total. It is an idealization: if the *same* fixed S samples are reused at every step (a fixed total budget of S), the empirical gradient is a deterministic function of w and equation 87 does not follow from the mere availability of S samples; we do not claim that it does, and the bounds below are not fixed-budget statistical guarantees. "Sufficient data" corresponds to $S \to \infty$ (equivalently $\sigma_0^2/S \to 0$), the exact-gradient adaptation with direction d_{new}.

Assumption H.41 (Idealized preconditioned stochastic adaptation)**.** Instead of the plain update equation 85, the adaptation step is

$$w_{\mathrm{new},k+1} = w_{\mathrm{new},k} - \alpha\, P_k\, \tilde{d}_{\mathrm{new}}(w_{\mathrm{new},k}), \tag{88}$$

where, writing $\mathcal{F}_k$ for the σ-field generated by the iterate history up to $w_{\mathrm{new},k}$ (before the step-k samples are drawn):

(a) the preconditioner P_k is $\mathcal{F}_k$-measurable, i.e. it is fixed before the step-k stochastic direction is sampled;

(b) P_k is symmetric with uniformly bounded eigenvalues, $0 < p_-\, I \preceq P_k \preceq p_+\, I$;

(c) the update uses the instantaneous direction $\tilde{d}_{\mathrm{new}}(w_{\mathrm{new},k})$ (no momentum);

(d) there is no weight decay ($\lambda_{\mathrm{wd}} = 0$).

Remark H.42 (Relation to AdamW). The experiments adapt with AdamW. The idealized process of Assumption H.41 differs from it in three respects, each of which would require additional analysis that we do not carry out. *(1) Momentum.* AdamW replaces $\tilde{d}_{\mathrm{new}}(w_{\mathrm{new},k})$ by an exponential moving average of past stochastic gradients evaluated at past iterates; its conditional mean given $\mathcal{F}_k$ is in general *not* $\nabla\psi_{\mathrm{new}}(w_{\mathrm{new},k})$, so the "same mean, no larger variance" shortcut is not available and the descent step in Lemma H.43 cannot be taken as written. *(2) Preconditioner measurability.* AdamW's second-moment estimate v_k includes the current stochastic gradient, so $P_k = \mathrm{diag}(1/(\sqrt{v_k} + \epsilon))$ is correlated with $\tilde{d}_{\mathrm{new}}(w_{\mathrm{new},k})$ and is not $\mathcal{F}_k$-measurable; the factorization $\mathbb{E}[g_k^\top P_k \tilde{d}_{\mathrm{new}} \mid \mathcal{F}_k] = g_k^\top P_k g_k$ used in the proof then fails. Condition (a) holds for a variant that uses the previous estimate v_{k-1}. *(3) Decoupled weight decay.* AdamW's update is $w^+ = w - \alpha P_k g - \alpha\lambda_{\mathrm{wd}} w$, whereas the preconditioned step on the regularized objective $F(w) + \frac{\lambda_{\mathrm{wd}}}{2}\|w\|^2$ is $w^+ = w - \alpha P_k g - \alpha\lambda_{\mathrm{wd}} P_k w$. The two coincide only for $P_k \propto I$, so decoupled weight decay is *not* equivalent to running the same iteration on an L_2-regularized objective, and Lemma H.43 does not transfer to it by a change of target. Consequently the guarantees of this subsection are statements about the idealized process, and their relevance to the AdamW runs in Section 4 is empirical, not proved.

Lemma H.43 (Contraction of the idealized preconditioned stochastic adaptation)**.** *Under Assumptions H.39, H.40 and H.41, the updates equation 88 with step size $0 < \alpha \leq p_-/(L_0 p_+^2)$ satisfy the one-step contraction*

$$\mathbb{E}\big[\psi_{\mathrm{new}}(w_{\mathrm{new},k+1}) - \psi^*_{\mathrm{new}}\big] \;\leq\; (1 - \alpha\mu_{\mathrm{PL}} p_-)\, \mathbb{E}\big[\psi_{\mathrm{new}}(w_{\mathrm{new},k}) - \psi^*_{\mathrm{new}}\big] + \frac{L_0 \alpha^2 p_+^2}{2} \cdot \frac{\sigma_0^2}{S}, \tag{89}$$

and hence, telescoping from $w_{\mathrm{new},0} = w_0$,

$$\mathbb{E}\big[\psi_{\mathrm{new}}(w_{\mathrm{new},K}) - \psi^*_{\mathrm{new}}\big] \;\le\; (1 - \alpha\mu_{\mathrm{PL}}p_-)^K\big(\psi_{\mathrm{new}}(w_0) - \psi^*_{\mathrm{new}}\big) + \frac{L_0\alpha p_+^2}{2\mu_{\mathrm{PL}}p_-}\cdot\frac{\sigma_0^2}{S}. \tag{90}$$

Both bounds hold in expectation over the sampling of the adaptation directions; they are not per-run statements.

