THE SAME TARGET CAN HELP OR HURT: REGIME-DEPENDENT SHARPNESS UTILITY IN KNOWLEDGE DISTILLATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Should teacher targets be softened or sharpened in knowledge distillation? We study this question with a fixed prediction-preserving intervention, $z_T \mapsto \alpha z_T$, that changes target sharpness while preserving logit ordering and the teacher's top-1 prediction. In a controlled ImageNet-100 DeiT setting, we fix the dataset, architecture pair, teacher, KD objective, and continuation protocol, and vary only the student starting state. The preferred direction reverses between the predeclared endpoints: softening is favored from the untrained state, whereas sharpening is favored from the trained state, with a planned endpoint contrast of $+3.05$ pp. Broader regimes exhibit the same directional non-invariance: controlled CIFAR is softening-favored, matched scratch DeiT shows no measurable directional preference, and pretrained DeiT is sharpening-favored. Gradient, difficulty, target-component, and functional-geometry diagnostics constrain simple explanations but yield no regime-independent directional rule. Target-sharpness utility is therefore relational rather than target-intrinsic: the same intervention can favor opposite directions under different supervision regimes.

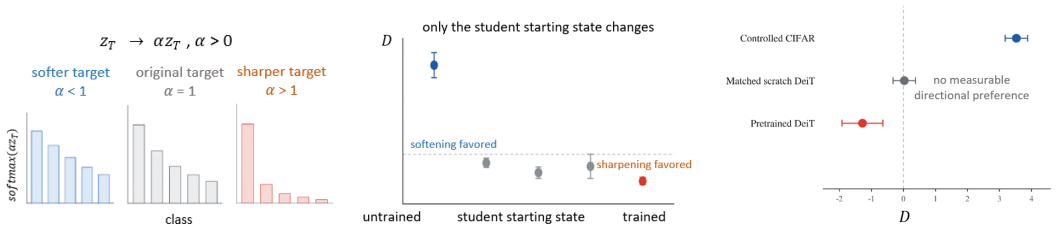
1 INTRODUCTION

Knowledge distillation transfers more than the teacher's predicted class. The teacher's output distribution determines how supervision is distributed across alternative classes Hinton et al. (2015); Ba & Caruana (2014); Zhao et al. (2022), and its confidence structure shapes the signal received by the student Harutyunyan et al. (2023); Kim et al. (2025). Temperature scaling and logit transformations can substantially alter this structure Li et al. (2023); Zheng & Yang (2024); Sun et al. (2024). Teacher quality and training state also affect distillation outcomes: a stronger or later teacher is not always more useful to the student Cho & Hariharan (2019); Wang et al. (2022); Harutyunyan et al. (2023). These findings establish that the effectiveness of teacher supervision depends on the teacher, the student, and the form of the target. They leave a distinct question unresolved: if the target intervention itself is fixed, does its preferred direction remain invariant across teacher–student supervision regimes? Variation in effect magnitude is not the same as variation in effect direction. If the same intervention favors softening in one regime and sharpening in another, then target sharpness cannot be assigned a regime-independent preferred direction from the transformation alone.

Answering this question requires a controlled comparison. Differences across teachers, checkpoints, architectures, datasets, or training histories can reveal changes in distillation behavior, but they cannot by themselves show whether the response to the same intervention can reverse. We therefore use positive logit rescaling, $z_T \mapsto \alpha z_T$ with $\alpha > 0$, as a controlled probe of target sharpness. The transformation changes the concentration of the teacher distribution while preserving logit ordering and the teacher's top-1 prediction Hinton et al. (2015); Li et al. (2023); Zheng & Yang (2024); Sun et al. (2024). Figure 1a illustrates this fixed prediction-preserving intervention. We then ask whether its preferred direction changes when the teacher–student supervision state changes.

We first test directional invariance in a controlled ImageNet-100 DeiT setting. We fix the dataset, architecture pair, teacher, KD objective, and continuation protocol, and vary only the student starting state. The same intervention favors softening at the predeclared untrained endpoint and sharpening at the trained endpoint, with a planned endpoint contrast of $+3.05$ pp. The intermediate states do not

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.



(a) Prediction-preserving sharpness intervention.

(b) Controlled reversal under a fixed setup.

(c) Broader regimes span distinct directions.

Figure 1: **Preferred sharpness direction varies across supervision regimes.** **(a)** Positive logit rescaling, $z_T \mapsto \alpha z_T$, changes target sharpness while preserving logit ordering and the teacher's top-1 prediction. **(b)** Under a fixed ImageNet-100 DeiT setup, changing only the student starting state reverses the preferred direction between the predeclared untrained and trained endpoints; intermediate states are shown for context and do not imply a gradual or monotonic transition. **(c)** Broader regimes span softening-favored, no measurable directional preference, and sharpening-favored outcomes. Here, $D > 0$ favors softening and $D < 0$ favors sharpening.

support a gradual or monotonic transition over training progress. The controlled result is therefore an endpoint reversal: within this fixed family, changing only the student starting state is sufficient to reverse the observed sharpness preference. Figure 1b summarizes this result.

We then examine whether this directional variation extends beyond the controlled bridge. The same probe is softening-favored in controlled CIFAR, shows no measurable directional preference in matched scratch DeiT, and is sharpening-favored in pretrained DeiT; Figure 1c summarizes these regimes. These comparisons establish breadth rather than identify a common cause. We therefore test several natural explanations for the observed variation. Gradient magnitude, example difficulty and correctness, target-component probes, and functional geometry constrain different aspects of the response, but none yields a regime-independent directional rule.

Together, these results show that target-sharpness utility has no regime-independent preferred direction across the tested supervision regimes. Whether softening or sharpening is favored depends on the teacher–student supervision regime in which the target acts. We therefore treat sharpness utility as relational rather than target-intrinsic.

Our main contributions are:

- **Controlled directional reversal.** We show that the preferred direction of a fixed prediction-preserving sharpness intervention can reverse when only the student starting state changes within a fixed ImageNet-100 DeiT family.
- **Distinct response directions across broader regimes.** The same probe is softening-favored in controlled CIFAR, shows no measurable directional preference in matched scratch DeiT, and is sharpening-favored in pretrained DeiT.
- **No regime-independent directional rule.** Gradient magnitude, difficulty and correctness, target-component probes, and functional geometry constrain different aspects of the observed response, but none determines sharpness direction across regimes.

