BACKWARD TRUNCATION CHANGES THE TARGET GEOMETRY OF LOCAL CREDIT ASSIGNMENT

Anonymous authors
Paper under double-blind review

ABSTRACT

Adaptive states can keep updating in the forward pass while their derivatives are truncated. We show that the local supervision matched to the full task gradient depends on these retained derivatives: the adjustment can be directional rather than a global loss rescaling. We derive a direction-wise matching rule and test its consequences with controlled interventions. Changing only retained state derivatives reverses the average preference between fixed supervision rules. Directional controls separate the correction from scalar weakening; Quotient reduces mean test error by 9.4% relative to Point. With known dynamics, we also construct supervision before training, obtaining mean gains across two tracker variants and two optimizers. In a coherent optical receiver using recordings from a 1125-km link, gradient-matching preferences vary with horizon and trainable space, while training effects depend on configuration and retained derivatives. With dispersion, nonlinear, and residual compensation trainable under truncation, fixed Complex Quotient supervision improves receiver quality factor (Q-factor) over Point at all 20 tested power–channel workpoints (five powers, four channels), with gains in all 40 paired training-capture comparisons and a mean of 0.045 dB. These findings establish backward differentiation policy as an explicit design variable for local supervision in adaptive systems.

1 INTRODUCTION

Changing only retained state derivatives reverses the average training preference between two fixed supervision rules in our controlled adaptive system. We demonstrate this reversal while holding the task, inputs, initialization, and forward recursion fixed. Local supervision can therefore depend on the backward paths used for training, even with a fixed forward task.

Consider a learned predictor followed by an adaptive tracker that estimates a nuisance offset and corrects the prediction. At fixed parameters, inputs, and initial state, truncating differentiation through the tracker preserves its state trajectory and corrected outputs, but changes the gradient reaching the predictor. This separation arises in differentiable filters and controllers (Kloss et al., 2021; Amos et al., 2018), where computational limits often motivate truncating backward paths through recurrent or iterative computations (Aicher et al., 2020; Shaban et al., 2019). The forward computation alone therefore does not determine which local supervision supplies the appropriate upstream correction.

The reason is geometric. Downstream adaptation may absorb one component of a prediction error while leaving another largely unaffected. Full differentiation accounts for this selective compensation. Truncating the state derivatives can instead ask the predictor to correct too much of the absorbed component. Scaling down the whole loss also weakens supervision on the unaffected component. Matching the full task gradient can therefore require changing the relative penalties across error directions (Figure 1). We call this directional penalty structure **target geometry**; the reference label can stay fixed while the geometry changes.

The matched target geometry depends jointly on **task relevance, forward residual, and exposed derivative**. The task specifies which residual errors matter; adaptation determines which errors survive; and the *backward closure* specifies which state derivatives return upstream. Under the shared-basis assumptions, the matching weight is task relevance times forward residual divided by exposed derivative, direction by direction.

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

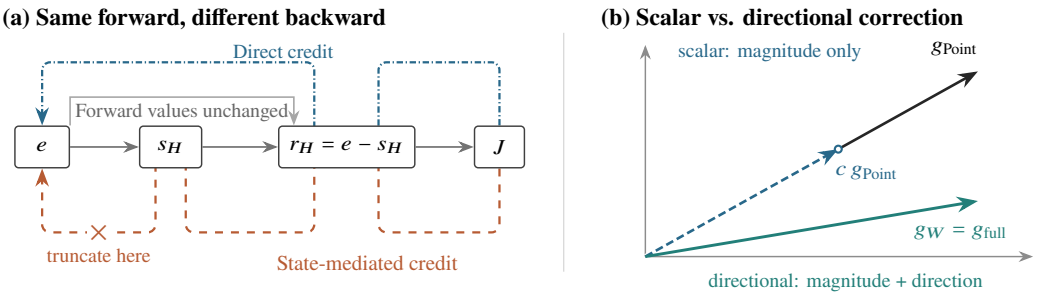

Figure 1: **Changing the backward graph can change matched local supervision without changing the forward computation.** (a) Forward values and state updates are identical at fixed parameters and entry state. Direct credit remains; truncation blocks the state-mediated derivative through s_H at the marked segment. (b) In this two-direction example, scalar scaling preserves the Point-gradient direction, whereas directional weighting matches the full task gradient.

Much prior work approximates or localizes credit assignment (e.g., Aicher et al., 2020; Jaderberg et al., 2017); we ask how local supervision should change to match the downstream task under a specified backward closure.

We make three contributions.

First, we characterize how backward closure changes matched target geometry. A scalar tracker and shared-basis extension give exact matching rules. Beyond that basis, we identify a symmetry obstruction to exact matching and quantify what general PSD approximation retains in coupled 2D and 4D systems.

Second, we isolate the mechanism and test prospective supervision construction. Changing only retained derivatives reverses the average preference between Point and the same fixed Partial. Directional controls distinguish the correction from scalar weakening; Quotient lowers mean test error by **9.4%** relative to Point. With known dynamics, constructing supervision before training also improves over a development-selected lookup across two tracker variants and two optimizers. A two-weight family captures most of the general PSD gain.

Third, we demonstrate consistent gains across physical receiver operating conditions. On recordings from a 1125-km link (Fan et al., 2021a;b), the same fixed Complex Quotient rule improves receiver quality factor (Q-factor) over Point at **all 20 tested workpoints spanning five launch powers and four channels**, with gains in **40/40 paired training-capture comparisons** and a mean of **+0.045 dB** when dispersion, nonlinear, and residual compensation are trainable under Detached closure. Gradient-matching preferences vary with horizon and trainable space; training effects depend on configuration and retained derivatives.

Together, these results establish a mechanism and a design principle: controlled systems identify when retained derivatives change matched supervision geometry, while the physical receiver demonstrates consequences for gradient matching and training under the tested conditions.

2 The Forward Task and the Backward Training Signal

A stateful system's forward computation and its training derivatives are distinct objects: truncating derivative paths can preserve the numerical forward values while changing the correction propagated upstream.

2.1 Keeping Adaptation While Truncating Its Derivatives

An upstream model produces u_t, a downstream module updates its state s_t, and the task observes o_t:

$$u_t = G_\theta(y_t), \qquad s_{t+1} = F(s_t, u_t), \qquad o_t = O(s_t, u_t). \tag{1}$$

For the task objective

$$J_{\text{task}}(\theta) = \sum_{t=1}^{T} \ell_t(o_t),$$

the gradient with respect to the interface variable u_t can be decomposed into a direct contribution and a contribution mediated by later adaptive states. Throughout, "full" denotes differentiation through the stated adaptive recursion; the differentiated variable or parameter space is indicated by context or subscript.

$$g_{u_t}^{\text{full}} = g_{u_t}^{\text{direct}} + \underbrace{\sum_{\tau>t} \left(\frac{\partial s_\tau}{\partial u_t}\right)^\top \nabla_{s_\tau} \ell_\tau}_{g_{u_t}^{\text{state}}}. \tag{2}$$

The direct contribution is defined with s_t held fixed:

$$g_{u_t}^{\text{direct}} := \left(\frac{\partial o_t}{\partial u_t}\right)^\top \nabla_{o_t} \ell_t.$$

Here, $\nabla_{s_\tau} \ell_\tau$ denotes the gradient of the downstream loss with respect to the state s_τ, including its effect through the output at time τ. The derivatives in the sum hold the other interface inputs fixed. The direct term measures how u_t affects the current loss through o_t; the state-mediated term measures how u_t changes later losses through the adaptive state dynamics.

We use *backward closure* for the choice of state-update derivative paths retained in this calculation. A *Detached* closure omits the state-mediated term while still executing the state updates in Eq. (1). A *Full* closure retains those state paths over the specified differentiation horizon. Changing the closure therefore changes the derivative calculation while leaving the forward adaptation active.

Three choices should be kept distinct. The *closure* specifies which state paths are differentiated; the *differentiation horizon* specifies how far the comparison extends; and the *trainable configuration* specifies which parameter blocks receive the resulting gradient. The controlled closure intervention varies the first axis, while the receiver studies examine all three.

2.2 Local Supervision Separates the Reference from Its Geometry

A *reference* specifies the desired output; *local supervision* specifies the comparison rule. Its *target geometry* is the directional penalty structure, and a *loss* is the scalar objective implementing that rule. The reference can remain fixed while the directional penalties change. We call the geometry that matches the full task gradient for a specified task and closure the *matched target geometry*.

For a supervised error r, a quadratic example is

$$L_W(r) = \frac{1}{2} r^\top W r, \qquad W \succeq 0.$$

Point supervision uses $W = I$; Partial supervision reduces selected directional penalties; and Quotient-like supervision removes selected directions. These supervision rules change how deviations from the same reference are penalized without changing the task itself.

The two-direction example in Figure 1 makes the distinction explicit. Suppose the truncated Point gradient and the full task gradient are

$$g_{\text{Point}}(e) = \begin{pmatrix} e_1 \\ e_2 \end{pmatrix}, \qquad g_{\text{full}}(e) = \begin{pmatrix} e_1 \\ \rho e_2 \end{pmatrix}, \qquad 0 < \rho < 1. \tag{3}$$

For $e_1 e_2 \neq 0$, matching the first component by scalar rescaling requires scale one, while matching the second requires scale ρ. A single scalar cannot satisfy both conditions. By contrast,

$$W = \text{diag}(1, \rho)$$

changes the relative correction across directions and exactly matches the full task gradient in this example.

Section 3 derives such directional weights from the interaction among the task, the forward adaptive dynamics, and the derivatives retained during training.

3 MATCHING TARGET GEOMETRY UNDER A SPECIFIED BACKWARD CLOSURE

Section 2 showed why changing the retained derivatives can change the supervision needed upstream. We now derive the matched target geometry under a specified backward closure. We first isolate the rule in a scalar tracker, extend it in a shared spectral basis, and examine its boundary when task and adaptation directions couple.

3.1 TRUNCATION CHANGES THE MATCHED LOCAL WEIGHT

Let a scalar parameter produce errors $e_c(\theta) = \theta - 1$ and $e_r(\theta) = \theta + 1$. Adaptation leaves e_c unchanged and tracks e_r through $s_{h+1} = (1-\kappa)s_h + \kappa e_r$, with $s_0 = 0$. After H steps, $e_r - s_H = R_H e_r$, where $R_H = (1-\kappa)^H$.

The task assigns relevance $\alpha \geq 0$ to this residual, giving the objective and its full gradient with respect to θ:

$$J_{\text{task}} = \frac{1}{2}e_c^2 + \frac{\alpha}{2}(R_H e_r)^2, \qquad g_\theta^{\text{full}} = e_c + \alpha R_H^2 e_r. \tag{4}$$

Now consider a local surrogate that uses the same numerical residual but replaces the residual weight α by λ, while differentiating the adaptive state according to backward closure Γ. Define

$$d_H(\Gamma) := \frac{\partial^\Gamma s_H}{\partial e_r}.$$

Since $\partial e_c/\partial\theta = \partial e_r/\partial\theta = 1$, the resulting parameter gradient is

$$g_{\theta;\lambda,\Gamma} = e_c + \lambda R_H\big[1 - d_H(\Gamma)\big]e_r. \tag{5}$$

The two factors multiplying e_r play different roles. The factor R_H determines the surviving forward residual and remains unchanged when derivative paths are truncated. By contrast, $1 - d_H(\Gamma)$ is the residual derivative exposed by the backward closure. The same forward residual can therefore expose a different gradient under a different backward closure.

Matching the coefficient of e_r in Eqs. (4)–(5) gives

$$\lambda^\star(H,\Gamma) = \frac{\alpha R_H}{1 - d_H(\Gamma)}, \qquad R_H > 0, \quad 1 - d_H(\Gamma) > 0. \tag{6}$$

Under Full differentiation, $1 - d_H(\text{Full}) = R_H$, so $\lambda^\star = \alpha$. Under Detached differentiation, $d_H(\text{Detached}) = 0$, giving $\lambda^\star = \alpha R_H$. The task relevance α has not changed. What changes is how much of the adaptive compensation is already represented in the backward signal. Under Full differentiation, the chain rule supplies that contribution. Under Detached differentiation, the local supervision must supply the attenuation missing from the derivative.

For unit task relevance and a finite contractive horizon, $0 < R_H < 1$, the Detached matched weight lies strictly between full Point penalty and complete removal. Thus a direction that is partly absorbed by adaptation need not be task-null. Appendix A.1 gives the closed-form derivation and analytic illustrations.

3.2 THE MATCHING RULE EXTENDS DIRECTION BY DIRECTION

The scalar calculation separates three quantities: task relevance, the surviving forward residual, and the derivative exposed by the backward closure. Under a shared-basis linearization, the same separation applies direction by direction.

At a fixed linearization, let $R_H e$ denote the residual seen by the task, let $A \succeq 0$ denote its task metric, and let $C_{H,\Gamma}$ denote the residual derivative exposed by backward closure Γ. The full task gradient and the gradient induced by a local metric $W \succeq 0$ are

$$g_e^{\text{full}} = R_H^\top A R_H e, \qquad g_e^W = C_{H,\Gamma}^\top W R_H e. \tag{7}$$

These operators generalize the scalar factors α, R_H, and $1 - d_H(\Gamma)$: A specifies task relevance, R_H determines the surviving forward residual, and $C_{H,\Gamma}$ is the exposed derivative.

Proposition 1 (Directional matching in a shared basis)**.** *Suppose A, R_H, and $C_{H,\Gamma}$ are symmetric positive-semidefinite operators sharing an orthogonal eigenbasis U, with eigenvalues a_i, r_i, and c_i, respectively. Assume $c_i > 0$ whenever $a_i r_i > 0$.*

For every direction with $r_i > 0$ and $c_i > 0$, an exact matching metric has weights

$$w_i^\star = \frac{a_i r_i}{c_i}, \qquad W^\star = U\ \mathrm{diag}(w_i^\star)\, U^\top. \tag{8}$$

Weights on all remaining inactive directions may be set to zero. A scalar metric $W = wI$ matches the full interface gradient for every e only if all active required weights $w_i^\star$ coincide.

Equation (8) is the direction-wise counterpart of Eq. (6): a_i, r_i, and c_i play the roles of task relevance, forward residual factor, and exposed derivative, respectively. Changing closure alters the denominator without changing the forward residual. For two closures with positive weights, $w_i^{(1)}/w_i^{(2)} = c_i^{(2)}/c_i^{(1)}$. If this ratio varies across directions, a geometry matched to one closure cannot be transferred to the other by a global scalar rescaling.

Under Full differentiation, $C_{H,\Gamma} = R_H$, so the matching metric reduces to $W^\star = A$ on the active subspace. The identity metric is the special case of a unit task metric.

Beyond a shared basis. Under Detached closure with symmetric $R_H \succ 0$, exact PSD matching is possible if and only if $AR_H = R_H A$ (Lemma 1). For isotropic errors and zero ridge, a closed-form PSD optimum exists (Lemma 2): at 45° task rotation in a fixed 2D tracker, relative operator errors are 0.124 (PSD), 0.358 (best diagonal), and 0.413 (best scalar). Appendix A.5 reports the full angle sweep, 100 constructed 4D rotations, and actual local task decreases, including the nonzero mismatch. Section 4.3 then uses this objective to construct supervision before training nonlinear predictors.

Exact interface matching transfers through the same upstream Jacobian; approximate matching depends on the active parameter space. The receiver therefore measures candidate gradients in the parameters being optimized. Appendices A.2–A.4 give proofs and extensions.

4 Controlled Tests of Closure-Dependent Supervision

The theory makes two experimentally separable predictions. **P1: closure changes the training effect of fixed supervision.** Changing only retained state derivatives can change which rule is preferred under the same forward recursion. **P2: the correction is directional.** When matching weights differ across directions, global scalar weakening cannot reproduce the required correction. We test each prediction with a controlled intervention.

The P1/P2 interventions are semi-synthetic: measured MIT Push trajectories (Yu et al., 2016) supply the physical inputs, while translations, cues, and tracker recursions provide known experimental structure. The five closure worlds are resampled constructions from one trajectory pool. Appendix B gives the construction and control matrix.

4.1 P1: Retained Derivatives Reverse the Supervision Preference

An affine predictor feeds a two-direction tracker. Across backward closures, we hold fixed the task, data order, parameter initialization, entry state, forward recursion, and supervision pair; only retained derivatives change. Point and the same fixed Partial rule are compared under Detached and Full. The direction-wise matching rule gives unequal weights under Detached and unit weights under Full; we use the same Detached-matched Partial under both closures. Training uses a gradient-equivalent reduced quadratic; explicit recursion checks verify its derivative interpretation (Appendix B.2).

Figure 2 shows the mean contrast $\Delta_\Gamma = J_{\mathrm{Point},\Gamma} - J_{\text{fixed Partial},\Gamma}$ reversing from $+0.00217$ under Detached to -0.00126 under Full. The mean interaction is $+0.00343$, positive in all five worlds; four worlds follow each endpoint's mean sign. On average, the same fixed Partial improves learning when state-mediated derivatives are omitted but worsens it when they are retained. Under Full, those derivatives already supply the compensating contribution, so Detached-matched attenuation changes the signal again. Appendix B.5 gives the additional controls and complete candidate comparison.

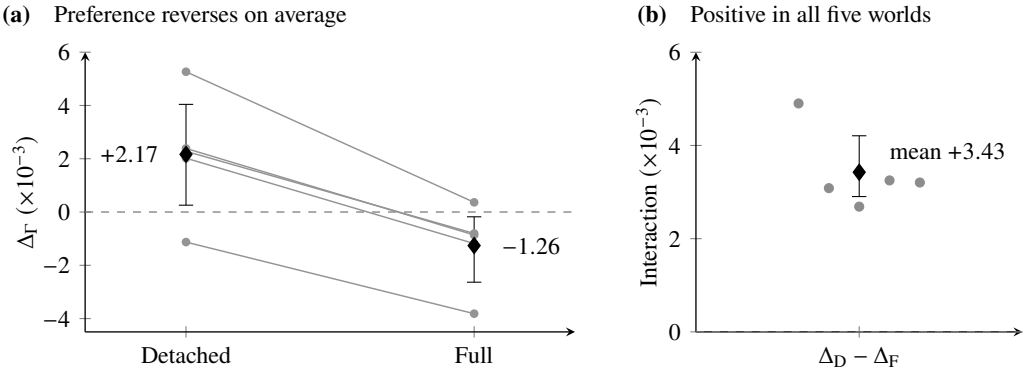

Figure 2: **P1:** Changing only retained derivatives reverses the average preference between two fixed supervision rules. The task, inputs, initial conditions, forward recursion, and supervision pair are shared across closures. **(a)** $\Delta_\Gamma = J_{\text{Point}} - J_{\text{fixed Partial}}$ is positive under Detached and negative under Full on average. Lines connect the five paired worlds. **(b)** The interactions $\Delta_D - \Delta_F$ are positive in all five worlds. Black diamonds and bars show means and the archived 95% paired-bootstrap intervals.

4.2 P2: Directional Correction Is Not Scalar Weakening

The second intervention tests the directional prediction of Proposition 1 with scalar and subspace controls. A translation of the reference-tip sensor input drives a pretrained frontend; a fixed tracker estimates the reference bias to correct its object output. The relative-displacement loss is invariant to a common translation of the two predictions, identifying a direction to relax; a same-rank rotated subspace supplies the directional control.

The five original branches share Detached differentiation and the task term:

$$L_{\text{train}} = J_{\text{task}} + 0.5 L_{\text{local}}. \tag{9}$$

The local term is Point, Scalar Point, Wrong Partial, Partial, or Quotient. Let $g_{\text{miss}} = g_{\text{task,Full}} - g_{\text{task,Detached}}$ denote the missing task parameter gradient. Exact matching requires $0.5\nabla_\theta L_{\text{local}} = g_{\text{miss}}$; cosine tests direction. Scalar Point uses frozen calibration (Appendices B.3–B.4).

Translation-quotient gradients align more strongly with missing credit; Partial outperforms Wrong Partial in all five training seeds (Figure 3). A post-primary task-only control reuses the protocol without local supervision. Quotient lowers mean test error by 9.4% versus Point and 8.33% versus task-only, improving all five paired seeds. Point and calibrated Scalar Point are slightly worse than task-only on average; arbitrary auxiliary supervision need not help (Appendix B.4).

4.3 Known Dynamics Enable Supervision Construction Before Training

We next construct a PSD metric from known task and tracker operators by minimizing Eq. (16) with $\Sigma = I$ and $\beta = 0$. Using known operators alone, we freeze the metric before generating training and test samples, without Full parameter-gradient probes. We train nonlinear predictors on 32 new synthetic tasks, split equally between fixed and rotated tracker variants, with three horizons, two task orientations, and both normalized SGD and Adam.

Under Detached closure, construction reduces mean normalized test loss by 0.37–0.55% relative to a development-selected candidate lookup indexed by tracker variant, optimizer, closure, horizon, and task orientation. The four groups improve in 15/16, 15/16, 15/16, and 14/16 task means (Appendix B.6). A structured two-weight family captures most of the general PSD improvement; PSD adds 0.03–0.14% over that construction. In a post-primary same-objective control, the two-weight construction lowers mean loss by 1.05–1.70% relative to $W = \alpha A$, with lower task-mean loss in 15/16–16/16 evaluation tasks across the four groups (Appendix B.6). These results concern known dynamics, not identification of an unknown receiver metric.