Proof. Write $F := \psi_{\mathrm{new}}$, $g_k := \nabla F(w_{\mathrm{new},k})$. By Assumption H.39(i),(iii) F is L_0-smooth along the segment from $w_{\mathrm{new},k}$ to $w_{\mathrm{new},k+1}$, so the descent lemma applied to equation 88 gives

$$F(w_{\mathrm{new},k+1}) \le F(w_{\mathrm{new},k}) - \alpha\, g_k^\top P_k\, \tilde{d}_{\mathrm{new}} + \frac{L_0\alpha^2}{2}\|P_k\,\tilde{d}_{\mathrm{new}}\|^2.$$

Take conditional expectation given $\mathcal{F}_k$. Since P_k is $\mathcal{F}_k$-measurable (Assumption H.41(a)) it is constant under this expectation, so by unbiasedness equation 87 $\mathbb{E}[g_k^\top P_k\,\tilde{d}_{\mathrm{new}} \mid \mathcal{F}_k] = g_k^\top P_k g_k \ge p_-\|g_k\|^2$ (using $P_k \succeq p_- I$), and by $P_k \preceq p_+ I$ together with the bias–variance split, $\mathbb{E}[\|P_k\,\tilde{d}_{\mathrm{new}}\|^2 \mid \mathcal{F}_k] \le p_+^2\,\mathbb{E}[\|\tilde{d}_{\mathrm{new}}\|^2 \mid \mathcal{F}_k] \le p_+^2(\|g_k\|^2 + \sigma_0^2/S)$. Therefore

$$\mathbb{E}\big[F(w_{\mathrm{new},k+1}) \mid \mathcal{F}_k\big] \le F(w_{\mathrm{new},k}) - \alpha\Big(p_- - \tfrac{L_0\alpha p_+^2}{2}\Big)\|g_k\|^2 + \frac{L_0\alpha^2 p_+^2}{2}\cdot\frac{\sigma_0^2}{S}.$$

Since $\alpha \le p_-/(L_0 p_+^2)$ we have $p_- - \frac{L_0\alpha p_+^2}{2} \ge \frac{p_-}{2}$, so $-\alpha(p_- - \frac{L_0\alpha p_+^2}{2})\|g_k\|^2 \le -\frac{\alpha p_-}{2}\|g_k\|^2$. Applying the PL inequality equation 86 of Assumption H.39(ii) at $w_{\mathrm{new},k} \in \mathcal{U}$, $\|g_k\|^2 \ge 2\mu_{\mathrm{PL}}(F(w_{\mathrm{new},k}) - F^*)$, gives $-\frac{\alpha p_-}{2}\|g_k\|^2 \le -\alpha\mu_{\mathrm{PL}}p_-(F(w_{\mathrm{new},k}) - F^*)$; subtracting F^* from both sides and taking total expectation yields equation 89. Iterating equation 89 and summing the geometric series,

$$\begin{aligned}\mathbb{E}[F(w_{\mathrm{new},K}) - F^*] &\le (1 - \alpha\mu_{\mathrm{PL}}p_-)^K(F(w_0) - F^*) + \frac{L_0\alpha^2 p_+^2}{2}\cdot\frac{\sigma_0^2}{S}\sum_{j=0}^{K-1}(1 - \alpha\mu_{\mathrm{PL}}p_-)^j \\ &\le (1 - \alpha\mu_{\mathrm{PL}}p_-)^K(F(w_0) - F^*) + \frac{L_0\alpha p_+^2}{2\mu_{\mathrm{PL}}p_-}\cdot\frac{\sigma_0^2}{S},\end{aligned}$$

using $\sum_{j\ge0}(1 - \alpha\mu_{\mathrm{PL}}p_-)^j = 1/(\alpha\mu_{\mathrm{PL}}p_-)$. This is equation 90. □