2 Related Work

Distillation utility is conditional. A more accurate, higher-capacity, or more fully trained teacher does not necessarily produce a better student, especially when teacher and student differ in capacity, optimization state, or the supervision that the student can effectively exploit Cho & Hariharan (2019); Stanton et al. (2021); Harutyunyan et al. (2023); Mirzadeh et al. (2020); Qian et al. (2025); Huang et al. (2025). More broadly, theoretical and empirical studies have shown that distillation behavior can depend strongly on data and model geometry as well as seemingly implicit training-design choices Phuong & Lampert (2019); Beyer et al. (2022). This observation has motivated methods that select intermediate teacher checkpoints or exploit teacher trajectories instead of re-

lying only on the final converged model Wang et al. (2022); Panigrahi et al. (2025). These studies establish that distillation utility depends on the relation between teacher and student states rather than on teacher quality alone. They characterize how utility changes across teachers or training states, but they do not test whether the response to one fixed target intervention preserves its direction as the supervision regime changes.

Target transformations are usually optimized for performance. Temperature scaling has long been used to control the concentration of teacher predictions in knowledge distillation Hinton et al. (2015). Later work adapts temperature during training, calibrates teacher confidence, transforms or normalizes logits, or restructures logit-based supervision to improve transfer Li et al. (2023); Kim et al. (2025); Zheng & Yang (2024); Sun et al. (2024); Zhao et al. (2022); Jin et al. (2023). These approaches treat the form of the teacher target as a design variable. We use the transformation differently: we hold a prediction-preserving sharpness intervention fixed and use it as an experimental probe. This design separates the question of which target transformation performs best from the question of whether the same intervention favors the same direction across teacher–student supervision regimes.

Compatibility measures describe supervision state. Prior work has used gradient similarity, supervision complexity, prediction geometry, and interactions between task and distillation gradients to characterize or adapt how teacher supervision interacts with the current student state Zhu & Wang (2021); Harutyunyan et al. (2023); Stanton et al. (2021); Hu et al. (2024); Hayder et al. (2026). Recent work has also treated compatibility as a student-specific teacher-selection problem, using student-gradient statistics to guide teacher and temperature choices Panigrahi et al. (2026). These measures can reveal structure in teacher–student compatibility and, in some settings, guide how distillation is applied. Our CE–KD geometry has a narrower role. We use it as a diagnostic readout of supervision state rather than as an adaptation rule or an assumed mechanism. A compatibility measure can organize supervision states without, by itself, specifying whether softening or sharpening should be preferred.

Taken together, these lines of work establish that teacher usefulness depends on the teacher–student relation, that target structure can be deliberately modified, and that supervision compatibility can be characterized. They do not by themselves answer whether a fixed target intervention has an invariant preferred direction across supervision regimes. We study this question as a directional invariance problem: the intervention is held fixed, and the object of interest is whether the sign of its response is preserved as the surrounding teacher–student supervision regime changes.

3 A Controlled Test of Target-Sharpness Direction

3.1 A Prediction-Preserving Intervention

A directional test must change target sharpness without changing the teacher's predicted class. We therefore rescale the teacher logits by a positive factor,

$$z_T^{(\alpha)} = \alpha z_T, \qquad \alpha > 0. \tag{1}$$

Because $\alpha > 0$, the transformation preserves the ordering of all teacher logits and therefore leaves the teacher's top-1 prediction unchanged. It changes the concentration of the induced target distribution: $\alpha < 1$ produces a softer target, whereas $\alpha > 1$ produces a sharper one. With the KD temperature fixed at $T = 4$ throughout the directional comparisons, this rescaling is equivalent to changing the teacher-side effective temperature,

$$\mathrm{softmax}\Big(\frac{\alpha z_T}{T}\Big) = \mathrm{softmax}\Big(\frac{z_T}{T/\alpha}\Big). \tag{2}$$

We use this transformation as a controlled probe, not as a new distillation method. Within each comparison, we keep the teacher–student setup, KD formulation, optimization protocol, and training budget fixed, and vary only target sharpness.

3.2 Separating Direction from Absolute KD Benefit

Directional preference is distinct from absolute distillation benefit. For each sharpness level, we first measure KD relative to a matched CE continuation,

$$G(\alpha) = \mathrm{Acc}_{\mathrm{KD},\alpha} - \mathrm{Acc}_{\mathrm{CE,ref}}. \tag{3}$$

The matched reference separates the KD outcome from improvement associated with continued training alone. We then define the directional contrast between the fixed endpoint interventions, $\alpha = 0.5$ and $\alpha = 4.0$,

$$D = G(0.5) - G(4.0). \tag{4}$$

Because both gains use the same CE reference, D reduces algebraically to $\mathrm{Acc}_{\mathrm{KD},0.5} - \mathrm{Acc}_{\mathrm{KD},4.0}$. We retain $G(\alpha)$ because this cancellation does not make absolute KD benefit and directional preference interchangeable. A positive D means that softening is favored over sharpening, whereas a negative D means that sharpening is favored over softening. Neither sign implies that the favored condition itself outperforms the matched CE reference. When uncertainty does not resolve either direction, we report no measurable directional preference rather than define neutrality through an arbitrary numerical threshold.

4 A Controlled Reversal Across Student Starting States

Section 3 defines sharpness preference through the sign of a fixed intervention. We now test whether that sign can reverse while holding the teacher, dataset, architecture pair, KD objective, and continuation protocol fixed. This comparison isolates student starting state within a single teacher–student family.

4.1 Varying Only the Student Starting State

We use ImageNet-100 with a pretrained DeiT-S teacher, a DeiT-T student, a fixed KD objective, and a common 20-epoch local-reset continuation protocol that carries over only the student weights while reinitializing the optimizer and schedule. We evaluate five predeclared student starting states drawn from three independently trained CE trajectories, corresponding to CE epochs 0, 30, 60, 90, and 120, denoted S_0–S_4. Here, S_0 is the seeded random initialization and S_4 is the trained endpoint. Within each state, the softened and sharpened conditions share the same starting weights, data order, augmentation randomness, optimizer initialization, schedule, and training budget. Thus, across the five state-level comparisons, only the student starting state changes, while within each paired comparison only target sharpness changes.

