REPRESENTATION-CORRECT MULTI-OBJECTIVE LORA

Anonymous authors
Paper under double-blind review

ABSTRACT

Modern post-training often requires balancing multiple objectives, such as helpfulness and safety, while low-rank adaptation (LoRA) provides a parameter-efficient mechanism for model adaptation. However, LoRA factors are non-identifiable: infinitely many factor pairs represent the same effective update. We show that this redundancy makes multi-objective optimization representation-dependent. Equivalent factorizations of the same adapter can change cross-objective gradient geometry, even reversing conflict signs and altering the Pareto weights selected by gradient-based methods. Moreover, even when the same effective combined gradient is fixed, updating equivalent raw factors can produce different next effective models. Thus, representation dependence arises separately in both multi-objective decision making and update execution. We formalize *representation correctness* for the complete optimizer and give conditions under which equivalent LoRA representations induce identical effective trajectories. Based on this principle, we develop *Quotient Pareto Adaptation* (QPA), which performs Pareto optimization directly in the effective fixed-rank geometry. We further derive *Stiefel Pareto Adaptation* (SPA) as a complementary representation-correct construction under a distinct geometry. Both extend to transported first-moment momentum with pathwise guarantees. We evaluate QPA and SPA on helpfulness-safety Pareto frontiers across Qwen and Llama models and on external safety benchmarks.

1 INTRODUCTION

Language-model post-training increasingly requires balancing objectives that are individually desirable but not always aligned, such as improving helpfulness while preserving safety (Bai et al., 2022; Ouyang et al., 2022; Dai et al., 2024). Multi-objective optimization provides a natural way to manage such trade-offs by using the geometry of objective gradients to select a compromise update (Désidéri, 2012; Sener & Koltun, 2018). At the same time, low-rank adaptation (LoRA) has become a standard parameter-efficient mechanism for adapting large models by representing each weight update through two small factors (Hu et al., 2022). Their combination raises a basic requirement: if two LoRA factorizations represent exactly the same effective model, they should induce the same multi-objective optimization behavior.

This requirement is not automatically satisfied. For an effective LoRA update $Z = cBA$, every invertible matrix Q yields an equivalent factorization $(BQ, Q^{-1}A)$ with exactly the same Z. These different factors are therefore only alternative coordinates for the same model update. Nevertheless, under the standard Euclidean geometry of the factors, equivalent representations can induce different multi-objective decisions. In particular, a change of factorization can alter cross-objective gradient inner products, even reversing their signs, and can therefore change the Pareto weights selected by gradient-based multi-objective methods. Thus, two representations of the same model can disagree on whether objectives conflict and on how that conflict should be resolved.

The problem does not end with gradient combination. Even if the same effective combined direction is fixed, applying the corresponding update directly to different but equivalent factor pairs can produce different next effective matrices. Hence, representation dependence arises at two distinct stages: *decision*, where objectives are compared and combined, and *execution*, where the chosen update is realized in the LoRA factors. Correcting the former alone is therefore insufficient. For

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.

stateful optimizers, the same issue extends across time: optimizer memory and its transport must also carry a representation-consistent meaning if equivalent runs are to remain equivalent.

We therefore formulate *representation correctness* for the complete optimizer, requiring equivalent LoRA representations to induce identical effective trajectories. Building on this principle, we develop *Quotient Pareto Adaptation* (QPA), which performs Pareto optimization directly in the effective fixed-rank geometry. We also derive *Stiefel Pareto Adaptation* (SPA) as a complementary representation-correct construction under a distinct geometry. Together, these results show that representation correctness constrains admissible optimizers without prescribing a unique optimization geometry.

Our main contributions are:

- **Representation-dependent multi-objective LoRA.** We show that equivalent factorizations can alter gradient conflicts and Pareto decisions, while raw factor execution can yield different next effective models even for the same effective direction, revealing distinct *decision* and *execution* failure modes.
- **Complete-map representation correctness.** We formulate correctness jointly over decision making, execution, control flow, and optimizer state, and give sufficient conditions for equivalent representations to follow identical effective trajectories.
- **Representation-correct Pareto adaptation.** We develop *Quotient Pareto Adaptation* (QPA), which optimizes directly in the effective fixed-rank geometry, and derive *Stiefel Pareto Adaptation* (SPA) as a complementary construction under a distinct geometry. Both admit stateless and transported-momentum variants with pathwise representation guarantees.

2 Theory of Representation-Correct Multi-Objective LoRA

2.1 Setup and representation equivalence

Low-rank adaptation (LoRA) represents a weight update through two low-dimensional factors and is a widely used form of parameter-efficient adaptation (Hu et al., 2022; He et al., 2022). Consider K differentiable objectives

$$F_k(B, A) = \widetilde{F}_k(Z),$$

where a LoRA module represents the effective update

$$Z = cBA, \quad B \in \mathbb{R}^{m\times r}, \quad A \in \mathbb{R}^{r\times n}, \quad c > 0. \tag{1}$$

Throughout the theory, we assume that B has full column rank and A has full row rank. Under this assumption, two factor pairs represent the same effective update if and only if they differ by an invertible change of basis:

$$(B, A) \sim (BQ, Q^{-1}A), \quad Q \in \mathrm{GL}(r), \tag{2}$$

where $\sim$ denotes equivalence of factor representations with the same product BA, and $\mathrm{GL}(r)$ is the group of invertible $r \times r$ real matrices.

Thus, the factors are non-identifiable coordinates for the same effective adapter Z. Prior work has shown that factor parameterization, balancing, and transformation-aware updates can materially affect LoRA optimization (Hayou et al., 2024; Liu et al., 2024; Yen et al., 2025; Castin et al., 2026). Our question is whether multi-objective optimization is well defined at the level of the effective adapter: do equivalent factorizations produce the same objective weights and the same next effective update?

For objective k, let $G_k = \nabla_Z \widetilde{F}_k(Z)$ denote its effective gradient. Differentiating through $Z = cBA$ gives the ordinary factor gradients

$$G_{B,k} = cG_kA^\top, \quad G_{A,k} = cB^\top G_k. \tag{3}$$

Under the transformation in Equation (2), these gradients transform as

$$G_{B,k} \mapsto G_{B,k}Q^{-\top}, \quad G_{A,k} \mapsto Q^\top G_{A,k}. \tag{4}$$

We use the multiple-gradient descent algorithm (MGDA) as a canonical multi-objective decision rule (Fliege & Svaiter, 2000; Désidéri, 2012; Sener & Koltun, 2018). Writing $y_k = (G_{B,k}, G_{A,k})$, raw-factor MGDA solves

$$\lambda^* \in \arg\min_{\lambda \in \Delta_K} \lambda^\top H \lambda, \quad H_{ij} = \langle G_{B,i}, G_{B,j} \rangle_F + \langle G_{A,i}, G_{A,j} \rangle_F, \tag{5}$$

where $\Delta_K = \{\lambda \geq 0 : \mathbf{1}^\top \lambda = 1\}$.

The resulting weights define the combined factor gradient

$$y_{\lambda^*} = \sum_{k=1}^{K} \lambda_k^* y_k. \tag{6}$$

Equivalently, define the combined effective gradient

$$G_{\lambda^*} = \sum_{k=1}^{K} \lambda_k^* G_k. \tag{7}$$

By linearity of Equation (3), its corresponding factor gradients are

$$G_{B,\lambda^*} = c G_{\lambda^*} A^\top, \quad G_{A,\lambda^*} = c B^\top G_{\lambda^*}. \tag{8}$$

Thus λ^* determines both the effective gradient combination and the factor-space directions used for the update. Non-unique minimizers are resolved by minimum Euclidean norm within the minimizer set. For multiple LoRA modules indexed by ℓ, we use the global Gram matrix

$$H_{ij} = \sum_{\ell} \left(\langle G_{B,i}^{(\ell)}, G_{B,j}^{(\ell)} \rangle_F + \langle G_{A,i}^{(\ell)}, G_{A,j}^{(\ell)} \rangle_F \right), \tag{9}$$

and a single shared objective-weight vector λ^* is applied across all modules.

2.2 Representation-dependent decision

Many multi-objective methods use the relative geometry of objective gradients to determine how competing objectives should be combined (Désidéri, 2012; Sener & Koltun, 2018; Yu et al., 2020; Liu et al., 2021a; Navon et al., 2022). For LoRA, however, the equivalence in Equation (2) preserves the effective model exactly while generally changing the product-Euclidean geometry of its factor gradients.

Theorem 1 (Representation-dependent Pareto geometry)**.** *Let $P = QQ^\top, Q \in \mathrm{GL}(r)$. Under the equivalent factorization $(BQ, Q^{-1}A)$, the raw product-Euclidean Gram matrix becomes*

$$H_{ij}(Q) = \mathrm{tr}\left(P^{-1} G_{B,i}^\top G_{B,j}\right) + \mathrm{tr}\left(P G_{A,j} G_{A,i}^\top\right). \tag{10}$$

Hence the Gram matrix is invariant under orthogonal changes of basis, but is generally not invariant under the full $\mathrm{GL}(r)$ equivalence of LoRA factors. Therefore equivalent representations of the same effective adapter may induce different finite MGDA decisions.

The distortion is already visible under the scalar gauge $Q = sI, s > 0$, which replaces (B, A) by $(sB, s^{-1}A)$ without changing Z. Define

$$\alpha_{ij} = \langle G_{B,i}, G_{B,j} \rangle_F, \quad \beta_{ij} = \langle G_{A,i}, G_{A,j} \rangle_F. \tag{11}$$

Then Equation (10) reduces to

$$H_{ij}(s) = s^{-2} \alpha_{ij} + s^2 \beta_{ij}. \tag{12}$$

Corollary 2 (Scaling can reverse objective conflict)**.** *If $\alpha_{12}\beta_{12} < 0$, then $H_{12}(s)$ has the unique positive zero*

$$s_\star = \left(-\frac{\alpha_{12}}{\beta_{12}} \right)^{1/4}, \tag{13}$$

and has opposite signs on the two sides of $s_\star$. Thus two factorizations of exactly the same effective adapter can disagree on whether the same pair of objectives is aligned or conflicting.

Because MGDA determines its weights from this Gram matrix, the same effective model can consequently receive different Pareto weights under different factor representations. Unlike methods that explicitly adapt the optimization metric, such as MAdam (Liu et al., 2026), this ambiguity arises even under a fixed raw product-Euclidean rule, solely from equivalent factor representations.

The dependence nevertheless has a sharp boundary.

Proposition 3 (Exact Pareto stationarity is representation-invariant)**.** *Let* $y_k(Q) = \left(G_{B,k}Q^{-\top}, Q^\top G_{A,k}\right), Q \in \mathrm{GL}(r)$. *For every* $\lambda \in \Delta_K$,

$$\sum_{k=1}^{K} \lambda_k y_k(Q) = 0 \quad \Longleftrightarrow \quad \sum_{k=1}^{K} \lambda_k y_k(I) = 0. \tag{14}$$

Hence the simplex set producing an exactly zero combined factor gradient is identical across all equivalent full-rank factorizations.

The distinction is therefore precise: exact first-order Pareto stationarity is representation-invariant, whereas the nonzero gradient geometry used to choose a Pareto update generally is not. Thus equivalent factorizations agree on exact stationarity, but can make different finite Pareto decisions away from stationarity. Raw-factor MGDA therefore does not define a unique multi-objective decision for a given effective adapter. We next show that correcting this decision alone is still insufficient because update execution has a separate representation dependence.

2.3 Representation-dependent execution

The previous subsection shows that equivalent factorizations can produce different Pareto decisions. This alone does not capture the full representation dependence. Transformation-aware LoRA optimizers such as LoRA-RITE address representation dependence in update execution (Yen et al., 2025). Here we isolate this issue as the second stage of a broader multi-objective failure, after the Pareto decision has already been fixed, so that all equivalent factorizations are assigned the same effective combined gradient

$$G_\lambda = \sum_{k=1}^{K} \lambda_k G_k. \tag{15}$$

We now ask whether executing this same effective combined gradient through the raw LoRA factors produces the same next effective adapter.

Consider the scalar-equivalent representatives $(sB, s^{-1}A), s > 0$. For the fixed effective gradient G_λ, their corresponding factor gradients are

$$G_B^{(s)} = cs^{-1}G_\lambda A^\top, \quad G_A^{(s)} = csB^\top G_\lambda. \tag{16}$$

Theorem 4 (Representation-dependent execution)**.** *For each representative* $(sB, s^{-1}A)$*, apply the same Euclidean SGD rule with step size* η *to both factor blocks:*

$$B_s^+ = sB - \eta G_B^{(s)}, \quad A_s^+ = s^{-1}A - \eta G_A^{(s)}.$$

The resulting effective update is

$$\begin{aligned} Z_s^+ = Z - \eta c^2 \left(s^{-2}G_\lambda A^\top A + s^2 BB^\top G_\lambda\right) \\ + \eta^2 c^3 G_\lambda A^\top B^\top G_\lambda. \end{aligned} \tag{17}$$

Hence Z_s^+ *generally depends on* s*, even though both the starting effective adapter* Z *and the effective combined gradient* G_λ *are identical.*

This isolates execution as a separate source of representation dependence. Even when equivalent factorizations are forced to use the same effective combined gradient, raw factor-space execution can produce different next effective models. Representation dependence therefore arises at two distinct stages: the *decision* stage determines how objectives are combined, while the *execution* stage determines how the chosen combination is realized as a new effective adapter.