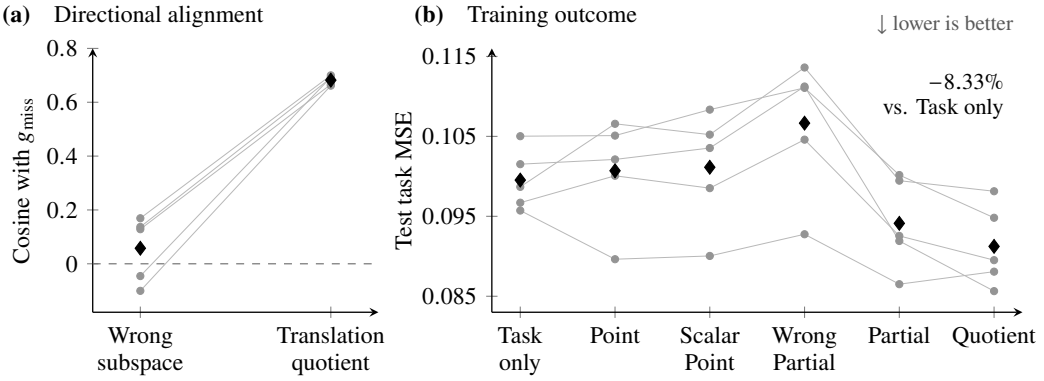

Figure 3: **P2:** Direction, not just loss strength, matters. **(a)** Pure local-gradient cosine with $g_{\text{miss}} = g_{\text{task,Full}} - g_{\text{task,Detached}}$, averaged over eight batches per seed. Translation-quotient gradients align more strongly than the wrong-subspace control. **(b)** The five original branches use Eq. (9); the post-primary Task only branch omits local supervision. Quotient reduces mean test error by 9.4% versus Point and 8.33% versus Task only, beating Task only in all five seeds. Gray traces show paired seeds; black diamonds show means.

5 Consistent Gains and Configuration Dependence in a Physical Adaptive Receiver

We find consistent gains across tested operating conditions when dispersion, nonlinear, and residual compensation are jointly trainable under Detached closure. We characterize how supervision effects vary with trainable configuration and retained derivatives, separating gradient diagnostics from training outcomes where the matched geometry is unknown.

5.1 Receiver and Experimental Axes

We use LabPtPTm2 recordings (Fan et al., 2021b) from a 1125-km coherent optical link (Fan et al., 2021a). The adaptive-filter closure *AF-Full* retains derivatives through the early adaptive MIMO, frequency-offset estimation, and final adaptive MIMO state updates within the specified horizon, with the entry state held fixed. *Detached* omits these state-update derivatives while preserving forward adaptation. Three nested trainable configurations update residual compensation (R; FDBP), nonlinear plus residual (N+R; EDBP), or dispersion plus nonlinear plus residual (D+N+R; GDBP). Point compares recovered symbols with transmitted references; Complex Quotient (CQ) fits a shared complex scalar before comparison; Partial mixes $0.6L_{\text{Point}} + 0.4L_{\text{CQ}}$. Receiver quality factor (Q, in dB) is derived from bit error rate (BER) (Fan et al., 2021a). Models are trained separately at each power–channel workpoint; Appendix C specifies coverage and aggregation.

5.2 Candidate Gradient Matching Depends on Horizon and Trainable Space

At fixed Point checkpoints and a shared entry state, we compare Detached candidate gradients with a same-horizon AF-Full Point reference—the gradient of a differentiable symbol-error surrogate, not of BER or Q. Let $E_k(H)$ be relative parameter-gradient error after the best nonnegative global scale fit. The matching gap

$$D_H = E_{\text{CQ}}(H) - E_{\text{Partial}}(H) \tag{10}$$

is positive when Partial is closer and negative when CQ is closer. Appendix C.5 specifies the checkpoint/window protocol.

Across all 27 model–workpoint means, D_H decreases from 500 to both 2,000 and 4,000 symbols (Figure 4(a–c)). N+R and D+N+R cross from Partial to CQ at all nine workpoints; R-only crosses at four, while the other five already favor CQ at 500 symbols. Substantial mismatch remains (Appendix C.6).

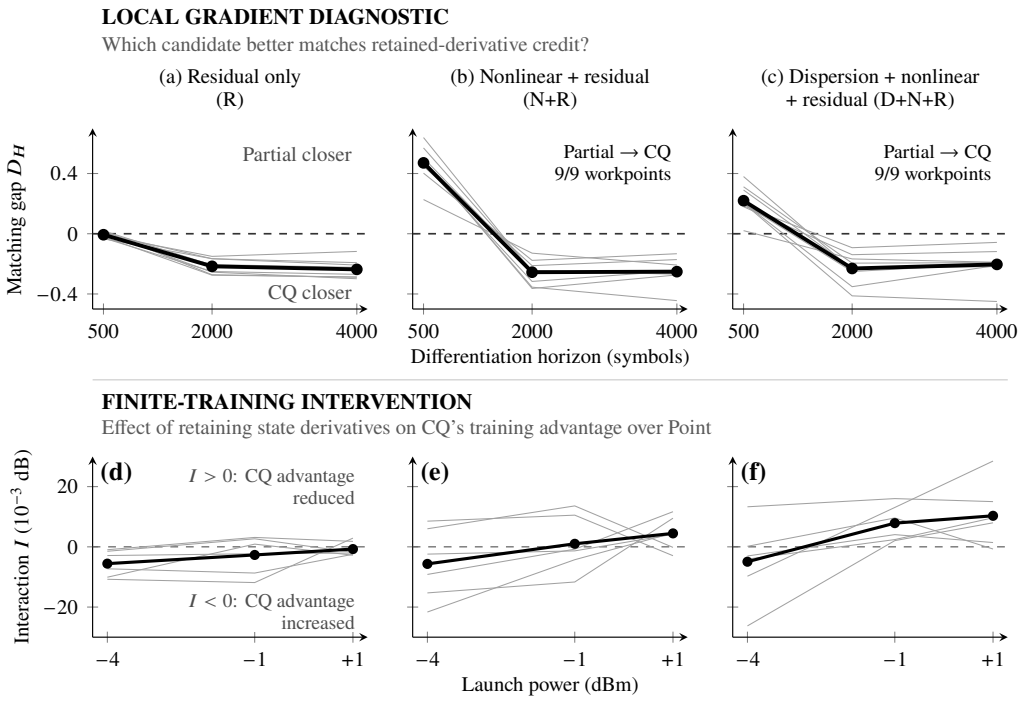

Figure 4: **Local gradient matching and finite-training effects reveal distinct configuration dependencies.** Top: fixed-checkpoint diagnostics with horizon-specific loss windows; bottom: a separate equal-cadence training intervention. (a–c) Longer horizons shift D_H toward CQ in all 27 configuration–workpoint means; N+R and D+N+R cross from Partial to CQ at 9/9 workpoints each. (d–f) The interaction I measures how retaining derivatives changes CQ's training advantage over Point; its sign varies with configuration and power. Thin gray lines retain all nine workpoint means (top) or six channel/capture combinations (bottom). Black lines with markers show configuration means.

5.3 Consistent Gains Across Powers and Channels Depend on Training Configuration

With D+N+R trainable under Detached closure, fixed CQ improves Q over Point at **all 20 workpoints (five powers, four channels)** and in **40/40 paired training-capture comparisons**, averaging $+0.04508$ dB (Appendix C.6, Figure 10). For scale, a hypothetical increase from 8.5 to 8.545 dB gives approximately **4% lower BER** (Appendix C.4). N+R and R yield $+0.00799$ and $+0.00094$ dB, with 36/40 and 26/40 positive pairs. The separate Condition-B selection study in Appendix D.1 uses a different operating point and protocol and is not included in these systematic coverage statistics.

Separately, equal-cadence training isolates retained state derivatives from update cadence. Point and CQ use Detached or AF-Full differentiation at the same 2,000-symbol cadence, with interaction

$$I = (Q_{\mathrm{CQ}} - Q_{\mathrm{Point}})_{\mathrm{Detached}} - (Q_{\mathrm{CQ}} - Q_{\mathrm{Point}})_{\mathrm{AF\text{-}Full}}. \tag{11}$$

Retaining derivatives reduces CQ's advantage when $I > 0$ and increases it when $I < 0$. Of 54 interactions, 26 are positive and 28 negative (Figure 4(d–f)). R-only has negative means at all three powers; D+N+R changes from negative at -4 dBm to positive at -1 and $+1$ dBm; N+R is mixed.

5.4 Local Task Descent and Finite Training

Matched-step controls in six receiver contexts distinguish precision and step-size effects (Appendix C.8). At relative displacements 10^{-6} and 10^{-5}, CQ gives greater task decrease than Point in 6/6 double-precision contexts, versus 5/6 and 3/6 in single precision; at 10^{-3}, both give 2/6.

Higher precision does not eliminate finite-step reversals. Matching diagnoses local descent (Appendix A.3), without arbitrary-step or endpoint guarantees. We do not attribute the broader 54/54 gradient-matching versus 45/54 training gap (Appendix C.6) to single precision.

6 Relation to Credit Assignment and Structured Supervision

Truncated and approximate credit assignment. Truncated differentiation and adaptive horizons (Aicher et al., 2020; Shaban et al., 2019; Bravo et al., 2024) trade gradient accuracy against computational cost; UORO and SnAp approximate recurrent gradients (Tallec & Ollivier, 2018; Menick et al., 2021). Burn-in methods study truncation in recurrent training (Schiller et al., 2026), while meta-gradient methods approximate long differentiation paths in nested optimization (Zhang et al., 2026).

Local credit assignment and gradient matching. Synthetic gradients predict downstream signals (Jaderberg et al., 2017; Czarnecki et al., 2017); target propagation and local learning construct intermediate objectives (Lee et al., 2015; Manchev & Spratling, 2020; Belilovsky et al., 2019; Nøkland & Eidnes, 2019; Dalm et al., 2023; As-Saquib et al., 2025). Objective parameterization (Ishikawa et al., 2025) and BP-derived corrections (Ren & Li, 2025) address related questions. We isolate the joint dependence of matched geometry on adaptation and retained state derivatives.

Task-aware and structured supervision. Task-aware learning incorporates downstream consequences (Donti et al., 2017; Wilder et al., 2019; Elmachtoub & Grigas, 2022), while invariant and quotient representations encode transformation structure (Cohen & Welling, 2016; Kondor & Trivedi, 2018; Bronstein et al., 2021; Mehr et al., 2018; Marchetti et al., 2023; Garrido et al., 2023). Such structure identifies directions to treat differently; our matching rule determines relaxation magnitude jointly from task relevance, forward adaptation, and retained derivatives.

7 Discussion

Differentiation and supervision are joint design choices. The task specifies relevance, adaptation the forward residual, and closure the exposed derivative. Retaining derivatives restores the chain rule; reshaping supervision corrects the surrogate under that closure. With known operators, a two-weight construction captures most of the general PSD improvement (Section 4.3).

Task relevance and transformation structure play different roles. Invariance identifies directions that may be relaxed, but their penalties depend on task relevance, forward residual, and exposed derivative. A partly absorbed direction can retain task relevance at a finite horizon; quotienting it out is therefore stronger than identifying it as absorbable.

Local matching and optimization trajectories are different objects. Local matching describes one parameter/state context; training integrates signals along a changing trajectory. Local descent need not imply a better endpoint (Section 5.4).

Evidence scope. Construction assumes known 2D tracker operators; receiver estimation remains open. Workpoints are correlated within one physical platform. Appendices D–E give scope studies, negative results, and reproducibility.

8 Conclusion

Local supervision is part of credit assignment in adaptive systems. Changing retained derivatives can change matched geometry and reverse fixed-rule training preferences while preserving forward computation. The adjustment can be directional, and known dynamics permit its construction before training. In the receiver, CQ improves Q at all 20 tested workpoints (40/40 pairs) with D+N+R trainable under Detached closure; effects vary with configuration and operating condition. Backward differentiation policy is therefore part of local-supervision design.

REPRODUCIBILITY STATEMENT

Appendices A–E provide complete proofs, data construction details, receiver configurations, training and evaluation protocols, full result strata, and a claim-to-artifact map. The supplementary material includes figure inputs, analysis scripts, and task ledgers distinguishing newly computed, reused, and data-blocked results. Appendix E specifies which reported results can be reproduced directly from the supplied artifacts and which require the archived experimental outputs.

AI USE STATEMENT

Generative AI tools were used to assist with writing and polishing the manuscript, literature retrieval and discovery, and developing and checking mathematical arguments and proofs. All AI-assisted content was reviewed and verified by the authors, who take responsibility for the final manuscript, references, and mathematical claims.

REFERENCES

Christopher Aicher, Nicholas J. Foti, and Emily B. Fox. Adaptively truncating backpropagation through time to control gradient bias. In *Proceedings of the 35th Conference on Uncertainty in Artificial Intelligence*, volume 115, pp. 799–808, 2020.

Brandon Amos, Ivan Jimenez, Jacob Sacks, Byron Boots, and J. Zico Kolter. Differentiable MPC for end-to-end planning and control. In *Advances in Neural Information Processing Systems*, volume 31, 2018.

Nazmus Saadat As-Saquib, A. N. M. Nafiz Abeer, Hung-Ta Chien, Byung-Jun Yoon, Suhas Kumar, and Su-in Yi. Forward target propagation: A forward-only approach to global error credit assignment via local losses. *arXiv preprint arXiv:2506.11030*, 2025. URL https://arxiv.org/abs/2506.11030.

Eugene Belilovsky, Michael Eickenberg, and Edouard Oyallon. Greedy layerwise learning can scale to ImageNet. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97, pp. 583–593, 2019.

João Bravo, Jacopo Bono, Pedro Saleiro, Hugo Ferreira, and Pedro Bizarro. Mind the truncation gap: Challenges of learning on dynamic graphs with recurrent architectures. *Transactions on Machine Learning Research*, 2024.

Michael M. Bronstein, Joan Bruna, Taco Cohen, and Petar Veličković. Geometric deep learning: Grids, groups, graphs, geodesics, and gauges. *arXiv preprint arXiv:2104.13478*, 2021. URL https://arxiv.org/abs/2104.13478.

Taco Cohen and Max Welling. Group equivariant convolutional networks. In *Proceedings of the 33rd International Conference on Machine Learning*, volume 48, pp. 2990–2999, 2016.

Wojciech Marian Czarnecki, Grzegorz Świrszcz, Max Jaderberg, Simon Osindero, Oriol Vinyals, and Koray Kavukcuoglu. Understanding synthetic gradients and decoupled neural interfaces. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70, pp. 904–912, 2017.

Sander Dalm, Nasir Ahmad, Luca Ambrogioni, and Marcel van Gerven. Gradient-adjusted incremental target propagation provides effective credit assignment in deep neural networks. *Transactions on Machine Learning Research*, 2023.

Priya Donti, Brandon Amos, and J. Zico Kolter. Task-based end-to-end model learning in stochastic optimization. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

Adam N. Elmachtoub and Paul Grigas. Smart "predict, then optimize". *Management Science*, 68 (1):9–26, 2022. doi: 10.1287/mnsc.2020.3922.

Qirui Fan, Chao Lu, and Alan Pak Tao Lau. Combined neural network and adaptive DSP training for long-haul optical communications. *Journal of Lightwave Technology*, 39(22):7083–7091, 2021a. doi: 10.1109/JLT.2021.3111437.

Qirui Fan, Chao Lu, and Alan Pak Tao Lau. Dataset: 1125km 7-channel DP-16QAM WDM transmission using quantum random bit source. Figshare dataset, June 2021b. URL https://doi.org/10.6084/m9.figshare.14843037.

Quentin Garrido, Laurent Najman, and Yann LeCun. Self-supervised learning of split invariant equivariant representations. In *Proceedings of the 40th International Conference on Machine Learning*, volume 202, pp. 10975–10996, 2023.

Satoki Ishikawa, Rio Yokota, and Ryo Karakida. Local loss optimization in the infinite width: Stable parameterization of predictive coding networks and target propagation. In *International Conference on Learning Representations*, 2025.

Max Jaderberg, Wojciech Marian Czarnecki, Simon Osindero, Oriol Vinyals, Alex Graves, David Silver, and Koray Kavukcuoglu. Decoupled neural interfaces using synthetic gradients. In *Proceedings of the 34th International Conference on Machine Learning*, volume 70, pp. 1627–1635, 2017.

Alina Kloss, Georg Martius, and Jeannette Bohg. How to train your differentiable filter. *Autonomous Robots*, 45(4):561–578, 2021. doi: 10.1007/s10514-021-09990-9.

Risi Kondor and Shubhendu Trivedi. On the generalization of equivariance and convolution in neural networks to the action of compact groups. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80, pp. 2747–2755, 2018.

Dong-Hyun Lee, Saizheng Zhang, Asja Fischer, and Yoshua Bengio. Difference target propagation. In *Machine Learning and Knowledge Discovery in Databases*, volume 9284, pp. 498–515, 2015. doi: 10.1007/978-3-319-23528-8_31.

Nikolay Manchev and Michael Spratling. Target propagation in recurrent neural networks. *Journal of Machine Learning Research*, 21(7):1–33, 2020.

Giovanni Luca Marchetti, Gustaf Tegnér, Anastasiia Varava, and Danica Kragic. Equivariant representation learning via class-pose decomposition. In *Proceedings of the 26th International Conference on Artificial Intelligence and Statistics*, volume 206, pp. 4745–4756, 2023.

Éloi Mehr, André Lieutier, Fernando Sanchez Bermudez, Vincent Guitteny, Nicolas Thome, and Matthieu Cord. Manifold learning in quotient spaces. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 9165–9174, 2018. doi: 10.1109/CVPR.2018.00955.

Jacob Menick, Erich Elsen, Utku Evci, Simon Osindero, Karen Simonyan, and Alex Graves. Practical real time recurrent learning with a sparse approximation. In *International Conference on Learning Representations*, 2021.

Arild Nøkland and Lars Hiller Eidnes. Training neural networks with local error signals. In *Proceedings of the 36th International Conference on Machine Learning*, volume 97, pp. 4839–4850, 2019.

Lianhai Ren and Qianxiao Li. BP-modified local loss for efficient training of deep neural networks. In *International Conference on Learning Representations*, 2025.

Julian D. Schiller, Malte Heinrich, Victor G. Lopez, and Matthias A. Müller. Tuning the burn-in phase in training recurrent neural networks improves their performance. In *International Conference on Learning Representations*, 2026.

Amirreza Shaban, Ching-An Cheng, Nathan Hatch, and Byron Boots. Truncated back-propagation for bilevel optimization. In *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics*, volume 89, pp. 1723–1732, 2019.

Corentin Tallec and Yann Ollivier. Unbiased online recurrent optimization. In *International Conference on Learning Representations*, 2018.

Bryan Wilder, Bistra Dilkina, and Milind Tambe. Melding the data-decisions pipeline: Decision-focused learning for combinatorial optimization. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pp. 1658–1665, 2019. doi: 10.1609/aaai.v33i01.33011658.

Kuan-Ting Yu, Maria Bauza, Nima Fazeli, and Alberto Rodriguez. More than a million ways to be pushed: A high-fidelity experimental dataset of planar pushing. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 30–37, 2016. doi: 10.1109/IROS.2016.7758091.

Yilang Zhang, Abraham Jaeger Mountain, Bingcong Li, and Georgios B. Giannakis. Binomial gradient-based meta-learning for enhanced meta-gradient estimation. In *International Conference on Learning Representations*, 2026.

APPENDIX GUIDE AND TERMINOLOGY

The appendices follow the order in which the main claims can be checked: the matching calculation and coupled-direction scope test (A), the two controlled training interventions and prospective supervision construction (B), the physical receiver and its full results (C), additional scope studies (D), and the artifact map (E). Figure and equation numbers continue from the main text.

Term or symbol	Meaning used throughout this paper
Task objective	The evaluation goal, held fixed across a supervision comparison.
Local supervision	The rule comparing an output with its reference; it can change without replacing the reference label.
Reference	The label or desired output used in a comparison.
Loss	The scalar objective implementing a supervision rule.
Target geometry	The directional penalty structure; a PSD metric W represents it in a fixed quadratic model.
Matched target geometry	Target geometry that matches the full task gradient for a specified task and backward closure.
Backward closure Γ	Which state-update derivative paths are retained. The numerical state updates continue.
Differentiation horizon H	The interval over which loss and reference gradients are evaluated. In the receiver, changing H changes the reference as well as the differentiated interval.
Trainable configuration	Which batch-optimized parameter blocks are active; fixed blocks still participate in forward signal processing.
R_H, $C_{H,\Gamma}$, A	Forward residual operator, exposed residual derivative, and task relevance metric in the fixed linearization.
Point / Partial / Quotient	Full pointwise penalty / partial directional relaxation / removal of specified directions in the idealized metric.
CQ	The receiver's complex-gain fitted loss, with the numerical ridge and normalization specified in Appendix C.4.
Full / AF-Full	Full in the controlled tracker covers the stated recursion. AF-Full in the receiver retains specified early and final adaptive paths inside the stated unroll. In both cases, the horizon and entry state define the derivative boundary.
D / N / R	Dispersion, nonlinear, and residual compensation parameter blocks. Original model names are mapped in Appendix C.2.
D_H, I	CQ-minus-Partial matching error / Detached-minus-AF-Full CQ training gain. These are distinct contrasts.