4.2 The Preferred Direction Reverses

The two predeclared endpoints favor opposite sharpness directions. At the untrained endpoint S_0, the mean directional contrast is $D = +2.35$ pp across three trajectories, with values 2.84, 2.48, and 1.72 pp. At the trained endpoint S_4, the mean is $D = -0.70$ pp, with values -0.52, -0.86, and -0.72 pp. The planned endpoint contrast is

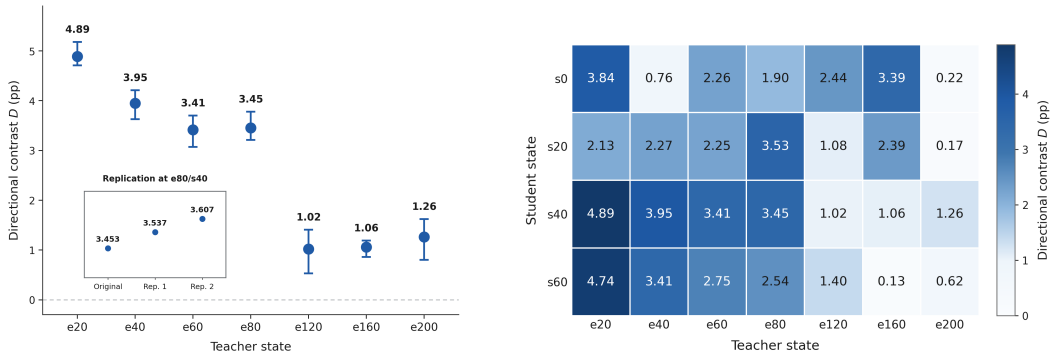
$$R = D(S_0) - D(S_4) = +3.05 \text{ pp}, \tag{5}$$

with a 95% hierarchical paired-bootstrap interval of $[2.17, 3.91]$ pp and a trajectory-level Student-t interval of $[1.74, 4.35]$ pp. The same prediction-preserving intervention is therefore softening-favored at the untrained endpoint and sharpening-favored at the trained endpoint.

Table 1 reports the absolute KD gains at the two predeclared endpoints and clarifies what the directional reversal means. At S_0, $G(0.5) = +1.23$ pp and $G(4) = -1.12$ pp, so strong sharpening underperforms the matched CE continuation. At S_4, $G(0.5) = -0.10$ pp and $G(4) = +0.60$ pp. The softened condition at S_4 is not measurably below CE. The negative value of D at this endpoint therefore indicates that sharpening is preferred relative to softening; it does not show that softening itself is harmful.

Table 1: Controlled endpoint reversal under the fixed ImageNet-100 DeiT setup. The student starting state is the only varying component. All values are reported in percentage points.

Student state	$G(0.5)$	$G(4)$	D
S_0 (untrained)	$+1.23$	-1.12	$+2.35$
S_4 (trained)	-0.10	$+0.60$	-0.70



(a) Teacher-state sweep.

(b) Full state map.

Figure 2: **Controlled CIFAR remains softening-favored while response magnitude varies across teacher–student states.** **(a)** With the student fixed at $s40$, D varies substantially across seven teacher states but remains positive; the inset shows replication at $e80/s40$ across three independently trained teacher trajectories. **(b)** Across the full 7×4 teacher–student state map, all 28 measured combinations have positive mean D, despite substantial variation in magnitude. Points denote means across three student seeds; vertical ranges show the minimum and maximum across seeds.

4.3 An Endpoint Reversal, Not a Training Law

The intermediate states limit the claim to an endpoint reversal. Mean D across S_0–S_4 is $+2.35$, -0.22, -0.48, -0.31, and -0.70 pp. All trained states have non-positive mean D, but their uncertainty differs: the intervals for S_1 and S_3 span zero under both schemes, while S_2 is supported by the bootstrap interval but not by the Student-t interval. These results do not establish a gradual or monotonic sign transition. They establish opposite sharpness preferences at the predeclared S_0 and S_4 endpoints.

A symmetric log-scale contrast, $D_{0.5,2} = G(0.5) - G(2)$, gives $+0.99$, -0.54, and -0.85 pp at S_0, S_2, and S_4, with the same sign across all three trajectories at each state. The bootstrap intervals preserve these directions, whereas the Student-t intervals span zero. We therefore treat this contrast as a qualitative robustness check rather than an independent inferential result. Within this controlled family, changing the student starting state is sufficient to reverse the endpoint preference, but this result does not imply that student state alone explains the broader cross-regime differences.

5 Broader Regimes Span Distinct Sharpness Directions

Section 4 shows that the preferred sharpness direction can reverse within a tightly controlled teacher–student family. We next apply the same probe across broader supervision regimes. Controlled CIFAR uses a ResNet-34/ResNet-18 pair, while both DeiT regimes use DeiT-Small/DeiT-Tiny. These settings differ in dataset, architecture, or training history, so they establish breadth rather than isolate a common cause.

5.1 Controlled CIFAR: Stable Direction, Variable Magnitude

Controlled CIFAR uses 30-epoch continuations from frozen student states on the unrestarted global schedule. Across all 28 measured conditions, mean D is positive.

Table 2: Directional response across the broader supervision regimes. For scratch and pretrained DeiT, brackets denote 95% confidence intervals for the within-regime state contrast.

Regime	Evaluation	D (pp)	Within-regime evidence
Controlled CIFAR	7×4 state map	$+0.133$ to $+4.890$	$28/28$ positive
Scratch DeiT	$e40,\ e120$	$-0.003,\ +0.057$	$-0.060\ [-0.38, +0.25]$
Pretrained DeiT	$e1,\ e30$	$-1.917,\ -0.640$	$-1.277\ [-1.560, -1.050]$

With the student fixed at $s40$, D varies non-monotonically across teacher states, from 4.890 pp at $e20$ to 1.020 pp at $e120$ and 1.263 pp at $e200$. The $e80/s40$ response replicates across three independent teacher trajectories, with $D = 3.453$, 3.537, and 3.607 pp (Figure 2a). Across the full 7×4 state map, all 28 teacher–student combinations remain positive across three student seeds, spanning 0.133–4.890 pp (Figure 2b). Controlled CIFAR is therefore reproducibly softening-favored despite substantial variation in response magnitude.

5.2 Matched Scratch DeiT: Direction Unresolved

Matched scratch DeiT does not reproduce the softening-favored direction observed in controlled CIFAR. At the two evaluated teacher states,

$$D_{e40} = -0.003 \text{ pp}, \qquad D_{e120} = +0.057 \text{ pp}, \qquad D_{e40} - D_{e120} = -0.060 \text{ pp}. \tag{6}$$

The planned between-state contrast has a 95% confidence interval of $[-0.38, +0.25]$ pp. These measurements do not resolve either direction. Their near-zero values should not be interpreted as evidence of equivalence between softening and sharpening.

5.3 Pretrained DeiT: Sharpening-Favored

Pretrained DeiT exhibits the opposite directional preference from controlled CIFAR. At the mature evaluated state,

$$D_{\text{mature}} = -0.640 \text{ pp}, \qquad 95\% \text{ CI} = [-0.88, -0.433] \text{ pp}. \tag{7}$$

The earlier evaluated state has the same sign with larger magnitude,

$$G(0.5) = +0.18 \text{ pp}, \qquad G(4.0) = +2.10 \text{ pp}, \qquad D_{\text{early}} = -1.917 \text{ pp}, \tag{8}$$

with a 95% confidence interval of $[-2.117, -1.643]$ pp. Both sharpness conditions improve over the matched CE reference, but sharpening yields the larger gain.