Correcting the decision stage alone is therefore insufficient. A representation-correct optimizer must control the complete optimization step rather than gradient combination alone. We formalize this requirement next.

2.4 Representation correctness of the complete optimizer

An optimizer is *representation-correct* if equivalent initial factorizations, with compatibly transformed optimizer states and shared randomness, produce identical effective trajectories. A sufficient condition is equivariance of the decision, execution, and state-update maps, together with invariant scalar controls and branch choices. Appendix B gives the formal definition, complete-map criterion, and proof.

3 Representation-Correct Pareto Adaptation

We now instantiate representation correctness with *Quotient Pareto Adaptation* (QPA), which performs both the Pareto decision and its execution in the effective fixed-rank geometry. We also give *Stiefel Pareto Adaptation* (SPA) as a complementary representation-correct construction in constrained factor coordinates.

3.1 QPA: Pareto optimization in effective space

QPA uses the standard fixed-rank manifold

$$\mathcal{M}_r = \{Z \in \mathbb{R}^{m\times n} : \operatorname{rank}(Z) = r\}$$

with the Frobenius metric (Absil et al., 2008; Vandereycken, 2013; Boumal, 2023). For a thin SVD $Z = U\Sigma V^\top$, the tangent projection is

$$\mathcal{P}_Z(X) = UU^\top X + XVV^\top - UU^\top XVV^\top. \tag{18}$$

For objective k, let $g_k = \mathcal{P}_Z(G_k)$. QPA solves

$$\lambda^* \in \operatorname*{arg\,min}_{\lambda\in\Delta_K} \lambda^\top H^Q \lambda, \quad H^Q_{ij} = \langle g_i, g_j\rangle_F, \tag{19}$$

and executes

$$d = -\sum_{k=1}^{K} \lambda^*_k g_k, \quad Z^+ = \mathcal{P}_r(Z + \eta d), \tag{20}$$

where $\mathcal{P}_r$ is the best rank-r Frobenius approximation. By the minimum-norm property of MGDA,

$$\langle g_k, d\rangle_F \le -\|d\|_F^2 \quad \text{for every } k, \tag{21}$$

so every nonzero stateless QPA direction is a common first-order descent direction. Although defined in effective space, QPA uses only ordinary LoRA backward quantities and low-rank linear algebra. Algorithm 1 gives one step; the compact identities are derived in Appendix C.

Theorem 5 (Representation correctness of QPA)**.** *Assume full-rank factors, representation-invariant tie breaking and control flow, shared randomness, and a strict rank-r spectral cutoff gap for every accepted trial. Then arbitrary $\mathrm{GL}(r)$-equivalent factorizations induce identical effective QPA trajectories.*

A sufficient condition for the cutoff gap is: $\eta\|d\|_{\mathrm{op}} \le \rho\,\sigma_r(Z), \quad 0 < \rho < \frac{1}{2}$.

3.2 SPA: a complementary construction

SPA removes the non-orthogonal factor ambiguity through product-preserving Stiefelization:

$$(B, A) \mapsto (\bar{B}, \bar{A}) = (BH_B^{-1}, H_B A), \quad H_B = (B^\top B)^{1/2}. \tag{22}$$

This preserves cBA. For an equivalent factorization $(B', A') = (BQ, Q^{-1}A)$ with $Q \in \mathrm{GL}(r)$, let

$$H'_B = (Q^\top H_B^2 Q)^{1/2}, \quad O = H_B Q (H'_B)^{-1}.$$

Then $O \in \mathrm{O}(r)$ and the Stiefelized representatives satisfy $(\bar{B}', \bar{A}') = (\bar{B}O, O^\top \bar{A})$. Thus Stiefelization reduces the original $\mathrm{GL}(r)$ ambiguity to a residual orthogonal change of basis. SPA then uses the product Frobenius geometry on $\mathrm{St}(m, r) \times \mathbb{R}^{r\times n}$ (Edelman et al., 1998; Absil et al., 2008). Define

$$\operatorname{sym}(X) = \frac{1}{2}(X + X^\top), \quad \operatorname{polar}(Y) = Y(Y^\top Y)^{-1/2}.$$

Algorithm 2 gives the resulting update.

Algorithm 1 Quotient Pareto Adaptation (QPA)

Require: Full-rank (B, A), LoRA scale c, step size η

Recover $Z = cBA = U\Sigma V^\top$ in thin-SVD form
$R_B \leftarrow U^\top B, \quad R_A \leftarrow AV$
for $k = 1, \dots, K$ **do**
  Obtain LoRA gradients $(G_{B,k}, G_{A,k})$
  $T_{B,k} \leftarrow c^{-1} G_{B,k} R_A^{-\top}, \quad T_{A,k} \leftarrow c^{-1} G_{A,k}^\top R_B^{-1}$
  $M_k \leftarrow U^\top T_{B,k}$
  $K_k \leftarrow (I - UU^\top) T_{B,k}, \quad L_k \leftarrow (I - VV^\top) T_{A,k}$
end for
$H^Q_{ij} \leftarrow \langle M_i, M_j \rangle_F + \langle K_i, K_j \rangle_F + \langle L_i, L_j \rangle_F$
$\lambda^* \leftarrow \arg\min_{\lambda \in \Delta_K} \lambda^\top H^Q \lambda$
$(M, K, L) \leftarrow -\sum_k \lambda_k^* (M_k, K_k, L_k)$
Compute thin QR factorizations $K = Q_K R_K$ and $L = Q_L R_L$
Form

$$C = \begin{bmatrix} \Sigma + \eta M & \eta R_L^\top \\ \eta R_K & 0 \end{bmatrix}$$

Compute rank-r truncated SVD $C_r = \widetilde{U}_r \Sigma^+ \widetilde{V}_r^\top$
$U^+ \leftarrow [U, Q_K] \widetilde{U}_r, \quad V^+ \leftarrow [V, Q_L] \widetilde{V}_r$
$B^+ \leftarrow c^{-1/2} U^+ (\Sigma^+)^{1/2}, \quad A^+ \leftarrow c^{-1/2} (\Sigma^+)^{1/2} (V^+)^\top$
return (B^+, A^+)

Algorithm 2 Stiefel Pareto Adaptation (SPA)

Require: Full-rank (B, A), step size η

$H_B \leftarrow (B^\top B)^{1/2}$
$(B, A) \leftarrow (BH_B^{-1}, H_B A)$
for $k = 1, \dots, K$ **do**
  Obtain LoRA gradients $(G_{B,k}, G_{A,k})$
  $r_{B,k} \leftarrow G_{B,k} - B \, \mathrm{sym}(B^\top G_{B,k})$
  $r_{A,k} \leftarrow G_{A,k}$
end for
$H^S_{ij} \leftarrow \langle r_{B,i}, r_{B,j} \rangle_F + \langle r_{A,i}, r_{A,j} \rangle_F$
$\lambda^* \leftarrow \arg\min_{\lambda \in \Delta_K} \lambda^\top H^S \lambda$
$(q_B, q_A) \leftarrow \sum_k \lambda_k^* (r_{B,k}, r_{A,k})$
$B^+ \leftarrow \mathrm{polar}(B - \eta q_B), \quad A^+ \leftarrow A - \eta q_A$
return (B^+, A^+)

Theorem 6 (Representation correctness of SPA)**.** *Under residual-*$\mathrm{O}(r)$*-invariant decisions and control flow, shared randomness, full-column-rank accepted polar trials, and full-row-rank* A *blocks along the trajectory, SPA is representation-correct for arbitrary* $\mathrm{GL}(r)$*-equivalent initial factorizations.*

QPA and SPA can nevertheless select different Pareto weights, showing that representation correctness does not uniquely determine the optimization geometry.

Transported optimizer memory. Because tangent spaces vary with the iterate, persistent optimizer state must be transported consistently (Bonnabel, 2013; Kasai et al., 2019; Kong et al., 2023). Let q_t be the stateless gradient-like Pareto combination and M_t the stored tangent state:

$$v_t = \beta_m M_t - (1 - \beta_m) q_t. \tag{23}$$

The executor uses v_t. After an accepted step, QPA projects it onto the new effective tangent space, while SPA projects its left-factor component onto the new Stiefel tangent space and retains the right-factor component. Rejected or stationary proposals preserve both model and state.

Corollary 7 (Pathwise correctness with momentum)**.** *Under the corresponding stateless assumptions and equivalent initial optimizer states, the transported first-moment variants of QPA and SPA remain representation-correct under shared randomness.*

This is a pathwise representation guarantee; the momentum-smoothed direction is not claimed to satisfy the stateless common-descent bound in Equation (21).

All proofs and auxiliary derivations for this section, including compact QPA recovery and retraction, SPA equivariance, transported optimizer state, and the QPA-SPA geometry separation, are deferred to Appendices C, D, E, and D.3.

4 Experiments

Conflict judgments under equivalent representations. We compare raw MGDA, QPA, and SPA on Qwen2.5-7B and Llama-3.2-3B. For each backbone, 63 checkpoints and eight minibatches per checkpoint yield 504 matched checkpoint–minibatch pairs. The checkpoints comprise seven shared warm starts and the step-1,000 and step-2,000 iterates of the four proposed methods. We apply equivalent changes of factors $(B, A) \mapsto (BQ, Q^{-1}A)$, keeping the effective adapter $Z = cBA$ fixed. The non-orthogonal grid includes four scalar rescalings, $Q = sI$ with $s \in \{1/4, 1/2, 2, 4\}$, and three fixed orientations at each condition number $\kappa \in \{3, 10\}$. A pair counts as flipped if any of ten non-orthogonal transformations reverses the global objective-gradient cosine sign in the corresponding decision geometry relative to the balanced reference, with both magnitudes exceeding 10^{-6}.

Figure 1 shows that raw MGDA reverses its conflict judgment in 47/504 pairs (9.33%) on Qwen and 146/504 pairs (28.97%) on Llama. QPA and SPA exhibit no flips across the tested transformations. Table 1 breaks down the raw-MGDA flips by checkpoint source and training stage. Perceived helpfulness-safety conflict can therefore arise from the choice of LoRA coordinates, rather than disagreement intrinsic to the effective model. Scalar rescaling alone produces flips in 39/504 Qwen pairs and 137/504 Llama pairs, whereas all three decision rules have zero flips under the three orthogonal controls. This distinction agrees with the raw-Gram transformation law: orthogonal changes preserve the factor-gradient inner products, while non-orthogonal changes can alter their relative contributions. Increasing the ambiguity threshold to 10^{-3} retains 47/504 flips on Qwen and 138/504 on Llama, with none for QPA or SPA. The observed sensitivity therefore persists when near-zero cosine comparisons are excluded more conservatively. Appendix H provides the complete protocol, numerical verification, and additional analyses.

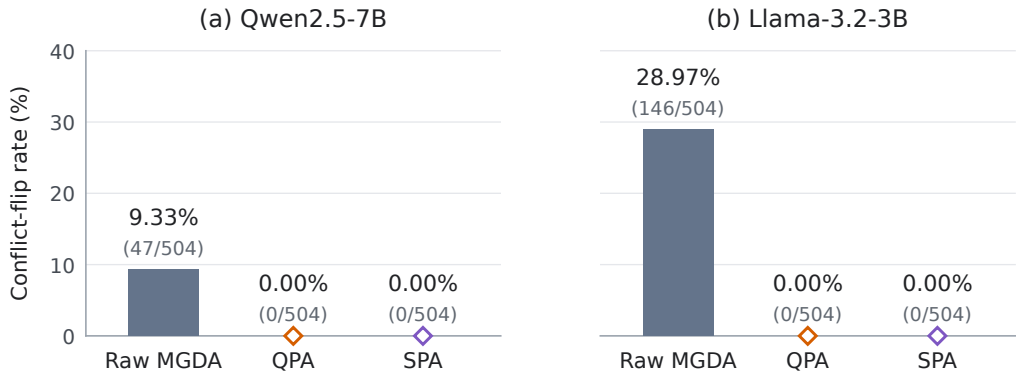

Figure 1: Conflict-flip rates under equivalent LoRA representations. Each panel summarizes 504 checkpoint–minibatch pairs. Labels show rates and counts; open markers denote zero observed flips.

Separating decisions from execution. We study representation sensitivity by independently varying the decision rule and the executor. For both Qwen2.5-7B-Instruct and Llama-3.2-3B-Instruct,

Table 1: Raw-MGDA flips at checkpoints generated by each method. Each cell reports flipped pairs out of 56 (7 initializations $\times$ 8 minibatches). The shared warm starts contribute an additional 3/56 flips per backbone. QPA and SPA decisions have zero flips in every stratum.