Proposition H.44 (Gap: initializer based few-shot adaptation vs. data-rich local training)**.** *Under Assumptions H.1–H.3, H.34 (applied to ψ_{new} at w^*_{new}), H.39, H.40 and H.41, with $0 < \alpha \le p_-/(L_0 p_+^2)$, consider a new client that evaluates the same objective functions l_j as the cluster, whose preference $\boldsymbol{\lambda}_{\mathrm{new}}$ lies in cluster C_n ($\|\boldsymbol{\lambda}_{\mathrm{new}} - \boldsymbol{\lambda}_i\| \le \delta_\lambda$ for $i \in C_n$) and that did not participate in initializer training. Let*

- *w_K be the* Pipeline 1 *output: K idealized fresh-sample preconditioned stochastic steps equation 88 with batch size S per step, started from a point w_0 in the region $\mathcal{U}$ of Assumption H.39; and*
- *$w^{(2)} := w^*_{\mathrm{new}}$ be the* Pipeline 2 *output: the data-rich local model that trains equation 5 to convergence with the same preference, so $\psi_{\mathrm{new}}(w^{(2)}) = \psi^*_{\mathrm{new}}$ and $\beta_{\mathrm{new}}(w^{(2)}) = 0$.*

Write $\psi^*_{\mathrm{new}} = \min_w \psi_{\mathrm{new}}(w)$, denote the starting-point suboptimality by $\Delta_0 = \Delta_0(w_0) := \psi_{\mathrm{new}}(w_0) - \psi^*_{\mathrm{new}}$, and assume β_{new} is locally L_β-Lipschitz in the loss vector on the adaptation region. Then:

(a) Chebyshev indicator. *Pipeline 1 obeys, in expectation over the sampling of the adaptation directions,*

$$\mathbb{E}\big[\psi_{\mathrm{new}}(w_K) - \psi^*_{\mathrm{new}}\big] \;\le\; (1 - \alpha\mu_{\mathrm{PL}}p_-)^K\,\Delta_0(w_0) + \frac{L_0\alpha p_+^2}{2\mu_{\mathrm{PL}}p_-}\cdot\frac{\sigma_0^2}{S}, \tag{91}$$

while Pipeline 2 attains $\psi_{\mathrm{new}}(w^{(2)}) - \psi^*_{\mathrm{new}} = 0$. For a starting point $w_0 \in \{w_{\mathrm{I}}, w_{\mathrm{II}}\}$ produced by the idealized shared iteration (Section H.5; w_{II} is covered by Corollary H.32), if the hypotheses of Corollary H.37 hold with L^* covering w^*_{new} as in equation 80, then

$$\Delta_0(w_0) \;\le\; G_{\mathrm{opt}} + \delta_\lambda\|l(w_0)\|_\infty + \delta_\lambda L^* \;\le\; \bar{\epsilon}_0 + \delta_\lambda\big(\|l(w_0)\|_\infty + L^*\big), \tag{92}$$

where for $w_0 = w_{\mathrm{II}}$ the cluster-average band radius $\bar{\epsilon}_0$ is that of the two-phase radius equation 74 (step size $\max(\eta_{\mathrm{I}}, \eta_{\mathrm{II}})$). In particular the bound applies to the point Algorithm 1 actually returns, within the idealization of Corollary H.32 ($K = 1$ inner step, population gradients, trajectory hypothesis on the Phase-2 rounds).

(b) Parameter space. *If, in addition, w_K lies in the neighbourhood of w^*_{new} where the local quadratic growth $\frac{\mu}{2}\|w_K - w^*_{\mathrm{new}}\|^2 \le \psi_{\mathrm{new}}(w_K) - \psi^*_{\mathrm{new}}$ of Assumption H.34 holds, then*

$$\mathbb{E}\|w_K - w^*_{\mathrm{new}}\|^2 \;\le\; \frac{2}{\mu}\Big[(1 - \alpha\mu_{\mathrm{PL}}p_-)^K\,\Delta_0 + \frac{L_0\alpha p_+^2}{2\mu_{\mathrm{PL}}p_-}\cdot\frac{\sigma_0^2}{S}\Big], \qquad \|w^{(2)} - w^*_{\mathrm{new}}\| = 0, \tag{93}$$

hence, by Jensen's inequality, $\mathbb{E}\|w_K - w^*_{\mathrm{new}}\| \le \big(\mathbb{E}\|w_K - w^*_{\mathrm{new}}\|^2\big)^{1/2}$, and the preference uniformity of Pipeline 1 satisfies $\mathbb{E}\,\beta_{\mathrm{new}}(w_K) \le L_\beta\,\mathbb{E}\|w_K - w^*_{\mathrm{new}}\|$ while $\beta_{\mathrm{new}}(w^{(2)}) = 0$. These are mean-square and mean statements; no per-run bound on $\|w_K - w^*_{\mathrm{new}}\|$ is claimed.