The between-state contrast is

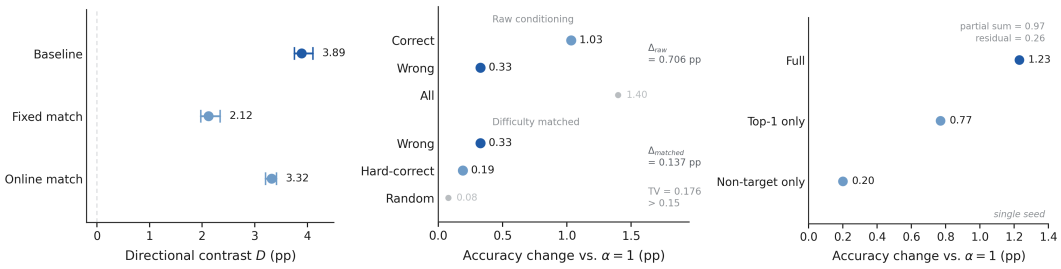
$$D_{\text{early}} - D_{\text{mature}} = -1.277 \text{ pp}, \qquad 95\% \text{ CI} = [-1.560, -1.050] \text{ pp}. \tag{9}$$

Both evaluated states remain sharpening-favored. These results therefore show state-dependent variation in response magnitude within the pretrained regime, not a monotonic state law or a second within-regime directional reversal.

Table 2 places the three broader regimes on the same directional scale. Controlled CIFAR remains positive across all 28 measured states, matched scratch DeiT resolves no directional preference, and pretrained DeiT remains negative at both evaluated states. Because these regimes differ in dataset, architecture, and training history, the comparison establishes breadth rather than a common source of the differences.

6 Constraining Natural Single-Factor Explanations

The preceding sections establish directional non-invariance without identifying its source. We therefore test three natural reductions of the softening-favored response in controlled CIFAR: CE–KD gradient magnitude, example correctness and difficulty, and isolated components of the teacher target. Figure 3 shows these diagnostics, and Table 3 summarizes their quantitative outcomes.



(a) Gradient matching. (b) Difficulty matching. (c) Target-component probes.

Figure 3: **Natural single-factor diagnostics constrain but do not resolve the controlled CIFAR response. (a)** Gradient matching changes the magnitude of D while preserving its positive direction. **(b)** Difficulty matching reduces the correctness gap from 0.706 to 0.137 pp, but the predefined class-balance gate is not satisfied ($\mathrm{TV} = 0.176 > 0.15$). **(c)** Full softening yields 1.23 pp, compared with 0.77 and 0.20 pp for the top-1-only and non-target-only probes; the partial probes are non-additive.

Table 3: Single-factor diagnostics in controlled CIFAR. Values are reported in percentage points unless noted otherwise.

Probe	Reference	Controlled / partial	Residual / gate
Gradient magnitude	$D = 3.887$	2.12 (fixed), 3.32 (online)	$D > 0$
Difficulty matching	0.706 gap	0.137 gap	$\mathrm{TV} = 0.176 > 0.15$
Target components	1.23 (full)	0.77 (top-1), 0.20 (non-target)	0.26 residual

6.1 Gradient Matching Changes Magnitude, Not Direction

The softening-favored response remains positive under both gradient controls. The strict matched baseline gives $D = 3.887$ pp across three seeds, with values in $[3.75, 4.10]$ pp. Fixed gradient matching reduces the mean to 2.12 pp, with values in $[1.97, 2.34]$ pp, while online matching gives $D = 3.32$ pp, with values in $[3.20, 3.41]$ pp. Figure 3a summarizes these controls. Gradient magnitude therefore modulates response magnitude, but these controls do not support it as a sufficient explanation of the observed direction.

6.2 Difficulty Narrows the Correctness Gap

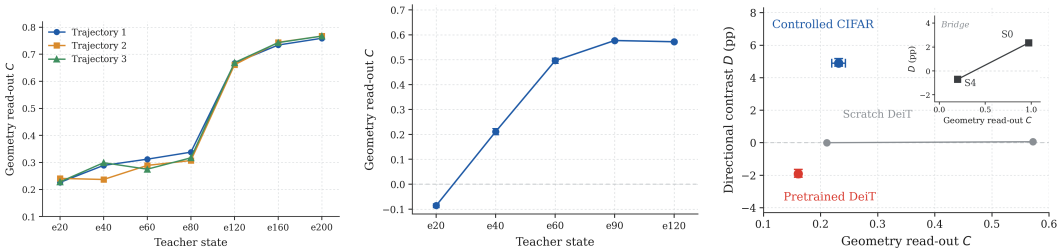
Teacher-correct and teacher-wrong examples initially differ by 0.706 pp under softening: $+1.033$ versus $+0.327$ pp relative to the original target at $\alpha = 1$, with $+1.397$ pp over all examples. Difficulty matching reduces this gap to 0.137 pp. The hard-correct subset yields $+0.190$ pp, compared with $+0.327$ pp on teacher-wrong examples and $+0.077$ pp on a random matched subset. Figure 3b shows the raw and difficulty-matched comparisons. The residual gap falls below the preregistered 0.40 pp equivalence threshold, but the matched subsets fail the predefined class-balance gate, with $\mathrm{TV} = 0.176 > 0.15$. We therefore interpret the result as evidence that difficulty explains much of the apparent correctness difference, not as a clean equivalence result.

6.3 Target-Component Probes Are Non-Additive

Full softening yields $+1.23$ pp relative to $\alpha = 1$, versus $+0.77$ and $+0.20$ pp for the top-1-only and non-target-only probes. Their sum, $+0.97$ pp, leaves a 0.26 pp residual (Figure 3c). These single-seed probes are non-additive, so we do not interpret them as a decomposition; neither component alone reproduces the full response.

7 Functional Geometry Tracks State Without Fixing Direction

A useful description of supervision state need not determine the response to an intervention. We therefore characterize the functional relation between CE and KD supervision using the pooled



(a) Controlled CIFAR geometry readout.

(b) Scratch DeiT geometry progression.

(c) Geometry and directional response across regimes.