	Qwen2.5-7B		Llama-3.2-3B	
Checkpoint source	Step 1,000	Step 2,000	Step 1,000	Step 2,000
QPA	7/56	7/56	27/56	21/56
Stateful QPA	7/56	0/56	30/56	18/56
SPA	7/56	3/56	14/56	5/56
Stateful SPA	9/56	4/56	18/56	10/56

the decision rule can operate in the raw (R) or intrinsic (I) space, and the executor can update parameters in the raw (R) or effective (I) space, giving four combinations: RR, RI, IR, and II. All experiments use FP32. For each combination, we use three seeds with matched effective initializations and minibatches. We compare the identity transformation ($Q = I$) with scalar rescaling ($Q = 0.75I$) and a nonuniform transformation Q having condition number 3. Step sizes are calibrated so that the initial relative displacement in effective space is 10^{-3}. We measure representation drift as $\|Z_t(Q) - Z_t(I)\|_F / \|Z_t(I)\|_F$, aggregated across modules. As shown in Figure 2, at step 300, RR and IR retain a median drift of about 2-3%. In contrast, RI reduces the drift to approximately 10^{-5}, while II reduces it further to below 5×10^{-7}, consistently across both models. The persistent drift in IR shows why invariant decisions alone are insufficient; the further reduction from RI to II supports combining a representation-correct decision rule with a compatible executor.

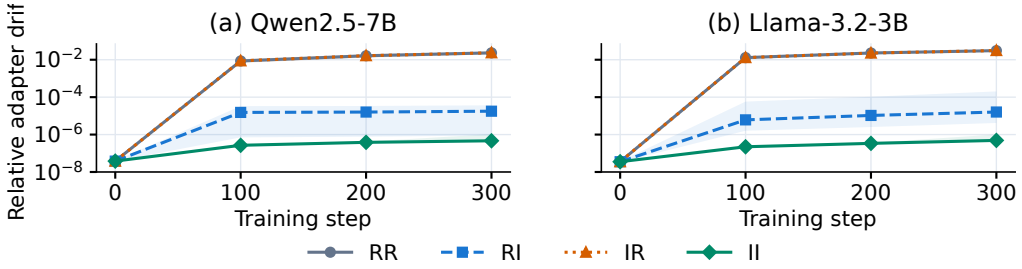

Figure 2: Decision-executor representation sensitivity (log scale). Curves show the median relative adapter drift, with bands spanning three seeds and two non-identity representations.

Setup. We fine-tune Qwen2.5-7B-Instruct (Qwen Team, 2024) and Llama-3.2-3B-Instruct (Meta, 2024) on the helpfulness and safety preferences from PKU-SafeRLHF (Ji et al., 2024). We use DPO (Rafailov et al., 2023) with a coefficient of 0.1, keeping the pretrained backbone frozen as the reference model and using 16 preference pairs per objective at each step. LoRA is applied to the query and value projections of the final 12 transformer blocks (24 modules), with rank $r = 8$, scaling factor $\alpha = 16$, and no dropout. We initialize seven shared models from 100-step weighted-DPO warm starts, using helpfulness weights of $0, 1/6, \ldots, 1$. These initializations represent different helpfulness-safety trade-offs rather than independent random seeds.

Methods. We compare against five baselines: weighted DPO with AdamW, raw MGDA with AdamW, MGDA with LoRA-RITE (Yen et al., 2025), MGDA with AdamW and BaLoRA (Castin et al., 2026), and MAdam (Liu et al., 2026). All nine methods are trained for 2,000 steps with a nominal learning rate of 10^{-3}. For QPA, we use retractions based on the observed singular-value gap and rank checks, while SPA uses polar retraction. Stateful variants use $\beta_m = 0.5$ with bias correction. After warm-starting, we reset the optimizer state and use random seed 13 for formal training.

Evaluation. Each method is evaluated at seven held-out preference-accuracy points. We report the arithmetic mean of these points for Help and Safe. Hypervolume (HV) is computed using $(0.5, 0.5)$ as the reference point (Zitzler & Thiele, 1999). For external evaluation, we use the harmful completion rate from HarmBench (Mazeika et al., 2024) and the balanced accuracy of XSTest (Röttger

et al., 2024), which averages safe compliance and unsafe refusal. External scores are averaged across the seven initializations.

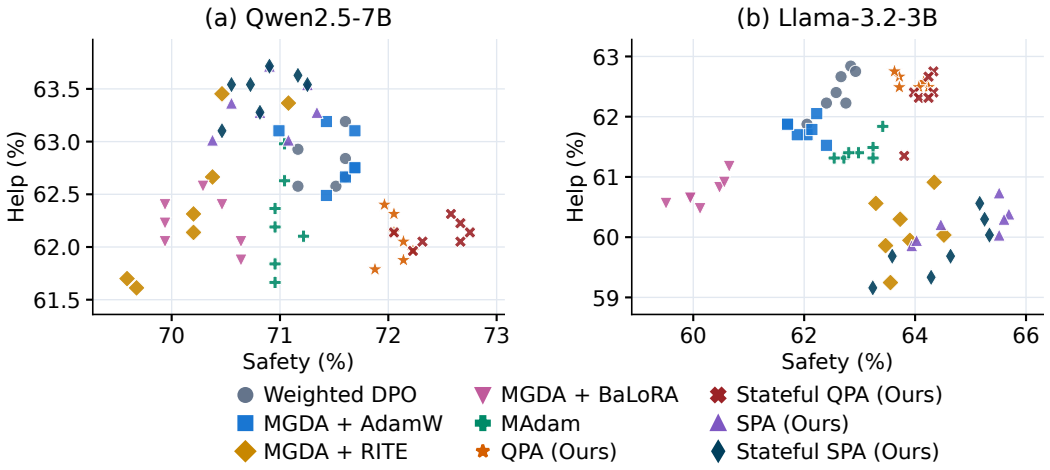

Figure 3: Helpfulness-safety trade-offs. Each point is the endpoint of one initialization; each method contributes seven points.

Table 2: Results are reported across the two backbones. HV summarizes the seven points shown in Figure 3, while Help and Safe report their arithmetic means. HarmBench (HB) and XSTest (XS) scores are also averaged across the seven initializations. All scores are reported as percentages except HV.

	Qwen2.5-7B					Llama-3.2-3B				
Method	HV↑	Help↑	Safe↑	HB↓	XS↑	HV↑	Help↑	Safe↑	HB↓	XS↑
Weighted DPO	0.02850	62.78	71.47	19.69	82.00	0.01660	62.43	62.60	19.69	85.55
MGDA + AdamW	0.02861	62.86	71.50	19.06	81.75	0.01494	61.79	62.01	19.38	84.85
MGDA + RITE	0.02835	62.46	70.23	18.44	82.80	0.01583	60.12	63.83	**16.25**	87.20
MGDA + BaLoRA	0.02595	62.23	70.26	19.38	81.50	0.01190	60.75	60.17	20.00	84.40
MAdam	0.02752	62.25	71.02	18.13	82.50	0.01589	61.44	62.99	19.38	86.80
QPA (Ours)	0.02756	62.08	72.06	18.66	81.67	0.01815	**62.56**	63.97	17.59	**87.96**
Stateful QPA (Ours)	0.02802	62.13	**72.47**	18.71	81.18	**0.01827**	62.31	64.14	18.35	87.46
SPA (Ours)	**0.02926**	63.32	70.90	**17.63**	**83.78**	0.01684	60.21	**64.97**	19.73	85.24
Stateful SPA (Ours)	0.02915	**63.48**	70.84	17.77	83.77	0.01619	59.82	64.50	19.60	85.31

Task performance. Figure 3 and Table 2 show that representation correctness leaves a meaningful choice of task trade-off. SPA attains the highest HV on Qwen (0.02926), whereas stateful QPA leads on Llama (0.01827). On Qwen, QPA variants favor mean safety and SPA variants favor mean helpfulness; on Llama, SPA variants instead trade helpfulness for greater safety relative to QPA. Thus, removing representation dependence does not make the two geometries interchangeable. External benchmarks also distinguish preference accuracy from generated behavior: SPA leads XSTest on Qwen (83.78%), while QPA leads on Llama (87.96%), despite SPA's higher mean ID safety. Higher preference accuracy alone therefore does not ensure better external safety behavior.

5 Conclusion

Large-model post-training requires trade-offs that reflect the underlying objectives rather than arbitrary LoRA coordinates. Linear scalarization is sensitive to objective scaling, so fixed coefficients can induce very different effective trade-offs across objectives. MGDA avoids fixing these trade-offs a priori by adapting weights through gradient geometry, yet equivalent LoRA factorizations can alter that geometry and even reverse apparent conflicts. QPA and SPA address this representation dependence through compatible decisions, execution, and optimizer state. Together, they show that multi-objective LoRA optimization can be made well defined at the level of the effective adapter, independent of arbitrary factor coordinates.

AI USE STATEMENT

Generative AI tools were used for language editing and polishing of the manuscript. They also assisted with implementing experiments designed by the authors and organizing and summarizing the resulting experimental data. The authors take full responsibility for the content, mathematical claims, experimental code, and reported results.

ETHICS STATEMENT

This work studies multi-objective optimization for language-model adaptation using existing preference datasets and safety benchmarks. The evaluation includes harmful prompts and model responses to assess safety-related behavior. Results reflect performance on the evaluated datasets and benchmarks. Use and redistribution of the models and datasets remain subject to their respective licenses and access conditions.

REPRODUCIBILITY STATEMENT

Sections 2-3 define the optimization maps and their mathematical assumptions. Appendices A-E give complete proofs, and Appendix F specifies the preference objectives, shared initialization procedure, optimizer settings, and evaluation rules. The experimental design pairs initial effective updates and data schedules across methods and evaluates the seven final outputs using the same scoring rules. An anonymized implementation is provided in the supplementary material. We will release a non-anonymous code repository upon acceptance.

REFERENCES

P.-A. Absil and Jérôme Malick. Projection-like retractions on matrix manifolds. *SIAM Journal on Optimization*, 22(1):135–158, 2012. doi: 10.1137/100802529.

P.-A. Absil, Robert Mahony, and Rodolphe Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2008. URL https://press.princeton.edu/absil.

Yuntao Bai, Andy Jones, Kamal Ndousse, Amanda Askell, Anna Chen, Nova DasSarma, Dawn Drain, Stanislav Fort, Deep Ganguli, Tom Henighan, Nicholas Joseph, Saurav Kadavath, Jackson Kernion, Tom Conerly, Sheer El-Showk, Nelson Elhage, Zac Hatfield-Dodds, Danny Hernandez, Tristan Hume, Scott Johnston, Shauna Kravec, Liane Lovitt, Neel Nanda, Catherine Olsson, Dario Amodei, Tom Brown, Jack Clark, Sam McCandlish, Chris Olah, Ben Mann, and Jared Kaplan. Training a helpful and harmless assistant with reinforcement learning from human feedback, 2022. URL https://arxiv.org/abs/2204.05862.

Vladimir Bogachev, Vladimir Aletov, Alexander Molozhavenko, Denis Bobkov, Vera Soboleva, Aibek Alanov, and Maxim Rakhuba. LoRA meets Riemannion: Muon optimizer for parametrization-independent low-rank adapters. In *International Conference on Learning Representations*, 2026. URL https://openreview.net/forum?id=WtbXgc9GVA.

Silvère Bonnabel. Stochastic gradient descent on Riemannian manifolds. *IEEE Transactions on Automatic Control*, 58(9):2217–2229, 2013. doi: 10.1109/TAC.2013.2254619. URL https://arxiv.org/abs/1111.5280.

Nicolas Boumal. *An Introduction to Optimization on Smooth Manifolds*. Cambridge University Press, 2023. doi: 10.1017/9781009166164. URL https://www.nicolasboumal.net/book/.

Valérie Castin, Kimia Nadjahi, Pierre Ablin, and Gabriel Peyré. Balanced LoRA: Removing parameter invariance to accelerate convergence, 2026. URL https://arxiv.org/abs/2605.31484.

Zhao Chen, Vijay Badrinarayanan, Chen-Yu Lee, and Andrew Rabinovich. GradNorm: Gradient normalization for adaptive loss balancing in deep multitask networks. In *Proceedings of the 35th International Conference on Machine Learning*, 2018. URL https://arxiv.org/abs/1711.02257.

Josef Dai, Xuehai Pan, Ruiyang Sun, Jiaming Ji, Xinbo Xu, Mickel Liu, Yizhou Wang, and Yaodong Yang. Safe RLHF: Safe reinforcement learning from human feedback. In *International Conference on Learning Representations*, 2024. URL https://arxiv.org/abs/2310.12773.

Jean-Antoine Désidéri. Multiple-gradient descent algorithm (MGDA) for multiobjective optimization. *Comptes Rendus Mathématique*, 350(5–6):313–318, 2012. doi: 10.1016/j.crma.2012.03.014.

Tim Dettmers, Artidoro Pagnoni, Ari Holtzman, and Luke Zettlemoyer. QLoRA: Efficient finetuning of quantized LLMs. In *Advances in Neural Information Processing Systems*, volume 36, 2023. URL https://arxiv.org/abs/2305.14314.

Carl Eckart and Gale Young. The approximation of one matrix by another of lower rank. *Psychometrika*, 1:211–218, 1936. doi: 10.1007/BF02288367.