(c) Gap. *The expected suboptimality of the collaborative data-scarce pipeline relative to data-rich local training splits into two independent sources,*

$$\mathbb{E}[\Delta\psi] := \mathbb{E}\big[\psi_{\mathrm{new}}(w_K) - \psi_{\mathrm{new}}(w^{(2)})\big] \;\le\; \underbrace{(1 - \alpha\mu_{\mathrm{PL}}p_-)^K\,\Delta_0}_{\textit{initialization + finite-}K} \;+\; \underbrace{\frac{L_0\alpha p_+^2}{2\mu_{\mathrm{PL}}p_-}\cdot\frac{\sigma_0^2}{S}}_{\textit{statistical scarcity}}, \tag{94}$$

and $\mathbb{E}[\Delta\beta] := \mathbb{E}\big[\beta_{\mathrm{new}}(w_K) - \beta_{\mathrm{new}}(w^{(2)})\big] = O\big(\sqrt{(1 - \alpha\mu_{\mathrm{PL}}p_-)^K\Delta_0} + \sqrt{\sigma_0^2/S}\big)$. The initialization term depends on the starting point only through $\Delta_0(w_0)$ and decays geometrically in K. For $w_0 \in \{w_{\mathrm{I}}, w_{\mathrm{II}}\}$ its prefactor is controlled by equation 92, which is independent of the round counts R, R' (horizon-independent by Corollary H.37) but whose $\bar{\epsilon}_0$ contains the tolerance term $\frac{1}{|C_n|}\sum_i b_i/c_d$ that is not shown to be small (Theorem H.31), so $\Delta_0(w_0) \to 0$ as $\bar{\Delta}_i, \delta_\lambda, \eta_{\mathrm{I}}, \eta_{\mathrm{II}} \to 0$ is asserted only in the strict form $b_i = 0$. The statistical term vanishes as $S \to \infty$.

Proof. Every adaptation step of Pipeline 1 is an idealized preconditioned stochastic step on the Chebyshev objective ψ_{new} with the balancing direction of equation 5, whose finite-sample error obeys Assumption H.40. Applying the one-step contraction of Lemma H.43 with $F = \psi_{\mathrm{new}}$ over the K adaptation steps started from $w_{\mathrm{new},0} = w_0$ and telescoping the geometric series gives equation 90 with $F(w_0) - F^* = \psi_{\mathrm{new}}(w_0) - \psi^*_{\mathrm{new}} = \Delta_0(w_0)$, which is equation 91. For $w_0 \in \{w_{\mathrm{I}}, w_{\mathrm{II}}\}$: because the new client shares the cluster's objective functions and $\boldsymbol{\lambda}_{\mathrm{new}}$ lies in cluster C_n, Corollary H.37 applied to $i = \mathrm{new}$ (so that L^* covers w^*_{new}, cf. equation 80; Term (c) of its proof then reads $\psi^* - \psi^*_{\mathrm{new}} \le \delta_\lambda\|l(w^*_{\mathrm{new}})\|_\infty \le \delta_\lambda L^*$) bounds $\Delta_0(w_0)$ exactly as in equation 79–equation 81, giving equation 92; for $w_0 = w_{\mathrm{II}}$ the training-gap bound equation 81 is supplied by Corollary H.32 with the two-phase radius equation 74. Pipeline 2 trains equation 5 locally with abundant data ($\sigma_0^2/S = 0$) to convergence, so its transient vanishes and it attains $\psi_{\mathrm{new}}(w^{(2)}) = \psi^*_{\mathrm{new}}$, i.e. $w^{(2)} = w^*_{\mathrm{new}}$. Subtracting the deterministic value $\psi_{\mathrm{new}}(w^{(2)}) = \psi^*_{\mathrm{new}}$ inside the expectation in equation 91 gives equation 94: an initialization/finite-K term and a statistical floor.