Figure 4: **Functional geometry tracks supervision state but does not fix sharpness direction. (a)** Controlled CIFAR shows reproducible state-dependent geometry. **(b)** Scratch DeiT shows a large shift followed by a late plateau. **(c)** Geometry does not uniquely determine direction: scratch DeiT remains unresolved, the controlled bridge reverses, and broader regimes show opposite preferences at low to low-moderate C.

cosine similarity

$$C = \frac{\sum_i g_{\mathrm{CE}}^{(i)} \cdot g_{\mathrm{KD}}^{(i)}}{\sqrt{\sum_i \|g_{\mathrm{CE}}^{(i)}\|^2}\sqrt{\sum_i \|g_{\mathrm{KD}}^{(i)}\|^2}}, \tag{10}$$

computed over a fixed diagnostic set of 2,000 examples. Here, g_{CE} and g_{KD} are gradients with respect to student logits, and C is a pooled cosine rather than an average of per-example similarities. We use C as a diagnostic readout of supervision geometry, not as a causal or universal predictor of sharpness direction.

7.1 Geometry Provides a Reproducible State Readout

In controlled CIFAR with the student fixed at $s40$, C increases from 0.232 at $e20$ to 0.321 at $e80$ and 0.764 at $e200$. Three independent teacher trajectories at $e80$ give $C = 0.338$, 0.307, and 0.317, showing reproducible state-dependent structure (Figure 4a). Under leave-one-teacher-trajectory-out evaluation, C also captures the measured state variation better than teacher epoch:

$$\mathrm{RMSE}_C = 0.685 < 0.776 = \mathrm{RMSE}_{\mathrm{epoch}}, \qquad \rho_C = 0.869 > 0.814 = \rho_{\mathrm{epoch}}. \tag{11}$$

These results support C as an informative within-regime descriptor of supervision state, not as a determinant of sharpness utility.

Scratch DeiT shows a similar within-regime pattern. With the student fixed at $s60$, C evolves as $-0.086 \to 0.211 \to 0.496 \to 0.577 \to 0.572$ over $e20, e40, e60, e90, e120$, with a substantial shift followed by a late plateau (Figure 4b). We do not compare the absolute scale of C across CIFAR and DeiT because the settings differ in dataset, architecture, and output dimensionality.

7.2 State Geometry Is Not a Directional Rule

Scratch DeiT separates supervision geometry from directional preference: C changes from 0.211 to 0.572 between $e40$ and $e120$, while D remains unresolved at -0.003 and $+0.057$ pp. In the controlled bridge, D reverses from $+2.35$ to -0.70 pp as geometry changes between the predeclared endpoints. We interpret this association descriptively because the bridge was designed to establish directional reversal, not to identify C as its cause or predictor.

Across regimes, controlled CIFAR at $e20/s40$ has $(C, D) = (0.232, +4.890\ \mathrm{pp})$, whereas early pretrained DeiT has $(C, D) = (0.161, -1.917\ \mathrm{pp})$. The two settings therefore exhibit opposite sharpness preferences at low to low-moderate values of C. Because C is not calibrated across architectures, datasets, or output dimensionalities, this comparison supports only the narrower conclusion that low C does not imply a common sharpness direction (Figure 4c).

Table 4 makes the distinction explicit: C organizes supervision state but does not uniquely map to sharpness direction. Functional geometry is therefore a useful state descriptor, not a regime-independent directional rule.

Table 4: Functional geometry and directional response at selected states. Values of C are interpreted within regimes and are not treated as calibrated across architectures, datasets, or output dimensionalities.

Setting	State	C	D (pp)
Scratch DeiT	$e40$	0.211	-0.003
Scratch DeiT	$e120$	0.572	$+0.057$
Controlled CIFAR	$e20/s40$	0.232	$+4.890$
Pretrained DeiT	early	0.161	-1.917

8 Discussion and Scope

Our results support a relational view of sharpness utility. The same prediction-preserving intervention does not exhibit a regime-independent preferred direction across the tested supervision regimes. In the controlled bridge, changing only the student starting state reverses the response between the predeclared untrained and trained endpoints. The broader regimes span softening-favored, unresolved, and sharpening-favored outcomes. The diagnostic analyses constrain several simple reductions without yielding a directional rule: gradient matching changes response magnitude while D remains positive, difficulty matching narrows the correctness gap but fails the predefined balance gate, target-component probes are non-additive, and functional geometry tracks supervision state without fixing the sign of D.

These results establish directional non-invariance, not a general theory of sharpness utility. The controlled bridge establishes an endpoint reversal rather than a gradual or monotonic training law, and it does not imply that student state alone explains the broader cross-regime differences. The broader comparisons vary in dataset, architecture, and training history. Scratch DeiT shows no measurable directional preference rather than equivalence, and the evaluated pretrained states do not establish a monotonic teacher-state law. We therefore identify one controlled axis along which the preferred direction can reverse, while the conditions governing such reversals remain open.

9 Conclusion

We study target sharpness through a fixed prediction-preserving intervention and show that its preferred direction is not invariant across the tested supervision regimes. Under a fixed ImageNet-100 DeiT setup, changing only the student starting state reverses the preferred direction between the predeclared untrained and trained endpoints. Broader regimes span softening-favored, no measurable directional preference, and sharpening-favored responses, while the diagnostic analyses yield no regime-independent directional rule. These results support a relational view of sharpness utility: the response to the same target intervention depends on the supervision regime in which it acts rather than on target sharpness alone. The conditions governing such directional reversals remain open.

AI Use Statement

Generative AI tools were used to assist with manuscript drafting and language polishing, literature retrieval and discovery, and research ideation and execution. In particular, these tools were used to help refine scientific writing, identify potentially relevant prior work, discuss experimental designs and analyses, and assist with research workflows. All AI-assisted outputs were reviewed and verified by the authors before use. The authors independently evaluated the scientific claims, experimental results, citations, and final manuscript, and take full responsibility for the content of the paper.

Reproducibility Statement

We provide the experimental protocols, continuation schemes, replication structure, and statistical procedures needed to reproduce the reported comparisons in Appendix A. Appendix B reports the complete controlled-bridge results, including state-level contrasts, sharpness-response measurements, and robustness analyses. Appendix C provides the full controlled-CIFAR state map and additional details for the diagnostic and functional-geometry analyses. The main text defines the

directional intervention and evaluation quantities in Section 3, presents the controlled reversal in Section 4, and reports the broader and diagnostic analyses in Sections 5–7.

REFERENCES

Lei J Ba and Rich Caruana. Do deep nets really need to be deep? *Advances in neural information processing systems*, 27, 2014.

Lucas Beyer, Xiaohua Zhai, Amélie Royer, Larisa Markeeva, Rohan Anil, and Alexander Kolesnikov. Knowledge distillation: A good teacher is patient and consistent. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10915–10924. IEEE, 2022.

Jang Hyun Cho and Bharath Hariharan. On the efficacy of knowledge distillation. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 4793–4801. IEEE, 2019.