Alan Edelman, Tomás A. Arias, and Steven T. Smith. The geometry of algorithms with orthogonality constraints. *SIAM Journal on Matrix Analysis and Applications*, 20(2):303–353, 1998. doi: 10.1137/S0895479895290954.

J"org Fliege and Benar Fux Svaiter. Steepest descent methods for multicriteria optimization. *Mathematical Methods of Operations Research*, 51:479–494, 2000. doi: 10.1007/s001860000043.

Soufiane Hayou, Nikhil Ghosh, and Bin Yu. LoRA+: Efficient low rank adaptation of large models. In *Proceedings of the 41st International Conference on Machine Learning*, 2024. URL https://arxiv.org/abs/2402.12354.

Junxian He, Chunting Zhou, Xuezhe Ma, Taylor Berg-Kirkpatrick, and Graham Neubig. Towards a unified view of parameter-efficient transfer learning. In *International Conference on Learning Representations*, 2022. URL https://arxiv.org/abs/2110.04366.

Neil Houlsby, Andrei Giurgiu, Stanislaw Jastrzebski, Bruna Morrone, Quentin de Laroussilhe, Andrea Gesmundo, Mona Attariyan, and Sylvain Gelly. Parameter-efficient transfer learning for NLP. In *Proceedings of the 36th International Conference on Machine Learning*, 2019. URL https://arxiv.org/abs/1902.00751.

Edward J. Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022. URL https://arxiv.org/abs/2106.09685.

Jiaming Ji, Donghai Hong, Borong Zhang, Boyuan Chen, Juntao Dai, Boren Zheng, Tianyi Qiu, Jiayi Zhou, Kaile Wang, Boxuan Li, Sirui Han, Yike Guo, and Yaodong Yang. PKU-SafeRLHF: Towards multi-level safety alignment for LLMs with human preference, 2024. URL https://arxiv.org/abs/2406.15513.

Damjan Kalajdzievski. A rank stabilization scaling factor for fine-tuning with LoRA, 2023. URL https://arxiv.org/abs/2312.03732.

Hiroyuki Kasai, Pratik Jawanpuria, and Bamdev Mishra. Riemannian adaptive stochastic gradient algorithms on matrix manifolds. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97 of *Proceedings of Machine Learning Research*, pp. 3262–3271, 2019. URL https://proceedings.mlr.press/v97/kasai19a.html.

Lingkai Kong, Yuqing Wang, and Molei Tao. Momentum Stiefel optimizer, with applications to suitably-orthogonal attention, and optimal transport. In *International Conference on Learning Representations*, 2023. URL https://arxiv.org/abs/2205.14173.

Brian Lester, Rami Al-Rfou, and Noah Constant. The power of scale for parameter-efficient prompt tuning. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, 2021. URL https://arxiv.org/abs/2104.08691.

Xiang Lisa Li and Percy Liang. Prefix-tuning: Optimizing continuous prompts for generation. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing*, 2021. URL https://arxiv.org/abs/2101.00190.

Zhizhong Li, Sina Sajadmanesh, Jingtao Li, and Lingjuan Lyu. StelLA: Subspace learning in low-rank adaptation using Stiefel manifold. In *Advances in Neural Information Processing Systems*, 2025. URL https://arxiv.org/abs/2510.01938.

Xi Lin, Hui-Ling Zhen, Zhenhua Li, Qingfu Zhang, and Sam Kwong. Pareto multi-task learning. In *Advances in Neural Information Processing Systems*, volume 32, 2019. URL https://arxiv.org/abs/1912.12854.

Bo Liu, Xingchao Liu, Xiaojie Jin, Peter Stone, and Qiang Liu. Conflict-averse gradient descent for multi-task learning. In *Advances in Neural Information Processing Systems*, volume 34, 2021a. URL https://arxiv.org/abs/2110.14048.

Fengbei Liu, Rachit Saluja, Sunwoo Kwak, Ruibo Wang, Ruining Deng, Heejong Kim, Johannes C. Paetzold, and Mert R. Sabuncu. MAdam: Metric-aware multi-objective Adam, 2026. URL https://arxiv.org/abs/2606.03904v1.

Liyang Liu, Yi Li, Zhanghui Kuang, Jing-Hao Xue, Yimin Chen, Wenming Yang, Qingmin Liao, and Wei Zhang. Towards impartial multi-task learning. In *International Conference on Learning Representations*, 2021b. URL https://openreview.net/forum?id=IMPnRXEWpvr.

Shih-Yang Liu, Chien-Yi Wang, Hongxu Yin, Pavlo Molchanov, Yu-Chiang Frank Wang, Kwang-Ting Cheng, and Min-Hung Chen. DoRA: Weight-decomposed low-rank adaptation. In *Proceedings of the 41st International Conference on Machine Learning*, 2024. URL https://arxiv.org/abs/2402.09353.

Debabrata Mahapatra and Vaibhav Rajan. Multi-task learning with user preferences: Gradient descent with controlled ascent in pareto optimization. In *Proceedings of the 37th International Conference on Machine Learning*, volume 119 of *Proceedings of Machine Learning Research*, pp. 6597–6607, 2020. URL https://proceedings.mlr.press/v119/mahapatra20a.html.

Mantas Mazeika, Long Phan, Xuwang Yin, Andy Zou, Zifan Wang, Norman Mu, Elham Sakhaee, Nathaniel Li, Steven Basart, Bo Li, David Forsyth, and Dan Hendrycks. HarmBench: A standardized evaluation framework for automated red teaming and robust refusal. In *Proceedings of the 41st International Conference on Machine Learning*, 2024. URL https://arxiv.org/abs/2402.04249.

Fanxu Meng, Zhaohui Wang, and Muhan Zhang. PiSSA: Principal singular values and singular vectors adaptation of large language models. In *Advances in Neural Information Processing Systems*, volume 37, 2024. URL https://arxiv.org/abs/2404.02948.

Meta. Llama 3.2 model card, 2024. URL https://github.com/meta-llama/llama-models/blob/main/models/llama3_2/MODEL_CARD.md.

Aviv Navon, Aviv Shamsian, Idan Achituve, Haggai Maron, Kenji Kawaguchi, Gal Chechik, and Ethan Fetaya. Multi-task learning as a bargaining game. In *Proceedings of the 39th International Conference on Machine Learning*, 2022. URL https://arxiv.org/abs/2202.01017.

Long Ouyang, Jeff Wu, Xu Jiang, Diogo Almeida, Carroll L. Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, John Schulman, Jacob Hilton, Fraser Kelton, Luke Miller, Maddie Simens, Amanda Askell, Peter Welinder, Paul Christiano, Jan Leike, and Ryan Lowe. Training language models to follow instructions with human feedback. In *Advances in Neural Information Processing Systems*, volume 35, 2022. URL https://arxiv.org/abs/2203.02155.

Juneyoung Park, Minjae Kang, Seongbae Lee, Haegang Lee, Seongwan Kim, and Jaeho Lee. Riemannian optimization for LoRA on the Stiefel manifold. In *Findings of the Association for Computational Linguistics: EMNLP 2025*, 2025. URL https://arxiv.org/abs/2508.17901.

Qwen Team. Qwen2.5 technical report, 2024. URL https://arxiv.org/abs/2412.15115.

Rafael Rafailov, Archit Sharma, Eric Mitchell, Stefano Ermon, Christopher D. Manning, and Chelsea Finn. Direct preference optimization: Your language model is secretly a reward model. In *Advances in Neural Information Processing Systems*, volume 36, 2023. URL https://arxiv.org/abs/2305.18290.

Paul Röttger, Hannah Rose Kirk, Bertie Vidgen, Giuseppe Attanasio, Federico Bianchi, and Dirk Hovy. XSTest: A test suite for identifying exaggerated safety behaviours in large language models. In *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 2024. URL https://arxiv.org/abs/2308.01263.

Ozan Sener and Vladlen Koltun. Multi-task learning as multi-objective optimization. In *Advances in Neural Information Processing Systems*, volume 31, 2018. URL https://arxiv.org/abs/1810.04650.

Bart Vandereycken. Low-rank matrix completion by Riemannian optimization. *SIAM Journal on Optimization*, 23(2):1214–1236, 2013. URL https://arxiv.org/abs/1209.3834.

Shaowen Wang, Linxi Yu, and Jian Li. LoRA-GA: Low-rank adaptation with gradient approximation. In *Advances in Neural Information Processing Systems*, volume 37, 2024. URL https://papers.nips.cc/paper_files/paper/2024/hash/62c4718cc334f6a0a62fb81c4a2095a1-Abstract-Conference.html.

Jui-Nan Yen, Si Si, Zhao Meng, Felix Yu, Sai Surya Duvvuri, Inderjit S. Dhillon, Cho-Jui Hsieh, and Sanjiv Kumar. LoRA done RITE: Robust invariant transformation equilibration for LoRA optimization. In *International Conference on Learning Representations*, 2025. URL https://arxiv.org/abs/2410.20625.

Tianhe Yu, Saurabh Kumar, Abhishek Gupta, Sergey Levine, Karol Hausman, and Chelsea Finn. Gradient surgery for multi-task learning. In *Advances in Neural Information Processing Systems*, volume 33, 2020. URL https://arxiv.org/abs/2001.06782.

Longteng Zhang, Lin Zhang, Shaohuai Shi, Xiaowen Chu, and Bo Li. LoRA-FA: Efficient and effective low rank representation fine-tuning, 2023a. URL https://arxiv.org/abs/2308.03303.

Qingru Zhang, Minshuo Chen, Alexander Bukharin, Nikos Karampatziakis, Pengcheng He, Yu Cheng, Weizhu Chen, and Tuo Zhao. Adaptive budget allocation for parameter-efficient fine-tuning. In *International Conference on Learning Representations*, 2023b. URL https://arxiv.org/abs/2303.10512.

Eckart Zitzler and Lothar Thiele. Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach. *IEEE Transactions on Evolutionary Computation*, 3(4):257–271, 1999. doi: 10.1109/4235.797969.

A FACTOR REPRESENTATIONS AND RAW OPTIMIZATION

All mathematical statements are in exact arithmetic on the full-rank domain. The Frobenius inner product is $\langle X, Y\rangle_F = \mathrm{tr}(X^\top Y)$. For a tuple of adapters it is the sum of the corresponding module-wise inner products. Throughout, objectives depend on factors through their effective updates, the scaling c is fixed and positive, and external randomness is shared when comparing equivalent representations. We use the notation of Sections 2-3; SPA denotes Stiefel Pareto Adaptation, defined in Section 3.2. Appendices A and B establish the representation and complete-map results. Appendices C-E give the compact constructions and the stateless and stateful correctness proofs. We retain $F_k(B, A) = \widetilde{F}_k(Z)$, $G_k = \nabla_Z \widetilde{F}_k(Z)$, and $g_k = \mathcal{P}_Z(G_k)$ from the main text. The MGDA solution is λ^*, its gradient-like combination is q, and the stateless descent direction is $d = -q$. The compact block M_k in Algorithm 1 is distinct from the transported optimizer moment M_t in Equation (23).

A.1 THE FULL-RANK FIBER

Lemma A.1. *If $BA = B'A'$ and all four factors have rank r, there is a unique $Q \in \mathrm{GL}(r)$ such that $B' = BQ$ and $A' = Q^{-1}A$.*

Proof. Since A has full row rank, $\mathrm{col}(BA) = \mathrm{col}(B)$. The same holds for $B'A'$, so the two left factors span the same r-dimensional column space. There is therefore a unique invertible Q satisfying $B' = BQ$. Substituting into $BA = BQA'$ and multiplying by a left inverse of B gives $A = QA'$. □

A.2 FACTOR GRADIENTS AND PROOF OF THEOREM 1

The chain rule and the change of factors give

$$G_{B,k} = cG_kA^\top, \quad G_{A,k} = cB^\top G_k, \qquad G'_{B,k} = G_{B,k}Q^{-\top}, \quad G'_{A,k} = Q^\top G_{A,k}. \tag{A.1}$$

The differential of the effective update is $\mathrm{d}Z = c(\mathrm{d}B)A + cB(\mathrm{d}A)$. For an objective with ambient derivative G_k,

$$\mathrm{d}F_k = \langle G_k, \mathrm{d}Z\rangle_F = \langle cG_kA^\top, \mathrm{d}B\rangle_F + \langle cB^\top G_k, \mathrm{d}A\rangle_F.$$

This proves the factor-gradient identities. At $(BQ, Q^{-1}A)$, the ambient derivative is unchanged, yielding the transformation law in Equation (A.1). Writing $P = QQ^\top$ as in Theorem 1, we obtain

$$\begin{aligned}\langle G_{B,i}Q^{-\top}, G_{B,j}Q^{-\top}\rangle_F &= \mathrm{tr}(Q^{-1}G_{B,i}^\top G_{B,j}Q^{-\top}) = \mathrm{tr}(P^{-1}G_{B,i}^\top G_{B,j}),\\ \langle Q^\top G_{A,i}, Q^\top G_{A,j}\rangle_F &= \mathrm{tr}(G_{A,i}^\top P G_{A,j}) = \mathrm{tr}(PG_{A,j}G_{A,i}^\top).\end{aligned}$$

Their sum is Equation (10). In the orthogonal case $P = I$, both inner products are preserved. For $Q = sI$, the formula becomes $s^{-2}\alpha_{ij} + s^2\beta_{ij}$, using the notation of Equation (12). When $\alpha_{12}\beta_{12} < 0$, multiplying by $s^2 > 0$ gives the unique positive root $s^4 = -\alpha_{12}/\beta_{12}$ and opposite signs on its two sides. This also proves the conflict-reversal corollary and Equation (13).