For (b), local quadratic growth (Assumption H.34) converts the ψ_{new}-suboptimality of the scarce run into a parameter-space deviation, $\frac{\mu}{2}\|w_K - w^*_{\mathrm{new}}\|^2 \le \psi_{\mathrm{new}}(w_K) - \psi^*_{\mathrm{new}}$; taking expectations and substituting equation 91 gives equation 93, and $\|w^{(2)} - w^*_{\mathrm{new}}\| = 0$ since Pipeline 2 reaches the optimum. Note that the adaptation begins at $w_{\mathrm{new},0} = w_0$, where Assumption H.34 gives the starting distance $\|w_0 - w^*_{\mathrm{new}}\|^2 \le \frac{2}{\mu}\Delta_0(w_0)$ (for $w_0 = w_{\mathrm{I}}$ this is equation 82 of Corollary H.37); equation 93 is exactly this starting distance after the $(1 - \alpha\mu_{\mathrm{PL}}p_-)^K$ contraction, plus the statistical floor. Because w^*_{new} lies on the $\boldsymbol{\lambda}^{-1}_{\mathrm{new}}$ ray, $\beta_{\mathrm{new}}(w^*_{\mathrm{new}}) = 0$; Lipschitzness of β_{new} in the loss vector (hence, through L_0-smoothness, in w) then gives $\beta_{\mathrm{new}}(w_K) \le L_\beta\|w_K - w^*_{\mathrm{new}}\|$ pointwise, and taking expectations together with Jensen's inequality $\mathbb{E}\|w_K - w^*_{\mathrm{new}}\| \le (\mathbb{E}\|w_K - w^*_{\mathrm{new}}\|^2)^{1/2}$ and equation 93 yields $\mathbb{E}[\Delta\beta] = O(\sqrt{(1 - \alpha\mu_{\mathrm{PL}}p_-)^K\Delta_0} + \sqrt{\sigma_0^2/S})$. For $w_0 \in \{w_{\mathrm{I}}, w_{\mathrm{II}}\}$ the initialization term inherits horizon-independence from Corollary H.37 (its bound depends on the cluster's $\bar{\Delta}_i, \delta_\lambda, b_i$ and the step sizes, not on the round counts R, R'), completing (c). □

Remark H.45 (Local adaptation contracts the initializer gap). This makes precise how Pipeline 1 closes the gap that the subsection "The Initializer Approximates Each Client's Individual Optimum"

had already measured at the output of the shared iteration. *Before adaptation*, when the starting point is $w_0 \in \{w_{\mathrm{I}}, w_{\mathrm{II}}\}$, Corollary H.37 (with Corollary H.32 for w_{II}) bounds the gap to the new client's own optimum w^*_{new} in both spaces: in objective space

$$\psi_{\mathrm{new}}(w_0) - \psi^*_{\mathrm{new}} \;=\; \Delta_0(w_0) \;\le\; G_{\mathrm{opt}} + \delta_\lambda \|l(w_0)\|_\infty + \delta_\lambda L^* \qquad \text{(Eq. equation 79)},$$

and in parameter space $\|w_0 - w^*_{\mathrm{new}}\|^2 \le \frac{2}{\mu}\Delta_0(w_0)$ (Eq. equation 82). *The adaptation does not create this gap—it inherits it as the initial value and shrinks it.* Writing the objective gap before and after the K local steps,

$$\underbrace{\psi_{\mathrm{new}}(w_{\mathrm{new},0}) - \psi^*_{\mathrm{new}} = \Delta_0(w_0)}_{\text{start (at } w_0)} \;\xrightarrow{K \text{ steps}}\; \underbrace{\mathbb{E}\big[\psi_{\mathrm{new}}(w_K) - \psi^*_{\mathrm{new}}\big] \le (1 - \alpha\mu_{\mathrm{PL}} p_-)^K \Delta_0 + \frac{L_0 \alpha p_+^2}{2\mu_{\mathrm{PL}} p_-}\frac{\sigma_0^2}{S}}_{\text{after (Eq. equation 91)}},$$