Hrayr Harutyunyan, Ankit Singh Rawat, Aditya Krishna Menon, Seungyeon Kim, and Sanjiv Kumar. Supervision complexity and its role in knowledge distillation. *arXiv preprint arXiv:2301.12245*, 2023.

Zeeshan Hayder, Ali Cheraghian, Lars Petersson, Mehrtash Harandi, and Richard Hartley. Dto-kd: Dynamic trade-off optimization for effective knowledge distillation. In *The Fourteenth International Conference on Learning Representations*, 2026.

Geoffrey Hinton, Oriol Vinyals, and Jeff Dean. Distilling the knowledge in a neural network. *arXiv preprint arXiv:1503.02531*, 2015.

Chengming Hu, Haolun Wu, Xuan Li, Chen Ma, Xi Chen, Boyu Wang, Jun Yan, and Xue Liu. Less or more from teacher: Exploiting trilateral geometry for knowledge distillation. In *International Conference on Learning Representations*, volume 2024, pp. 24503–24518, 2024.

Tao Huang, Shan You, Fei Wang, Chen Qian, and Chang Xu. Dist+: Knowledge distillation from a stronger adaptive teacher. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 47 (7):5571–5585, 2025.

Ying Jin, Jiaqi Wang, and Dahua Lin. Multi-level logit distillation. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 24276–24285. IEEE, 2023.

Suyoung Kim, Seonguk Park, Junhoo Lee, and Nojun Kwak. The role of teacher calibration in knowledge distillation. *IEEE Access*, 13:115548–115557, 2025.

Zheng Li, Xiang Li, Lingfeng Yang, Borui Zhao, Renjie Song, Lei Luo, Jun Li, and Jian Yang. Curriculum temperature for knowledge distillation. In *Proceedings of the AAAI conference on artificial intelligence*, volume 37, pp. 1504–1512, 2023.

Seyed Iman Mirzadeh, Mehrdad Farajtabar, Ang Li, Nir Levine, Akihiro Matsukawa, and Hassan Ghasemzadeh. Improved knowledge distillation via teacher assistant. In *Proceedings of the AAAI conference on artificial intelligence*, volume 34, pp. 5191–5198, 2020.

Abhishek Panigrahi, Bingbin Liu, Sadhika Malladi, Andrej Risteski, and Surbhi Goel. Progressive distillation induces an implicit curriculum. In *International Conference on Learning Representations*, volume 2025, pp. 23694–23742, 2025.

Abhishek Panigrahi, Bingbin Liu, Sadhika Malladi, Sham Kakade, and Surbhi Goel. In good graces: Principled teacher selection for knowledge distillation. In *International Conference on Learning Representations*, volume 2026, pp. 54779–54822, 2026.

Mary Phuong and Christoph Lampert. Towards understanding knowledge distillation. In *International conference on machine learning*, pp. 5142–5151. PMLR, 2019.

Chengyao Qian, Trung Le, and Mehrtash Harandi. A good teacher adapts their knowledge for distillation. In *2025 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 1239–1248. IEEE, 2025.

Samuel Stanton, Pavel Izmailov, Polina Kirichenko, Alexander A Alemi, and Andrew G Wilson. Does knowledge distillation really work? *Advances in neural information processing systems*, 34:6906–6919, 2021.

Shangquan Sun, Wenqi Ren, Jingzhi Li, Rui Wang, and Xiaochun Cao. Logit standardization in knowledge distillation. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 15731–15740. IEEE, 2024.

Chaofei Wang, Qisen Yang, Rui Huang, Shiji Song, and Gao Huang. Efficient knowledge distillation from model checkpoints. *Advances in Neural Information Processing Systems*, 35:607–619, 2022.

Borui Zhao, Quan Cui, Renjie Song, Yiyu Qiu, and Jiajun Liang. Decoupled knowledge distillation. In *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 11943–11952. IEEE, 2022.

Kaixiang Zheng and En-Hui Yang. Knowledge distillation based on transformed teacher matching. In *International Conference on Learning Representations*, volume 2024, pp. 44963–44979, 2024.

Yichen Zhu and Yi Wang. Student customized knowledge distillation: Bridging the gap between student and teacher. In *2021 IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 5037–5046. IEEE, 2021.

A Experimental Protocol and Statistical Details

This appendix specifies the protocols and statistical units used in the directional comparisons reported in the main text.

A.1 Directional Probe

All directional comparisons use the prediction-preserving intervention

$$z_T^{(\alpha)} = \alpha z_T, \qquad \alpha > 0, \tag{12}$$

with KD temperature fixed at $T = 4$. The primary directional statistic compares the fixed endpoint interventions $\alpha = 0.5$ and $\alpha = 4.0$:

$$D = G(0.5) - G(4.0), \qquad G(\alpha) = \mathrm{Acc}_{\mathrm{KD},\alpha} - \mathrm{Acc}_{\mathrm{CE,ref}}. \tag{13}$$

Intermediate values of α appear only in designated response-curve and diagnostic analyses.

Controlled CIFAR and the broader DeiT regimes use

$$\mathcal{L} = \mathcal{L}_{\mathrm{CE}} + 0.5\,T^2 \mathrm{KL}\left(\mathrm{softmax}\Big(\frac{\alpha z_T}{T}\Big) \,\Big\|\, \mathrm{softmax}\Big(\frac{z_S}{T}\Big)\right). \tag{14}$$

The intervention acts only on the teacher logits. CIFAR uses standard cross-entropy, whereas the DeiT runs use label smoothing 0.1. The controlled bridge uses the same sharpness intervention and $T = 4$ under its separately frozen continuation protocol.

A.2 Controlled ImageNet-100 Bridge

The controlled bridge uses ImageNet-100, a pretrained DeiT-S teacher, and a DeiT-T student. Three independently trained CE student trajectories define five predeclared starting states,

$$S_0, S_1, S_2, S_3, S_4 \quad \leftrightarrow \quad 0, 30, 60, 90, 120 \text{ CE epochs}. \tag{15}$$

The state S_0 is the seeded random initialization, and S_4 is the trained endpoint.

Each starting state uses a 20-epoch local-reset continuation. The continuation carries over only the student weights. AdamW is reinitialized with a peak learning rate of 5×10^{-4}, followed by a two-epoch warmup and cosine decay over the 20-epoch window. The optimizer and scheduler are therefore reset rather than resumed from the CE trajectory.

Table 5: Primary directional experimental protocols.