A.3 PROOF OF PROPOSITION 3

For the zero condition, define the invertible linear map

$$\mathcal{T}_Q(X_B, X_A) = (X_BQ^{-\top}, Q^\top X_A).$$

Linearity gives $\sum_k \lambda_k y_k(Q) = \mathcal{T}_Q(\sum_k \lambda_k y_k(I))$. Invertibility proves, for every simplex weight vector,

$$\sum_k \lambda_k y_k(Q) = 0 \quad \Longleftrightarrow \quad \sum_k \lambda_k y_k(I) = 0. \tag{A.2}$$

In particular, membership of zero in the convex hull of the gradients is unchanged. This establishes the invariant first-order stationarity condition even when the nonzero minimum-norm vector or its weights change.

A.4 Proof of Theorem 4

Fix $G_\lambda = \sum_k \lambda_k G_k$. The factor gradients at $(sB, s^{-1}A)$ for the fixed ambient combination G_λ are $cs^{-1}G_\lambda A^\top$ and $csB^\top G_\lambda$. Thus

$$\begin{aligned} Z_s^+ &= c\big(sB - \eta cs^{-1}G_\lambda A^\top\big)\big(s^{-1}A - \eta csB^\top G_\lambda\big) \\ &= Z - \eta c^2\big(s^{-2}G_\lambda A^\top A + s^2 BB^\top G_\lambda\big) \\ &\quad + \eta^2 c^3 G_\lambda A^\top B^\top G_\lambda. \end{aligned} \tag{A.3}$$

The formula already holds for fixed weights, so the source of representation dependence is distinct from the Gram distortion.

B The complete-map criterion

The preceding results show that representation dependence can enter both the multi-objective decision and its execution. Correcting either stage in isolation is therefore insufficient. For a stateful optimizer, the requirements extend further: optimizer memory must evolve consistently, and equivalent runs must make consistent control-flow choices. We capture these requirements by viewing one optimization step as a map on the augmented state (x, S).

Let x denote a parameter representation and let $\pi(x)$ denote its effective readout. For LoRA,

$$x = (B, A), \quad \pi(x) = cBA.$$

Let S denote optimizer memory, and let ω collect the minibatch and any external randomness shared between equivalent runs. We write one optimization step as

$$\begin{aligned} \delta &= D(x, S, \omega), \\ x^+ &= E(x, S, \delta, \omega), \\ S^+ &= T(x, S, \delta, x^+, \omega), \end{aligned} \tag{B.1}$$

where D is the decision map, E executes the selected update, and T updates the optimizer state. The decision object δ may include, for example, Pareto weights, an update direction, and representation-invariant control quantities. Step-size rules, tie breaking, proposal rejection, fallbacks, and other branches are regarded as part of these maps rather than as external logic.

Let a representation group $\mathcal{G}$ act on x by $x \mapsto x \cdot g$ while preserving the effective readout:

$$\pi(x \cdot g) = \pi(x), \quad g \in \mathcal{G}. \tag{B.2}$$

For LoRA, $\mathcal{G} = \mathrm{GL}(r)$ and $(B, A) \cdot Q = (BQ, Q^{-1}A)$. Suppose $\mathcal{G}$ also acts on the optimizer-state space and decision space through

$$S \mapsto \sigma_g(S), \quad \delta \mapsto \rho_g(\delta),$$

respectively. These actions specify how the same optimizer memory and decision are represented after a change of parameter coordinates. We define two augmented states to be *equivalent* if, for some $g \in \mathcal{G}$,

$$(x', S') = \big(x \cdot g, \sigma_g(S)\big). \tag{B.3}$$

Thus, equivalent augmented states represent the same effective model together with the same optimizer memory in different coordinates.

Definition B.1 (Representation correctness)**.** An optimizer is *representation-correct* if, for every $g \in \mathcal{G}$, every pair of equivalent initial augmented states

$$x_0' = x_0 \cdot g, \quad S_0' = \sigma_g(S_0),$$

and the same sequence of external randomness, the resulting runs satisfy

$$\pi(x_t') = \pi(x_t) \quad \text{for all } t. \tag{B.4}$$

Representation correctness is therefore a property of the complete optimization trajectory, not merely of the gradient-combination rule.

To state a sufficient condition, let $\mathcal{G}$ also act on decision objects through $\delta \mapsto \rho_g(\delta)$, describing how the same decision is represented after a change of parameter coordinates.

Theorem B.2 (Complete-map criterion for representation correctness)**.** *Suppose that, for every $g \in \mathcal{G}$, the maps in Equation* (B.1) *satisfy*

$$D(x \cdot g, \sigma_g(S), \omega) = \rho_g\left(D(x, S, \omega)\right), \tag{B.5}$$

$$E(x \cdot g, \sigma_g(S), \rho_g(\delta), \omega) = E(x, S, \delta, \omega) \cdot g, \tag{B.6}$$

$$T(x \cdot g, \sigma_g(S), \rho_g(\delta), x^+ \cdot g, \omega) = \sigma_g\left(T(x, S, \delta, x^+, \omega)\right). \tag{B.7}$$

Then the optimizer is representation-correct. Moreover, equivalent initial augmented states remain equivalent pathwise:

$$x'_t = x_t \cdot g, \quad S'_t = \sigma_g(S_t), \quad \pi(x'_t) = \pi(x_t) \quad \textit{for all } t. \tag{B.8}$$

The theorem separates three requirements. The decision map must represent the same decision consistently across equivalent parameterizations; the executor must map equivalent decisions to equivalent next representations; and the optimizer state must preserve augmented-state equivalence across steps. Any component of the decision or control flow that is invariant under the representation action, such as Pareto weights, step lengths, acceptance decisions, or deterministic tie-breaking outcomes, must therefore agree across equivalent runs. The stateless case is recovered by omitting S and T.

Proof of Theorem B.2. Assume $(x', S') = (x \cdot g, \sigma_g(S))$. Equation (B.5) implies $\delta' = \rho_g(\delta)$. Equation (B.6) then gives $x'^+ = x^+ \cdot g$, and Equation (B.7) gives $S'^+ = \sigma_g(S^+)$. The relation is therefore preserved by one update. Induction from the assumed initial relation proves Equation (B.8). Finally $\pi(x_t \cdot g) = \pi(x_t)$ gives equal effective trajectories. □

The proof includes a rejected proposal if the branch predicate is invariant and the branch leaves both states unchanged. A deterministic tie rule based solely on the invariant Gram matrix also meets the criterion. In the stateless case, remove S and T; the same argument proves correctness of the decision-executor composition. Explicitly, $D(x \cdot g) = \rho_g(D(x))$ and $E(x \cdot g, \rho_g(\delta)) = E(x, \delta) \cdot g$ imply

$$\pi(E(x \cdot g, D(x \cdot g))) = \pi(E(x, D(x)) \cdot g) = \pi(E(x, D(x))).$$

For an effective-state implementation such as QPA, equality of the effective objects at one iteration gives equality of all subsequent effective operations directly.

C QPA: RECOVERY, DESCENT, AND EXECUTION

C.1 COMPACT TANGENT RECOVERY

Let $Z = U\Sigma V^\top$ be a thin SVD. Its tangent space admits the orthogonal decomposition

$$T_Z\mathcal{M}_r = \{UMV^\top + KV^\top + UL^\top : U^\top K = 0,\ V^\top L = 0\}. \tag{C.1}$$

Indeed, in orthonormal bases extending U and V, a rank-r matrix near Z has an invertible leading $r \times r$ block. Its lower-right block equals the product of its lower-left block, the inverse leading block, and its upper-right block. At Z both off-diagonal blocks vanish, so the derivative of the lower-right block is zero. The other three derivative blocks are unrestricted, which gives Equation (C.1) and the projector in Equation (18). Since the factors have full rank, write $B = UR_B$ and $A = R_AV^\top$ with invertible R_B, R_A satisfying $cR_BR_A = \Sigma$. Substitution of $B = UR_B$ and $A = R_AV^\top$ into the chain rule yields

$$G_{B,k} = cG_kVR_A^\top, \qquad G_{A,k} = cR_B^\top U^\top G_k.$$

Multiplication by the inverse coordinate factors gives

$$T_{B,k} = c^{-1}G_{B,k}R_A^{-\top} = G_kV, \qquad T_{A,k} = c^{-1}G_{A,k}^\top R_B^{-1} = G_k^\top U. \tag{C.2}$$

Define $M_k = U^\top T_{B,k}$, $K_k = (I - UU^\top)T_{B,k}$, and $L_k = (I - VV^\top)T_{A,k}$, as in Algorithm 1. Then

$$\begin{aligned} g_k &= UM_kV^\top + K_kV^\top + UL_k^\top, \\ H_{ij}^Q &= \langle M_i, M_j\rangle_F + \langle K_i, K_j\rangle_F + \langle L_i, L_j\rangle_F. \end{aligned} \tag{C.3}$$

The compact components expand as

$$\begin{aligned} UM_kV^\top &= UU^\top G_k VV^\top, \\ K_kV^\top &= (I - UU^\top)G_k VV^\top, \\ UL_k^\top &= UU^\top G_k (I - VV^\top). \end{aligned}$$

Their sum is $\mathcal{P}_Z(G_k)$. The conditions $U^\top K_k = 0$ and $V^\top L_k = 0$ make these three components pairwise orthogonal. Because U and V have orthonormal columns, their pairwise inner products reduce to those of the compact blocks. This proves Equation (C.3). For multiple adapters, sum these block inner products over all adapted modules before solving for a single λ^*.

C.2 The minimum-norm direction

Let $\mathcal{C} = \mathrm{conv}\{g_1, \ldots, g_K\}$. It is compact and convex, so the Euclidean projection $q = \mathrm{Proj}_{\mathcal{C}}(0) = \sum_k \lambda_k^* g_k$ is unique. The projection inequality gives $\langle -q, z - q\rangle_F \le 0$ for every $z \in \mathcal{C}$. Taking $z = g_k$ gives $\langle q, g_k\rangle_F \ge \|q\|_F^2$. With $d = -q$, this proves Equation (21), since $\|q\|_F = \|d\|_F$. The vector q is unique even when several weight vectors express it. The minimum-norm tie rule selects a unique weight vector from that convex minimizer set.

For two objectives and either Gram matrix $H \in \{H^Q, H^S\}$, write $\lambda = (w, 1-w)$ and $h = H_{11} + H_{22} - 2H_{12}$. This is the squared distance between the two gradients in the chosen decision geometry; for QPA, $h = \|g_1 - g_2\|_F^2$. If $h > 0$, the solution is

$$w^* = \mathrm{clip}_{[0,1]}\left(\frac{H_{22} - H_{12}}{h}\right), \qquad \lambda^* = (w^*, 1 - w^*). \tag{C.4}$$

If $h = 0$, the gradients coincide and the minimum-norm tie rule gives $w^* = 1/2$.

C.3 Small-core retraction and spectral gap

For a tangent direction $d = UMV^\top + KV^\top + UL^\top$, factor $K = Q_K R_K$ and $L = Q_L R_L$. Choose Q_K and Q_L with orthonormal columns spanning the nonzero residual subspaces. They satisfy $U^\top Q_K = V^\top Q_L = 0$. The trial matrix has the compact factorization

$$Z + \eta d = [U, Q_K] \begin{bmatrix} \Sigma + \eta M & \eta R_L^\top \\ \eta R_K & 0 \end{bmatrix} [V, Q_L]^\top. \tag{C.5}$$

Expanding the block product in Equation (C.5) gives $U\Sigma V^\top + \eta UMV^\top + \eta KV^\top + \eta UL^\top$, proving the identity. The outer factors are isometries, so the nonzero singular values and best rank-r approximation are determined by the central matrix. If one residual block is rank deficient, a basis of its actual column space reduces the corresponding core dimension. The central matrix is at most $2r \times 2r$. After truncation, a thin SVD $Z^+ = U^+\Sigma^+(V^+)^\top$ is stored with balanced factors $B^+ = c^{-1/2}U^+(\Sigma^+)^{1/2}$ and $A^+ = c^{-1/2}(\Sigma^+)^{1/2}(V^+)^\top$, as in Algorithm 1.

A sufficient condition for the strict cutoff gap required by Theorem 5 is

$$\eta\|d\|_{\mathrm{op}} \le \rho\, \sigma_r(Z), \qquad 0 < \rho < \tfrac{1}{2}. \tag{C.6}$$

The singular-value perturbation inequality gives

$$\sigma_r(Z + \eta d) \ge \sigma_r(Z) - \eta\|d\|_{\mathrm{op}}, \qquad \sigma_{r+1}(Z + \eta d) \le \eta\|d\|_{\mathrm{op}},$$

where the second bound uses $\sigma_{r+1}(Z) = 0$. Equation (C.6) therefore implies

$$\sigma_r(Z + \eta d) - \sigma_{r+1}(Z + \eta d) \ge (1 - 2\rho)\sigma_r(Z) > 0. \tag{C.7}$$

The strict gap makes the truncated-SVD matrix unique by the best low-rank approximation theorem (Eckart & Young, 1936). In particular, the retained singular values are positive, so the output has rank exactly r. The bound in Equation (C.6) is sufficient, not necessary: directly checking a strict observed cutoff gap and positive retained singular values certifies the same unique projection without imposing that bound.