each step multiplies the expected population gap by the contraction factor $1 - \alpha\mu_{\mathrm{PL}} p_- \in (0,1)$, so K steps reduce the initialization gap by $(1 - \alpha\mu_{\mathrm{PL}} p_-)^K$ down to the statistical floor. Equivalently, the (root-mean-square) distance to the true optimum contracts from $\|w_0 - w^*_{\mathrm{new}}\| \le \sqrt{\frac{2}{\mu}\Delta_0}$ to $\mathbb{E}\|w_K - w^*_{\mathrm{new}}\| \le \big(\mathbb{E}\|w_K - w^*_{\mathrm{new}}\|^2\big)^{1/2} \le \sqrt{\frac{2}{\mu}\big[(1 - \alpha\mu_{\mathrm{PL}} p_-)^K \Delta_0 + \frac{L_0 \alpha p_+^2}{2\mu_{\mathrm{PL}} p_-}\frac{\sigma_0^2}{S}\big]}$ (Eq. equation 93). Since Pipeline 2 sits exactly at w^*_{new} (zero gap, $\beta_{\mathrm{new}} = 0$), the Pipeline 1 vs. Pipeline 2 gap $\mathbb{E}[\Delta\psi]$ of equation 94 is precisely the *residual* of this contraction: the starting point sets *where* the contraction starts ($\Delta_0(w_0)$, bounded by Cor. H.37 in terms of cluster tightness, step sizes and descent tolerance for $w_0 \in \{w_{\mathrm{I}}, w_{\mathrm{II}}\}$), and local adaptation sets *how far* it travels (the factor $(1 - \alpha\mu_{\mathrm{PL}} p_-)^K$), leaving only the irreducible $O(\sigma_0^2/S)$ scarcity floor.

Remark H.46 (Interpretation). The gap in equation 94 isolates the two prices Pipeline 1 pays relative to a data-rich client that self-trains to its own optimum; viewed through Remark H.45, these are the two components of the residual after adaptation contracts the initializer gap. *(i) An initialization/finite-K term* $(1 - \alpha\mu_{\mathrm{PL}} p_-)^K \Delta_0$. A data-scarce client cannot self-train to w^*_{new}; it instead borrows the shared starting point w_0 and refines it with only K adaptation steps. For $w_0 \in \{w_{\mathrm{I}}, w_{\mathrm{II}}\}$ the prefactor $\Delta_0 \le \bar{\epsilon}_0 + \delta_\lambda(\|l(w_0)\|_\infty + L^*)$ is controlled by the cluster tightness ($\bar{\Delta}_i, \delta_\lambda$), by the step sizes ($\eta_{\mathrm{I}}$, and η_{II} for w_{II} via Corollary H.32), and by the descent-tolerance term $\frac{1}{|C_n|}\sum_i b_i/c_d$ inside $\bar{\epsilon}_0$; the last term has no proved vanishing rate (Theorem H.31), so we do not claim that Δ_0 is small, only that it is explicitly bounded under the stated conditions (Corollary H.37), which for w_{II} include the one-inner-step, population-gradient idealization of Phase 2. Whatever its size, this term decays geometrically in K. *(ii) A statistical floor* $O(\sigma_0^2/S)$ from adapting on only S samples. Pipeline 2 pays neither: it has abundant data (no floor) and optimizes its own objective directly (no initialization gap), reaching ψ^*_{new} with $\beta_{\mathrm{new}} = 0$. The expected gap therefore vanishes as $K \to \infty$, $S \to \infty$, and, in the strict form $b_i = 0$, as the cluster tightens ($\bar{\Delta}_i, \delta_\lambda \to 0$ with $\eta \to 0$)—i.e. once $S \gtrsim L_0 \alpha p_+^2 \sigma_0^2/(\mu_{\mathrm{PL}}^2 p_- \varepsilon)$, K is a few dozen adaptation steps, and H_{C_n}, δ_λ are small. The same two terms govern the parameter space: by Proposition H.44(b), $\mathbb{E}\|w_K - w^*_{\mathrm{new}}\| = O\big(\sqrt{(1 - \alpha\mu_{\mathrm{PL}} p_-)^K \Delta_0} + \sqrt{\sigma_0^2/S}\big)$, so in expectation the collaborative data-scarce client matches the data-rich local optimum up to an initialization term and an $O(\sqrt{1/S})$ deviation. All of this is for the idealized adaptation process of Assumption H.41; see Remark H.42 for the gap to AdamW.