Setting	Teacher → Student	Evaluated states	Continuation	Replication
Bridge	DeiT-S → DeiT-T	S_0–S_4	20 ep, local reset	3 student traj.
Controlled CIFAR	ResNet-34 → ResNet-18	7×4 map	30 ep, global	3 student seeds/cell
Scratch DeiT	DeiT-S → DeiT-T	$e40, e120$; $s60$	20 ep, resumed	2 teacher × 3 student
Pretrained DeiT	DeiT-S → DeiT-T	$e1, e30$; $s60$	20 ep, resumed	2 teacher × 3 student

Within each $(\text{trajectory}, \text{state})$ pair, the compared sharpness conditions share the same starting weights, data order, augmentation randomness, optimizer initialization, schedule, and training budget. The primary $\alpha = 0.5$ and $\alpha = 4.0$ conditions are evaluated at all five states. The additional $\alpha = 1$ and $\alpha = 2$ conditions are evaluated only at S_0, S_2, and S_4. The state set, three trajectories, endpoint contrast, and additional sharpness conditions were fixed before the main confirmation analysis.

A.3 Broader Supervision Regimes

Table 5 summarizes the primary directional protocols. The controlled bridge resets the local optimizer and schedule, whereas the broader CIFAR and DeiT experiments continue from frozen states under their existing schedules.

Controlled CIFAR evaluates teacher states $e20, e40, e60, e80, e120, e160, e200$ against student states $s0, s20, s40, s60$. Each state continues for 30 epochs on the unrestarted global schedule. Each cell uses three student seeds. The $e80/s40$ setting is also evaluated with two independently retrained teacher trajectories in addition to the canonical trajectory.

The scratch and pretrained DeiT regimes use ImageNet-100 with a DeiT-S teacher and a DeiT-T student fixed at $s60$. Scratch DeiT evaluates teacher states $e40$ and $e120$ using two teacher seeds and three student seeds. Pretrained DeiT evaluates $e1$ and $e30$ with two teacher seeds and three student seeds. Its two teacher seeds are fine-tuning replicates of one pretrained checkpoint rather than independent pretraining runs.

A.4 Statistical Procedures

For the controlled bridge, the planned endpoint statistic is

$$R = D(S_0) - D(S_4). \tag{16}$$

The hierarchical paired bootstrap uses 10,000 resamples. It first resamples the three student trajectories and then resamples paired validation examples within each sampled trajectory. The Student-t interval treats the three trajectory-level values of R as the observations.

Controlled CIFAR reports means over three student seeds. The vertical ranges in Figure 2 show the seed minimum and maximum rather than confidence intervals. For scratch DeiT, the reported confidence interval applies only to the planned $e40 - e120$ contrast. A 10,000-resample percentile bootstrap is applied to the six paired teacher-seed × student-seed differences.

For pretrained DeiT, each evaluated state contains six teacher–student cells from two teacher fine-tuning replicates and three student seeds. The reported state-level intervals use a 10,000-resample hierarchical percentile bootstrap. The procedure first resamples teacher replicates and then resamples student seeds within each selected teacher. The resulting intervals therefore describe the observed fine-tuning and student-seed hierarchy rather than variation across independently pretrained models.

B Extended Bridge Evidence

This appendix exposes the complete evidence underlying the controlled bridge in Section 4. We report all state-level directional contrasts, the full evaluated sharpness responses, and the symmetric contrast. These analyses support the endpoint reversal without implying a gradual or monotonic training law.

Table 6: Complete directional results for the controlled bridge. Values are in percentage points.

State	Traj. 0	Traj. 1	Traj. 2	Mean D	Bootstrap 95% CI	Student-t 95% CI
S_0	$+2.84$	$+2.48$	$+1.72$	$+2.35$	$[+1.54, +3.12]$	$[+0.93, +3.77]$
S_1	-0.44	-0.02	-0.20	-0.22	$[-0.64, +0.21]$	$[-0.74, +0.30]$
S_2	-0.18	-0.62	-0.64	-0.48	$[-0.91, -0.02]$	$[-1.13, +0.17]$
S_3	-0.60	$+0.32$	-0.66	-0.31	$[-0.91, +0.35]$	$[-1.68, +1.05]$
S_4	-0.52	-0.86	-0.72	-0.70	$[-1.14, -0.25]$	$[-1.12, -0.28]$

Table 7: Matched KD gain $G(\alpha)$ over the full evaluated sharpness response. Values are means in percentage points with 95% bootstrap intervals.

State	$\alpha = 0.5$	$\alpha = 1$	$\alpha = 2$	$\alpha = 4$
S_0	$+1.23\,[+0.60, +1.85]$	$+1.35\,[+0.64, +2.07]$	$+0.23\,[-0.39, +0.89]$	$-1.12\,[-1.89, -0.29]$
S_2	$+0.23\,[-0.16, +0.63]$	$+0.32\,[-0.07, +0.71]$	$+0.77\,[+0.23, +1.37]$	$+0.71\,[+0.19, +1.21]$
S_4	$-0.10\,[-0.51, +0.29]$	$+0.32\,[-0.15, +0.78]$	$+0.75\,[+0.17, +1.31]$	$+0.60\,[+0.21, +1.01]$

B.1 Complete State-Level Directional Results

Table 6 reports $D = G(0.5) - G(4.0)$ for all five predeclared student states and three independent CE trajectories. All three trajectories have $D > 0$ at S_0 and $D < 0$ at S_4. At S_1 and S_3, both interval constructions span zero; at S_2, the bootstrap interval excludes zero but the Student-t interval does not. The intermediate states therefore delimit the endpoint result rather than establish a gradual sign transition.

The trajectory-level endpoint contrasts are 3.36, 3.34, and 2.44 pp, giving $R = D(S_0) - D(S_4) = +3.05$ pp. The 95% paired-bootstrap interval is $[2.17, 3.91]$ pp and the trajectory-level Student-t interval is $[1.74, 4.35]$ pp.

B.2 Full Sharpness Response

The primary statistic compares the fixed endpoint interventions $\alpha = 0.5$ and $\alpha = 4.0$. We additionally evaluate $\alpha = 1$ and $\alpha = 2$ at S_0, S_2, and S_4 to characterize the response between these endpoints (Table 7).

At S_0, $G(0.5)$ is close to $G(1)$ and the response falls toward $\alpha = 4$; at S_2 and S_4, the largest observed mean gain occurs at $\alpha = 2$. Thus, D should be read as a fixed directional contrast, not as evidence that either endpoint is universally optimal.

B.3 Symmetric-Contrast Robustness

We additionally evaluate the log-symmetric contrast

$$D_{0.5,2} = G(0.5) - G(2). \tag{17}$$

Table 8 reports the three states for which the full sharpness response is available.