Local retraction property. For a trial matrix W with a strict cutoff gap, let $P_L(W)$ be the spectral projector of $WW^\top$ onto its leading r eigenvectors. The separated eigenspaces make $P_L(W)$ smooth in a neighborhood of W, even if eigenvalues within the retained subspace coincide. The truncated matrix is $\mathcal{P}_r(W) = P_L(W)W$, hence it is smooth there. At a rank-r point Z, $\mathcal{P}_r(Z) = Z$. For any $\xi \in T_Z\mathcal{M}_r$, choose a smooth rank-r curve γ with $\gamma(0) = Z$ and $\gamma'(0) = \xi$. Differentiating $\mathcal{P}_r(\gamma(t)) = \gamma(t)$ at zero gives $D\mathcal{P}_r(Z)[\xi] = \xi$. Therefore $\mathcal{R}_Z(\xi) = \mathcal{P}_r(Z + \xi)$ is a local retraction (Absil & Malick, 2012; Vandereycken, 2013). Consequently, for $q \neq 0$ and each smooth objective,

$$\widetilde{F}_k(\mathcal{R}_Z(\eta d)) = \widetilde{F}_k(Z) + \eta\langle g_k, d\rangle_F + o(\eta) < \widetilde{F}_k(Z)$$

for sufficiently small positive η. The spectral condition certifies the rank projection; a finite step's objective change also depends on local curvature.

C.4 Proof of Theorem 5

Any equivalent full-rank factors have the same effective Z. The projector $\mathcal{P}_Z$ depends only on its row and column spaces, giving identical tangent gradients and the same intrinsic Gram. The deterministic weight rule gives identical λ^* and d. Effective-object control selects the same step and branch. On an accepted branch, both runs project the same trial $Z + \eta d$, and the strict cutoff gap guarantees an identical projected matrix. An invariant rejected or stationary branch returns the same unchanged effective matrix. The argument repeats at the next effective state. Internal SVD basis choices do not affect this reasoning: both the tangent projector and the final projected matrix are basis independent.

Computational representation. Given the LoRA factors, thin QR decompositions of B and $A^\top$ reduce the SVD of cBA to an $r \times r$ core. For fixed K, tangent recovery, Gram construction, and retraction require $O((m+n)r^2 + r^3)$ arithmetic per module outside backpropagation and $O((m+n)r)$ storage per compact tangent vector. These are algebraic costs; the language-model forward and backward passes remain part of the total training cost.

D SPA representation correctness and metric separation

D.1 Stiefelization of equivalent factors

For any full-rank input, the product-preserving initialization is

$$H_B = (B^\top B)^{1/2}, \qquad \bar{B} = BH_B^{-1}, \qquad \bar{A} = H_B A. \tag{D.1}$$

For a full-rank transformed input $(BQ, Q^{-1}A)$, let $H'_B = (Q^\top H_B^2 Q)^{1/2}$, following Equation (22). Define

$$O = H_B Q (H'_B)^{-1}.$$

Then $O^\top O = (H'_B)^{-1}Q^\top H_B^2 Q(H'_B)^{-1} = I$. The transformed left factor after Stiefelization is

$$\bar{B}' = BQ(H'_B)^{-1} = BH_B^{-1}O = \bar{B}O.$$

For the right factor, $(H'_B)^2 = Q^\top H_B^2 Q$ implies

$$H'_B Q^{-1} = (H'_B)^{-1}Q^\top H_B^2 = O^\top H_B, \qquad \bar{A}' = H'_B Q^{-1}A = O^\top \bar{A}.$$

Thus Stiefelization maps one full $\mathrm{GL}(r)$ fiber to one residual $\mathrm{O}(r)$ orbit. Conversely, Lemma A.1 shows that any two full-rank constrained factors with the same product differ by an invertible right action; orthonormality forces that action to be orthogonal.

D.2 Equivariance of gradients, polar retraction, and transport

Define the Stiefel tangent projection $\Pi_B(X) = X - B\,\mathrm{sym}(B^\top X)$. For $O \in \mathrm{O}(r)$,

$$\begin{aligned}\Pi_{BO}(XO) &= XO - BO\,\mathrm{sym}(O^\top B^\top XO)\\ &= \big[X - B\,\mathrm{sym}(B^\top X)\big]O = \Pi_B(X)O.\end{aligned}$$

The residual action $(\xi_B, \xi_A) \mapsto (\xi_B O, O^\top \xi_A)$ preserves the product Frobenius inner product. The objective partial derivatives follow this action by Equation (A.1), and so do their tangent projections. The resulting Gram and deterministic MGDA weights are invariant.

For any full-column-rank Y, the principal inverse square root obeys

$$(O^\top Y^\top Y O)^{-1/2} = O^\top (Y^\top Y)^{-1/2} O,$$

which gives $\mathrm{polar}(YO) = \mathrm{polar}(Y)O$. If ξ_B is a Stiefel tangent, then

$$(B + \eta\xi_B)^\top (B + \eta\xi_B) = I + \eta^2 \xi_B^\top \xi_B \succ 0.$$

Thus a finite tangent trial has a well-defined polar factor in exact arithmetic. The right factor must retain full row rank to stay in the full-rank LoRA domain at subsequent iterations.

Proof of Theorem 6. Stiefelization gives representatives related by a fixed $O \in \mathrm{O}(r)$. The gradient and MGDA arguments above give directions related by the residual action. Invariant scalar control gives the same step and branch. Polar equivariance gives $B'^{+} = B^{+}O$, and the linear right-factor update gives $A'^{+} = O^\top A^{+}$. Their products agree. The same relation is preserved at every subsequent full-rank iterate, proving the trajectory claim. □

D.3 Derivation of the rank-one metric example

For $c = 1$, $Z = \sigma u v^\top$, take $B = u$ and $A = \sigma v^\top$. Write an ambient gradient as its tangent component $g = auv^\top + kv^\top + u\ell^\top$ plus a normal component N satisfying $Nv = N^\top u = 0$. The normal component does not affect the factor partials. They are $G_B = \sigma(au + k)$ and $G_A = av^\top + \ell^\top$. Projection to the sphere tangent removes the component $\sigma a u$, leaving

$$r_B = \sigma k, \qquad r_A = av^\top + \ell^\top.$$

For two tangent gradients with components (a_i, k_i, ℓ_i), the pairings are therefore

$$\begin{aligned} H^Q_{ij} &= a_i a_j + \langle k_i, k_j \rangle + \langle \ell_i, \ell_j \rangle, \\ H^S_{ij} &= a_i a_j + \sigma^2 \langle k_i, k_j \rangle + \langle \ell_i, \ell_j \rangle. \end{aligned} \tag{D.2}$$

Set $\sigma = 2$, $\ell_1 = \ell_2 = 0$, and let w be a unit vector orthogonal to u. Choosing $(a_1, k_1) = (2, 3w)$ and $(a_2, k_2) = (3, 2w)$ gives

$$H^Q = \begin{bmatrix} 13 & 12 \\ 12 & 13 \end{bmatrix}, \quad H^S = \begin{bmatrix} 40 & 30 \\ 30 & 25 \end{bmatrix}, \quad \lambda^{Q,*} = (\tfrac{1}{2}, \tfrac{1}{2}), \quad \lambda^{S,*} = (0, 1). \tag{D.3}$$

The superscripts Q and S identify QPA and SPA, respectively. Equation (C.4) gives $w^*_Q = (13 - 12)/(13 + 13 - 24) = 1/2$ and $w^*_S = \mathrm{clip}_{[0,1]}((25 - 30)/(40 + 25 - 60)) = 0$. Both are decisions at the same effective matrix. The mapped negative SPA gradient also weights the $kv^\top$ component by σ^2, making the distinction visible in effective directions as well as in the weights.

E Transported optimizer state

E.1 State, bias correction, and transport

Let M_t denote the stored tangent first moment, as in Equation (23), and let n_t count accepted updates before iteration t. Starting from $M_0 = 0$ and $n_0 = 0$, form

$$v_t = \beta_m M_t - (1 - \beta_m) q_t, \qquad \widehat{v}_t = \frac{v_t}{1 - \beta_m^{n_t+1}}, \qquad 0 \le \beta_m < 1. \tag{E.1}$$

Here q_t is the gradient-like MGDA combination; the minus sign makes v_t an update direction. Without bias correction, set $\widehat{v}_t = v_t$. The correction is a shared scalar and does not change any transformation law.

For QPA, $M_t \in T_{Z_t}\mathcal{M}_r$ and the trial is $Z_t + \eta_t \widehat{v}_t$. After acceptance of its unique rank-r projection, transport the uncorrected moment by

$$Z_{t+1} = \mathcal{P}_r(Z_t + \eta_t \widehat{v}_t), \qquad M_{t+1} = \mathcal{P}_{Z_{t+1}}(v_t). \tag{E.2}$$

The accepted trial must have a strict cutoff gap and positive retained singular values; the sufficient bound in Equation (C.6) may be applied to $\widehat{v}_t$ or replaced by direct gap and rank checks.

For SPA, $M_t = (M_{B,t}, M_{A,t})$ with $M_{B,t} \in T_{B_t}\mathrm{St}(m, r)$. Equation (E.1) is applied componentwise to $q_t = (q_{B,t}, q_{A,t})$. The accepted update is

$$B_{t+1} = \mathrm{polar}(B_t + \eta_t \widehat{v}_{B,t}), \qquad A_{t+1} = A_t + \eta_t \widehat{v}_{A,t}.$$

The accepted transport is

$$M_{B,t+1} = \Pi_{B_{t+1}}(v_{B,t}), \qquad M_{A,t+1} = v_{A,t}. \tag{E.3}$$

The state action is $(M_B, M_A) \mapsto (M_B O, O^\top M_A)$. Acceptance sets $n_{t+1} = n_t + 1$. On rejection or a stationary no-op, the factors, moments, and accepted-update counter are retained.

E.2 Complete stateful algorithms

Algorithms 3 and 4 give the complete training loops with bias correction. All factor operations are module-wise; Gram entries sum over modules before one shared MGDA solve. Every MGDA solve uses the minimum-Euclidean-norm tie rule. Stationarity and validity tests use the representation-consistent controls of Appendix E; in exact arithmetic, stationarity means $q_t = 0$. Each loop defaults to an unchanged next state, so a skipped or rejected proposal preserves both parameters and memory.

Algorithm 3 Stateful QPA

Require: Full-rank (B_0, A_0), scale $c > 0$, momentum $0 \le \beta_m < 1$
Require: N iterations, shared inputs ω_t, nominal steps $\eta_t > 0$
Require: Backtracking factor $0 < \gamma < 1$ and maximum reductions J

$M_0 \leftarrow 0, \quad n_0 \leftarrow 0$
for $t = 0, \dots, N-1$ **do**
  $(B_{t+1}, A_{t+1}, M_{t+1}, n_{t+1}) \leftarrow (B_t, A_t, M_t, n_t)$
  Recover $Z_t = cB_tA_t = U_t\Sigma_t V_t^\top$ in thin-SVD form
  for $k = 1, \dots, K$ **do**
    Obtain $(G_{B,k}, G_{A,k})$ at (B_t, A_t) using ω_t
    Recover compact g_k via Equations (C.2)-(C.3)
  end for
  $H^Q_{ij} \leftarrow \langle g_i, g_j \rangle_F$ summed over modules
  $\lambda^*_t \leftarrow \arg\min_{\lambda \in \Delta_K} \lambda^\top H^Q \lambda$
  $q_t \leftarrow \sum_k \lambda^*_{k,t} g_k$
  If the decision is invalid or stationary, **continue** to the next t
  $v_t \leftarrow \beta_m M_t - (1 - \beta_m) q_t$
  $\widehat{v}_t \leftarrow v_t / (1 - \beta_m^{n_t+1})$
  for $b = 0, \dots, J$ **do**
    $\eta \leftarrow \gamma^b \eta_t, \quad W \leftarrow Z_t + \eta \widehat{v}_t$
    If trial data are nonfinite, **continue** to the next b
    Compute the small-core SVD via Equation (C.5); on failure, **continue**
    if all modules pass finite, observed cutoff-gap, and retained-rank checks **then**
      $Z^+ \leftarrow \mathcal{P}_r(W) = U^+\Sigma^+(V^+)^\top$
      $B^+ \leftarrow c^{-1/2} U^+ (\Sigma^+)^{1/2}$
      $A^+ \leftarrow c^{-1/2} (\Sigma^+)^{1/2} (V^+)^\top$
      $M^+ \leftarrow \mathcal{P}_{Z^+}(v_t)$
      $(B_{t+1}, A_{t+1}, M_{t+1}, n_{t+1}) \leftarrow (B^+, A^+, M^+, n_t + 1)$
      break out of the backtracking loop
    end if
  end for
end for
return (B_N, A_N, M_N, n_N)

QPA retries keep the current state, minibatch, and moment fixed; only the shared step size changes. Acceptance uses positive retained singular values and strict observed gaps, without the additional bound in Equation (C.6). Compact operations avoid dense effective matrices.