A THEORY DETAILS AND PROOFS

A.1 TRACKER MATCHING AND ITS ANALYTIC ILLUSTRATION

The scalar recursion in Section 3 solves to

$$s_H = \big(1 - (1-\kappa)^H\big)e_r, \qquad e_r - s_H = (1-\kappa)^H e_r. \tag{12}$$

The residual depends numerically on e_r through the factor $R_H = (1-\kappa)^H$. Under closure Γ, its exposed derivative with respect to e_r is $1 - d_H(\Gamma)$. Matching the coefficient of e_r in Eqs. (4)–(5) therefore gives $\lambda R_H[1 - d_H(\Gamma)] = \alpha R_H^2$. Dividing by the positive factors yields Eq. (6). Full differentiation gives $1 - d_H = R_H$; Detached differentiation gives $d_H = 0$.

The unit Point weight is recovered under Full when $\alpha = 1$. If $\alpha = 0$, the matched weight is zero for every closure satisfying the denominator condition. A genuinely task-null direction is therefore different from a direction that is only partly absorbed over a finite horizon. For $0 < \kappa < 1$ and fixed $\alpha > 0$, the Detached matched weight decreases with H in this contractive model. The monotonicity is specific to this contractive recursion.

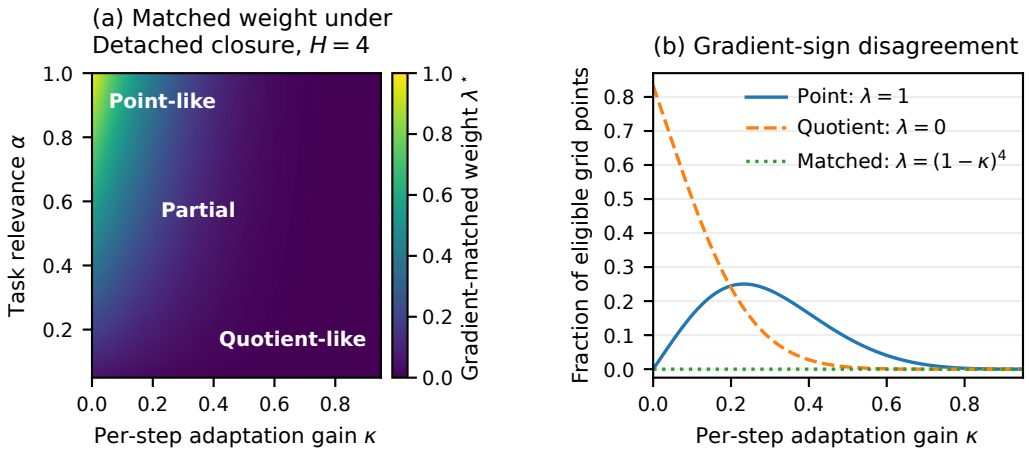

Figure 5: **Analytic matching in the contractive tracker.** (a) The Detached matched weight at $H = 4$. Region names describe a continuous range of penalties. (b) The fraction of eligible scalar-parameter grid points where the exposed and full gradients have opposite signs. The original plotted values are retained; the grid and method-specific denominator are defined below.

Figure 5 grids and gradient-sign disagreement. The calculation uses $H = 4$ and Detached closure (D). Panel (a) uses 101 uniform values of $\kappa \in [0, 0.95]$ and 101 uniform values of $\alpha \in [0.05, 1]$; $\lambda^\star = \alpha(1-\kappa)^4$. Its labels describe limiting regimes, not strict region boundaries. Panel (b) fixes $\alpha = 1$ and uses 20,001 uniform $\theta \in [0, 1.2]$ values, with $e_c = \theta - 1$ and $e_r = \theta + 1$. Writing $R = (1-\kappa)^4$, the parameter gradients are $g_\theta^{\text{full}} = e_c + R^2 e_r$ and $g_{\theta;\lambda,\mathrm{D}} = e_c + \lambda R e_r$. For method m with $\lambda_m \in \{1, 0, R\}$, define

$$\mathcal{E}_{m,\kappa} = \{\theta : \ |g_\theta^{\text{full}}(\theta)| > 10^{-10}, \ |g_{\theta;\lambda_m,\mathrm{D}}(\theta)| > 10^{-10}\}, \tag{13}$$

$$f_{m,\kappa} = \frac{\#\{\theta \in \mathcal{E}_{m,\kappa} : g_{\theta;\lambda_m,\mathrm{D}}(\theta) g_\theta^{\text{full}}(\theta) < 0\}}{\#\mathcal{E}_{m,\kappa}}. \tag{14}$$

Thus the denominator is method- and gain-specific; neither training trials nor finite-step increases are being counted. The accompanying analysis material records the grid, fractions, and eligible/excluded counts.

A.2 Spectral characterization and the approximation objective

In the shared orthogonal basis of Proposition 1, a diagonal metric produces the coefficient $c_i w_i r_i$, whereas the full gradient has coefficient $a_i r_i^2$. Matching every interface error requires

$$c_i w_i r_i = a_i r_i^2. \tag{15}$$

For $r_i > 0$ and $c_i > 0$, this yields $w_i = a_i r_i / c_i \geq 0$. The resulting metric is feasible and exactly matches the two operators. If $r_i = 0$, that direction contributes to neither operator and its weight can be set to zero. If $r_i > 0$ but $c_i = 0$ with positive a_i, the exposed graph cannot reproduce the nonzero full coefficient by a finite weight; that case is outside the proposition's active-support assumption.

A scalar metric wI must use the same w on every active direction, so exact matching requires the weights to coincide. For two closures with positive weights, their ratio is $c_i^{(2)}/c_i^{(1)}$. Unequal ratios across directions exclude transferring the exact matching metric by a global scalar. With Full closure $C_{H,\Gamma} = R_H$, $W = A$ is sufficient; the identity is the special case of a unit task metric.

For general fixed operators, error second moment $\Sigma = \mathbb{E}[ee^\top] \succeq 0$, and $\beta \geq 0$, consider the structured approximation problem

$$W^\star = \arg\min_{W \succeq 0} \|(C^\top W R - R^\top A R)\Sigma^{1/2}\|_F^2 + \beta \|W\|_F^2, \quad C = C_{H,\Gamma}, \ R = R_H. \tag{16}$$

The map $W \mapsto C^\top W R$ is linear, so the objective is convex on the closed PSD cone. For $\beta > 0$, coercivity and strong convexity guarantee a unique minimizer. At $\beta = 0$, attainment is not automatic for arbitrary operators: $W^\star$ denotes a minimizer only when the infimum is attained. Attainment holds in the shared-basis setting below and in the setting of Lemma 2. The spectral match attains zero residual under its assumptions, without imposing uniqueness on inactive directions.

When $\Sigma = U \operatorname{diag}(\sigma_i) U^\top$ also shares the basis, each diagonal quadratic yields

$$w^\star_{i,\beta} = \frac{\sigma_i a_i c_i r_i^3}{\sigma_i c_i^2 r_i^2 + \beta}. \tag{17}$$

This expression applies when its denominator is positive; zero-ridge directions unobserved by the objective are unconstrained. Off-diagonal terms contribute only additional nonnegative squared terms and can be set to zero. The PSD assumptions make the diagonal solution nonnegative. Choosing the exact matching weights on the active directions gives the main-text result.

A feasibility boundary outside the shared basis.

Lemma 1 (Detached matching with a positive-definite residual)**.** *Let $C = I$, $R = R^\top \succ 0$, and $A = A^\top \succeq 0$. A symmetric PSD metric matches the full gradient for every e if and only if $AR = RA$. When it exists, it is uniquely $W = RA$.*

Proof. Matching requires $WR = RAR$. Invertibility of R gives $W = RA$. Its symmetry is equivalent to $RA = (RA)^\top = AR$. Conversely, commuting real symmetric R and A have a shared orthogonal basis; their product has nonnegative eigenvalues, so $RA \succeq 0$ and matches exactly. □

This obstruction concerns a fixed symmetric quadratic metric under Detached closure. It does not cover every nonsymmetric residual operator, singular residual, or other supervision family. With Full closure, $C = R$ and $W = A$ still matches exactly without a shared basis. Thus changing only closure can change even the feasibility of exact PSD matching.

A closed-form PSD optimum outside the shared basis. For $C = I, \Sigma = I, \beta = 0, R = R^\top \succ 0$, and $A \succeq 0$, the unregularized objective is strictly convex on symmetric matrices because $W \mapsto WR$ is injective. Its stationary equation is

$$WR^2 + R^2 W = TR + RT, \qquad T = RAR.$$

In an orthogonal eigenbasis of R, with hats denoting coordinates in that basis, the unique symmetric minimizer is

$$\widehat{W}^{\mathrm{sym}}_{ij} = \frac{r_i + r_j}{r_i^2 + r_j^2} \widehat{T}_{ij}, \qquad \min_{W \succeq 0} \|WR - T\|_F^2 = \sum_{i<j} \widehat{T}_{ij}^2 \frac{(r_i - r_j)^2}{r_i^2 + r_j^2}. \tag{18}$$

The first expression follows by solving the stationary equation entrywise; substitution gives the residual over symmetric matrices. The following lemma establishes that this is also the exact PSD-constrained optimum.

Lemma 2 (Automatic PSD feasibility)**.** *Under the conditions above, $W^{\mathrm{sym}} \succeq 0$ and is the unique minimizer of Eq.* (16)*. If $A \succ 0$, then $W^{\mathrm{sym}} \succ 0$. No commutation between A and R is required.*

Proof. Let $K_{ij} = (r_i + r_j)/(r_i^2 + r_j^2)$. Since $r_i > 0$,

$$K_{ij} = 2 \int_0^\infty e^{-(r_i + r_j)t} \cos((r_i - r_j)t)\, dt.$$

This is the Gram matrix of the real two-component functions $\sqrt{2} e^{-r_i t}(\cos(r_i t), \sin(r_i t))$ in $L^2([0,\infty); \mathbb{R}^2)$, so $K \succeq 0$. Also $\widehat{T} \succeq 0$, because $T = RAR \succeq 0$. The Schur product theorem gives $\widehat{W}^{\mathrm{sym}} = K \circ \widehat{T} \succeq 0$. If $A \succ 0$, choose $m = \lambda_{\min}(\widehat{T}) > 0$; then $K \circ \widehat{T} \succeq m \operatorname{diag}(K_{ii}) \succ 0$, since $K_{ii} = 1/r_i$. Strict convexity gives uniqueness, and feasibility transfers the symmetric minimum and its residual to the PSD problem. □

Thus the positive definiteness of all 107 cases in Appendix A.5 follows analytically from their positive-definite task metrics. In the noncommuting cases, the nonzero residual is a symmetry obstruction, not an active PSD boundary or a numerical failure. This guarantee concerns the stated Detached, isotropic, zero-ridge problem; general C, Σ, or regularization still require the general formulation in Eq. (16).

A.3 Parameter gradients, approximation, and descent

For probe q with upstream Jacobian $J_{\theta,q}$, the parameter-gradient measurement is

$$M_q(W) = J_{\theta,q}^\top C_q^\top W R_q e_q. \tag{19}$$

Thus exact interface matching transfers through the same Jacobian. Approximate matching is sensitive to the map: an interface error δg_e has absolute parameter error at most $\|J_\theta^\top\|\|\delta g_e\|$, while its relative error also depends on $\|J_\theta^\top g_e^{\text{full}}\|$, which can be small or zero. The same local mismatch can therefore have different significance in different active parameter spaces.

For a nonzero full parameter gradient, suppose $\|g_\theta^W - g_\theta^{\text{full}}\| \le r\|g_\theta^{\text{full}}\|$ with $r < 1$. Then

$$\langle g_\theta^{\text{full}}, g_\theta^W\rangle = \|g_\theta^{\text{full}}\|^2 + \langle g_\theta^{\text{full}}, g_\theta^W - g_\theta^{\text{full}}\rangle \tag{20}$$

$$\ge (1-r)\|g_\theta^{\text{full}}\|^2 > 0. \tag{21}$$

The negative surrogate gradient is a strict first-order descent direction for that differentiable task at the same parameter/state context. A finite-step decrease additionally needs regularity and an appropriate step size. The certificate requires unnormalized parameter gradients at that same context. Block-normalized scores and endpoint test Q after Adam training measure different quantities.

From candidate matching to local task decrease. This connection does not require the shared spectral basis used to derive an analytic metric. Let $g = \nabla_\theta J \neq 0$ and let $v_k \neq 0$ be a candidate parameter gradient, both evaluated at the same fixed entry state and window. Use real parameter coordinates (real and imaginary parts for complex parameters). Define

$$a_k = \frac{[\langle g, v_k\rangle]_+}{\|v_k\|^2}, \qquad d_k = a_k v_k, \qquad E_k = \frac{\|d_k - g\|}{\|g\|}, \qquad [z]_+ = \max(z, 0).$$

Orthogonal projection onto the candidate ray gives

$$\langle g, d_k\rangle = \|d_k\|^2 = (1 - E_k^2)\|g\|^2, \qquad 1 - E_k^2 = [\cos(g, v_k)]_+^2. \tag{22}$$

Indeed, when $\langle g, v_k\rangle > 0$, substitution of a_k gives both equalities; otherwise $d_k = 0$ and $E_k = 1$. If J has an L-Lipschitz gradient on the step segment, the descent lemma yields

$$J(\theta) - J(\theta - \eta d_k) \ge \eta(1 - L\eta/2)(1 - E_k^2)\|g\|^2, \qquad 0 < \eta < 2/L. \tag{23}$$

Thus substantial relative mismatch can coexist with useful local descent; exact recovery of missing credit is not necessary. This is a local bound, not an ordering theorem for arbitrary finite steps. Its scale fit uses the reference gradient and is therefore an oracle diagnostic.

An equal-displacement comparison removes this scale fit. With $u_k = v_k/\|v_k\|$ and common $h > 0$,

$$J(\theta) - J(\theta - h u_k) = h\|g\|\cos(g, v_k) + O(h^2). \tag{24}$$

For two candidates with cosine gap $\delta > 0$, the better aligned candidate has greater decrease whenever $h < \|g\|\delta/L$, under the same smoothness condition. A finite-candidate calibration rule can therefore select the largest mean raw cosine on calibration windows, then freeze that supervision rule for a held-out window. Generalization to the held-out context must be measured; it is not implied by the same-context bound. Decision changes and finite-precision arithmetic can also limit the usable step range. Appendix C.8 tests these distinctions with saved receiver contexts. Averaging Eq. (22) requires averaging $1 - E_k^2$ per context, not squaring an already averaged error.

For a metric perturbation relative to $W^\star$, the mean squared interface mismatch is

$$\mathbb{E}\|g_e^W - g_e^{W^\star}\|_2^2 = \|C^\top(W - W^\star)R\Sigma^{1/2}\|_F^2. \tag{25}$$

The expression selects directions present in the data, surviving forward adaptation, and exposed by the closure. It connects the directional interpretation to the parameter-space diagnostic at the specified context.

A.4 Output-dependent geometry and finite-probe identification

The fixed-metric view extends locally to output-dependent transformation families. If $U(y)$ contains tangent directions and $\Lambda \succeq 0$ regularizes an inner displacement, define

$$\ell_\Lambda(e; y) = \min_\delta \tfrac{1}{2}\|e - U(y)\delta\|^2 + \tfrac{1}{2}\delta^\top \Lambda \delta. \tag{26}$$

At a fixed U, put $M_\Lambda = U^\top U + \Lambda$. A minimum-norm minimizer and the effective metric are

$$\delta^\star = M_\Lambda^\dagger U^\top e, \qquad W_\Lambda = I - U M_\Lambda^\dagger U^\top. \tag{27}$$

Here $\dagger$ denotes the Moore–Penrose inverse; it may be needed even for nonzero $\Lambda \succeq 0$. Indeed, $\ker M_\Lambda = \ker U \cap \ker \Lambda$, so $U^\top e$ lies in $\operatorname{range} M_\Lambda$. All minimizers are $\delta^\star + \ker M_\Lambda$ and have the same fitted displacement $U\delta^\star$. Substitution gives $\ell_\Lambda(e; y) = \frac{1}{2} e^\top W_\Lambda e$, with $0 \preceq W_\Lambda \preceq I$. When $M_\Lambda \succ 0$, the pseudoinverse is the ordinary inverse. For $\Lambda = tI$, $t \to \infty$ recovers Point; finite regularization gives a Partial metric, and $\Lambda = 0$ gives the local quotient projector. Scaling a singular regularizer need not recover Point in unpenalized directions. When U depends on the output, this metric describes the objective locally; gradients are taken from the original output-dependent loss. At a regular unique optimum of a differentiable inner objective $f(y, \delta)$, the envelope gradient is $\partial_y f(y, \delta^\star(y))$, retaining all explicit output dependence while holding the optimizer fixed.

For a finite probe set $\mathcal{Q}$, write $M_\mathcal{Q}(W) = (M_q(W))_{q \in \mathcal{Q}}$.

Proposition 2 (Finite-probe identifiability)**.** *An unrestricted symmetric PSD metric is uniquely identifiable over a feasible class only if $M_\mathcal{Q}$ is injective on feasible differences. If its kernel contains a nonzero symmetric ΔW and $W_0 \succ 0$, then sufficiently small $W_0 \pm \eta \Delta W$ are feasible and give identical probe measurements.*

Proof. Injectivity means equal measurements imply zero difference. Conversely, let $0 \neq \Delta W = \Delta W^\top \in \ker M_\mathcal{Q}$. Choosing $0 < \eta < \lambda_{\min}(W_0)/\|\Delta W\|_2$ keeps both perturbations positive definite. Linearity then gives $M_\mathcal{Q}(W_0 + \eta \Delta W) = M_\mathcal{Q}(W_0 - \eta \Delta W) = M_\mathcal{Q}(W_0)$. Thus noninjective probes identify an equivalence class modulo the common symmetric null space. Structural restrictions may make the restricted map injective; a finite number of probes alone does not rule out identification. □

The restricted candidate losses in the receiver and the selector in Appendix D.1 use this distinction: they compare structured objectives rather than recover an unrestricted time-varying metric.

A.5 Beyond the shared basis: coupled target–adaptation directions

This constructed experiment tests how much of the full-task gradient can be recovered by Eq. (16) when task and adaptation directions no longer share a basis. The matrices, all seven angles, three step sizes, direction grid, 4D sample count, seed, and metrics were fixed in a hashed protocol before executing this experiment. It uses no receiver data or training endpoints.

Fixed adaptation, rotated task relevance. The two-dimensional tracker is

$$s_{h+1} = (I - K)s_h + Ke, \quad s_0 = 0, \quad K = \operatorname{diag}(0.20, 0.03), \quad H = 4.$$

The residual operator $R = (I - K)^H = \operatorname{diag}(0.8^4, 0.97^4)$ and forward recursion remain fixed. The task is $J_\phi(e) = \frac{1}{2}(Re)^\top A_\phi Re$, where

$$A_\phi = Q_\phi \operatorname{diag}(1, 0.25) Q_\phi^\top, \qquad Q_\phi = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}, \qquad \phi \in \{0, 5, 10, 15, 20, 30, 45\}^\circ.$$

Within each angle, all supervision comparisons use the same task. Across angles the task metric changes deliberately, while adaptation and Detached closure remain fixed. This is a test of the matching rule's scope, distinct from the closure-only intervention P1. The commutator vanishes at 0° and is nonzero at every other tested angle; Lemma 1 then excludes exact symmetric PSD matching.

Nested geometry classes and an infeasible oracle. Write $T_\phi = R^\top A_\phi R$. We compare Point ($W = I$), the best nonnegative scalar wI, the best diagonal PSD matrix, and the general symmetric PSD solution of Eq. (16), always with $\Sigma = I$ and $\beta = 0$. Diagonal metrics are diagonal in the fixed tracker coordinates. The scalar and diagonal weights minimize the same operator objective as general PSD; they are not hand-tuned controls. An unconstrained matrix $W_{\mathrm{uncon}} = RA_\phi$ achieves zero operator error but is nonsymmetric when $\phi > 0$. It is only an oracle lower bound, never a valid quadratic loss metric or an additional update rule.

The primary measures are

$$M(W;\phi) = \frac{\|WR - T_\phi\|_F}{\|T_\phi\|_F}, \qquad \chi(\phi) = \frac{\|A_\phi R - RA_\phi\|_F}{\|A_\phi R\|_F}.$$

The first compares complete linear operators, without a subsequent scalar fit; it is not a finite-sample or worst-direction error. Equivalently, its squared numerator is the expected squared gradient error for $\mathbb{E}[ee^\top] = I$. The nested feasible classes guarantee that the optimum cannot worsen on moving from scalar to diagonal to general PSD. The experiment measures the size of this benefit and the irreducible residual, not a statistical test of that weak ordering.