The sign agrees with the primary contrast in all three trajectories at each evaluated state. The bootstrap intervals preserve these directions, whereas the corresponding Student-t intervals span zero. We therefore treat this analysis as qualitative robustness evidence rather than a second inferential result.

C Additional Supporting Analyses

This appendix reports the complete controlled-CIFAR state map and provides additional details for the single-factor diagnostics and functional-geometry analysis.

C.1 Controlled CIFAR State Map and Replication

Table 9 reports the complete controlled CIFAR state map. We evaluate all $7 \times 4 = 28$ teacher–student combinations under the same 30-epoch continuation protocol, with three student seeds per

Table 8: Symmetric sharpness contrast $D_{0.5,2}$. Values are in percentage points.

State	Traj. 0	Traj. 1	Traj. 2	Mean	Bootstrap 95% CI
S_0	$+1.18$	$+0.40$	$+1.40$	$+0.99$	$[+0.21, +1.71]$
S_2	-0.38	-0.90	-0.34	-0.54	$[-1.01, -0.10]$
S_4	-0.18	-1.16	-1.22	-0.85	$[-1.47, -0.15]$

Table 9: Complete directional contrast D for the controlled CIFAR state map. Values are means over three student seeds in percentage points.

	$e20$	$e40$	$e60$	$e80$	$e120$	$e160$	$e200$
$s0$	3.840	0.757	2.257	1.900	2.440	3.390	0.217
$s20$	2.127	2.270	2.253	3.530	1.080	2.387	0.170
$s40$	4.890	3.947	3.413	3.453	1.020	1.060	1.263
$s60$	4.743	3.410	2.747	2.540	1.403	0.133	0.620

cell and no interpolation. Mean D is positive in every cell and ranges from 0.133 pp at $s60/e160$ to 4.890 pp at $s40/e20$.

At $e80/s40$, the canonical teacher trajectory gives $D = 3.453$ pp, while two independently retrained teacher trajectories give $D = 3.537$ and 3.607 pp. The dense map shows that the positive direction extends across the measured state space, while the trajectory comparison tests replication at the canonical operating point.

C.2 Single-Factor Diagnostic Details

The diagnostics use two distinct response quantities. Gradient matching reports the directional contrast D. The correctness, difficulty, and target-component probes instead report recovery relative to the unmodified $\alpha = 1$ KD target because these interventions were not evaluated at $\alpha = 4$. Table 10 summarizes the corresponding measurements.

Gradient magnitude. The strictly matched comparison uses the same three student seeds in all conditions. Fixed matching computes the KD-to-CE parameter-gradient norm scaling at initialization and holds it fixed. Online matching recomputes the scaling at every optimizer step. A dedicated validation run records zero relative norm-matching error at every logged step under online matching. Both controls change the response magnitude while preserving positive D.

Correctness and difficulty. We define teacher correctness by whether the teacher top-1 prediction matches the ground-truth label. The teacher-wrong and teacher-correct subsets contain $5{,}744$ and $44{,}256$ examples, respectively. Their raw recovery difference is 0.706 pp, but the subsets differ in both size and difficulty, so this comparison does not isolate correctness.

We define the pre-KD difficulty score as

$$d_i = z(\text{CE loss}_i) + z(\text{error probability}_i) + z(-\text{teacher margin}_i). \tag{18}$$

The score uses CE-only reference trajectories and the teacher margin; it does not use student KD outputs. We construct a teacher-correct subset by greedy nearest-difficulty matching without replacement to the teacher-wrong subset, and we use a random size-matched teacher-correct subset as a control. The teacher-wrong, hard-correct, and random matched subsets recover 0.327, 0.190, and 0.077 pp, respectively. The wrong-versus-hard-correct gap is 0.137 pp, below the predeclared 0.40 pp threshold. However, the matched subsets have class-balance total variation 0.176, which exceeds the predefined 0.15 gate. This failure prevents an equivalence claim. The local matching result also does not account for the 1.397 pp recovery obtained by softening all examples.

Target components. The top-1-only probe preserves the softened probability assigned to the teacher's top class and redistributes the remaining probability mass uniformly across the non-target classes. The non-target-only probe keeps the top-class probability at its unmodified value and changes only the relative non-target distribution. The full softened target recovers 1.23 pp, while the top-1-only and non-target-only probes recover 0.77 and 0.20 pp. The two restricted probes sum

Table 10: Summary of the single-factor diagnostics. Recovery denotes accuracy change relative to the $\alpha = 1$ KD target and is distinct from D.

Diagnostic	Quantity	Measured response
Gradient magnitude	D	$3.887 \to 2.12$ (fixed), 3.32 (online)
Correctness	Recovery	wrong $+0.327$, correct $+1.033$
Difficulty matching	Recovery	hard-correct $+0.190$, random $+0.077$
Target components	Recovery	full $+1.23$, top-1 $+0.77$, non-target $+0.20$

to 0.97 pp, leaving a 0.26 pp residual relative to the full response. These single-seed probes are non-additive and provide diagnostic rather than decompositional evidence.

C.3 Functional-Geometry Details

We define the functional-geometry readout as

$$C = \frac{\sum_i g_{\mathrm{CE}}^{(i)} \cdot g_{\mathrm{KD}}^{(i)}}{\sqrt{\sum_i \|g_{\mathrm{CE}}^{(i)}\|^2}\sqrt{\sum_i \|g_{\mathrm{KD}}^{(i)}\|^2}}. \tag{19}$$

We compute C on a fixed diagnostic set of $2{,}000$ examples using gradients with respect to the student logits. The statistic is a pooled cosine over the diagnostic set rather than an average of per-example cosine similarities.

In controlled CIFAR with the student fixed at $s40$, C changes from 0.232 at $e20$ to 0.321 at $e80$ and 0.764 at $e200$. Three independently trained teacher trajectories at $e80$ give $C = 0.338$, 0.307, and 0.317. Leave-one-teacher-trajectory-out evaluation gives

$$\mathrm{RMSE}_C = 0.685 < 0.776 = \mathrm{RMSE}_{\mathrm{epoch}}, \qquad \rho_C = 0.869 > 0.814 = \rho_{\mathrm{epoch}}. \tag{20}$$

With the scratch DeiT student fixed at $s60$, C evolves across teacher states $e20, e40, e60, e90, e120$ as

$$-0.086 \to 0.211 \to 0.496 \to 0.577 \to 0.572. \tag{21}$$

The sequence shows a substantial shift followed by a late plateau. We interpret C as a within-regime descriptor of supervision state. We do not treat its numerical scale as calibrated across datasets, architectures, or output dimensionalities, and the analysis does not establish C as a causal or regime-independent predictor of sharpness direction.