Algorithm 4 Stateful SPA

Require: Full-rank (B_0, A_0), scale $c > 0$, momentum $0 \le \beta_m < 1$
Require: N iterations, shared inputs ω_t, nominal steps $\eta_t > 0$
$H_B \leftarrow (B_0^\top B_0)^{1/2}$
$(B_0, A_0) \leftarrow (B_0 H_B^{-1}, H_B A_0)$
$M_0 = (M_{B,0}, M_{A,0}) \leftarrow (0, 0), \quad n_0 \leftarrow 0$
for $t = 0, \dots, N-1$ **do**
  $(B_{t+1}, A_{t+1}, M_{t+1}, n_{t+1}) \leftarrow (B_t, A_t, M_t, n_t)$
  for $k = 1, \dots, K$ **do**
    Obtain $(G_{B,k}, G_{A,k})$ at (B_t, A_t) using ω_t
    $r_{B,k} \leftarrow G_{B,k} - B_t \operatorname{sym}(B_t^\top G_{B,k})$
    $r_{A,k} \leftarrow G_{A,k}$
  end for
  $H^S_{ij} \leftarrow \langle r_{B,i}, r_{B,j} \rangle_F + \langle r_{A,i}, r_{A,j} \rangle_F$ summed over modules
  $\lambda_t^* \leftarrow \arg\min_{\lambda \in \Delta_K} \lambda^\top H^S \lambda$
  $q_t = (q_{B,t}, q_{A,t}) \leftarrow \sum_k \lambda^*_{k,t} (r_{B,k}, r_{A,k})$
  If the decision is invalid or stationary, **continue** to the next t
  $v_t \leftarrow \beta_m M_t - (1 - \beta_m) q_t$ componentwise
  $\widehat{v}_t \leftarrow v_t / (1 - \beta_m^{n_t+1})$
  $Y_B \leftarrow B_t + \eta_t \widehat{v}_{B,t}, \quad A^+ \leftarrow A_t + \eta_t \widehat{v}_{A,t}$
  If the trial is nonfinite or Y_B lacks full column rank, **continue**
  $B^+ \leftarrow \operatorname{polar}(Y_B)$
  if all modules pass finite, orthogonality, and full-row-rank A^+ checks **then**
    $M^+ \leftarrow \big(\Pi_{B^+}(v_{B,t}), v_{A,t}\big)$
    $(B_{t+1}, A_{t+1}, M_{t+1}, n_{t+1}) \leftarrow (B^+, A^+, M^+, n_t + 1)$
  end if
end for
return (B_N, A_N, M_N, n_N)

SPA uses one polar/linear trial at η_t and does not use QPA's spectral test or backtracking loop. In both algorithms, n_t counts accepted updates rather than minibatches, and the stored moment is the transported, uncorrected v_t. Transport neither multiplies v_t by the accepted step size nor applies bias correction a second time. The initialization lines apply to fresh warm starts; continuation restores the saved factors, moment, accepted-update counter, and data/randomness position instead of resetting them.

E.3 Equivariance of projection transport

For QPA, the transport of any ambient tangent vector X is $\mathcal{P}_{Z^+}(X)$. Both its argument and the projector depend only on effective objects, so equivalent effective states produce identical transported vectors. For SPA, define

$$\mathcal{T}_{(B,A)\to(B^+,A^+)}(X_B, X_A) = (\Pi_{B^+}(X_B), X_A).$$

If the endpoints are related by $O \in \mathrm{O}(r)$, the projector identity proved in Appendix D gives

$$\mathcal{T}_{(BO,O^\top A)\to(B^+O,O^\top A^+)}(X_B O, O^\top X_A) = \big(\Pi_{B^+}(X_B) O, O^\top X_A\big). \tag{E.4}$$

Thus transport obeys the same residual action as the parameters and tangent directions.

E.4 Proof of Corollary 7

QPA. Suppose two runs have identical Z_t, M_t, and scalar metadata, including n_t, but arbitrary full-rank representatives of the same Z_t. The frontend in Theorem 5 supplies identical q_t. Equation (E.1) then gives identical v_t and $\widehat{v}_t$. Effective-only control selects the same step and branch. On acceptance, both runs project the same trial, so the strict cutoff gap gives the same Z_{t+1}. The tangent projector at that common matrix is identical, and Equation (E.2) gives equal M_{t+1}. Their counters advance together; any additional metadata updated from effective objects and shared randomness

also remains equal. On an invariant no-op, the previous model and optimizer states are retained. Induction from the identical initial effective state proves the QPA claim. □

SPA. Suppose $(B'_t, A'_t) = (B_t O, O^\top A_t)$ and $(M'_{B,t}, M'_{A,t}) = (M_{B,t}O, O^\top M_{A,t})$ with identical scalar metadata. The stateless gradients transform by the same residual action. Linearity of Equation (E.1) and its scalar bias correction preserve that action for both proposed directions. Invariant control selects the same step and branch. Polar equivariance and the linear right-factor update preserve it for the accepted endpoints. Equation (E.4) then preserves it for the transported moments. The accepted counters and invariant metadata evolve identically. An invariant no-op preserves the relation directly. Induction proves the augmented-state correspondence, and $cB'_tA'_t = cB_tOO^\top A_t = cB_tA_t$ gives equality of the effective trajectories. Stiefelization maps arbitrary equivalent full-rank inputs to the required residual orbit; zero initial moments satisfy the state relation automatically. □

E.5 Scope of the pathwise guarantees

The stored moment is the transported uncorrected first moment. A checkpoint therefore contains both model factors and optimizer memory, together with this counter and the data/randomness state needed to continue the same trajectory. The proofs apply in exact arithmetic on the full-rank smooth stratum, with shared exogenous randomness and representation-consistent choices throughout the complete map. QPA requires a unique accepted rank-r projection; SPA requires well-defined polar trials and full-row-rank right factors along the trajectory. Step selection, clipping, numerical tie rules, nonfinite checks, and fallbacks must satisfy the same effective or residual invariance conditions as the main update. These are pathwise representation guarantees: they do not assert stochastic or global convergence, common descent for momentum directions, or simultaneous objective improvement at arbitrary finite step sizes. Finite-precision executions are assessed by numerical representation drift, rather than exact equality. Elementwise squared-gradient buffers generally do not transform as tangent vectors under orthogonal mixing: squaring a rotated array is different from rotating its elementwise squares. The two stateful constructions use tangent first moments with the transformation laws proved above.

F Experimental definitions and evaluation

F.1 Preference objectives and shared initialization

For an objective-specific preferred pair (x, y^+, y^-), define the policy-to-reference score

$$a_Z(x, y) = \beta_{\mathrm{DPO}}\big[\log p_Z(y \mid x) - \log p_{\mathrm{ref}}(y \mid x)\big], \qquad \beta_{\mathrm{DPO}} = 0.1. \tag{F.1}$$

The reference p_{ref} is the frozen pretrained backbone with adapters disabled. The helpfulness and safety losses use the corresponding preference labels in PKU-SafeRLHF:

$$\widetilde{F}_k(Z) = -\mathbb{E}_{(x,y^+,y^-)\sim\mathcal{D}_k} \log \mathrm{sigmoid}\big(a_Z(x, y^+) - a_Z(x, y^-)\big), \quad k \in \{H, S\}. \tag{F.2}$$

Each step uses 16 preference pairs per objective. For initialization $j \in \{1, \ldots, 7\}$, a 100-step weighted-DPO warm start uses helpfulness weight $w_j = (j-1)/6$ and loss $w_j\widetilde{F}_H + (1-w_j)\widetilde{F}_S$. The seven initial effective adapters are shared across all methods and span preference trade-offs; they are not seven independent random training seeds. Each adapted module must have rank eight before formal training. Optimizer states are reset after warm-starting, and formal training uses seed 13.

Both Qwen2.5-7B-Instruct and Llama-3.2-3B-Instruct use query and value projections in their final 12 transformer blocks, giving 24 LoRA modules with $r = 8$, $\alpha = 16$, and no dropout. The LoRA scale in the theory is $c = \alpha/r = 2$. All nine methods train for 2,000 formal steps with nominal learning rate 10^{-3}. The weighted-DPO baseline uses equal helpfulness and safety weights during formal training.

F.2 Decisions, execution, and optimizer state

Weighted DPO applies AdamW to its fixed scalarized loss. The three raw-MGDA compositions use the global factor-gradient Gram in Equation (5) with, respectively, AdamW, LoRA-RITE,

and AdamW followed by BaLoRA balancing. MAdam is the fifth comparison method, using its preference-conditioned adaptive metric (Liu et al., 2026). QPA and SPA use the effective-tangent and product-tangent Grams derived in Appendices C and D, respectively. The stateful variants retain those decisions and add the first moments of Appendix E, with $\beta_m = 0.5$ and bias correction.

QPA checks the observed singular-value cutoff gap and retained rank of each proposed retraction. These checks certify the unique effective projection used in the correctness proof. The sufficient perturbation bound in Equation (C.6) is not imposed as an additional conservative step cap. SPA uses the polar left-factor update and linear right-factor update. For both constructions, any rejection or reduction of the proposed step must use the invariant controls specified in the corresponding theorem.

F.3 Decision-executor representation sensitivity

The first and second letters in RR, RI, IR, and II identify the decision space and execution space, respectively. R denotes a raw-factor decision or executor; I denotes an intrinsic decision or effective-space executor. The four combinations are evaluated on both backbones in FP32 for 300 steps. Three seeds pair the effective initializations and minibatches across equivalent representations. The reference representation is $Q = I$; the two non-identity transformations are $Q = 0.75I$ and a nonuniform Q with condition number 3. Step sizes are calibrated to an initial relative effective displacement of 10^{-3}.

For module ℓ, write its effective adapter at step t as $Z_t^{(\ell)}$. The drift uses the product Frobenius norm over the adapted modules:

$$D_t(Q) = \frac{\left(\sum_\ell \|Z_t^{(\ell)}(Q) - Z_t^{(\ell)}(I)\|_F^2\right)^{1/2}}{\left(\sum_\ell \|Z_t^{(\ell)}(I)\|_F^2\right)^{1/2}}. \tag{F.3}$$

The curves in Figure 2 report the pointwise median across the three seeds and two non-identity representations. The bands span their minimum and maximum; they are not confidence intervals.

F.4 Preference accuracy, means, and hypervolume

For a held-out preference pair, compare the two scores in Equation (F.1). The accuracy credit is one if the preferred response has the larger score, zero if it has the smaller score, and one half for an exact tie. Let $S_{H,\mu,j}$ and $S_{S,\mu,j}$ be the resulting helpfulness and safety accuracies for method μ and initialization j at step 2,000. Figure 3 plots the seven pairs per method, with safety on the horizontal axis. Table 2 reports

$$\overline{S}_{H,\mu} = \frac{1}{7}\sum_{j=1}^{7} S_{H,\mu,j}, \qquad \overline{S}_{S,\mu} = \frac{1}{7}\sum_{j=1}^{7} S_{S,\mu,j}. \tag{F.4}$$

Using coordinates $(\mathrm{Safe}, \mathrm{Help})$ and reference point $(0.5, 0.5)$, the hypervolume is

$$\mathrm{HV}_\mu = \mathrm{area}\left(\bigcup_{\substack{j:\, S_{S,\mu,j}\geq 0.5,\\ S_{H,\mu,j}\geq 0.5}} [0.5, S_{S,\mu,j}] \times [0.5, S_{H,\mu,j}]\right). \tag{F.5}$$

It measures the union of the seven attained rectangles, not the rectangle associated with the mean point. Dominated points do not enlarge the union. Hypervolume is computed from accuracies in $[0, 1]$; the two mean accuracies are displayed as percentages.

F.5 External benchmark aggregation

HarmBench reports the fraction of evaluated requests judged to have elicited a harmful completion, for which lower is better (Mazeika et al., 2024). For XSTest, safe compliance is the fraction of

safe requests answered without refusal, and unsafe refusal is the fraction of unsafe requests refused (Röttger et al., 2024). For each initialization, balanced accuracy is

$$\mathrm{XS}_{\mu,j} = \tfrac{1}{2}\big(\,\mathrm{SafeCompliance}_{\mu,j} + \mathrm{UnsafeRefusal}_{\mu,j}\,\big). \tag{F.6}$$

The HB and XS columns average the corresponding scores across the seven final initialization outputs:

$$\overline{\mathrm{HB}}_{\mu} = \frac{1}{7}\sum_{j=1}^{7}\mathrm{HB}_{\mu,j}, \qquad \overline{\mathrm{XS}}_{\mu} = \frac{1}{7}\sum_{j=1}^{7}\mathrm{XS}_{\mu,j}\,. \tag{F.7}$$

Both are displayed as percentages. These summaries are arithmetic means over the seven shared trade-off initializations, not scores of a single selected checkpoint.