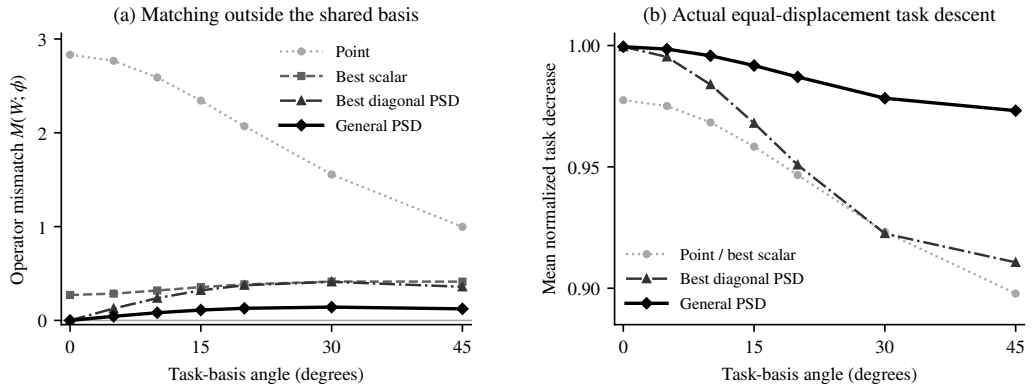

Figure 6: **Coupled directions admit useful but imperfect PSD matching.** **(a)** Relative operator mismatch for all seven fixed task-basis angles. The general PSD metric is exact at 0° and has nonzero residual elsewhere. The zero line is the unconstrained operator lower bound; that oracle is generally not a valid loss metric. **(b)** Actual task decrease divided by $h\|g_{\mathrm{full}}\|$, averaged over 360 uniformly spaced unit-circle directions, with common displacement $h = 10^{-3}$. Positive scalar scaling and Point give identical normalized updates and share one curve. Every loss is reevaluated through the forward tracker. The other two fixed step sizes are retained in artifact A10. Directions are deterministic quadrature points, not independent samples.

Table 1: **Complete 2D coupling sweep.** All quantities are dimensionless. The general PSD solution has exactly zero recorded symmetry error at each angle and a strictly positive minimum eigenvalue, as guaranteed by Lemma 2. Its relative stationarity residual is below 5×10^{-15}; the acceptance tolerance is 10^{-11}.

ϕ (deg.)	χ	M_{scalar}	M_{diag}	M_{PSD}	$\lambda_{\min}(W_{\mathrm{PSD}})$
0	0.0000	0.2713	0.0000	0.0000	0.2213
5	0.0932	0.2856	0.1270	0.0438	0.2208
10	0.1783	0.3189	0.2372	0.0818	0.2192
15	0.2496	0.3550	0.3205	0.1105	0.2162
20	0.3041	0.3841	0.3748	0.1292	0.2117
30	0.3636	0.4144	0.4114	0.1419	0.1975
45	0.3548	0.4134	0.3584	0.1236	0.1673

Matching accuracy and its boundary. Table 1 retains every angle. At 45°, general PSD reduces the relative operator error to 0.1236, compared with 0.3584 for diagonal and 0.4134 for scalar. Across the six nonzero angles, its error ranges from 0.0438 to 0.1419. Neither normalized coupling χ nor matching error is monotone in the angle: both are larger at 30° than at 45° for these fixed spectra. Allowing off-diagonal penalties therefore recovers much, but not all, of the task geometry over this sweep.

Actual local descent and numerical checks. We evaluate $e(\omega) = (\cos\omega, \sin\omega)^\top$ at 360 uniform angles in $[0, 2\pi)$, and update

$$e' = e - h\frac{WRe}{\|WRe\|}, \qquad h \in \{10^{-4}, 3\times10^{-4}, 10^{-3}\}.$$

Each method receives the same Euclidean displacement. All computations use float64, and every perturbed loss restarts the tracker at $s_0 = 0$. At 45° and $h = 10^{-3}$, mean normalized actual decrease is 0.9732 for general PSD, 0.9107 for diagonal, and 0.8978 for Point/scalar. The corresponding mean cosines are 0.9740, 0.9114, and 0.8983. The PSD metric has greater mean decrease than the diagonal metric at every nonzero angle for all three steps; the two tie at 0°. This is not pointwise dominance: at 45° it gives greater decrease than diagonal in 306 of the 360 directions, and its minimum cosine is 0.6827. All four legal metrics decrease the task in all tested directions and steps. These are local deterministic comparisons, not finite-training or statistical-generalization results.

Reverse steps verify $[J(e+hu) - J(e-hu)]/(2h) = g_{\text{full}}^\top u$ for the quadratic task. The maximum discrepancy divided by $\|g_{\text{full}}\|$ is 3.36×10^{-12}; direct decreases agree with the exact quadratic formula to 1.38×10^{-16} in absolute value. Independent finite differences of the local loss, with the forward state frozen, verify $g_W = WRe$. Full closure with $W = A_\phi$ matches at every angle, as a separate analytic sanity check.

4D rotation scope check. We keep $A = \mathrm{diag}(1, 0.6, 0.3, 0.15)$ fixed and set

$$R = QR_0Q^\top, \qquad R_0 = \mathrm{diag}(0.8^4, 0.88^4, 0.94^4, 0.97^4).$$

The 100 orthogonal matrices Q are generated by Gaussian QR with the diagonal-sign correction (seed 20260926). This is equivalent to rotating the adaptation gains while holding their spectrum fixed. Diagonal controls use the fixed task coordinates, and all methods share A and R within a realization. No local-step or training experiment is added in 4D. Table 2 reports all realizations, without significance tests or selection by outcome.

Table 2: **Operator mismatch over 100 constructed 4D rotations.** Percentiles describe this matrix ensemble, not independent physical systems. General PSD strictly improves over both diagonal and scalar in 100/100 cases, but retains nonzero error in every case.

Metric	10th percentile	Median	90th percentile
Point	1.1494	1.4992	1.9618
Best scalar	0.4460	0.4955	0.5641
Best diagonal PSD	0.1218	0.1774	0.2219
General PSD	0.0452	0.0691	0.1014

Optimization certificate and scope. All 107 constructed cases have a strictly positive-definite symmetric least-squares minimizer from Eq. (18), so it is the global PSD optimum; no iterative SDP tolerance explains the residual. The smallest eigenvalue over all cases is 0.06490, the maximum relative stationarity residual is 2.15×10^{-15}, and a second solve in an explicit symmetric-matrix basis agrees within 7.17×10^{-15} in Frobenius norm. The analytic irreducible-residual formula also agrees numerically. Artifact A10 includes the frozen protocol, all matrices and step records, independent checks, and figure source. These tests make Eq. (16) constructive beyond a shared basis for the specified known operators. They do not establish feasibility for arbitrary C, R, A, identify an unknown receiver metric, or supply a universal supervision selector.

B CONTROLLED INTERVENTIONS

B.1 WHAT IS HELD FIXED AND WHAT IS CHANGED?

Both experiments isolate a supervision decision within a fixed numerical system. The closure experiment crosses the same supervision pair with two backward closures; the directional experiment changes the local penalty under one closure. Equality of forward values refers to a common parameter value and entry state; trained branches subsequently follow their own trajectories.

Quantity	Closure intervention	Direction intervention
Inputs and task	Shared within each world	Shared within each seed
Initial parameters/state	Shared within each world	Shared pretrained parameters; zero tracker bias
Forward recursion	Same tracker rule	Same tracker rule
Backward closure	Detached versus Full; intermediate truncations also logged	Detached for plotted training arms
Supervision comparison	Same Point/fixed-Partial pair across closures	Scalar, wrong-subspace, and correct-subspace local terms
Evaluation	Fixed 400-update endpoint	Validation-selected checkpoint after 18 epochs

Measured data and protocol-derived quantities. Both interventions reuse measured MIT Push recordings (Yu et al., 2016). The loaders use timestamped planar tip/object poses and planar force–torque measurements (f_x, f_y, τ_z). They use trajectory summaries or phase-normalized sequences, not rendered images. File selection, splits, translations, cue noise where applicable, and downstream recursions are our protocol constructions. Normalization uses training data.

B.2 CUMULATIVE MATCHING AND THE CLOSURE-TRAINING IMPLEMENTATION

Summary inputs and predictor. The 15 summary features are tip/object displacements (four), initial relative position (two), wrench means/standard deviations (six), duration, and two path lengths (three). They use recorded series without asserting synchronization. The normalized terminal object–tip relative position defines target a; $b = a + \delta$. Two injected cues, δ plus noise, complete the 17-input vector, standardized using training data. A single affine layer maps it to two outputs. This differs from the sequence frontend in Section 4.2.

Five worlds reuse the same pool of 420 trajectory summaries. Each world creates its own 70/15/15 train/validation/test split and resamples injected translation and cue noise. Initialization is shared across targets and closures within each world. Full-batch Adam is run for 400 updates, and the terminal model is tested; validation measurements are logged but do not select this experiment's endpoint. These worlds test the intervention on resampled constructions of one trajectory pool, not on independently collected datasets. Each coordinate of δ is sampled from $N(0, 0.8^2)$ and clipped to $[-1.5, 1.5]$; independent cue noise is $N(0, 0.25^2)$. For world seed w, the split uses seed w, perturbations use $w + 100{,}000$, and parameter initialization uses $w + 777$. The affine layer uses PyTorch's default initialization: weights and biases uniform on $[-1/\sqrt{17}, 1/\sqrt{17}]$. Adam uses learning rate 0.02, $(\beta_1, \beta_2) = (0.9, 0.999)$, $\epsilon = 10^{-8}$, and weight decay 10^{-6}.

Let h be the predictor output, a the private target, b the shared target, and $e_r = h - b$. For direction i, put $q_i = 1 - \kappa_i$, where $(\kappa_1, \kappa_2) = (0.20, 0.03)$. The tracker uses $s_0 = 0$ and

$$u_{t,i} = e_{r,i} - s_{t,i}, \qquad s_{t+1,i} = s_{t,i} + \kappa_i u_{t,i}, \qquad u_{t,i} = r_{t,i} e_{r,i}, \quad r_{t,i} = q_i^t. \tag{28}$$

The residual for time t is computed *before* the state update. A TBPTT-K boundary stops the derivative after every K updates, without resetting the state value. Therefore, the residual derivative exposed at time t is

$$c_{t,i}^{(\Gamma)} = \begin{cases} 1, & \Gamma = \text{Detached}, \\ q_i^{\,t \bmod K}, & \Gamma = \text{TBPTT-}K, \\ q_i^t, & \Gamma = \text{Full}. \end{cases} \tag{29}$$

The tracker runs for 64 steps with an eight-step burn-in. The shared task averages $t = 8, \dots, 63$ and has unit task weight in both directions. Its full-gradient coefficient on $e_{r,i}$ is proportional to $\sum_{t=8}^{63} r_{t,i}^2$; the weight-$w_i$ surrogate under closure Γ exposes $w_i \sum_{t=8}^{63} r_{t,i} c_{t,i}^{(\Gamma)}$. Matching these coefficients gives

$$\boxed{w_{\Gamma,i}^{\star} = \frac{\sum_{t=8}^{63} r_{t,i}^2}{\sum_{t=8}^{63} r_{t,i} c_{t,i}^{(\Gamma)}}.} \tag{30}$$

The normalization by 56 scored times cancels. The weights are computed directly from Eq. (30), without fitting. The matching sequence, in the order Detached, TBPTT-4, TBPTT-16, and Full, is

$$(0.093, 0.470) \longrightarrow (0.119, 0.491) \longrightarrow (0.500, 0.595) \longrightarrow (1, 1). \tag{31}$$

In Full closure the numerator and denominator agree; the unit result here relies on the experiment's unit task weight. At seven-decimal precision, the weights in closure order are

$$(0.0932071, 0.4701037), \quad (0.1190216, 0.4912257),$$
$$(0.5000057, 0.5947599), \quad (1, 1).$$

Explicit recursion versus gradient-equivalent training. Define

$$\gamma_{\Gamma,i}(w_i) = \frac{1}{56} \sum_{t=8}^{63} w_i r_{t,i} c_{t,i}^{(\Gamma)}. \tag{32}$$

These coefficients are computed directly from Eq. (32). For training, we differentiate the gradient-equivalent quadratic

$$L_{\text{fast}} = \frac{1}{2} \operatorname{mean}_{n,i} (h_{n,i} - a_{n,i})^2 + \frac{1}{2} \operatorname{mean}_{n,i} \big[\gamma_{\Gamma,i}(w_i) (h_{n,i} - b_{n,i})^2 \big]. \tag{33}$$

At a fixed parameter value, this objective has the same parameter gradient as the explicit state-recursion implementation under the selected truncated graph. In general the two *scalar surrogate values* differ, since stop-gradient operators preserve forward values while changing derivatives. The reduction therefore justifies parameter-gradient and update equivalence, not scalar-loss equality. Explicit checks validate forward equality and matched/full gradient agreement. Thus the reported training results use the reduced objective, while the explicit recursion checks its derivative interpretation.

Heatmap and paired effects. For each displayed closure Γ and candidate k, first compute $\overline{J}_{\Gamma,k} = \frac{1}{5} \sum_w J_{w,\Gamma,k}$ and then $E_{\Gamma,k} = \overline{J}_{\Gamma,k} - \min_{\ell \in \mathcal{K}_{\text{shown}}} \overline{J}_{\Gamma,\ell}$. The minimum is over the five displayed candidates, not an additional tuning step. Figure 2(a) uses the same Detached-matched weights in both endpoint closures. Dots show all five paired worlds and diamonds their means; bars are the 95% bootstrap intervals of the paired contrasts. Their interpretation is restricted to the five resampled worlds, which share one finite trajectory pool.

The corresponding archived numerical values are

$$\Delta_{\text{D}} = 0.002165425941, \qquad \Delta_{\text{F}} = -0.001260203496,$$
$$\Delta_{\text{D}} - \Delta_{\text{F}} = 0.003425629437, \qquad 95\%\ \text{CI} = [0.002903490141, 0.004207772017].$$

B.3 Translation-gauge targets, mixed training, and distinct scalar controls

The gauge experiment reuses 420 measured trajectories, excluding the 240 trajectory records used in the preceding development run. The split is 294/63/63, with training-data normalization. Each sensor series is independently interpolated onto 128 normalized-phase samples, aligning normalized phase rather than absolute time. Angles are unwrapped before interpolation and rewrapped afterward; the timestep is derived from tip duration.

Perturbation, observation vector, and frontend. Tip and object positions share training-derived position normalization. For trajectory n, one two-dimensional offset δ_n is added to the normalized

reference-tip input and held fixed over all 128 times. Its coordinates are independent $N(0, 0.8^2)$ draws clipped to $[-1.5, 1.5]$:

$$p^{\text{sensor}}_{n,t} = p^{\text{ref}}_{n,t} + \delta_n. \tag{34}$$

The eight inputs are translated tip position (two), sine/cosine of tip orientation (two), normalized wrench (three), and timestep (one). Unshifted tip and object positions supply world-reference targets; no translation is explicitly added after either prediction head.

At each timestep, the shared frontend applies $8 \to 96$ linear, GELU, $96 \to 12$ linear, and Tanh, followed by separate $12 \to 2$ linear reference-tip and object heads. The invariance of L_Q concerns common shifts of the two predictions; their response to an input offset is determined by the frontend.

Tracker and task. The recursive tracker uses gain $\kappa = 0.25$, starts from zero bias, and scores the corrected object position after the first eight times:

$$\bar{p}^{\text{obj}}_t = \hat{p}^{\text{obj}}_t - b_t, \qquad b_0 = 0, \tag{35}$$

$$b_{t+1} = b_t + \kappa(\hat{p}^{\text{ref}}_t - p^{\text{ref}}_t - b_t), \tag{36}$$

$$J_{\text{task}} = \text{mean}_{n,t\geq 8,j}(\bar{p}^{\text{obj}}_{n,t,j} - p^{\text{obj}}_{n,t,j})^2. \tag{37}$$

The correction is evaluated before the update. Adding the same constant δ to both predicted positions changes the corrected output by $(1-\kappa)^t\delta$ when the entry state remains zero. The relative local loss is exactly invariant, but this finite-horizon closed-loop task retains a translation transient.

Local terms and the actual training objective. Let $e^{\text{ref}} = \hat{p}^{\text{ref}} - p^{\text{ref}}$ and $e^{\text{obj}} = \hat{p}^{\text{obj}} - p^{\text{obj}}$. With brackets denoting componentwise means over batch, all 128 times, and both spatial coordinates, the implementation uses

$$L_P = \tfrac{1}{2}\big(\langle (e^{\text{ref}})^2\rangle + \langle (e^{\text{obj}})^2\rangle\big), \tag{38}$$

$$L_Q = \langle (e^{\text{obj}} - e^{\text{ref}})^2\rangle, \tag{39}$$

$$L_W = \langle (e^{\text{obj}} - Q_\perp e^{\text{ref}})^2\rangle, \qquad Q_\perp = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{40}$$

$Q_\perp$ is fixed, not learned. The five original local terms are

Training label	L_{local} in $J_{\text{task}} + \beta L_{\text{local}}$
Point	L_P
Scalar Point	$s_P L_P$
Wrong Partial	$\lambda_{\text{mix}} L_P + (1-\lambda_{\text{mix}}) s_W L_W$
Partial	$\lambda_{\text{mix}} L_P + (1-\lambda_{\text{mix}}) s_Q L_Q$
Quotient	$s_Q L_Q$

Here $\beta = 0.5$ and $\lambda_{\text{mix}} = 0.5$ are distinct settings: the first weights the additional local term, and the second mixes its Point and quotient components. The five original branches and the post-primary task-only control in Figure 3(b) use Detached closure. Wrong Partial uses the mixed local objective above. Correct and wrong Partial have the same mixture coefficient; the pure-subspace diagnostic in Figure 3(a) is a different comparison.

B.4 Training, calibration, and the distinct scalar comparisons

Common pretraining, frozen training scales, and model selection. For each seed in $\{7, 17, 27, 37, 47\}$, the frontend is pretrained for 10 epochs using the equally weighted objective

$$L_{\text{pre}} = \langle (\hat{p}^{\text{ref}} - p^{\text{sensor}})^2\rangle + \langle [(\hat{p}^{\text{obj}} - \hat{p}^{\text{ref}}) - (p^{\text{obj}} - p^{\text{ref}})]^2\rangle.$$

The first target is the shifted sensor position; the second uses unshifted world-reference displacement. All main-training branches for that seed start from the same pretrained parameter state and a new Adam optimizer, resetting both moments and the step counter. Linear layers use the default PyTorch initialization, with the stated seed.

On the first calibration batch from the training split, let $v_P = \nabla_\theta L_P$, $v_Q = \nabla_\theta L_Q$, and $v_W = \nabla_\theta L_W$. The positive scale calibration is

$$s_Q = \mathrm{clip}_{[0.1,10]} \frac{\|v_P\|_2}{\|v_Q\|_2 + \epsilon}, \qquad s_W = \mathrm{clip}_{[0.1,10]} \frac{\|v_P\|_2}{\|v_W\|_2 + \epsilon}, \tag{41}$$

$$s_P = \frac{\|\lambda_{\mathrm{mix}} v_P + (1 - \lambda_{\mathrm{mix}}) s_Q v_Q\|_2}{\|v_P\|_2 + \epsilon}, \qquad \epsilon = 10^{-30}. \tag{42}$$

These three scales are frozen per seed before the 18 main-training epochs. The same seed/epoch minibatch order is used across branches, with batch size 32, Adam learning rate 10^{-3}, weight decay 10^{-5}, and gradient-norm clipping at 1. Adam uses $(\beta_1, \beta_2) = (0.9, 0.999)$ and $\epsilon = 10^{-8}$. Pretraining uses learning rate 2×10^{-3}. The main-training checkpoint minimizes validation task MSE and is then evaluated on the fresh test split.

The archived frozen scales are below; neither clipped scale reaches a bound in any seed. Calibration matches the specified initial gradient norms on the calibration batch; it does not keep those norms matched throughout training.

Seed	s_P	s_Q	s_W
7	0.815234	1.380931	0.424536
17	0.809151	1.348352	0.440214
27	0.767749	1.974266	0.386612
37	0.793109	1.188962	0.363619
47	0.783788	1.697230	0.411663

Pure local direction versus oracle total-gradient comparison. The local-gradient oracle uses the Detached Point model for each seed, eight batches of six test sequences per seed. It computes both Full and Detached *task-only* gradients. Figure 3(a) averages within each seed the cosines of $s_W \nabla L_W$ and $s_Q \nabla L_Q$ with $g_{\mathrm{miss}} = g_{\mathrm{task,Full}} - g_{\mathrm{task,Detached}}$; the positive scales do not change these cosines. No L_P is added or subtracted in this panel.

A separate *diagnostic* scalar coefficient a_P^{diag} is fitted using four validation batches of up to eight sequences each:

$$a_P^{\mathrm{diag}} = \max\left\{0, \frac{\sum_b (v_{P,b})^\top (g_{\mathrm{task,Full},b} - g_{\mathrm{task,Detached},b})}{\sum_b \|v_{P,b}\|_2^2 + \epsilon}\right\}. \tag{43}$$

The oracle scalar candidate is $g_{\mathrm{task,Detached}} + a_P^{\mathrm{diag}} v_P$, whereas the training Scalar Point uses $\beta s_P v_P$ as its local correction. These coefficients define distinct scalar comparisons. For the oracle relative-error comparison, each candidate *total* gradient g_k additionally receives its own nonnegative overall scale

$$\rho_k^\star = \max\left\{0, \frac{\langle g_k, g_{\mathrm{task,Full}} \rangle}{\|g_k\|_2^2 + \epsilon}\right\}, \qquad \mathcal{E}_k = \frac{\|\rho_k^\star g_k - g_{\mathrm{task,Full}}\|_2}{\|g_{\mathrm{task,Full}}\|_2 + \epsilon}. \tag{44}$$

These test-batch fits are used only for diagnosis; training uses the frozen calibration scales. Quotient reduces $\mathcal{E}_k$ relative to the diagnostic scalar candidate in 40/40 batches and relative to Partial in 37/40. The pure Wrong Quotient is evaluated only in the oracle comparison; Figure 3(b) trains the mixed Wrong Partial defined above. The experiments separately evaluate local-direction alignment, oracle-scaled gradient error, and validation-selected training performance. The comparison evaluates the calibrated scalar strategy defined above; broader scalar families are outside this experiment. The post-primary control below separately tests the benefit over task-only training. Unlike replacing the residual metric in Proposition 1, this additive construction requires $\beta \nabla_\theta L_{\mathrm{local}} = g_{\mathrm{miss}}$ for exact total-gradient matching. Cosine alignment tests only its direction.

Post-primary task-only control. After the primary analyses, we added $L_{\mathrm{train}} = J_{\mathrm{task}}$ as a sixth branch. It reuses the original 420 trajectories, normalization, split, five pretrained states, fresh Adam, minibatch order, 18 epochs, and validation-selected checkpoint rule. Original methods and outcomes are unchanged. This is a post-primary addition, not an originally prespecified comparison. The objective computes the task directly, without a zero-weight auxiliary graph: parameters unused by

the Detached task retain `None` gradients and are skipped by Adam, including weight decay. The reference-tip head consequently stays at its pretrained values.

Task-only has mean test MSE 0.099531, versus 0.091238 for Quotient and 0.094120 for Partial (Table 4). Quotient improves all five paired seeds and reduces mean error by 8.33%; Partial improves four and reduces the mean by 5.44%, losing seed 27. Point, Scalar Point, and Wrong Partial have higher means than task-only and improve only seed 37. Percentages compare the means over seeds, not the mean of seed-specific percentages. The seeds share one data split. Broader validation-tuned scalar weights or schedules remain outside this comparison.

Recovered inputs and matched-backend verification. All five pretrained parameter hashes match the original records. Replaying the ten original Point/Quotient checkpoints on the reconstructed inputs reproduces their test MSEs to within 9.82×10^{-9}. The new task-only runs use PyTorch 2.11.0 on CPU; the original runs used PyTorch 2.11.0 on a T4 GPU. We therefore additionally replayed the five original Quotient runs on the same CPU with unchanged settings. The maximum endpoint difference from the GPU archive is 5.75×10^{-8}; the CPU-to-CPU comparison retains all five wins and an 8.33169% mean reduction. These are verification replays of an existing method, not new supervision conditions. Separate forward calculations verify all five saved task-only checkpoints, and all validation selections and adverse outcomes are retained in the post-primary control addendum (artifact A12).

B.5 All plotted worlds, seeds, and additional controls

Table 3 lists every paired world shown in Figure 2. Its endpoint sign uses $J_{\text{Point}} - J_{\text{fixed Partial}}$ for both closures; the interaction is their difference. The intervals are the archived 95% paired-bootstrap intervals.

Table 3: All five paired-world contrasts in the closure intervention. Positive endpoint values favor the fixed Partial.

World	Detached Δ	Full Δ	Interaction
101	$+0.005265992$	$+0.000364367$	$+0.004901625$
211	$+0.002280112$	-0.000803994	$+0.003084106$
307	-0.001128014	-0.003815532	$+0.002687518$
401	$+0.002388299$	-0.000861580	$+0.003249878$
503	$+0.002020741$	-0.001184279	$+0.003205020$
Mean	$+0.002165426$	-0.001260203	$+0.003425629$

Table 4 gives the five seeds for each plotted supervision rule, including the post-primary task-only control. Wrong Partial uses the same mixture coefficient as correct Partial. Figure 3 retains the corrected pure-local diagnostic and all original Detached training records.

Table 4: All six Detached training branches at validation-selected checkpoints. Task only is the post-primary addition; the original five branches retain their archived values.

Seed	Task only	Point	Scalar Point	Wrong Partial	Partial	Quotient
7	0.096705	0.100070	0.098524	0.104591	0.092548	0.089517
17	0.101522	0.102115	0.103543	0.111236	0.091917	0.085643
27	0.098682	0.106573	0.105218	0.113614	0.099460	0.098136
37	0.095724	0.089634	0.090042	0.092757	0.086511	0.088078
47	0.105022	0.105097	0.108344	0.111066	0.100164	0.094818
Mean	0.099531	0.100698	0.101134	0.106653	0.094120	0.091238

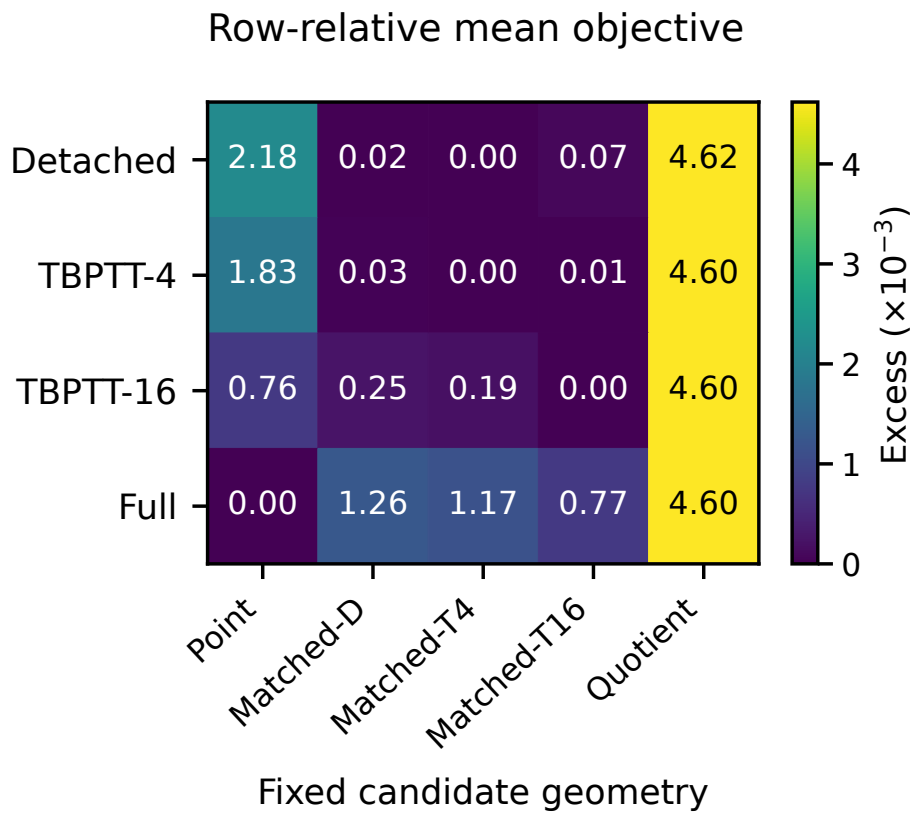

Figure 7: **The complete displayed candidate comparison.** Each cell first averages task error across five worlds, then subtracts the minimum of the displayed candidate means within its closure row. Matched-D, Matched-T4, and Matched-T16 are analytic weights fixed before training. This supplementary heatmap preserves the original candidate values while Figure 2 foregrounds the fixed-pair intervention.

Additional recorded comparisons. The swapped-weight and premature-Quotient comparisons in the closure study use all five worlds. Swapping the directional weights and removing the residual penalty both worsen test error in every world. Their mean excess test errors relative to the matched Detached Partial are 0.002863914147 and 0.004600899294, respectively. A scalar-controlled branch is within 0.000007215142 mean test error of matched Partial in that first system; this is one reason the separate sequence experiment supplies the explicit scalar-versus-direction training control. Analytic exact matching concerns the gradient relation, while these endpoint values record the actual finite optimizer runs.

B.6 Prospective supervision construction with known dynamics

This experiment uses Eq. (16) to construct a supervision metric before training a nonlinear predictor. It tests whether the construction has training value beyond choosing among fixed candidate rules. The original protocol, method definitions, primary comparison, random seeds, and code were frozen before its development and evaluation runs. Development outcomes determine the lookup baselines; all evaluation metrics and candidate choices are then frozen before evaluation training and test samples are generated. These are local timestamped records, not external preregistration. The protocol was developed after an exploratory candidate-selection pilot; only the new construction experiment is reported here. The operator-scalar control below was added after the primary analyses; it is explicitly post-primary and does not change the original frozen method set or outcomes.

Task and adaptive dynamics. For input $x \in \mathbb{R}^6$, a $6 \to 16 \to 2$ tanh network with biases produces $f_\theta(x)$; all 146 parameters are trainable. It serves a private reference y_c and an adaptively corrected reference y_r:

$$e_c = f_\theta(x) - y_c, \quad e_r = f_\theta(x) - y_r, \quad s_{h+1} = s_h + K(e_r - s_h), \quad s_0 = 0. \tag{45}$$

Each example resets the tracker and holds its error input fixed across the H updates. With $R = (I - K)^H$, the evaluation task is

$$J(\theta) = \tfrac{1}{2}\mathbb{E}\big[\|e_c\|^2 + (Re_r)^\top A(Re_r)\big]. \tag{46}$$

We use $H \in \{2, 4, 8\}$ and $A = Q_\phi \operatorname{diag}(1, 0.25)Q_\phi^\top$, $\phi \in \{0, 30\}^\circ$. The fixed-tracker variant has $K = \operatorname{diag}(0.2, 0.03)$; the rotated variant has $K = U \operatorname{diag}(0.2, 0.03)U^\top$, where U is a planar rotation with angle drawn uniformly from $[0, 180)^\circ$ once per task. Both variants therefore share the adaptation spectrum. Task and adaptation generally do not commute in the rotated variant.

The surrogate retains the private term and replaces the residual metric A by $W \succeq 0$. Its output gradient is $e_c + C^\top W R e_r$, where $C = I$ under Detached and $C = R$ under Full. All task scores replay the same explicit forward tracker. The main construction, called *operator PSD*, minimizes

$$W_{\mathrm{op}} = \arg\min_{W \succeq 0} \|C^\top W R - R^\top A R\|_F^2. \tag{47}$$

This is Eq. (16) with $\Sigma = I$, $\beta = 0$. It uses known A, R, C only, without calibration samples, Full parameter gradients, or training endpoints. The metric stays fixed during training.

Data and independent task draws. Each task draws a teacher and network initialization independently. Rows x are standard normal, and references are

$$y_c = \tanh(xM + b)V + 0.2xL + \epsilon, \qquad y_r = y_c + 0.7\tanh(xB + c) + 0.15xD.$$

Here $M \in \mathbb{R}^{6\times 8}$, $V \in \mathbb{R}^{8\times 2}$, $L, B, D \in \mathbb{R}^{6\times 2}$, $b \in \mathbb{R}^8$, and $c \in \mathbb{R}^2$. Their entries are independent zero-mean Gaussians with standard deviations $1/\sqrt{6}$, $0.7/\sqrt{8}$, $0.3/\sqrt{6}$, $1/\sqrt{6}$, $1/\sqrt{6}$, 0.2, and 0.2, respectively; $\epsilon \sim N(0, 0.05^2 I)$. The first and second network weight matrices have entry variances $1/6$ and $0.1^2/16$, with zero biases. Each task supplies 128 calibration, 512 training, 2,048 test, and 128 local-probe examples from separate random streams.

Each variant has eight development tasks and 16 new evaluation tasks. The six horizon/orientation conditions, two closures, and two optimizers are paired within task, sharing teacher, samples, initialization, and minibatch order. Thus there are 32 evaluation tasks; conditions and optimizer repetitions are not additional independent tasks.

Construction classes and lookup controls. Write $A^{1/2}$ for the symmetric positive square root and let U denote the known tracker eigenbasis. The *operator directional* control minimizes Eq. (47) over

$$W = A^{1/2} U \operatorname{diag}(w_1, w_2) U^\top A^{1/2}, \qquad w_1, w_2 \geq 0. \tag{48}$$

This is a continuous two-weight family, not a diagonal restriction in arbitrary coordinates. Both methods have the same known operators. The eight fixed candidates are I, A, the four metrics in Eq. (48) with $(w_1, w_2) \in \{(0.75, 1), (0.50, 1), (0.25, 1), (0, 1)\}$, and $0.5A$, $0.25A$. Even the fixed candidates receive the correct tracker basis in the rotated variant.

The primary comparator is a development-selected lookup that minimizes mean normalized endpoint loss among those eight candidates, separately for each variant, optimizer, closure, horizon, and task orientation. It is frozen before evaluation and selects no evaluation endpoint. Coarser lookups use one candidate per variant/optimizer or per variant/optimizer/closure. A gradient-selected candidate instead maximizes initialization-time parameter-gradient cosine with the Full reference, averaged over eight calibration minibatches of 16 examples.

Additional gradient-fit methods minimize the summed parameter-gradient squared mismatch at initialization, dividing each minibatch's squared error by its Full-reference squared norm. They respectively optimize a general PSD metric, Eq. (48), or $W = \alpha A$, $\alpha \geq 0$. The common private gradient is present on both sides of the fit; scalar fitting changes only the residual term. A closure-blind PSD fit pools the Full and Detached calibration objectives equally and uses the same metric for both. All fitted matrices are frozen before evaluation training; the operator constructions themselves do not use these gradient probes.

Training and prespecified comparison. Every branch takes 200 updates with batch size 64. Normalized SGD uses parameter displacement $0.025/\sqrt{1+t/50}$ at update t; Adam uses learning rate 0.01, $\beta_1 = 0.9$, $\beta_2 = 0.999$, and $\epsilon = 10^{-8}$. Hyperparameters and the final-update endpoint are fixed, without early stopping or evaluation-based tuning. Fourteen training branches per condition produce 10,752 evaluation trajectories; all are retained. Lookup and selection policies reuse the corresponding candidate trajectory.

For each condition we compute $z = J_{\mathrm{end}}/J_{\mathrm{init}}$ on the same test set, then average z within the reported group. Relative reduction against lookup is $100(1-\bar{z}_{\mathrm{method}}/\bar{z}_{\mathrm{lookup}})$. The primary contrast is operator PSD versus the fine lookup under Detached, fixed tracker, and normalized SGD. Rotated trackers, Adam, other methods, Full closure, and local steps are prespecified secondary comparisons. Table 5 reports both condition-level and task-mean wins; no independence or significance claim uses the 96 paired conditions as separate task draws.

Table 5: **Known dynamics guide supervision construction before training.** Detached closure. Lookup and PSD columns report mean normalized test loss; lower is better. Conditions and Tasks count strict improvements over the development-selected fine lookup out of 96 conditions and 16 task means. The first row is the prespecified primary comparison. Within each row, paired conditions not counted as strict wins are retained and have higher loss.

Tracker	Optimizer	Lookup	PSD	Reduction (%)	Conditions	Tasks
Fixed	Norm. SGD	0.205506	0.204528	0.476	87/96	15/16
Fixed	Adam	0.162512	0.161915	0.367	71/96	15/16
Rotated	Norm. SGD	0.198885	0.197806	0.542	85/96	15/16
Rotated	Adam	0.164582	0.163681	0.547	73/96	14/16

A two-weight family captures most of the PSD improvement. Operator PSD reduces mean normalized task loss by 0.37–0.55% across the four Detached groups. The operator-directional restriction already achieves reductions of 0.24–0.52%, leaving only 0.03–0.14% extra from the general PSD metric. Table 6 retains all policy controls: the reported gradient-fit scalar is worse than the fine lookup in every group, and candidate selection gives little or negative average gain. This scalar uses a parameter-gradient calibration objective; the post-primary operator-scalar control below instead uses the same objective as the operator constructions. Gradient-fit PSD performs similarly to operator PSD, without establishing a need for its additional reference-gradient probes. Figure 8 shows all task means, including losses.

Table 6: **Full controls for prospective construction.** Percent reduction in mean normalized test loss versus the same development-selected fine lookup under Detached. Negative values mean higher loss. SGD denotes normalized SGD. The fixed and closure-only lookups are also learned on development endpoints.

Method	Fixed / SGD	Fixed / Adam	Rotated / SGD	Rotated / Adam
Gradient-selected candidate	+0.038	-0.032	-0.045	-0.143
Operator PSD	+0.476	+0.367	+0.542	+0.547
Operator directional	+0.405	+0.243	+0.408	+0.515
Gradient-fit PSD	+0.418	+0.339	+0.545	+0.562
Gradient-fit directional	+0.359	+0.230	+0.402	+0.525
Gradient-fit scalar	-0.690	-1.490	-0.652	-1.027
Closure-blind PSD	+0.154	+0.155	+0.358	+0.247
Development fixed rule	-0.607	-2.339	-0.499	-0.759
Closure-only lookup	-0.607	-0.805	-0.499	-0.759

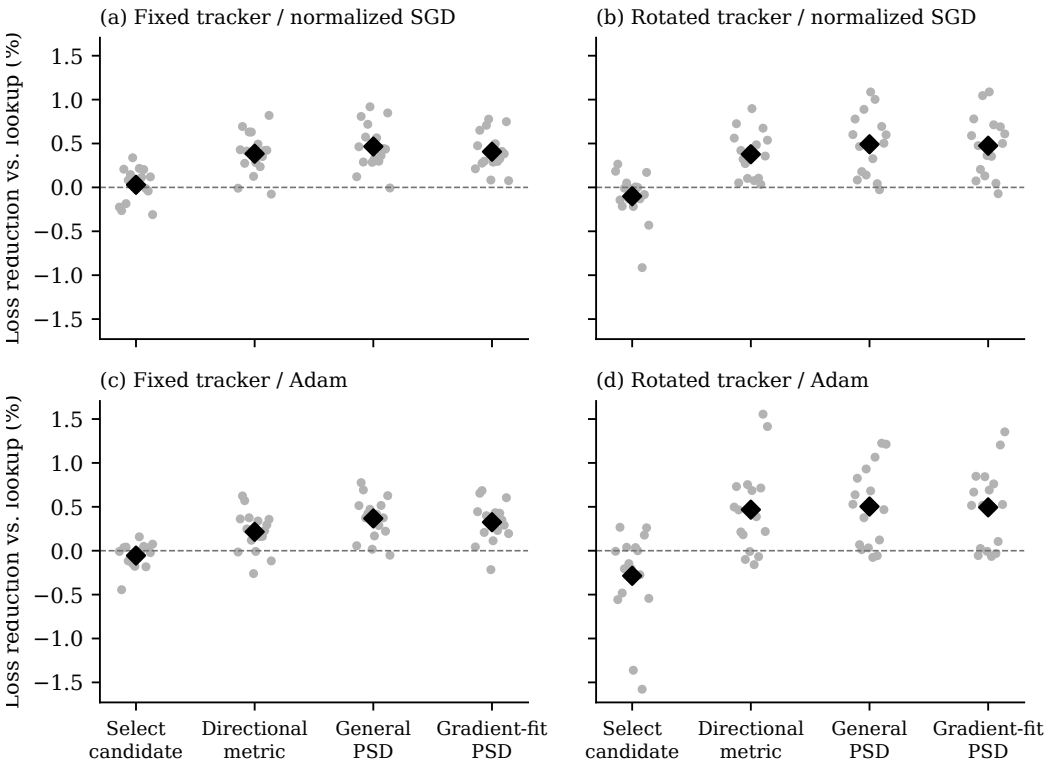

Figure 8: **Prospective construction improves most task means across two tracker variants and two optimizers.** All methods are frozen before training; all panels use Detached closure. Gray points show each of the 16 task means, aggregating its six horizon/orientation conditions; black diamonds average the task-specific percentage reductions. This differs slightly from the ratio of group means in Table 5. Zero is the fine lookup baseline. Directional and PSD construction use known operators; gradient-fit PSD additionally uses Full parameter-gradient probes. Candidate selection is retained as a control, including its losses.

Post-primary operator-scalar control. After the primary analyses, we added the restriction $W = \alpha A$, $\alpha \geq 0$, to the same operator objective in Eq. (47). Writing $B = C^\top AR$ and $T = R^\top AR$ gives

$$\min_{\alpha \geq 0} \|\alpha B - T\|_F^2, \qquad \alpha^\star = \max\left\{0, \frac{\langle B, T\rangle_F}{\|B\|_F^2}\right\}.$$

If $B = 0$, the objective is constant in α and we use $\alpha = 0$. This scalar receives the same known A, R, C as the directional and PSD constructions, without calibration samples or Full parameter-gradient probes. All 384 scalar metrics are frozen before the added training runs, while the original evaluation data, initializations, minibatch orders, optimizers, and 200-update endpoints are reused. The addition is post-primary, not part of the original prespecified method set; all 768 added endpoints are retained.