G SAME-Z DECISIONS AND TRAJECTORY SEPARATION

Representation correctness does not require QPA and SPA to choose the same optimization geometry. We examine this distinction at a shared effective adapter and then follow the resulting trajectories.

Setup. We use Llama-3.2-3B-Instruct and the fourth shared initialization, obtained from a 100-step weighted-DPO warm start with equal helpfulness and safety weights. The adapter uses the same 24 query/value modules, rank $r = 8$, scale $c = 2$, and zero dropout as in Section 4. The stateless QPA and SPA runs use seed 13, FP32 model computation, and ten matched non-test minibatches, with 16 preference pairs per objective. For this short comparison, both methods use $\eta = 10^{-5}$. QPA starts from balanced factors and SPA from their product-preserving Stiefelization. Their initial relative effective discrepancy is 3.66×10^{-8}. All ten QPA updates accept the full proposed step.

Decisions at the same effective adapter. Superscripts Q and S identify QPA and SPA, as in Appendix D.3. Using the tangent gradients and SPA combinations from Algorithms 1 and 2, their first-order effective descent directions are

$$d^{Q} = -\sum_{k}\lambda_{k}^{Q,*}g_{k}, \qquad d^{S} = -c(q_{B}\bar{A} + \bar{B}q_{A}). \tag{G.1}$$

Here $(\bar{B}, \bar{A})$ are the Stiefelized factors and (q_B, q_A) is the SPA gradient-like combination with weights $\lambda^{S,*}$. At the initial adapter, the weight difference is $\|\lambda^{Q,*} - \lambda^{S,*}\|_1 = 3.97 \times 10^{-3}$. The effective directions have cosine 0.95990, corresponding to an angle of $16.28°$. Inner products and squared norms are summed over the 24 modules, as in Appendix A.

Step-length control and subsequent trajectories. Write the relative separation of the effective adapters as

$$\varepsilon_t = \frac{\|Z_t^{Q} - Z_t^{S}\|_F}{\|Z_t^{Q}\|_F + 10^{-30}}, \tag{G.2}$$

where the norm is the product Frobenius norm over modules. With the shared learning rate, the first-step gap is 3.65×10^{-4}. In a separate first-step control, we keep QPA at 10^{-5} and set SPA's step size to 2.60445×10^{-6} to match the realized relative update norms. Both relative norms are approximately 1.2630×10^{-4}, differing by only 4.47×10^{-8} relative to QPA's norm, yet the endpoint gap remains 3.58×10^{-5}. Thus the first-step difference is not solely a difference in step magnitude. Continuing the original shared-learning-rate runs from their own iterates, the gap reaches 1.81×10^{-3} (0.181%) after ten steps. Figure 4 summarizes this example of distinct decisions and trajectories under the two representation-correct geometries.

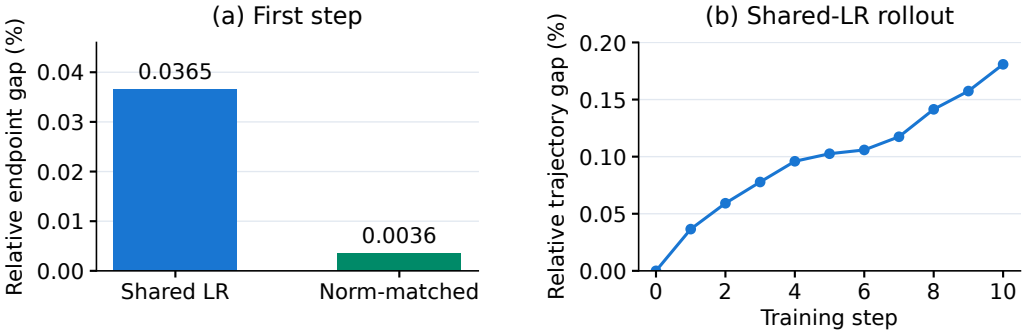

Figure 4: QPA-SPA separation on Llama-3.2-3B-Instruct. Left: first-step endpoint gaps under a shared learning rate and after matching realized relative update norms. Right: ten-step separation under the original shared learning rate. Both panels report Equation (G.2) as a percentage.

H CONFLICT JUDGMENTS UNDER EQUIVALENT LORA REPRESENTATIONS

We test whether a change of LoRA factors, with the effective adapter held fixed, changes the perceived conflict between helpfulness and safety. This diagnostic evaluates the raw-factor, QPA, and SPA decision geometries at the same checkpoints and on the same minibatches. Figure 1 summarizes the observed sign reversals.

Checkpoints and data. For each backbone, we use seven shared warm-start adapters and the step-1,000 and step-2,000 adapters from QPA, Stateful QPA, SPA, and Stateful SPA, yielding $7 + 4 \times 7 \times 2 = 63$ checkpoints. Eight minibatches, each containing 16 preference pairs per objective, are selected from the existing training schedule before inspecting gradients and reused at every checkpoint of that backbone. This gives $63 \times 8 = 504$ paired measurements per model. Tokenized inputs, reference log probabilities, and the 24 rank-eight LoRA modules are held fixed. The seven initializations span preference trade-offs rather than independent random seeds; the reported rates describe this fixed checkpoint–minibatch collection.

Equivalent representations. Each checkpoint is first expressed in balanced factors, preserving $Z = cBA$ with $c = 2$. For every adapted module ℓ, we apply $B^{(\ell)} \mapsto B^{(\ell)}Q_\ell$ and $A^{(\ell)} \mapsto Q_\ell^{-1}A^{(\ell)}$, where $Q_\ell \in \mathrm{GL}(8)$. The fixed grid contains the identity; four scalar transformations $Q_\ell = sI$, $s \in \{1/4, 1/2, 2, 4\}$; three orthogonal controls; and six nonuniform transformations. For the latter, singular values are geometrically spaced from $\sqrt{\kappa}$ to $1/\sqrt{\kappa}$, with three fixed random orientations at each $\kappa \in \{3, 10\}$. The matrices are determined by the module name and orientation seed and reused across checkpoints, minibatches, and decision rules. SPA includes product-preserving Stiefelization before computing its tangent Gram.

Conflict statistic. For each decision geometry, let $H(Q)$ be its global Gram matrix, summed over the 24 modules. We measure the cosine $\chi(Q) = H_{HS}(Q)/\sqrt{H_{HH}(Q)H_{SS}(Q)}$, where H and S index helpfulness and safety. A flip requires $\chi(I)$ and $\chi(Q)$ to have opposite signs and both magnitudes to exceed $\tau = 10^{-6}$. We count a checkpoint–minibatch pair once if any of the ten non-orthogonal transformations satisfies this criterion. No comparisons fall inside the ambiguity band at the primary threshold. The three orthogonal transformations are evaluated separately as controls.

Gradient computation and verification. One pair of objective gradients is computed in FP32 at each checkpoint and minibatch, with TF32 disabled. The exact transformation law in Equation (A.1) and FP64 small-matrix sufficient statistics then give the transformed Grams, without another model pass for every representation. For each backbone, independent forward/backward passes at the first minibatch of the equal-weight warm start check $Q = 0.5I$, one orthogonal transformation, and one nonuniform transformation with $\kappa = 3$. The largest relative discrepancies are below 8×10^{-6} for factor gradients and 9×10^{-6} for Gram matrices. The diagnostic performs no optimizer updates.

Results. Raw MGDA reverses its conflict judgment in 47/504 pairs (9.33%) on Qwen and 146/504 pairs (28.97%) on Llama. QPA and SPA exhibit no flips on either backbone across the tested grid. Scalar transformations alone produce flips in 39/504 Qwen pairs and 137/504 Llama pairs; all three decision rules have zero flips under the orthogonal controls. These observations agree with the raw-Gram transformation law: changing the relative scaling of the factor blocks can reverse their combined inner product without changing the effective model.

Table 1 separates the generating methods and training stages rather than treating their pooled rate as a universal property of a backbone. Increasing the ambiguity threshold to 10^{-3} leaves 47/504 flipped pairs on Qwen and 138/504 on Llama, with no flips for QPA or SPA. The observed raw-factor sensitivity therefore persists after excluding near-zero cosine comparisons more conservatively.

Relation to the nine-method comparison. In our implementation, MGDA with AdamW, MGDA with LoRA-RITE, MGDA with AdamW and BaLoRA, and the MAdam baseline all select preferences using the raw-factor Gram before applying their respective executors. At these identical inputs, they share the raw decision statistic; their executors are not invoked in this diagnostic. Weighted DPO uses fixed preferences, so a conflict-decision flip rate is not applicable. The stateful QPA and SPA variants share their respective instantaneous decision rules with the stateless methods. This experiment isolates decision geometry; the executors and transported optimizer states enter the separate trajectory comparisons.

I Extended related work

Parameter-efficient adaptation and initialization. Adapters insert compact trainable modules into a frozen network (Houlsby et al., 2019), while prefix tuning and prompt tuning optimize continuous conditioning vectors (Li & Liang, 2021; Lester et al., 2021). The unified view of He et al. (2022) organizes these choices by the hidden-state modifications they induce. LoRA instead learns additive low-rank weight updates (Hu et al., 2022). QLoRA combines such updates with a quantized frozen backbone (Dettmers et al., 2023), and LoRA-FA freezes one adapter factor to reduce training memory (Zhang et al., 2023a). These methods concern what is trained and how it is stored.

Other LoRA variants refine the learning dynamics within the adapter. AdaLoRA allocates a variable rank budget across weight matrices (Zhang et al., 2023b); rank-stabilized LoRA adjusts scaling with rank (Kalajdzievski, 2023); LoRA+ uses different learning rates for the two factors (Hayou et al., 2024). DoRA separates weight magnitude and direction (Liu et al., 2024). PiSSA initializes adaptation using principal singular components (Meng et al., 2024), whereas LoRA-GA aligns the initial effective gradient with full fine-tuning (Wang et al., 2024). Our experiments fix the rank, module set, and shared effective initializations so that the optimizer comparison has a common starting point.

Representation-aware optimization. LoRA-RITE uses matrix-valued adaptive preconditioning to address transformation invariance (Yen et al., 2025). BaLoRA selects balanced representatives after a step (Castin et al., 2026). Stiefel LoRA and StelLA use orthonormal constraints and subspace parameterizations (Park et al., 2025; Li et al., 2025). Riemannion treats low-rank adapters directly as fixed-rank matrices and develops a Riemannian Muon optimizer and geometry-informed initialization (Bogachev et al., 2026). These works motivate making the geometry of factorized training explicit. Here the complete-map criterion separates three requirements: an invariant multi-objective decision, a compatible executor, and compatible persistent state. The two constructions choose different metrics while satisfying this same representation requirement.

Gradient balancing and Pareto trade-offs. GradNorm adjusts task weights using gradient norms and relative training rates (Chen et al., 2018). PCGrad projects away conflicting components (Yu et al., 2020), CAGrad balances average descent with worst-task improvement (Liu et al., 2021a), and impartial multi-task learning balances projections of an aggregate direction onto task gradients (Liu et al., 2021b). Nash-MTL formulates the compromise as a bargaining game (Navon et al., 2022). Pareto MTL targets several preference regions (Lin et al., 2019), while preference-directed Pareto optimization combines descent and controlled ascent to seek user-specified trade-offs (Mahapatra & Rajan, 2020). These approaches complement the minimum-norm convex-hull construction of MGDA (Fliege & Svaiter, 2000; Désidéri, 2012; Sener & Koltun, 2018). MAdam further studies

how adaptive execution changes the metric associated with a multi-objective direction (Liu et al., 2026). Our analysis asks whether the gradient geometry and the resulting effective update remain the same when only the LoRA factorization changes.

Geometric computation and memory. The matrix-manifold literature develops embedded and quotient descriptions, Riemannian gradients, and retractions (Edelman et al., 1998; Absil et al., 2008; Boumal, 2023). Fixed-rank matrix optimization supplies compact tangent coordinates (Vandereycken, 2013). The best low-rank approximation theorem and projection-like retractions justify the local truncated-SVD executor (Eckart & Young, 1936; Absil & Malick, 2012). Riemannian stochastic descent (Bonnabel, 2013), adaptive methods (Kasai et al., 2019), and Stiefel momentum (Kong et al., 2023) provide the background for persistent state in changing tangent spaces. Our transported first moments preserve the relevant factor symmetry through the accepted update, bias correction, and subsequent transport.

Preference alignment and evaluation. Human-feedback training provides a route from preference judgments to language-model behavior (Ouyang et al., 2022; Bai et al., 2022). Safe RLHF explicitly separates helpfulness and harmlessness and optimizes a reward under a safety constraint (Dai et al., 2024). DPO expresses preference learning directly through policy-to-reference log-probability ratios (Rafailov et al., 2023); we use this loss for each of the two labels in PKU-SafeRLHF (Ji et al., 2024). Hypervolume summarizes an attained set relative to a fixed reference point (Zitzler & Thiele, 1999). HarmBench standardizes harmful-completion evaluation (Mazeika et al., 2024), while XSTest tests exaggerated refusal on safe requests alongside unsafe controls (Röttger et al., 2024). Together they distinguish unsafe compliance from unnecessary refusal across the seven final initialization outputs.