Table 7: **Post-primary scalar control under the same operator objective.** Detached closure. Reduction is $100(1 - \bar{z}_{\text{method}}/\bar{z}_{\text{scalar}})$, where $z = J_{\text{end}}/J_{\text{init}}$. Dir. denotes the two-weight directional family; PSD denotes the full PSD construction. Conditions count strict wins out of 96 repeated conditions; Tasks count wins after averaging each task's six horizon/orientation conditions. All other comparisons have higher loss; there are no ties. SGD is normalized SGD.

Tracker	Optimizer	Scalar $\bar{z}$	Reduction (%) Dir.	PSD	Conditions Dir., PSD	Tasks Both
Fixed	SGD	0.206984	1.116	1.187	92, 92	15/16
Fixed	Adam	0.164922	1.701	1.823	95, 95	16/16
Rotated	SGD	0.200184	1.054	1.187	87, 89	15/16
Rotated	Adam	0.166284	1.533	1.565	90, 87	15/16

The directional construction lowers mean normalized loss by 1.05–1.70% relative to this scalar, and PSD by 1.19–1.82% across the four groups (Table 7). Both improve all 24 tracker/optimizer/horizon/orientation stratum means, but not every task or condition. Both lose the task means for fixed task 4014 with SGD, rotated task 4024 with SGD, and rotated task 4031 with Adam. All individual-condition nonwins are listed in Table 8.

Under Full, $C = R$ yields $\alpha^\star = 1$; all 384 added Full endpoints match the archived Task branch exactly. Independent nonnegative least squares checks all scalar constructions, and a separate forward calculation checks all 768 endpoints. These results isolate the benefit of directional freedom under a shared operator objective and information source. They do not exclude every scalar weight schedule or validation-tuned training rule. Artifact A12 supplies all added records; A11's original records are unchanged.

Table 8: **All nonwins against Operator scalar under Detached.** Entries are percentage excess normalized endpoint loss over scalar for each condition. A dash denotes a strict win, rather than a missing result; all listed numeric entries are losses and there are no ties. Fixed tasks are numbered 4000–4015; rotated tasks 4016–4031. SGD is normalized SGD.

Task	Optimizer	H	ϕ (deg.)	Directional (%)	PSD (%)
4005	Adam	2	30	0.2507	0.2158
4006	SGD	2	30	0.1903	0.0859
4014	SGD	2	0	0.1135	0.1135
4014	SGD	2	30	0.3640	0.3861
4014	SGD	4	30	0.2749	0.2370
4017	Adam	2	0	0.2467	0.2422
4017	Adam	2	30	0.6990	0.7793
4017	Adam	4	30	0.3385	0.6994
4017	Adam	8	30	0.0946	0.4772
4017	SGD	2	30	0.1779	0.1837
4024	SGD	2	0	0.2829	0.2719
4024	SGD	2	30	0.3294	0.1943
4024	SGD	4	0	0.3668	0.2804
4024	SGD	4	30	0.3988	–
4024	SGD	8	0	0.2398	0.2887
4026	SGD	2	0	0.1176	0.0599
4026	SGD	2	30	0.4526	0.2549
4026	SGD	4	30	0.4662	–
4028	Adam	8	0	–	0.2422
4031	Adam	2	0	0.6906	0.7989
4031	Adam	2	30	–	0.0085
4031	Adam	4	0	0.4552	0.6314
4031	Adam	4	30	–	0.2202

A same-task check against the candidate-set limit. As a retrospective descriptive comparison, we give an oracle access to all eight fixed-candidate endpoints and let it select the lowest test loss separately in each evaluation condition. Operator PSD still has lower group-mean loss in all four groups, by 0.28–0.41% (Table 9). It does not win every condition. The oracle is never used for construction or baseline selection. This comparison establishes that useful constructed geometries extend beyond this finite candidate set; it does not establish an optimal geometry over all supervision rules or compare ceilings from different experiments.

Table 9: **Construction compared with an endpoint oracle on the same tasks.** Detached closure. The oracle takes the best of the eight fixed candidates separately in each condition, using test endpoints. It is evaluation-only. Columns report mean normalized losses, their relative reduction, and conditions where operator PSD has strictly lower loss.

Tracker	Optimizer	Candidate oracle	PSD	Reduction (%)	Conditions
Fixed	Norm. SGD	0.205118	0.204528	0.287	82/96
Fixed	Adam	0.162365	0.161915	0.277	68/96
Rotated	Norm. SGD	0.198630	0.197806	0.415	78/96
Rotated	Adam	0.164188	0.163681	0.309	68/96

Full closure, local descent, and scope. Under Full, operator PSD recovers $W = A$. Relative to fine lookup, mean loss changes by $+0.0059\%$ (fixed/SGD), -0.0013% (rotated/SGD), and zero for both Adam groups; here a positive sign denotes reduction. These are sanity comparisons, not part of the Detached gain. Under Detached, independent local-probe examples at the common pre-training parameter initialization give operator PSD mean parameter-gradient cosines 0.999544 and 0.999646 in the fixed and rotated variants, respectively. All 96 equal-displacement 10^{-3} local steps per variant reduce the actual task. The corresponding Full mean cosines are both 1, as exact matching predicts. These probes precede optimization; finite-training comparisons retain adverse cases.

All numerical records use float64. Independent finite differences over 24 gradients (146 coordinates each) have maximum relative error 2.44×10^{-11}. These use centered coordinate perturbations $\theta \pm 10^{-5} e_j$ and relative Euclidean gradient error. For Detached, the terminal state computed at the unperturbed parameter is held fixed for both evaluations of the scalar surrogate; for Full, it is recomputed. The checks cover two tracker variants, $H \in \{2, 8\}$, both closures, and Task, operator PSD, and gradient-fit PSD metrics at 30°. Recomputing all 10,752 endpoint losses with an independent forward implementation agrees within 1.39×10^{-17}. All 1,152 fitted PSD solutions have a projected-gradient optimality residual below 6.86×10^{-15}; their unconstrained symmetric solutions are already feasible, so this ensemble does not test active PSD-boundary cases. Artifact A11 includes all frozen matrices, samples, parameters, trajectories, controls, and numerical audits.

The construction is prospective and does not require a shared basis, but its operators are known and its downstream dynamics are linear and two-dimensional. It gives small average gains; the structured two-weight family captures most of the general PSD improvement. Estimating the required operators and validating such construction in an unknown physical receiver remain separate problems.

C Optical system and systematic experiments

C.1 Receiver architecture and state paths

The receiver combines batch-learned compensation and online adaptive signal processing as in Fan et al. (2021a). The task loss compares recovered symbols after the adaptive chain with their transmitted references.

The signal order is D/N, power normalization, an early adaptive frequency-offset module (RDE/MIMO and Kalman processing), R, and final adaptive MIMO/DDLMS. The FOE frame size is 100. The closure switch affects selected state-update derivatives while preserving numerical forward values at common parameters and entry state. "AF-Full" retains these paths inside the specified unroll. Entry states are fixed for differentiation, and pilot/tracking mode is a separate choice.

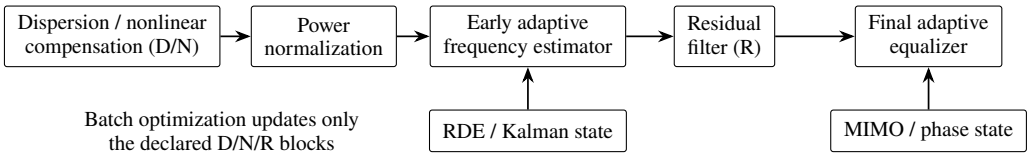

Figure 9: **The receiver separates trainable parameters from adaptive states.** D/N compensation precedes the early adaptive frequency-offset estimator; R follows it and precedes the final adaptive equalizer. Solid arrows indicate signal order. The state boxes update in the forward pass independently of whether their derivatives are retained. The diagram summarizes the inherited architecture of Fan et al. (2021a).

C.2 Configurations and parameter freezing

The receiver has three dispersion/nonlinearity steps, with 261-tap D-filters and a 61-tap R-filter. The nonlinear filter and active optimization blocks distinguish the configurations:

Configuration	D taps	N taps	Initial ξ	Batch-trainable blocks
CDC	261	1	0	R
DBP	261	1	0.15	R
FDBP	261	41	1.1	R
EDBP	261	41	1.1	N+R
GDBP	261	41	1.1	D+N+R

Physical D/N initialization and the delta R initialization are preserved. Fixed and active parameters are partitioned, then merged for the forward pass; signal derivatives through fixed-weight layers remain available. The optimizer manages only active leaves. Evaluation restores both the fixed and active parameters from the trained model. Adaptive DSP states follow their own update rules, separately from batch optimization.

C.3 Data provenance, coverage, and retained task accounting

The inherited data are received waveforms and transmitted references from LabPtPTm2 (Fan et al., 2021b), associated with the long-haul system in Fan et al. (2021a). SRC identifies the transmitted source sequence; REP identifies a capture within the source/configuration. The original physical transmission has seven WDM channels. The present analysis reuses the four received channels meeting the full data rule. Each model is trained and tested separately at each workpoint.

The plan contains 35 physical workpoints: five powers by seven received channels. Each requires SRC1/REP1–2 for training and SRC2/REP1–3 for testing. Raw-array presence is insufficient: references, usable length, and the matching supplement information for initialization must also be present. SRC2/REP3 supplements are unavailable or invalid at all five powers on channels 3, 5, and 6. No other capture's CD or frequency-offset initialization is substituted.

Twenty workpoints are complete. Coverage therefore executes $20 \times 5 \times 2 \times 2 = 400$ tasks. The horizon and closure studies use a prespecified subset of three powers and three channels. This mechanism subset uses nine workpoints, three models, two training captures, and four arms: $9 \times 3 \times 2 \times 4 = 216$. The 616 tasks comprise 604 newly trained results and 12 verified reused GDBP anchor arms; the original list of 916 retains 300 data-blocked tasks. All eligible training and required evaluation outcomes are complete, with no recorded numerical model failures. The computational outcome counts are separate from the archive-packaging receipt, whose status is preserved in the artifact records.

Within each configuration/workpoint/training capture, all compared branches share the same committed zero-step parameter and state values. Different captures retain their own data-dependent states. Reused anchor results retain their original experimental identity. Technical failures, numerical model failures, and missing data are separate states; no incomplete pair is zero-filled.

Study layer	Original plan	Complete-data execution
Point/CQ coverage, 500-symbol updates	700	400
Matched-cadence closure comparison	216	216
Total training/evaluation tasks	916	616
Data-blocked tasks retained in the ledger	—	300
Gradient diagnostic tasks	162	162

The original model–workpoint grid has 175 cells (five configurations, five powers, seven channels). Its complete-data panel has 100 cells (four channels), each with two training captures. The mechanism subset retains all 27 prespecified model–workpoint cells. These counts refer to the nested levels of the design.

C.4 Training cadence, objectives, and evaluation

The loader requests 1.1 million training symbols and 420,000 test symbols. Production branches use 2,000 updates of 500 symbols; mechanism branches use 500 updates, each averaging four consecutive 500-symbol microbatch losses with states advancing between them. Both see one million training symbols and use the fixed endpoint, rather than validation-selected checkpoints. Learning-rate levels are $10^{-4}, 10^{-5}, 10^{-6}$ with boundaries after 500 and 1,000 microbatches (250,000 and 500,000 symbols). The archived complex Adam update and parameter partition are unchanged between the paired targets. Complex parameters are represented as real/imaginary coordinate pairs. For a JAX gradient g, the optimizer updates them using $\Re g$ and $-\Im g$, respectively, with separate real Adam moments and elementwise squared coordinate gradients. Thus the second moments are not a shared $|g|^2$ accumulator. Both components use $(\beta_1, \beta_2) = (0.9, 0.999)$, $\epsilon = 10^{-8}$ and the

usual bias corrections. The optimizer implementation is included in the supplementary material. The segment boundaries truncate inter-segment derivatives, not the forward state values.

For a 500-symbol, two-polarization microbatch ($n = 1000$ complex entries), the systematic implementation uses $E_z = \|z^\star\|^2 + n \times 10^{-12}$ and $\epsilon_a = 10^{-12}$ in Eq. (50). Point is $\|\hat{z} - z^\star\|^2/E_z$. The diagnostic Partial is $0.6L_{\mathrm{Point}} + 0.4L_{\mathrm{CQ}}$. One complex alignment is fitted jointly across both polarizations *per 500 symbols*, even when $H = 2000$ or 4000. Thus alignment granularity is not the differentiated horizon. The residual-only stopped-alignment implementation and the complete minimized ridge objective have the same first-order parameter gradient: the stopped ridge term has zero outer derivative. Their scalar reported loss values need not be identical.

Test processing rebuilds the adaptive state, transfers the saved power-normalization state, and uses 200,000 pilot symbols before decision-directed operation. Scoring starts at effective output index 300,000, excludes a final 20,000-symbol guard, and uses complete 8,192-symbol windows. The original unaligned receiver outputs enter hard QAM decisions and BER-to-Q scoring; test outputs are not CQ-aligned to favor the CQ branch. For each trained model, window Q is averaged within each test capture and the three capture means are equally weighted. Point and CQ share the same valid output interval within a configuration. The shorter N-filter in CDC/DBP changes effective cropping; their corresponding source start differs by 30 symbols from FDBP/EDBP/GDBP for the same output-relative start. Cross-family source alignment therefore uses these recorded offsets rather than window numbers alone.

Quality factor and bit error rate. Following Fan et al. (2021a), receiver quality factor (Q-factor) is reported in decibels as a transform of bit error rate (BER):

$$Q = 20\log_{10}\left[\sqrt{2}\ \mathrm{erfc}^{-1}(2\,\mathrm{BER})\right], \qquad \mathrm{BER} = \tfrac{1}{2}\,\mathrm{erfc}\left(\frac{10^{Q/20}}{\sqrt{2}}\right). \tag{49}$$

Here erfc^{-1} is the inverse complementary error function. This logarithmic scale is not an absolute change in prediction accuracy, and the BER-derived Q-factor is not identified with physical signal-to-noise ratio. For scale, increasing Q from 8.5 to 8.54508 dB maps BER from 0.00389863 to 0.00374126: a 4.04% relative decrease, or about 157 fewer erroneous bits per million. Table 10 shows the baseline dependence. These are illustrative conversions, not additional receiver measurements. Window Q values are averaged before reporting; inverting their mean does not recover the mean BER. An aggregate BER effect would require the archived per-window errors and bit counts.

Table 10: **A Q-factor gain has a baseline-dependent BER interpretation.** Each row applies a hypothetical $+0.04508$-dB increment to the stated baseline using Eq. (49); these are metric-conversion examples, not measured aggregate BER effects.

Baseline Q (dB)	BER before	BER after	Relative reduction	Fewer errors/10^6 bits
7.1	0.011767	0.011410	3.03%	357
8.5	0.003899	0.003741	4.04%	157
9.1	0.002179	0.002079	4.57%	100

For historical scale, Fan et al. (2021a) reports average Q-factor gains of 0.13 and 0.36 dB for GDBP over FDBP and DBP, respectively, on the 1125-km platform. Those are architecture comparisons under the original protocol, not matched baselines for the present supervision intervention. Our $+0.04508$-dB mean is an incremental gain from changing supervision with the receiver architecture and inference-time DSP unchanged. Its evidence combines effect size with a positive sign in every tested D+N+R power–channel workpoint and both paired training-capture comparisons within each workpoint. It does not establish a decoding-threshold crossing or a transmission-reach increase.

The full task ledger records training, evaluation, reuse, and blocked conditions separately. Descriptive averages first preserve the paired training capture and the specified test-capture weighting. Counts and averages are descriptive summaries of correlated captures, workpoints, and windows within this platform.

The complete normalized, ridge-stabilized CQ objective used to define the parameter gradient is

$$L_{\mathrm{CQ}}(\hat{z}, z^\star) = \min_{a\in\mathbb{C}} \frac{\|a\hat{z} - z^\star\|^2 + \epsilon_a|a|^2}{E_z}, \qquad a^\star = \frac{\hat{z}^H z^\star}{\hat{z}^H\hat{z} + \epsilon_a}. \tag{50}$$

Reference normalization is output-independent. The first-order envelope gradient holds the fitted scalar fixed while differentiating all explicit occurrences of $\hat{z}$. The ridge stabilizes the numerical objective; exact complex-quotient invariance applies to the zero-ridge idealization.

C.5 Gradient diagnostic protocol and aggregation

Diagnostics use production-Point checkpoints after 0, 250,000, and 1,000,000 symbols. Probe starts are 1,025,000, 1,050,000, and 1,075,000; the reconstructed receiver state is warmed up for 20,000 symbols before each probe. The parameter checkpoint is fixed and no optimizer updates are executed. Each candidate/closure comparison starts from the same reconstructed and warmed-up state, separately from the original training-state history. BPN uses training-mode normalization, and the diagnostic equalizer uses pilots throughout.

Increasing H changes both the differentiated interval and its reference; the candidate loss definitions and complex-alignment fitting unit stay fixed. Each horizon $H \in \{500, 2000, 4000\}$ averages $H/500$ microbatch losses. The AF-Full reference differentiates Point through the specified early and final state paths over the requested horizon H. Candidate Point, Partial, and CQ gradients use Detached closure and the same horizon/data. Complex gradients are converted consistently to real coordinates, restricted to the actually trainable blocks.

For candidate vector g_k and nonzero reference r, raw matching uses

$$a_k^\star = \max\left(0, \frac{\langle g_k, r\rangle}{\|g_k\|^2 + 10^{-30}}\right), \qquad E_k = \frac{\|a_k^\star g_k - r\|}{\|r\|}. \tag{51}$$

The scale is fitted in the measured context for a diagnostic only, not deployed into training. In the block-normalized comparison, each active block is divided by $\max(\|r_b\|, 10^{-30})$ and the nonnegative scale is fitted again. Raw error already includes scale fitting; it is not unscaled Euclidean distance. FDBP has one active block, so this block normalization is an overall scalar that cancels from the optimally rescaled relative error.

There are $3\times3\times3\times2\times3 = 162$ model/workpoint/capture/checkpoint tasks, each with nine window–horizon contexts, for 1,458 contexts and 4,374 candidate rows. A workpoint mean in Figure 4 averages two captures, three checkpoints, and three windows (18 contexts per H). A task mean averages three windows. All denominators refer to these different, explicitly nested units.

C.6 Complete results, robustness, and adverse contexts

Table 11 gives all candidate means in the raw parameter metric. It includes the nonnegative scale fit in Eq. (51). The full window records also contain the block-normalized counterpart.

Table 11: All raw parameter-gradient relative-error means, including the optimal nonnegative scale fit. Each model/horizon entry averages 162 checkpoint–window contexts.

Model	Candidate	$H = 500$	$H = 2000$	$H = 4000$
FDBP	Point	0.7443	0.8888	0.9175
FDBP	Partial	0.6417	0.7984	0.8389
FDBP	CQ	0.6351	0.5819	0.6023
EDBP	Point	0.4725	0.9693	0.9836
EDBP	Partial	0.4285	0.9593	0.9766
EDBP	CQ	0.8992	0.7038	0.7248
GDBP	Point	0.6637	0.9752	0.9818
GDBP	Partial	0.6059	0.9632	0.9727
GDBP	CQ	0.8253	0.7324	0.7691

The workpoint-average matching gap moves toward CQ from $H = 500$ to 2000 and from 500 to 4000 in all 27 cells. Raw and block-normalized scores give the same workpoint crossover counts. The two comparisons share the $H = 500$ baseline; gaps between 2000 and 4000 vary by context. At the task level, simultaneous short-Partial/both-long-CQ crossovers are FDBP $18/54$, EDBP $53/54$, GDBP $51/54$, identically counted under the raw and block metrics. The remaining tasks favor CQ at all three horizons.

At the window level, FDBP has 70 Partial preferences and 92 CQ preferences at $H = 500$, with CQ preferred in all 162 checkpoint–window contexts at each longer horizon. EDBP has 159 Partial preferences at $H = 500$; each longer horizon has two three-candidate ties with fitted scale zero and error one. GDBP has 153 Partial preferences at $H = 500$ and two CQ non-wins at each longer horizon, at the -4 dBm/channel-4, one-million-symbol checkpoint's first window across the two captures. These candidates have errors close to one. Across all 1,458 contexts, the raw and block-normalized CQ-versus-Partial sign/tie classifications differ in four cases. All tied and adverse contexts remain in the provided tables.

Full training coverage. The five original configurations are all retained in Table 12. Table 13 gives the 200 paired gains as 40 workpoint/capture rows with five model columns. The main text focuses on the three nested configurations; the tables here include both additional R-only baselines. For D+N+R under Detached differentiation, Figure 10 shows every complete power–channel workpoint. All 20 workpoint means and all 40 paired training-capture comparisons are positive. The mean and median paired gains are $+0.04508$ and $+0.04475$ dB; the minimum is $+0.01243$ dB. Power-specific means range from $+0.03406$ to $+0.05219$ dB, and channel-specific means from $+0.01987$ to $+0.06387$ dB. The same fixed CQ rule is used throughout, with separate training at each workpoint; these results characterize consistency across tested operating conditions, not transfer of one trained model to unseen channels.

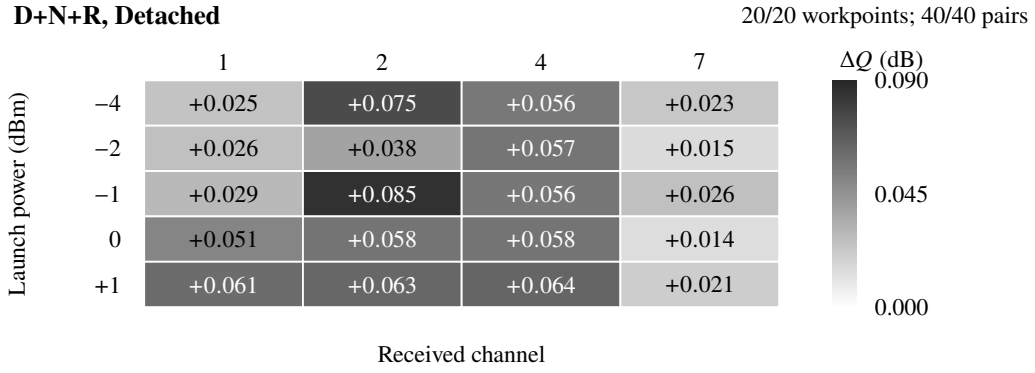

Figure 10: **CQ gains persist across all five tested launch powers and four complete received channels.** D+N+R is trainable under Detached differentiation. Each cell reports the CQ-minus-Point Q-factor gain in dB, averaged over the two paired training-capture comparisons at a fixed one-million-symbol endpoint with 500-symbol updates. Both individual comparisons are positive in every cell (40/40 pairs; 20/20 workpoints). Channels 3, 5, and 6 lack the required complete evaluation data and are excluded by the coverage rule, not by their performance. The grid describes correlated workpoints on one physical platform.

Table 12: Complete five-configuration coverage. Positive counts summarize the paired outcomes descriptively.

Model	Trainable	Mean gain (dB)	Positive pairs	Positive cells
CDC	R	+0.00136	30/40	14/20
DBP	R	+0.00012	21/40	11/20
FDBP	R	+0.00094	26/40	17/20
EDBP	N+R	+0.00799	36/40	20/20
GDBP	D+N+R	+0.04508	40/40	20/20

Table 13: All 200 production500 CQ-minus-Point Q gains (dB). Each entry is one training capture, evaluated by the prescribed equal-weight mean across three test captures.

Power	Ch.	Rep.	CDC	DBP	FDBP	EDBP	GDBP
-4	1	1	+0.00222	+0.00427	+0.00162	+0.00671	+0.03021
-4	1	2	+0.00207	+0.00570	+0.00130	+0.00513	+0.01921
-4	2	1	+0.00710	−0.00399	+0.00027	+0.01071	+0.08946
-4	2	2	+0.00283	+0.00177	−0.00370	−0.00038	+0.05974
-4	4	1	−0.00285	+0.00534	+0.00208	+0.00108	+0.05043
-4	4	2	+0.00071	+0.00164	+0.00366	+0.00266	+0.06175
-4	7	1	+0.00109	−0.00126	−0.00029	+0.00523	+0.01904
-4	7	2	+0.00107	−0.00009	+0.00231	+0.00146	+0.02786
-2	1	1	+0.00176	+0.00089	+0.00033	+0.01260	+0.03481
-2	1	2	+0.00135	−0.00067	+0.00017	+0.00298	+0.01738
-2	2	1	−0.00440	−0.00019	−0.00087	+0.00206	+0.03835
-2	2	2	−0.00208	+0.00207	+0.00214	−0.00075	+0.03794
-2	4	1	+0.00692	−0.00074	+0.00259	+0.01139	+0.07060
-2	4	2	+0.00460	+0.00298	−0.00121	+0.00325	+0.04370
-2	7	1	+0.00602	−0.00183	+0.00176	+0.00797	+0.01503
-2	7	2	+0.00002	−0.00211	−0.00012	−0.00041	+0.01468
-1	1	1	−0.00001	+0.00262	+0.00190	+0.01698	+0.03016
-1	1	2	−0.00138	+0.00041	−0.00134	+0.01771	+0.02813
-1	2	1	−0.00237	+0.00107	+0.00212	+0.00991	+0.08318
-1	2	2	+0.00055	+0.00310	+0.00162	+0.00390	+0.08767
-1	4	1	+0.00129	−0.00148	−0.00046	+0.00015	+0.04580
-1	4	2	+0.00437	−0.00089	+0.00067	+0.00101	+0.06717
-1	7	1	−0.00389	−0.00141	−0.00192	−0.00201	+0.02186
-1	7	2	+0.00172	+0.00487	−0.00200	+0.00566	+0.03053
+0	1	1	+0.00358	−0.00246	+0.00005	+0.02136	+0.04206
+0	1	2	+0.00341	−0.00157	+0.00097	+0.03328	+0.05952
+0	2	1	−0.00040	+0.00093	+0.00382	+0.00788	+0.05101
+0	2	2	+0.00310	+0.00040	+0.00548	+0.01303	+0.06469
+0	4	1	+0.00089	−0.00135	+0.00094	+0.00586	+0.05208
+0	4	2	+0.00078	−0.00167	−0.00049	+0.00827	+0.06327
+0	7	1	−0.00098	+0.00338	+0.00073	+0.00178	+0.01243
+0	7	2	−0.00193	−0.00663	−0.00171	+0.00108	+0.01607
+1	1	1	+0.00030	−0.00863	−0.00016	+0.01385	+0.05261
+1	1	2	+0.00043	+0.00174	+0.00164	+0.02733	+0.06952
+1	2	1	+0.00471	+0.00048	+0.00378	+0.01309	+0.05972
+1	2	2	+0.00396	+0.00054	−0.00033	+0.01867	+0.06691
+1	4	1	+0.00183	+0.00050	+0.00412	+0.00428	+0.05367
+1	4	2	+0.00172	−0.00486	+0.00460	+0.01997	+0.07385
+1	7	1	+0.00235	−0.00072	−0.00162	+0.00172	+0.01516
+1	7	2	+0.00193	+0.00252	+0.00318	+0.00320	+0.02608

The paired gain difference compares the same workpoint and training capture across configurations. GDBP exceeds EDBP in all 40 comparisons, with mean difference $+0.03709$ dB; EDBP exceeds FDBP in 34/40, with mean difference $+0.00705$ dB. These contrasts concern the training effect of supervision in whole trained configurations, including their distinct parameter and state trajectories. The complete paired differences are included in supplementary artifact A5 (Table 22).

All matched-cadence interactions. Table 14 retains all 54 capture-level interactions. Power-specific means, used in Figure 4(d–f), average the six channel/capture combinations at each power. Positive and negative interactions are both retained.

Table 14: All 54 segment2000 closure interactions (dB), retaining both training captures at every prespecified workpoint.

Power	Ch.	Rep.	FDBP	EDBP	GDBP
-4	1	1	-0.001470	$+0.008529$	$+0.013300$
-4	1	2	-0.000961	$+0.006000$	-0.009754
-4	4	1	-0.010771	-0.021648	$+0.000176$
-4	4	2	-0.010085	-0.009162	-0.026257
-4	7	1	-0.007287	-0.015298	-0.004080
-4	7	2	-0.002888	-0.002417	-0.003037
-1	1	1	$+0.002676$	$+0.010527$	$+0.016024$
-1	1	2	$+0.003128$	$+0.013596$	$+0.013074$
-1	4	1	-0.011864	-0.004271	$+0.009596$
-1	4	2	$+0.000903$	-0.000775	$+0.002064$
-1	7	1	-0.008701	-0.011670	$+0.002412$
-1	7	2	-0.002293	-0.001339	$+0.004058$
+1	1	1	-0.001843	-0.002942	$+0.015018$
+1	1	2	$+0.001820$	-0.000181	$+0.028489$
+1	4	1	$+0.002899$	$+0.004590$	-0.000720
+1	4	2	-0.002777	$+0.011671$	$+0.007984$
+1	7	1	-0.002472	$+0.009516$	$+0.009647$
+1	7	2	-0.002063	$+0.004156$	$+0.001429$

Local-to-training comparison. The diagnostic/endpoint join matches model, power, channel, and training capture. At the one-million-symbol checkpoint and $H = 2000$, all 54 task-mean CQ errors are below Point. Segment2000 Detached CQ training gains are positive in 11/18 FDBP, 16/18 EDBP, and 18/18 GDBP pairs; the corresponding positive closure interactions are 5/18, 8/18, and 13/18. The local gradients use 500-symbol-update Point checkpoints with reconstructed states. The endpoint comparison uses separately trained segment branches, which follow their own parameter and state trajectories. For R-only at -4 dBm/channel 4, both captures improve gradient matching but have negative training effects. Appendix A.3 gives the sufficient first-order descent condition at a single shared parameter/state context.

Table 15: Matched conditions, different evaluation objects: final Point-checkpoint diagnostics at $H = 2000$ versus separately segment-trained endpoints.

Configuration	CQ closer than Point	Positive D gain	Positive interaction
FDBP	18/18	11/18	5/18
EDBP	18/18	16/18	8/18
GDBP	18/18	18/18	13/18

C.7 MODULE-DEPENDENCY CHECK AT FIXED PARAMETERS

A separate check uses two completed Point endpoints, three predetermined windows, $H = 2000$, and Point loss. Three closures keep all state derivatives detached, open only final adaptive-state derivatives, or additionally open the early FOE/RDE paths. All three start with the same parameters, inputs, and warmed-up entry state; their stored forward outputs, Point losses, and final-state values coincide.

Opening the final paths changes D/N/R gradients. With those paths already open, additionally opening early paths changes D and N but not R beyond repeated-computation error. The block distance is $\|g_B^{(1)} - g_B^{(2)}\| / \max(\|g_B^{(1)}\|, \|g_B^{(2)}\|)$, without scale fitting. For the latter comparison, the six-context ranges are 0.0365–0.1518 (D), 0.4035–1.0643 (N), and 1.12×10^{-7}–1.97×10^{-7} (R). R gradients are nonzero; their maximum absolute cross-closure difference is 2.2526×10^{-8}, versus 2.6143×10^{-8} in repeated same-closure differentiation. This matches the serial location of R after the early adaptive module. This check establishes which retained derivatives affect each parameter group at fixed parameters and entry state; final-Q comparisons use the separately trained configurations.

C.8 SAME-CONTEXT TASK DESCENT AND NUMERICAL LIMITS

The gradient error in Section 5.2 need not be close to zero to imply local descent. Equation (22) provides a quantitative interpretation. Reaggregating the existing D+N+R scores at $H = 2000$, the mean fraction $[\cos(g, v_k)]_+^2$ is 0.4350 for CQ, 0.0632 for Partial, and 0.0424 for Point. At $H = 4000$, the CQ fraction is 0.3767. These are algebraic transformations of saved gradients, not new loss measurements or fractions of eventual Q improvement.

New receiver replay. We separately replay all six saved contexts from Appendix C.7: D+N+R trainable, channel 1 at -1 dBm, two Point checkpoints after one million symbols, and three 2,000-symbol windows per capture. Each replay restores the exact saved parameters, received/transmitted input segment, and warmed-up entry state. State updates remain active throughout the perturbed forward pass. We recompute Detached Point, Partial, and CQ gradients and the AF-Full Point reference, and measure the unaligned Point task loss after an update. This is a finite-horizon task-loss experiment, not a new training trajectory, Q evaluation, or BER measurement. The six contexts share one workpoint and are correlated within captures.

The first experiment uses float32/complex64 on CPU with the original pinned receiver source and JAX version. Detached and AF-Full forward outputs agree exactly on this backend. Relative differences from the original L4 records reach 1.78×10^{-4} for outputs and 9.77×10^{-4} for AF-Full gradients. The original strict elementwise cross-backend tolerances do not all pass; the CPU experiment is reported separately, with its own recomputed baseline and gradients, rather than as exact reproduction of the L4 run.

Fixed comparisons and calibration. For the primary comparison, each candidate receives the same Euclidean parameter displacement $h = \epsilon\|\theta\|$, using the conjugated JAX gradient for complex parameters. The step grid, fixed before perturbed-loss evaluation, is $\epsilon \in \{10^{-6}, 3\times10^{-6}, 10^{-5}, 3\times10^{-5}, 10^{-4}, 3\times10^{-4}, 10^{-3}\}$. AF-Full supplies a reference direction. Every trial resets the entry state; zero and reverse steps support a central-difference check. A secondary comparison uses $\theta - \eta a_k v_k$, with $\eta = h/\|g\|$ and the oracle scale a_k from Appendix A.3. No optimizer moments are used.

Before any perturbed loss is evaluated, each held-out window receives the candidate with greatest mean raw cosine on the other two windows of the same capture. All six folds select CQ. The held-out window's gradient and losses are not used for that choice. The folds overlap and are not independent; because the selected rule is always CQ, this check does not establish an advantage over always choosing CQ.

Small-step precision control. The float32 equal-displacement results have mixed orderings (Table 16); their central differences can disagree substantially with the local derivative. To diagnose this limitation, we subsequently promote the saved floating-point values to float64/complex128 and recompute both gradients and task losses, without retraining or replacing contexts. This is an exploratory numerical control, with a separately fixed grid $\{10^{-10}, 3\times10^{-10}, \ldots, 10^{-7}, 3\times10^{-7}, 10^{-6}\}$; it does not replace the original float32 results.

Table 16: **Actual task-loss decreases from receiver replay.** Each count combines six correlated contexts with the indicated step grid; it is not a count of independent experiments. Matched grids use the same seven relative displacements from 10^{-6} to 10^{-3}. The nine-step control uses smaller displacements. Both float64 controls were added after the primary float32 experiment; all outcomes are retained.

Replay	Steps	CQ > Point	CQ > Partial	CQ decreases task
Float32, matched grid	7	25/42	22/42	37/42
Float64, matched grid	7	31/42	32/42	40/42
Float32, oracle scale	7	38/42	38/42	38/42
Float64, small-step control	9	54/54	54/54	54/54

In this control, CQ produces greater actual decrease than Point and Partial in all six contexts at every step size. For steps through 3×10^{-7}, the maximum central-derivative error divided by $\|g\|$ is 1.24×10^{-4} across all four directions and all contexts. At 10^{-6}, the reverse-step check for Point and Partial fails in one context (normalized discrepancy 0.851), although their forward-step ordering remains unchanged. That boundary is retained in the supplied records. At 10^{-6}, CQ's mean relative task-loss decrease is only 1.16×10^{-5}: the control tests the local mechanism, not the size of a practical training gain.

Matched-step precision comparison. The preceding float64 control also uses a smaller step grid, so it does not by itself separate precision from step-size effects. We therefore evaluate float64 on the original seven-step grid and pair the results with the existing float32 trials at identical relative parameter displacements. The six contexts, saved inputs, parameters, entry states, candidate losses, and receiver source are shared. Each precision uses its own baseline and candidate gradients; the previously computed float64 gradients are reused. The shared 10^{-6} trials exactly reproduce the preceding float64 control.

At 10^{-6}, CQ gives greater task decrease than Point in 5/6 float32 and 6/6 float64 contexts; at 10^{-5} the counts are 3/6 and 6/6; at 10^{-3}, both give 2/6. Over the full seven-step matched grid, the counts are 25/42 and 31/42, respectively (Table 16). These repeated step-size comparisons share six contexts and are not independent samples. Higher precision improves agreement with the local ordering at small displacements, but does not remove finite-step reversals. This is consistent with the $O(h^2)$ remainder and step-size condition in Appendix A.3; it does not identify every source of finite-step discrepancy. We do not attribute the gap between the 54/54 gradient-matching diagnostics and the 45/54 finite-training gains to single-precision arithmetic.

Figure 11(b) compares the same-grid relative task-decrease advantage

$$B(\epsilon) = \frac{J(\theta - hu_{\mathrm{Point}}) - J(\theta - hu_{\mathrm{CQ}})}{J(\theta)}, \qquad h = \epsilon\|\theta\|.$$

Positive values favor CQ; negative values favor Point. Each curve averages the same six contexts, normalizing by each context's own baseline task loss.

These results connect candidate geometry to actual local task descent when the derivative is numerically resolved, while exposing limits to using that connection for a finite update. They do not derive CQ as the globally matched geometry or validate supervision selection across unseen operating conditions. Artifact A9 includes the original, small-step, and matched-step precision protocols, all perturbation records, calibration choices, source/input hashes, and replay code. The broader 20-workpoint coverage and the separate Condition-B study retain their original protocols and counts.

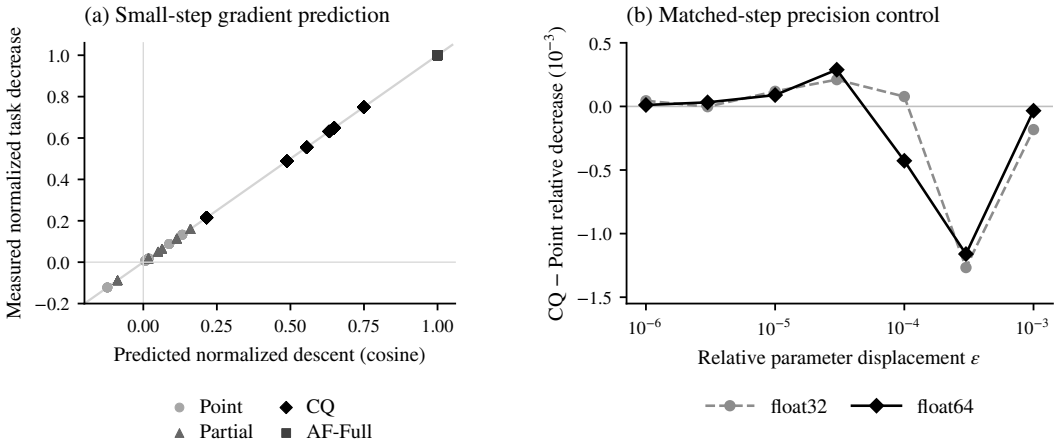

Figure 11: **Local descent has a measurable range of validity.** (a) In the double-precision small-step control, measured decrease divided by $h\|g\|$ follows the candidate cosine. All six contexts and nine small step sizes are shown; the diagonal is the first-order prediction. (b) Float32 and float64 use identical relative displacements. The mean CQ-minus-Point relative task decrease, $B(\epsilon)$, changes with precision; at larger steps, negative means occur under both precisions. Each mean uses the same six contexts. These replays measure Point task loss with active forward adaptation, not Q or an Adam training endpoint.

D Additional scope studies and negative results

The following studies answer questions beyond the main fixed-target, fixed-budget comparisons. They retain their own data partitions, objectives, and decision rules. These studies are separate from the systematic receiver panel in Section 5 and do not contribute to the workpoint counts, paired-comparison counts, or aggregate receiver gains reported there.

D.1 A frozen selection application

This separate study selects a target and scale within a predefined family before rollout on each acquisition. All six folds select CQ, so the reported comparison evaluates that frozen choice against gradient-calibrated, loss-scaled Point.

D.1.1 Family, parameter probes, and nominal decision gates

For the reference normalization used in the original selection protocol, one complex alignment is fitted jointly over both polarizations per 500-symbol microbatch ($n = 1000$, $E_z = \|z^\star\|^2 + n\epsilon$, $\epsilon = 10^{-12}$, and ridge $\epsilon_a = 10^{-12}$). The summed reference-energy floor is therefore $n\epsilon = 10^{-9}$. Its proximal family is

$$L_\tau(\hat{z}, z^\star) = \min_{a\in\mathbb{C}} \frac{\|a\hat{z} - z^\star\|^2 + \epsilon_a|a|^2}{E_z} + \tau|a-1|^2, \tag{52}$$

$$a_\tau^\star = \frac{\hat{z}^H z^\star + \tau E_z}{\hat{z}^H\hat{z} + \epsilon_a + \tau E_z}. \tag{53}$$

At $\tau = 0$ this is ridge-stabilized CQ. As $\tau \to \infty$, the Point gradient is recovered; the limiting objective differs by the output-independent constant ϵ_a/E_z. The historical stopped-alignment residual has the same first-order gradient, while the prospective implementation evaluates the full minimized objective. A 2,000-symbol objective averages four microbatches.

The evaluated grid is

$$\mathcal{T} = \{0, 10^{-4}, 3{\times}10^{-4}, 10^{-3}, 3{\times}10^{-3}, 10^{-2}, 3{\times}10^{-2}, 0.1, 0.3, 1, 3, 10, 30, 100, 300, 1000\}. \tag{54}$$

Algorithm 1 Paired endpoint-aware selection for held-out acquisition i

Require: Outer set S_i excluding i; candidates Point and L_τ; paired gradients
for each candidate k and acquisition $q \in S_i$ **do**
  Fit $s_k(S_i \setminus \{q\})$ and compute $E_q(k; S_i \setminus \{q\})$
end for
Bootstrap $d_q = E_q(\mathrm{CQ}) - E_q(\mathrm{Point})$
if nominal upper gate $U_{.95}(d) < 0$ and every condition mean is negative **then**
  $b \leftarrow \mathrm{CQ}$
else if nominal lower gate $L_{.95}(d) > 0$ and every condition mean is positive **then**
  $b \leftarrow \mathrm{Point}$
else
  return abstain (Point fallback; the full-success gate fails)
end if
Choose k_{int} minimizing condition-stratified cross-fitted error over $\tau \geq 0.01$
Bootstrap $E_q(k_{\mathrm{int}}) - E_q(b)$ using the same cross-fitted errors
if nominal upper gate is negative and every condition mean is negative **then**
  $k^\star \leftarrow k_{\mathrm{int}}$
else
  $k^\star \leftarrow b$
end if
Fit $s_{k^\star}(S_i), s_{\mathrm{Point}}(S_i), s_{\mathrm{CQ}}(S_i)$; freeze target, scales and aliases
return frozen configuration for acquisition i

We also include a separate Point endpoint and restrict interior upgrades to $\tau \geq 0.01$. For acquisition q, g_{qk} is the Detached candidate gradient and g_q^{ref} is the horizon-matched AF-Full-2000 Point gradient. Only batch-trainable D/N/R leaves are included. B_q normalizes each block by its reference-gradient norm. On a calibration set S,

$$s_k(S) = \max\left(0, \frac{\sum_{q\in S}\langle B_q g_{qk}, B_q g_q^{\mathrm{ref}}\rangle}{\sum_{q\in S}\|B_q g_{qk}\|^2 + 10^{-30}}\right), \tag{55}$$

$$E_q(k; S) = \frac{\|B_q(s_k(S) g_{qk} - g_q^{\mathrm{ref}})\|}{\|B_q g_q^{\mathrm{ref}}\| + 10^{-30}}. \tag{56}$$

Each outer calibration set internally leaves one acquisition out for scale fitting and error comparison. Emitted rollout scales must be finite and strictly positive; otherwise technical validation fails. The decision uses 20,000 deterministic paired bootstrap resamples, nominal one-sided 95% gates, and zero endpoint/interior margins. Finite-candidate risk ties use smaller τ, then candidate identifier.

Algorithm 1 separates a prespecified endpoint comparison from an interior-candidate search followed by a gate on the same errors. Both use nominal bootstrap decision thresholds. Formal coverage, including coverage after candidate selection, has not been established; the gates define the frozen decision rule.

D.1.2 Condition B: calibration, rollout, and outcomes

Condition B and its relation to the systematic study. Condition B is a separate frozen-selection application on channel 2 at -5 dBm, with six source–repetition units. It uses its own leave-one-acquisition-out calibration and rollout protocol and is not part of the systematic receiver panel in Section 5. Accordingly, these six units do not contribute to the 20-workpoint coverage, the 40 paired D+N+R comparisons, or the reported $+0.04508$-dB mean CQ-minus-Point gain in the systematic experiment.

"Previously unopened" refers to access under the study protocol to these previously recorded data. Before Condition-B waveform access, a B-shaped preflight replays the exact 5-of-6 procedure within each development condition. The pooled replay selects $\tau = 0.3$ at 500 symbols and CQ at 2,000 and 4,000; at 500, separate conditions select $\tau = 0.3$ on channel 4 and CQ on channel 5. This development replay uses its own protocol, separately from the systematic 162-task diagnostic.

For each Condition-B unit, its five siblings provide probes for target and scale calibration, and all six configurations are frozen before the 200-update Detached rollouts. Each acquisition is excluded from its own calibration, then used to train its assigned branch on a separate interval. The primary result is the natural-log NMSE ratio after update 200. Every unit remains in the analysis regardless of lock. The six calibration sets overlap. For each unit, evaluation computes

$$\mathrm{NMSE}_k = \frac{\mathrm{mean}\,|\hat{z}_k - z^\star|^2}{\mathrm{mean}\,|z^\star|^2 + 10^{-12}}, \qquad d = \log \frac{\mathrm{NMSE}_{\mathrm{selected}}}{\mathrm{NMSE}_{\mathrm{scaled\ Point}}}.$$

The means pool both polarizations over the complete scored interval in Table 17, after time alignment and pilot exclusion; no fitted complex amplitude/phase alignment is applied. Both NMSE and Q/BER use all $N_{\mathrm{NMSE}} = N_Q = 200{,}000$ symbols per polarization, including the 3,392-symbol tail beyond 24 complete 8,192-symbol windows. The complete-window count is a coverage requirement for lock, not a truncation of the metric support. Negative d favors the selected rule, and its unit-level fractional reduction is $1 - e^d$. The frozen rollout scale multiplies the scalar loss before differentiation and Adam; it does not multiply the Adam parameter update or learning rate. Adam normalizes its first and second moments, so a fixed positive loss scale does not in general scale the update. This control does not tune the learning rate or update magnitude.

Within-acquisition data flow. The six frozen branch configurations specify identical zero-based, half-open transmitted-symbol coordinates (Table 17). Probe, training, and evaluation raw supports are disjoint. A held-out unit is excluded from its own target/scale calibration but is used to train its assigned branch on a separate interval. Probe state is not transferred to rollout. Rollout uses a fresh initialization with zero optimizer state and its own 12-block warm-up; evaluation reconstructs fresh adaptive state and processes the pilot-to-decision-directed transition without resetting it at artificial chunks. The archived test routine transfers only the relevant trained power-normalization state. These details distinguish calibration exclusion from absence of subsequent training.

Table 17: Condition-B data flow. Aligned intervals are in transmitted-symbol coordinates; endpoints are excluded. Raw received-sample ranges are twice the corresponding sent-symbol supports.

Stage	Prespecified interval and state handling
Probe	Warm-up begins at batch 148; gradient targets $[80{,}247, 82{,}247)$. Total raw sent support $[74{,}000, 82{,}495)$.
Training	Warm-up begins at batch 240; 200 updates of 2,000 symbols. Aligned loss targets $[126{,}247, 526{,}247)$; warm-up/training raw support $[120{,}000, 526{,}495)$.
Evaluation	Fresh state; aligned pilots $[1{,}000{,}742, 1{,}200{,}742)$, then score $[1{,}200{,}742, 1{,}400{,}742)$. Raw sent support $[1{,}000{,}495, 1{,}400{,}990)$.
Scoring	NMSE and Q/BER use all 200,000 symbols per polarization, including the 3,392-symbol tail.
Lock coverage	24 complete windows of 8,192 symbols. Lock requires finite metrics, BER at most 0.02, Q at least 8 dB, and at least 20 complete windows.

Table 18: **Condition-B outcomes from the separate frozen-selection study.** This study is not included in the systematic receiver coverage or aggregate gains reported in Section 5. All folds select ridge-stabilized CQ before rollout. The table reports outcomes after 200 updates relative to gradient-calibrated, loss-scaled Point. All six prespecified units are retained regardless of receiver lock status.

Unit	Log NMSE ratio	Reduction	ΔQ (dB)	Point Q (dB)	Locked Point/selected
s1r1	-0.02009	1.989%	$+0.09466$	6.316	no/no
s1r2	-0.01888	1.871%	$+0.07965$	6.330	no/no
s1r3	-0.00455	0.454%	-0.00984	-14.092	no/no
s2r1	-0.01356	1.346%	$+0.06209$	6.388	no/no
s2r2	-0.01569	1.557%	$+0.07895$	6.362	no/no
s2r3	-0.01721	1.707%	$+0.06980$	6.372	no/no

The primary Condition-B outcome is the log-NMSE ratio at update 200. All six prespecified units improve on this metric (mean -0.014997). Receiver Q and lock status are reported alongside the primary outcome.

The s1r3 unit exhibits an extreme low-Q condition (Point $Q = -14.092$ dB) and fails the prespecified receiver-lock criterion for both Point and the selected method. We retain this unit according to the prespecified inclusion rule rather than excluding it post hoc. Its Q value is therefore visible in Table 18, but Condition B is analyzed separately and does not contribute to the systematic receiver aggregates in Section 5, including the 20-workpoint, 40-pair, and $+0.04508$-dB results.

Q improves in five of the six Condition-B units. Neither method meets the prespecified receiver-lock rule in any unit. The mean log-NMSE ratios at updates 20 and 100 are -0.002974 and -0.003525, respectively; one unit is slightly worse at update 20 (maximum adverse ratio $+0.000796$) before improving. Selected and CQ coincide because every fold retains that endpoint; this execution evaluates the selected endpoint under the specified within-condition protocol.

D.2 Failed or dominated selection alternatives

A preceding full-tree Euclidean search over a one-dimensional Point–Quotient mixture had a raw minimum near Quotient, but common gradient components diluted the contrast and a Pointward one-standard-error rule selected Point. A fixed-reference tangent family improved over Point in its own model class but was dominated by output-dependent ridge-stabilized CQ. These recorded outcomes motivate the endpoint-first, output-dependent family.

D.3 Earlier optical protocols and retained comparisons

The earlier single-anchor studies use protocols distinct from the systematic panel. The confirmation study uses five training-stream offsets within one capture and a shared physical initialization. D/N/R used the shared physical initial values; offsets at 0, 20,000, 40,000, 60,000, and 80,000 symbols produced one-million-symbol training intervals with 92–98% overlap. SRC1/REP2 validation selected checkpoints; SRC2/REP1–3 supplied testing. Their shared-capture, validation-selected outcomes are reported separately from the systematic fixed-endpoint panel.

D.3.1 Validation-selected multi-target confirmation

The historical table compared seven targets, unlike the new Point/CQ coverage. Its values are retained below. "Outage" denotes the fraction of scored windows below a frozen Q threshold; receiver lock is assessed separately. The threshold was the Point validation-window 10th percentile, 9.026291 dB. The original window protocol and validation-selected checkpoint differ from the new fixed-one-million-symbol protocol. Here Complex Partial is $0.6L_P + 0.4L_{\mathrm{CQ}}$, and Complex Quotient uses L_{CQ} alone. Scalar Partial and Scalar Quotient instead use $s_k L_P$, where $s_k = (\sum_{b=1}^{8} \|g_{k,b}\|^2 / \max(\sum_{b=1}^{8} \|g_{P,b}\|^2, 10^{-30}))^{1/2}$ is calibrated at the common initialization on eight successive training microbatches, with the adaptive state advancing, then frozen per stream.

For the Checker control, put $v_{t,p} = (-1)^t z^\star_{t,p}$ and $b^\star = v^H(z^\star - \hat{z})/(v^H v + 10^{-12})$. Its loss is the Point NMSE of $\hat{z} + b^\star v$, with $b^\star$ stopped in differentiation. Checker Partial mixes this loss with L_P using weights 0.4 and 0.6; Checker Quotient uses it alone. This relaxes the two-real-dimensional alternating reference direction rather than the shared complex-gain direction. These definitions refer to the historical study only.

Table 19: Historical validation-selected confirmation, not new fixed-endpoint coverage. Five paired, overlapping training-stream offsets each use three test captures.

Target	Mean Q (dB)	Q10 (dB)	CVaR10 (dB)	Window outage (%)
Point	9.0917	8.9998	8.9679	20.00
Scalar Partial	9.0926	8.9930	8.9635	20.00
Checker Partial	9.0910	8.9892	8.9573	21.11
Complex Partial	9.1053	9.0099	8.9785	12.78
Scalar Quotient	9.0929	8.9980	8.9699	18.33
Checker Quotient	9.0917	8.9892	8.9649	18.89
Complex Quotient	9.1169	9.0180	8.9983	12.78

D.3.2 Earlier horizon and closure diagnostics

The old horizon diagnostic used one checkpoint/capture, BPN test-mode normalization, all-pilot final equalization, and 128-symbol complex-alignment blocks. Its reference varied final-equalizer differentiation at 128, 512, 2,048, and 4,096 symbols. CQ wins 5/6 probe windows at 2,048 and 6/6 at 4,096. The systematic diagnostic instead retains all specified adaptive paths, uses training-mode normalization, and aligns every 500 symbols.

The original matched-cadence, correlated-stream closure study reported a mean-Q interaction of about $+0.00345$ dB, positive in four of five streams. The anchor result complements the 54 systematic comparisons reported in Section 5.3. A late-stage experiment, switching targets after 1,200 Point updates for 200 further updates, had a positive local gradient interaction but a negative final mean-Q interaction of -0.00639 dB. This target-switching result concerns a different optimization trajectory from the systematic fixed-target comparison.

D.4 Implementation-specific computational cost

The historical 4,096-symbol diagnostic on a T4 measured 212.61 s for Full and 89.23 s for Detached (2.38 times); TBPTT-4 reached cosine 0.972 to Full but took 214.48 s. The JIT-optimized segment worker measured 0.294 s/update for AF-Full-2000 versus 0.100 s for Detached (about 2.9 times). Condition-B probes consumed 575.5 s versus 3,563.2 s total rollout-worker time, a 16.2% ratio with that denominator. These ratios describe the stated implementations and timing denominators; the systematic-grid training and target-calibration costs require their own accounting.

E Reproducibility and evidence map

E.1 Experiment matrix

Table 21 separates the numerical objects compared in each study. In particular, a gradient probe at a saved Point checkpoint, a fully trained segment branch, and a cross-fitted selection rollout are different protocols. The matrix records the data, derivatives, and outcome measured by each protocol.

Table 20: **How the theory is tested.** Known tracker dynamics permit analytic matching or a certified PSD approximation; the receiver tests fixed structured candidates against a measured gradient reference.

Ingredient	Controlled systems	Optical receiver
Forward adaptation	Constructed tracker, known gains	Recorded signals, adaptive DSP
Backward derivatives	Same recursion, specified cuts	Specified state paths and horizons
Geometry	Analytic Partial; subspace controls; coupled PSD approximation; prospective construction	Fixed Point, Partial, and CQ losses
Question	How do matching geometry and learning change?	How does the dependence vary?

Table 21: Experiment matrix. D denotes Detached; state-path retention is always limited to the stated horizon or segment.

Study	Learned component	Local comparison	Derivative setting	Evaluation object
Closure control	Affine predictor on Push summaries	Point / fixed Partial; extra targets retained	D / TBPTT-4 / TBPTT-16 / Full	64-step cumulative task, fixed 400-update endpoint
Direction control	Two-head sequence frontend	Scalar and subspace variants	D in plotted training arms	128-step tracker task, validation-selected model
Coupled directions	Error interface; no trained model	Point / best scalar / diagonal / PSD	D; Full analytic check	2D and 4D operator error; 2D equal-displacement task decrease
Prospective construction	Nonlinear predictor, known tracker	Fixed candidates / constructed metrics	D / Full; two optimizers	Frozen metrics, new-task test loss after 200 updates
Optical diagnostic	R, N+R, D+N+R at Point checkpoints	Point / Partial / CQ	D candidates; same-horizon AF-Full reference	Gradient error, $H = 500, 2000, 4000$
Optical coverage	Five configurations	Point / CQ	D	Fixed 1M-symbol endpoint; 500-symbol updates
Optical closure	R, N+R, D+N+R	Point / CQ	D / AF-Full within segment	Fixed 1M-symbol endpoint; 2,000-symbol updates
Local descent replay	Saved D+N+R checkpoints	Point / Partial / CQ	D candidates; AF-Full reference	Same-context finite steps; CPU precision control
Selection application	Original Condition-B frontend	Frozen selected / loss-scaled Point / CQ	2,000-symbol reference; D rollout	Held-out calibration identity, then 200-update rollout

E.2 CLAIM-TO-ARTIFACT MAP

Table 22 links each claim to a numbered supplementary artifact. The plotting and aggregation scripts consume saved values and do not invoke a training notebook.

Table 22: Main claims and the saved records used to check them.

Claim or exhibit	Artifact	Reproduction
P1 reversal; Fig. 2	**A1** — Closure intervention results	Aggregation + Fig. 2 source
P2 direction control; Fig. 3	**A2** — Original direction-control training; **A3** — Corrected oracle diagnostic	Aggregation + Fig. 3 source
Gradient matching; Fig. 4(a–c)	**A4** — Optical gradient scores	Aggregation + Fig. 4 source
Training coverage; Table 12, Fig. 10	**A5** — Production coverage	Coverage aggregation
Closure interaction; Fig. 4(d–f)	**A6** — Receiver closure intervention	Aggregation + Fig. 4 source
Matching/endpoint join	**A7** — Diagnostic/training join	Aggregation check
Tracker coefficients	**A8** — Tracker implementation record	Eq. (30) check
Local descent; Fig. 11	**A9** — Receiver finite-step replay	New replay + precision control
Coupled scope; Fig. 6	**A10** — Coupled-direction matching	Matrix construction, certified solve + local steps
Prospective construction; Fig. 8	**A11** — Known-dynamics supervision construction	Frozen geometry, all training endpoints + audit
Post-primary controls	**A12** — Task-only and same-objective operator scalar	Saved inputs, new endpoints + backend verification

Exact relative paths, filenames, hashes, and reproduction commands for artifacts A1–A12 are indexed in the supplementary artifact manifest. The included A12 directory, `supplementary/post_primary_controls/`, has its own file inventory, dependencies, and reproduction commands. Original numerical records are preserved; anonymous packaging changes are documented separately.

E.3 TASK LEDGER, BINARIES, AND VALIDATION LEVELS

The task ledger retains all 916 original task identities, including 300 data-blocked tasks. The completion summary distinguishes 604 new training outcomes from 12 verified reused outcomes. The diagnostic ledger records 162 tasks, each covering nine window/horizon contexts. These ledgers preserve execution identity; they are not additional performance samples.

The manuscript bundle includes the controlled source CSVs, all 4,374 optical candidate-score rows, the paired training/interaction tables, and the scripts that prepare the figures. It also includes the prior evidence index and integrity reports. Full receiver checkpoints, received waveforms, scored output arrays, and the 1,458 raw gradient-array groups are separately archived. The large-evidence index records paths and hashes for separately archived binaries; these records do not imply that the corresponding binary data are included in the manuscript bundle. The analysis tables can be reaggregated here. Regenerating gradients or Q from receiver outputs requires the corresponding experimental binary/source archive.

The validation levels are kept separate: (i) a saved score can agree with its commit record; (ii) aggregation can reproduce a table or figure; (iii) saved vectors or outputs can regenerate scores; and (iv) the receiver can be rerun from waveforms. Artifacts A1–A8 perform level (ii) and source/data identity checks in this bundle; prior records state their own checked level. A9 additionally supplies the

six saved input segments, parameters and states, recomputes gradients, and runs new perturbed forward passes. Its CPU cross-backend discrepancies and post-primary precision control are reported explicitly in Appendix C.8. This limited replay does not regenerate the full systematic receiver experiments. A10 is fully runnable from its matrix-construction protocol without external data and includes an independent least-squares solve and finite-difference checks. A11 supplies all data and code for the synthetic construction study, including frozen choices before evaluation, independent endpoint checks, and the descriptive same-task candidate oracle. It requires no external data. Compiling the paper establishes none of the unperformed numerical steps.

A12 contains five task-only runs and five CPU replays of the original Quotient for backend verification, plus 768 operator-scalar endpoints. The normalized P2 inputs and original initial states are included, so these controls can be rerun without an external data download. Original outcomes were known when the controls were proposed; their post-primary status is separate from the original prospective protocol.

E.4 Diagnostic provenance and derivative scope

Controlled diagnostic correction. Figure 3(a) uses the corrected pure-local oracle diagnostics in artifact A3. An earlier wrong-subspace diagnostic degenerated to Point; the corrected diagnostic supplies the reported comparison. The training records in A2 are unchanged. Separate source-extracted unit checks verify both local-loss null spaces and the actual Wrong Partial training gradient; they do not execute the historical diagnostic wrapper or retrain models. The manifest supplies the check script and its results. Table 22 links both artifacts to the reported direction control.

Receiver consistency and metric scale. Figure 10 reaggregates all 40 GDBP pairs from the production coverage artifact A5 without resampling. The supplementary manifest identifies the source records and figure-generation procedure. The illustrative conversions in Table 10 use Eq. (49), not measured BER records; their numerical checks accompany A5.

Receiver derivative scope. The retained state paths and the differentiation horizon are specified separately: the same AF-Full scope is evaluated at $H = 500$, 2000, and 4000. Appendix C.5 states the differentiated paths and actual horizon; the supplementary artifact documentation records the corresponding execution identifier.

E.5 CLAIM SCOPE AND INTERPRETATION

Established quantity	Interpretation retained in the manuscript
Directional matching formula	Exact interface-gradient relation under the stated fixed-linearization/shared-basis assumptions; active directions are explicit.
Coupled PSD construction	Under Detached closure ($C = I$) with $R = R^\top \succ 0$ and $A \succeq 0$: exact feasibility criterion and, for isotropic errors and zero ridge, a closed-form PSD optimum; 2D/4D approximation and 2D local descent checks.
Prospective construction	Frozen supervision from known 2D operators; new synthetic task draws, two optimizers, and comparison with development-selected candidate lookup. A two-weight family captures most of the general PSD improvement.
Fixed-pair preference reversal	Average effect across five constructed worlds from one measured trajectory pool; paired points are all shown.
Directional training control	Calibrated scalar and correct/wrong Partial branches; post-primary task-only comparison. Pure wrong-Quotient is diagnostic only.
Gradient-matching preference	Ordering of predefined candidate gradients relative to each horizon's reference in the actual active parameter space.
Local task descent	Six correlated contexts at one workpoint; matched-step controls expose precision effects and finite-step reversals under both precisions, without attributing training endpoints to precision.
Training effect	Fixed-budget, within-model target contrasts; cross-model gain differences describe complete trained configurations.
Physical coverage	Four complete received channels from one seven-channel platform; nested captures, checkpoints, and windows remain correlated.
Selection rollout	A separate restricted-family procedure with overlapping calibration folds; Selected and CQ coincide, and neither restores the prescribed lock condition.